The University Of Sheffield.

Regulation of JAK/STAT Signalling by Endocytosis

By:

Rachel Moore

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Science
Department of Biomedical Sciences

September 2018

Acknowledgements

I would like to thank my supervisor, Professor Elizabeth Smythe, for securing funding for me to carry out a PhD in her lab, on a project that has fascinated and challenged me, and enabled me to develop skills in a vast range of experimental techniques. I am particularly grateful for her belief in my ability to grow as an independent researcher, and for her continued guidance and optimism in my work. I would also like to thank the rest of the Smythe lab, past and present. Your scientific input and technical knowledge has been vital throughout my PhD, and I have thoroughly enjoyed our Christmas lab walks. A special shout out to Patrick, Paul and Hannes for your friendship and for keeping me topped up on good Yorkshire cuppas, pies and pints.

I am also very grateful for fruitful discussions with Dr Martin Zeidler and Dr Elizabeth Seward, and for their encouragement throughout the duration of my PhD. I would also like to extend my gratitude to the staff at the Sheffield Microarray and Next Generation Sequencing Core Facility, and the Sheffield biOMICs Facility. An extra special thank you to Adelina McCosta-Martin for her patience and help with mass spectrometry, but also for being such an amazing advocate for women in science, a great friend and a delight to work alongside.

A huge thanks to all my friends in BMS; from the many Friday nights at the pub, to the Firth Court Mighty Ducks football team, I have many great memories with you all and have been blessed to make some amazing friends during my time in Sheffield. An extra special thanks to Becca and Cat for our date nights, and to Sarah, Kat and Harriet for many summery pints. I am also particularly grateful to my housemates, Ben and Alex, for their great friendship and support, and for listening to me practice many presentations in front of the TV!

I am also extremely appreciative of all my friends dotted around the country, for giving me plenty of opportunities to relax and escape from my PhD. Thanks also to my YCC and St. Thomas Crookes families for their continued love, support and prayers! I am also very appreciative of my new work colleagues at AstraZeneca, for being so welcoming in my first months and believing in my ability to thrive in a new area of research.
An enormous thanks to my boyfriend James for his encouragement and support throughout my PhD, for travelling miles to visit, and for understanding that when I ‘nip to the lab’ at the weekend I probably won’t return for a few hours. Thanks also for the great times, the amazing holidays and breaks away from science that has given my brain time to recharge!

Finally, I am incredibly grateful for my wonderful family, I love you all lots! Thanks to Leanne for the great times we had living together this last year, especially the times that involved takeaways or extending our gin collection! I am super proud of the hard work you put into your lesson plans and how you are blossoming as a teacher.
Words really can’t express how thankful I am to Mum and Dad for their unfailing love, encouragement and support throughout every stage of my life and education. I am especially appreciative of the past months during the final stages of my thesis; for keeping me very well fed and for carting me across the country for job interviews. Thanks for making what should have been one of the most stressful stages of my PhD relatively relaxing and a lot of fun!

Abstract

The JAK/STAT pathway is a highly evolutionarily conserved signal transduction pathway, whose activation can lead to a broad range of cellular outcomes. The pathway is used repeatedly during multiple developmental stages and in adult tissue, and therefore tight regulation is required to enable accurate responses in a context specific manner. Internalisation and endocytic trafficking of signalling components provides a mechanism whereby spatial compartmentalisation can enable distinct signalling outputs.

Within this study I have investigated the role of endocytosis in the regulation of the *Drosophila melanogaster* JAK/STAT pathway, and demonstrated that internalisation and endocytic trafficking differentially regulates target genes. Although the JAK/STAT pathway is transcriptionally competent and can regulate the expression of particular targets when the activated receptor is at the cell surface, receptor endocytosis and localisation to distinct endosomes is required for the expression of other targets. This appears to be context-dependent, as high levels of ligand stimulation overcomes endocytic regulation. STAT92E, the *Drosophila* JAK/STAT transcription factor, is a target of endocytic regulation. Although it is efficiently activated and undergoes nuclear translocation when endocytosis is perturbed, it is not capable of regulating a subset of target genes and therefore further STAT92E interacting partners and/or post translational modification must be required to fine-tune its transcriptional competency during endocytic trafficking. Utilising mass spectrometry I identified a novel STAT92E phosphorylation site, at threonine 702. Mutation of this threonine to prevent its phosphorylation, resulted in inhibition of STAT92E signalling and nuclear translocation, and also prevented phosphorylation of a highly conserved tyrosine residue at position 704, which is crucial for ligand activated JAK/STAT signalling outputs.

Therefore, this study has enhanced our understanding of mechanisms that can modulate JAK/STAT activity. I have revealed an important role for endocytosis in fine-tuning *Drosophila* JAK/STAT signalling outputs and also identified a novel phosphorylation site which is crucial in the activity of STAT92E.

Table of Contents

List of Figures

List of tables

Abbreviations

ARH	autosomal recessive hypercholesterolemia
AP2	adaptor protein 2
bp	base pair
CALM	clathrin assembly lymphoid myeloid leukaemia
CCP	clathrin coated pit
CCV	clathrin coated vesicle
cDNA	complementary Deoxyribonucleic acid
CBM	cytokine binding modules
CIP	calf intestine phosphatase
CM	conditioned medium
CME	clathrin mediated endocytosis
c-Met	mesenchymal-epithelial transition factor
CIE	clathrin independent endocytosis
CRISPR	*c*lustered regularly interspaced short palindromic repeat
DFz2	*Drosophila* Frizzled-2
D. melanogaster	*Drosophila melanogaster*
DNA	Deoxyribonucleic acid
dsRNA	double stranded ribonucleic acid
E. coli	*Escherichia coli*
EE	early endosome
EEA1	early endosome antigen 1
ELISA	enzyme linked immunosorbent assay
EDTA	ethylene diamine tetraacetic acid
EGFR	epidermal growth factor (receptor)
EGTA	ethylene glycol tetraacetic acid
EpoR	erythropoietin receptor
ESCRT	endosomal sorting complexes required for transport
et al.	et alia (latin for: and others)
g	gravity
GAP	GTPase-activating proteins
GEEC	GPI-AP-enriched early endosomal compartments
GFP	green fluoresent protein
GPCR	G protein coupled receptor
GTP	guanysyltriphophate
GEF	GTP exchange factor
Hrs	HGF-regulated tyrosine kinase substrate
HPLC	high pressure liquid chromatography
ILV	intraluminal vesicles

IFN	interferon
IFNAR1	interferon alpha receptor 1
IFNGR1	interferon gamma receptor 1
IP	immunopreciptiation
LE	late endosome
LB	Luria-Bertain
mAb	monoclonal antibody
MHC	major histocompatibility complex
MAPK	mitogen-activated protein kinase
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MVB	multivesicular bodies
NLS	nuclear localisation signal
O/N	over night
OM	oncostatin M
OD	optical density at x nm
PAGE	polyacryl amin gel electrophoresis
PBS	phosphorsalinebuffer
PCR	polymerase chain reaction
PI3P	phosphatidylinositol 3-phosphate
PI(4,5)P2	phosphatidylinositol 4,5-bisphosphate
PIP	phosphatidylinositol
PLC	phospholipase C
PTM	post translational modification
qPCR	quantitative polymerase chain reaction
RKT	receptor tyrosine kinase
RNA	Ribonucleic acid
rpm	rounds per minute
RT	room temperature
SARA	Smad anchor for receptor activation
SDS	Sodium Dodecyl Sulfate
SNARE	soluble NSF attachment protein receptor
SUMO	small ubiquitin-like modifier
TEMED	N-N-N'-N'' tetra methyl ethylene diamin
Tris	Trishydroxymethylaminoethan
V	volts
Wdp	Windpipe
Wg	Wingless
WT	wild type

Chapter 1. Introduction

Cell-surface receptors and their downstream signalling pathways play a variety of roles in cellular communication during embryonic development and adult homeostasis. The same receptors are used repeatedly in different tissue environments to produce distinct transcriptional responses and phenotypic outputs. Therefore, tight regulation of the signalling from transmembrane receptors is vital in enabling cells to produce specific intracellular responses and precise cellular outcomes in a context-dependent manner. The dysregulation and inappropriate activation of signalling pathways is associated with multiple diseases, facilitating cancer cell survival and tumour growth. Therefore, understanding the mechanisms that accurately control appropriate signalling, and how their dysregulation influences cell fate and disease, will enhance the development of therapeutics to effectively target signalling pathways. One mechanism by which precise control can occur is through endocytic regulation. Endocytosis was classically viewed as a process distinct from signalling, but it has become evident that endocytic uptake and trafficking of transmembrane receptors is tightly linked to signalling. Compartmentalisation during endocytosis provides a mechanism to spatially restrict molecules and facilitate protein interactions, enabling quantitative and qualitative regulation of signalling outputs. This project aims to expand this field of research by investigating the mechanism by which endocytosis regulates the transcriptional outcome of the JAK/STAT pathway in *Drosophila*.

1.1 The JAK/STAT signalling pathway

The Janus Kinase/Signal Transducer and activator of transcription (JAK/STAT) pathway is an example of a pleotropic signalling pathway which is employed during a variety of cellular events, from proliferation to apoptosis, throughout development and in immune responses. The pathway is ubiquitous and highly evolutionarily conserved from slime molds to humans. Due to the array of potential outcomes it is important for the pathway to be tightly regulated so that an accurate response occurs in the correct context. Dysregulation of the JAK/STAT pathway is involved in the pathogenesis of diseases such as gigantism, asthma, myocardial hypertrophy, leukaemia and severe combined immunodeficiency (SCID) (Devergne et al., 2007; Vidal et al., 2010). Identifying the details of this pathway and its regulatory mechanisms is therefore key

to our understanding of disease and will ultimately aid the development of new therapeutics.

The discovery of the JAK/STAT pathway began in 1957 by Isaacs and Lindenmann, who demonstrated that, upon infection with influenza virus, chick embryo cells secreted interferons (IFNs) which caused uninfected cells to become resistant to the infection. Subsequent studies linked IFNs to the rapid expression of specific genes, and during the 1980s-1990s key pathway components that enabled this signal transduction and transcriptional regulation were identified (Stark and Darnell, 2012). Today, the canonical model of the JAK/STAT signalling pathway (Figure 1.1) involves the activation of a cell-surface transmembrane receptor, by a cytokine, growth factor or hormone, which causes a conformation change in the cytoplasmic tail of the receptor. This stimulates activation of Janus kinases (JAKs) that are constitutively associated with the intracellular portion of the receptor. JAK activation leads to receptor phosphorylation at a specific tyrosine residue, subsequently allowing recruitment of signal transducer and activator of transcription (STAT) transcription factors through their src-homolgy 2 (SH2) domains. This association in turn allows JAK to phosphorylate STATs at a highly conserved tyrosine residue, causing their dimerization and translocation to the nucleus. Here they bind to specific DNA sequences to alter expression of target genes, resulting in developmental, haematological and immune-related response pathways (Stark and Darnell, 2012; O'Shea et al., 2015).

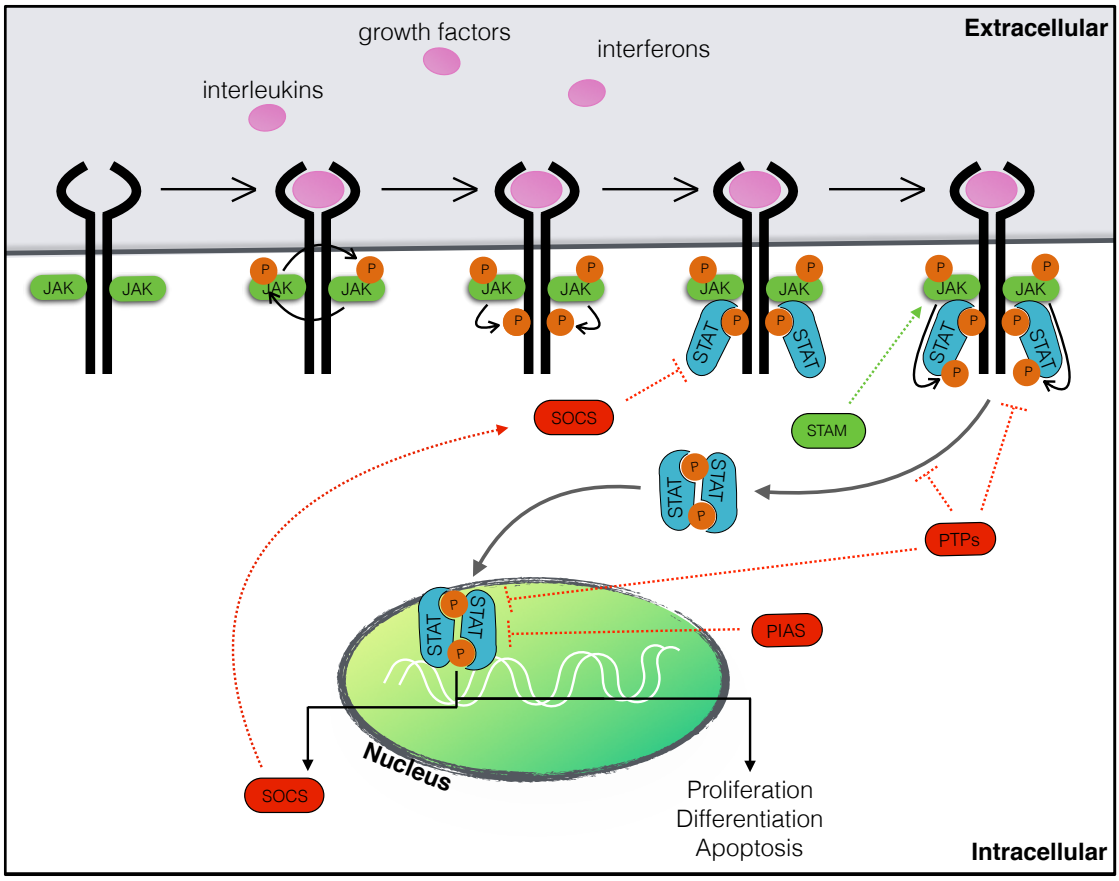


Figure 1.1: Schematic of JAK/STAT activation, signalling and regulators. After ligand binding JAKs transphosphorylate one another and the receptor tail. The transcription factor STAT is then recruited to the receptor resulting in JAK-mediated STAT phosphorylation. Phosphorylated STAT dimerises and translocates to the nucleus altering transcription. Regulators of JAK/STAT signalling include SOCS, PIAS and PTPs which negatively regulate the pathway, whereas STAM positively regulates the pathway.

1.1.1. The mammalian JAK/STAT signalling pathway

In mammals, a diverse range of transmembrane receptors activate the JAK/STAT pathway, including cytokine and growth factor receptors. The majority of ligands that activate the JAK/STAT pathway, such as interferons (IFNs) and interleukins, contain alpha-helices within their structure and interact with cytokine binding modules (CBM) on the extracellular portion of the cognate receptor (Mohr et al., 2012). Cytokine receptors are the major activators of JAK/STAT signalling, and are divided into two subtypes, class I and class II. CBMs of class I cytokine receptors contain two fibronectin type-III domains, four conserved cysteines and a consensus WSXWS motif (Bazan, 1990). The WSXWS plays roles in receptor dimerisation, uptake and trafficking (Hilton et al., 1996; Siupka et al., 2015), along with receptor activation (Dagil et al., 2012) and signal transduction (Chiba et al., 1992). Class 1 interact with IL-6 and

IL-12 type cytokines, whereas the class II receptors interact with interferons and IL-10 type cytokines (see Table 1.1) (Ferrao and Lupardus, 2017).

In the canonical model of JAK/STAT signalling, ligand binding results in receptor dimerisation. Although dimerisation of receptor subunits is required for JAK activity, numerous receptors pre-dimerise in the absence of ligand binding, but do not activate signalling (Mohr et al., 2012). Cytokine receptor subunits can form stable or transient homo or heterodimers, as well as higher order oligomers. The erythropoietin receptor (EpoR) homodimerises, whereas gp130 is a type 1 cytokine receptor that can heterodimerise with a range of other receptor subunits to elicit signalling from nine different ligands (Wang et al., 2009).

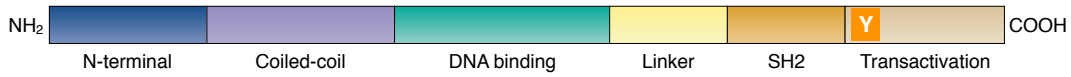
The precise mechanism by which the receptor-ligand interaction results in JAK activation remains unclear, but ligand binding results in rapid phosphorylation of JAKs causing an increase in their catalytic activity (Feng et al., 1997). JAKs can associate with the receptor through a proline rich “box 1” and/or through a hydrophobic “box 2” motif in the cytoplasmic portion of the receptor. The presence and location of these along the tail are suggested to provide specificity to JAK-receptor interactions (Usacheva et al., 2002). In mammals there are four JAK proteins, JAK1-3 and TYK2, that are ubiquitously expressed. They consist of a FERM domain and a SH2 domain, which are essential for receptor binding (Ferrao and Lupardus, 2017), a catalytically inactive pseudokinase domain and a kinase domain. Activation of JAKs causes them to transphosphorylate one another, and subsequently phosphorylate the receptor at a conserved tyrosine residue. Although ubiquitously expressed they have varying *in vivo* roles due to their specificity in the binding of receptors (Yamaoka et al., 2004). They contain the same receptor interacting modules, but the F2 subdomain of the FERM domain is subtlety different between the JAKs, which may provide binding specificity (Ferrao and Lupardus, 2017). Phosphorylation of the receptor tails facilitates interactions with STAT proteins, which in turn are phosphorylated by JAKs, causing STAT activation.

JAKs are commonly mutated in JAK/STAT related diseases, and are therefore a target of therapeutic intervention. Whilst SCID patients have JAK deficiencies, JAKs are commonly overactive in leukaemia. An acquired valine to phenylalanaine mutation at 617 within the pseudokinase domain of JAK2 has been identified at high rates in patients suffering from a range of myeloproliferative disorders (Mclornan et al., 2006). JAK2V617F causes constitutive activity of STAT3 and STAT5, resulting in the

upregulation of genes that provide a proliferation advantage to cancerous cells. In the last 13 years multiple JAK2 inhibitors have been developed, with ruxolitinib becoming increasingly used in the clinic (Gäbler et al., 2013; Bose and Verstovsek, 2017).

1.1.2. The STAT transcription factor

There are seven mammalian STAT proteins, STAT1, STAT2, STAT3, STAT4, STAT5a, STAT5b and STAT6, all encoded by different genes. STAT proteins homo- and heterodimerise in specific combinations. Interestingly, although they share similar structures and can be activated by the same ligands (Table 1.1), they have physiologically discrete roles (Ihle, 2001). All STATs contain a conserved tyrosine residue which is phosphorylated in response to ligand stimulation and is required for canonical 'parallel' dimerisation. STAT function is abolished by substitution of this specific tyrosine to phenylalanine in multiple species.



STAT1	$_{694}$DGPKGTG**Y**IKTELIS$_{708}$
STAT2	$_{683}$NLQERRK**Y**LKHRLIV$_{697}$
STAT3	$_{698}$DPGSAAP**Y**LKTKFICV$_{703}$
STAT4	$_{686}$TERGDKG**Y**VPSVFIP$_{700}$
STAT5a	$_{687}$LAKAVDG**Y**VKPQIKQ$_{701}$
STAT5b	$_{692}$TAKAADG**Y**VKPQIKQ$_{706}$
STAT6	$_{634}$MGKDGRG**Y**VPATIKM$_{648}$
STAT92E-C	$_{697}$RQDPVTG**Y**VKSTLHV$_{711}$
STAT92E-F	$_{704}$VLDPVTG**Y**VKSTLHV$_{718}$

Figure 1.2: Structure of STAT molecules and conservation of tyrosine across all STATs.

The crystal structure of various STAT monomers and dimers has provided insights into the roles of the modular domains present in all STAT proteins (Figure 1.2). Crystallisation of phosphorylated, DNA-bound STAT1 and STAT3 homodimers revealed that they form a clamp-like structure around the DNA double helix (Becker et al., 1998; Chen et al., 1998). The DNA binding domain is preceded by a coiled-coil domain that facilitates receptor binding and interactions with other proteins. Following the DNA binding domain is a linker region and an SH2 domain which is key for receptor association and binding the pTyr on the other STAT monomer (Lim and Cao, 2006).

Crystallisation of the N-terminal domain of STAT4 revealed that it can form dimers independently of other domains, suggesting a second interface for STAT dimerisation separate from tyrosine phosphorylation (Vinkemeier et al., 1998). The crystal structure of unphosphorylated STAT1 (Mao et al., 2005) and STAT5a (Neculai et al., 2005) confirmed that dimers can form between the N-terminal domains, which results in the formation of an 'antiparallel' dimer. Upon tyrosine phosphorylation STAT dimers are thought to undergo a large conformation change causing the N-terminal domains to separate (Neculai et al., 2005). However, the N-terminal domains also play key roles in tetramerization of STAT molecules on the DNA when there are multiple STAT consensus binding sites in promoters. Following the SH2 domain is a transactivation domain, which is the least conserved domain across the STAT family. This domain has been seen to undergo further post-translational modifications, which regulates protein-protein interactions and signalling outputs (Lim and Cao, 2006).

Despite their similar domain structure, STATs have distinct roles and transcriptional responses. STAT1 mediates anti-viral, anti-proliferative, pro-apoptotic and tumour suppressive responses in response to IFN-α/β and IFN-γ. Stimulation by IFN-α/β specifically causes STAT1 heterodimerisation with STAT2 and associates with IFN regulatory factor 9 (IRF9). This forms a different transcription factor complex to that of STAT1 homodimers, which is formed upon treatment with IFN-γ. Homodimers bind to gamma-activated sequence (GAS) elements, whereas the heterodimer complex binds to IFN-I-stimulated response element (ISRE) causing the regulation of different genes (Stark et al., 1998; Blaszczyk et al., 2016). Both hypo- and hyperactivation of STAT1 activity has been associated with cancers, such as colon cancer and leukaemia, respectively (Heppler and Frank, 2017). Defective STAT1 also leads to reduced immune responses, causing patients to have increased susceptibility to mycobacterial and viral diseases (Ihle, 2001). In contrast, STAT3 is activated by IL-6 family members and RTKs and mediates proliferation and survival. STAT3 is essential for early development, and plays crucial tissue specific roles such as wound healing in keratinocytes and liver regeneration (Lim and Cao, 2006). STAT5 activity is induced by IL-2/3/5/7/15, prolactin and growth hormone, and also mediates proliferation and survival. STAT3 and STAT5 have oncogenic roles, with persistent activity seen in both haematological malignancies and solid tumours (Heppler and Frank, 2017). STAT4 and STAT6 are predominantly activated by IL-12 and IL-4, respectively. They have specific physiological functions, regulating T helper cell differentiation (Ihle, 2001). STAT4 polymorphism is associated with immune diseases such as rheumatoid arthritis and systemic lupus erythematosus (Martínez et al., 2008). Due to vast number of roles

and transcriptional capabilities of the STAT proteins, tight and specific regulation is required.

1.1.2.1 Nuclear import

STATs are transcription factors and therefore require nuclear import to elicit their transcriptional functions. Hence, understanding the mechanism by which STATs enter the nucleus may provide a way to regulate STAT activity when it is dysregulated in disease. As with their function, there is variation in the mechanisms by which they are translocated to the nucleus.

Large proteins cannot enter the nucleus freely and so specific signal sequences are required to enable transit between cytoplasm and nucleus. Nuclear localisation signals (NLS) and nuclear export signals (NES) facilitate import and export, respectively, through nuclear pore complexes. Cargo harbouring an NLS is recognised by an importin-α/importin-β heterodimer which, through the hydrolysis of ras-related nuclear protein (Ran)GTP to RanGDP in the cytoplasm, enables translocation into the nucleus. In contrast, NES are recognised by the exportin CRM1. High levels of RanGTP in the nucleus results in dissociation of importins and cargo, enabling association with exportins (Meyer and Vinkemeier, 2004; L. Liu et al., 2005; Reich, 2013).

The canonical model for STAT activity is that tyrosine phosphorylation and parallel dimer formation is required for nuclear import. For STAT1 and STAT4, tyrosine phosphorylation does indeed have crucial roles in nuclear trafficking. Nuclear import of STAT1 occurs within minutes of IFN treatment, but this is transient and nuclear STAT1 levels are reduced after only a few hours. This transient cycle is dependent on the phosphorylation of the conserved tyrosine residue of STAT1. Phosphorylated parallel STAT1 dimers bind the importin-α5 adaptor, resulting in nuclear import. Upon DNA binding the associated importin-α5 is displaced (McBride et al., 2002). Dephosphorylation of STAT1 by the nuclear phosphatase TC45 (Hoeve et al., 2002) causes a change in the dimer conformation of STAT1, revealing an export sequence within STAT1's DNA binding domain. STAT1 then associates with an exportin receptor, causing its nuclear export (McBride et al., 2002). Although tyrosine phosphorylation is key in the cycling, unphosphorylated STAT1 is also seen in the nucleus (Reich, 2013).

In contrast, STAT2, STAT3, STAT5 and STAT6 undergo nuclear import that is independent of pathway activation and tyrosine phosphorylation. The association of STAT2 with IRF9 results in the constitutive shuttling of unphosphorylated STAT2 in and out of the nucleus (Banninger and Reich, 2004). STAT3/5/6 are all capable of nuclear translocation when unphosphorylated. This is interesting as roles of unphosphorylated STATs in DNA binding and gene expression has become evident (See Chapter 1.1.2.2). Both unphosphorylated and phosphorylated STAT3 are able to bind importin-α3 through a region within the coiled-coil domain (L. Liu et al., 2005), with phosphorylation increasing the rate of shuttling (Reich, 2013). STAT5 and STAT6 are also imported in an importin-α3 dependent manner, through association with distinct residues within their coiled-coil domains. Therefore STATs enter the nucleus through multiple, independent mechanisms that require different residues for association with importins.

1.1.2.2 Unphosphorylated STATs

Unphosphorylated STATs (U-STATs) were originally thought to be inactive and have no role in signalling. However, in the last ~20 years there has been increasing evidence to suggest important roles of U-STATs in regulating gene expression (Yang and Stark, 2008). Upon ligand induced formation of phosphorylated STAT3 dimers, the STAT3 gene itself is activated resulting in increased STAT3 expression, but also intriguingly in the concentration of U-STAT3. U-STAT3 can bind to DNA as both a monomer and a dimer (Timofeeva et al., 2012), and can also associate with unphosphorylated NF-Kβ to regulate a distinct subset of genes to those activated by phosphorylated STAT3 (Yang et al., 2007). This results in a prolonged response to cytokine activation that regulates differential gene expression. Similar results have been seen for STAT1, whereby the expression of U-STAT1 is increased upon ligand stimulation and regulates distinct processes, for example U-STAT1 is crucial in TNF-α induced apoptosis (Kumar et al., 1997; Chatterjee-Kishore et al., 2000). U-STAT6 has also been shown to regulate gene expression, but regulates the same genes as phosphorylated STAT6 therefore contributing to continuous expression of target genes (Cui et al., 2007). U-STATs have also been shown to be important in the stability of heterochromatin and hence are involved in the regulation of epigenetic gene silencing. Unphosphorylated STAT5a interacts with heterochromatin protein 1 (HP1), resulting in its stabilisation and promotes heterochromatin formation. Hu et al, 2013 demonstrated that this interaction inhibited colon cancer growth in xenograft mouse models and hence is important in tumour suppression (Hu et al., 2013).

STAT	Kinase	Ligand					
		Bind Type I cytokine receptor				Bind Type II cytokine receptor	
		IL-6 family	IL-12 Family	Common chain receptor family	Growth Hormones	IL-10 family	Interferons
STAT1	JAK1	IL-6, 11, 31, CNTF, CT-1, LIF, OSM, Leptin, G-CSF, CLF	IL-27	IL-2, 3, 4, 7, 9, 13, 15, 21, GM-CSF	GH, prolactin	IL-10, 22, 24, 26, 29	IFN-α/β, IFN-γ
	JAK2		IL-12, 23		Tpo, Epo	IL-28A	
STAT2	JAK2		IL-27	IL-15		IL-28A	IFN-α/β, IFN-γ
	JAK3			IL-7			
	Tyk2					IL-29	
STAT3	JAK1			IL-5		IL-19, 24	
	JAK2	IL-6, 11, 31, CNTF, CT-1, LIF, OSM, G-CSF, Leptin, CLF		IL-2, 3, 4, 13, GM-CSF	GH, Tpo, prolactin	IL-10, 20, 28A	
	JAK3			IL-7, 9, 15, 21	Epo		
	Tyk2		IL-12, 23, 27			IL-22, IL-29	IFN-α/β, IFN-γ
STAT4	JAK3			IL-2, 4, 15, 21			
	Tyk2		IL-12, 23, 27			IL-29	IFN-α/β
STAT5	JAK2	Leptin, IL-31		IL-5, GM-CSF	GH, prolactin	IL-20	
	JAK3			IL-2, 7, 13, 21	Epo		
	Tyk2	CT-1, LIF, OSM, G-CSF	IL-12, 23, 27	IL-3, 4, 9, 15	Tpo	IL-10, 22, 29	IFN-α/β, IFN-γ
STAT6	JAK2	Leptin		GM-CSF			
	JAK3			IL-2, 7,			
	Tyk2	OSM	IL-12, 27	IL-3, 4, 13, 15,			IFN-α/β, IFN-γ

Table 1.1: Ligands and their associated JAKs. which activate each STAT transcription factor. This table only shows STATs activated downstream of cytokine receptors and does not include growth factor receptor activation of the JAK/STAT pathway.

1.1.3. Regulation of the JAK/STAT pathway

Due to the importance of the JAK/STAT signalling pathway in a broad range of cellular outcomes, many different mechanisms are employed to tightly control signalling specificity, kinetics and intensity.

1.1.3.1 Protein regulators

Multiple molecular regulators of the JAK/STAT pathway have been identified to influence signalling in a quantitative manner (Figure 1.1). The Suppressor of Cytokine Signalling (SOCS) family of proteins harbour a C-terminal SOCS box and an SH2 domain, allowing binding to phosphorylated tyrosine residues. The SOCS box enables interactions with proteins, such as elongin and cullins, forming a ubiquitin ligase complex that labels proteins for proteasomal degradation (Jason S. Rawlings et al., 2004; Tamiya et al., 2011). Eight mammalian SOCS proteins have been identified and inhibit both the JAK/STAT and epidermal growth factor receptor (EGFR) pathways, play roles in development and are associated with various diseases, such as leukaemia and tuberculosis. SOCS1-3 primarily regulate JAK/STAT signalling from cytokines, whereas SOCS4-7 generally regulate signalling from RTKs, such as EGFR signalling (Trengove and Ward, 2013). They inhibit the JAK/STAT pathway using a variety of mechanisms which involve the specificity of the SH2 domain. The SH2 domain of SOCS1 facilitates binds to phosphorylated JAKs, whereas SOCS2 and SOCS3 bind to phosphotyrosines on activated receptors, preventing STAT recruitment. SOCS1 and SOCS3 can also bind through a pseudo-substrate domain which interacts with the JAK and blocking their kinase activity (Stec and Zeidler, 2011; Tamiya et al., 2011). SOCS proteins are also direct targets of the JAK/STAT pathway and their expression is increased rapidly upon STAT activation. Therefore, a negative-feedback system is formed, whereby activation of the JAK/STAT pathway increases SOCS expression which subsequently decreases JAK/STAT signalling (Krebs and Hilton, 2001).

Protein inhibitor of activated STAT (PIAS) also negatively regulates JAK/STAT signalling by altering the transcriptional activity of STATs. There are four mammalian PIAS proteins, that interact with STAT dimers after ligand stimulation and prevent DNA-binding (Jason S Rawlings et al., 2004). PIAS proteins can also recruit transcriptional corepressors, such as histone deacetylases (HDACs), which inhibit STAT-mediated transcription. PIAS proteins have SUMO-ligase functions, which may result in STAT SUMOylation and alter transcriptional of specific target genes (Shuai,

2006). Protein tyrosine phosphatases (PTPs) reverse STAT activation through removal of the phosphate group by a number of mechanisms (Jason S Rawlings et al., 2004). Dephosphorylation of STATs is key in the attenuation of signalling, and a variety of PTPs act at specific points along the signalling pathway. SHP-1 and SHP-2 associate with the receptor and counteract the phosphorylation of STAT by JAK at the receptor. Some PTPs, such as PTP1B, act in the cytoplasm to dephosphorylate STAT, whereas others are restricted to the nucleus. Several PTPs can also reduce the activity of JAKs, by dephosphorylating a tyrosine residue in their activation domain, and therefore preventing STAT activation (Böhmer and Friedrich, 2014). SLIM is a further example of a negative JAK/STAT pathway regulator. SLIM is a nuclear ubiquitin ligase that associates with STAT molecules and targets them for proteasomal degradation (Tanaka et al., 2005).

In contrast to the other regulators, Signal Transduction Adaptor Molecule (STAM) positively modulates the pathway by increasing the activity of JAKs (Hou et al., 2002). There are four human STAM proteins, which bind to JAKs through their ITAM region and are predicted to act as a scaffold to bring targets into close proximity of kinase activity. STAM also plays roles in endocytic trafficking, being a component of the ESCRT-0 complex (Section 1.2.2), and is suggested to integrate endocytosis with the actin cytoskeleton (Lohi and Lehto, 2001). Therefore, JAK/STAT pathway modulators are important for both positive and negative regulation, however regulation by confinement into intracellular organelles such as endosomes has been less well studied.

1.1.3.2 STAT post-translational modifications

Modulation of signalling is also regulated by controlling protein activity through post-translational modifications (PTMs). JAKs are understood to undergo ubiquitination and ISGylation (conjugation of Interferon-stimulated gene 15), which results in proteasomal degradation or prolonged activity, respectively (Shuai and Liu, 2003). Over recent years there has also been significant interest in STAT post-translational modifications and their association with disease.

Phosphorylation of STATs at a conserved tyrosine residue (residue 701 in STAT1) has long been considered a requirement for STAT transcriptional activity. Although this is the case for ligand-induced STAT targets, unphosphorylated STATs appear to have roles in transcription (Chapter 1.1.2.2). Further STAT phosphorylation sites have also been identified that alter STAT activity, such as a serine residue (727 in STAT1) within

the C-terminal transactivation domain of mammalian STAT1, STAT3, STAT4 and STAT5. Phosphorylation of STAT1 at serine 727 is required for maximum transcription of cytokine regulated genes (Wen et al., 1995). Serine phosphorylation also occurs under cellular stress and infection, and different kinases have been identified, such as PKC-δ (Uddin et al., 2002) and ERK1 (Zhang et al., 2004). Therefore, serine phosphorylation of STAT1 may present a point of integration of various signalling pathways (Zhang et al., 2004). Mutation of this site has been demonstrated to delay disease onset in mice studies, by enhancing the tumour surveillance activity of natural killer cells (Putz et al., 2013). Studies investigating the role of STAT3 serine 727 phosphorylation have been controversial, suggesting it as a negative modulator of tyrosine phosphorylation (Chung et al., 1997), or demonstrating its requirement for maximal transcription of specific genes (Costa-Pereira, 2011). A recent study by Huang et al, 2014, in mouse embryonic stem cells suggests that serine phosphorylation of STAT3 is important in determining cell fate; tyrosine phosphorylation is important in maintaining pluripotency, whereas serine phosphorylation overcomes this signal and initiates neuronal differentiation (Huang et al., 2014). STAT3 serine phosphorylation is a constitutive modification in a number of malignancies and S727 phosphorylation is required for Ras-mediated tumour formation (Gough et al., 2013). Serine phosphorylation of STAT5 has also been implicated in its oncogenic activity in STAT5-driven leukaemia's (Friedbichler and Kerenyi, 2010). STAT5a serine phosphorylation has also been associated with the suppression of prolactin driven expression of β-casein in the mammary gland. This reduces STAT5a DNA binding, which can be overcome through increased stimulation of glucocorticoid receptors to enable milk synthesis (Yamashita et al., 2001). A further serine, at residue 193 within the coiled-coil domain of STAT5b, is phosphorylated in response to cytokine stimulation, and its constitutive phosphorylation confers oncogenic activities (Mitra et al., 2012). Therefore, further phosphorylation of STAT proteins can specify their function, and due to their involvement in disease is a good candidate for therapeutic intervention.

Acetylation of transcription factors can also alter their activity, and the majority of STATs have been demonstrated to bind to CBP/p300, a histone acetyltransferase (HAT). CBP/p300 acetylates STAT3 at lysine 685 in response to cytokine treatment, which increases STAT3 DNA binding and transcriptional activity, (R. Wang et al., 2005). Three further acetylation sites have been identified in STAT3, all of which are in close proximity to the conserved tyrosine and have essential roles in tyrosine phosphorylation (Nie et al., 2009). STAT5 acetylation in close proximity to the

conserved tyrosine is also key for transcriptional activity and dimer formation (Zhuang, 2013). In contrast, acetylation of STAT1 opposes tyrosine phosphorylation, preventing target gene expression. CBP/p300-dependent acetylation leads to the recruitment of T-cell protein tyrosine phosphatases which dephosphorylate STAT1 (Krämer and Heinzel, 2010).

Arginine methylation of STAT1 at residue 31 was demonstrated to increase DNA binding and modulates tyrosine phosphorylation. R31 methylation is also reduced in multiple cancer cell lines preventing the transcription of multiple tumour suppressor genes (Mowen et al., 2001; Zhu et al., 2002). Lysine methylation at K140 of STAT3 occurs in the nucleus, when STAT3 is bound to subset of promoters and blocks DNA binding (Yang et al., 2010).

Ubiquitination is primarily associated with targeting proteins for proteasomal degradation. Ubiquitination involves the conjugation of an 8.5kDa regulatory protein to a substrate protein via the action of a ubiquitin-activating enzyme (E1), a ubiquitin conjugating enzyme (E2), and a ubiquitin ligase (E3). SLIM is a E3 ligase that ubiquitinates STAT proteins and promotes their degradation (Lim and Cao, 2006). Ubiquitin also plays essential roles in protein activity. Tumour necrosis factor receptor-associated factors (TRAFs) are involved in signalling from TNF receptors. TRAF6 has ubiquitin ligase activity and interacts with STAT3 leading to ubiquitination within its SH2 domain (Wei et al., 2012; Ruan et al., 2017). Wie et al., 2012, suggested that this negatively regulates STAT3 transcriptional activity, yet Ruan et al. 2017, provide data suggesting that TRAF6 ubiquitination promotes STAT3 membrane recruitment and phosphorylation.

SUMOylation involves the conjugation of a Small Ubiquitin-like Modifier (SUMO) protein to a lysine in the target protein via SUMO-specific E1, E2 and E3 ligases, in an enzymatic process similar to ubiquitination. SUMO proteins are ~11kDa in size and have been shown to modulate proteins implicated in signal transduction and transcription. PIAS, an E3 SUMO ligase, SUMOylates STAT1 at lysine 703 (Rogers et al., 2003). This SUMOylation inhibits the expression a subset of STAT1-regulated genes, suggesting it provides specificity to STAT1 transcription in a promotor dependent manner (Ungureanu et al., 2005), by altering the DNA-binding affinity of STAT1 (Grönholm et al., 2012).

PTMs at the same residue can antagonistically regulate the activity of STAT proteins. Acetylation and SUMOylation of lysis 696 of STAT5a have contrasting roles on tyrosine phosphorylation in lymphocytes. Acetylation at K969 is required for STAT5a phosphorylation, whereas SUMOylation at the same residue prevents STAT5a phosphorylation. Therefore, the study provides a model whereby acetylated STAT5a is phosphorylated, dimerises and enters the nucleus. Here STAT5a is deacetylated, allowing SUMOylation at K696, and dephosphorylated. Removal of the SUMO group from STAT5a is then required for STAT5a to exit then nucleus, where it is available for phosphorylation and acetylation again (Van Nguyen et al., 2012).

1.1.3.3 Regulation through cross-talk of signalling pathways

For a specific biological response, such as an immune response to infection, multiple ligands are involved in regulating the outcome. Although signalling pathways were classically regarded as linear protein cascades, in reality the situation is much more complex with considerable communication occurring between pathways. This increases the number of potential outputs available from a limited number of signalling pathways, and finetunes context-dependent regulation. Dsyregulation of signalling crosstalk is associated with multiple diseases (Pálfy et al., 2012).

Cross-talk between different cytokine pathways can have both synergistic or antagonist effects on signalling. For example, treatment with IFN-γ elicits an increase in STAT1 and IRF9 expression, sensitising cells to treatment with IFN-α. Conversely, IFN treatment results in increased SOCS1 expression, which inhibits the activation of STAT6 in response to IL-4 treatment (Dickensheet et al., 1999). Phosphorylation of STAT1 at serine 727 (discussed in section 1.1.3.2) is considered an integration point with MAPK signalling. UV-induced activation of a variety of MAPKs, such as ERKs, JNKs, p38 kinase and MEK1, and leads to STAT1 S727 phosphorylation in an EGFR-independent manner, but does not alter phosphorylation of the conserved tyrosine residue (Zhang et al., 2004).

STATs can also bind to proteins from distinct signal transduction pathways. The nuclear factor-κB (NF-κB) family are transcription factors crucial in immune regulation. NF-κB interacts with STAT6, synergistically inducing signalling by increasing the DNA-binding affinity of STAT6 (Shen and Stavnezer, 1998). NF-κB transcription factors p65 and p50 physically bind to STAT3, but have opposing functions. The binding of p65 inhibits the transcription function of STAT3 at GAS elements. In contrast, association

with p50 enhances the transcriptional activity of STAT3 and expression of GAS site dependent genes (Yoshida et al., 2004). Overactivation of STAT3 and NF-κB is associated with epithelial tumorigenesis and therefore understanding their crosstalk and cooperation may be key for treatments (Grivennikov and Karin, 2011). STAT3 is activated by Notch through the association of STAT3 with Notch target Hse1, and is essential in the developing nervous system (Kamakura et al., 2004).

1.1.4. The *Drosophila* JAK/STAT pathway

The JAK/STAT pathway is highly evolutionarily conserved, with invertebrates such as *Drosophila melanogaster* having a full complement of pathway components. However, in comparison to the complexity of the mammalian JAK/STAT pathway (demonstrated in table 1.1) the *Drosophila* system is much simpler. A single receptor, Domeless (Dome), one JAK, Hopscotch (Hop), and a one STAT, STAT92E, make up the signalling pathway. Therefore, *Drosophila* provides an excellent model in which to investigate the intricacies of JAK/STAT pathway regulation, without the difficulties of compensation and signalling crosstalk experienced in the mammalian system. In fact, investigating JAK/STAT signalling in *Drosophila* has led to key breakthroughs in understanding the impact of its dysregulation in human disease. Gain-of-function JAK mutants gave rise to leukaemia-like defects in flies, making the first connection between JAK/STAT signalling and cancer (Ekas et al., 2010).

As in the mammalian system, the *Drosophila* JAK/STAT pathway is important during multiple developmental processes, such as cell growth and patterning in the developing wing and eye imaginal discs. Tight regulation of the level of JAK/STAT signalling is required to maintain the correct tissue size, with overactivation leading to tissue overgrowth (Bach et al., 2003) and uncontrolled JAK/STAT pathway activity leading to tumour formation. Polycomb repressor complexes (PRCs) bind to DNA response elements and cause epigenetic gene silencing of *Drosophila* JAK/STAT pathway ligands. PRC mutants causes increased expression of these ligands and upregulation of JAK/STAT pathway activity that leads to tumour formation (Amoyel et al., 2014). Mutations of Endosomal Sorting Complexes Required for Transport (ESCRT) complexes (see Chapter 1.2.2) cause epithelial tumours in imaginal discs. Specificially, mutants of TSG101, an ESCRT-I component, resulted in increased JAK/STAT signalling causing cell cycle dysregulation and tissue organisation defects (Gilbert et al., 2009).

1.1.4.1 Drosophila JAK/STAT ligands

The *Drosophila* genome encodes only 3 molecules capable of activating the JAK/STAT pathway, Unpaired/Outstretched (Upd), Upd2 and Upd3. Upd was originally identified in a loss-of-function study, where the *upd* mutants produced a similar embryonic segmentation phenotype as *hop* and *stat92e* mutants. Molecular analysis demonstrated that *upd* encodes a IL-6-like glycosylated protein that is secreted, interacts with the extracellular matrix and causes Hop phosphorylation (Harrison et al., 1998). *In silico* searches identified two further Upd-like genes at the *upd* locus, which were named Upd2 and Upd3 (Agaisse et al., 2003). Upd and Upd3 are both capable of interacting with the ECM, whereas Upd2 is a freely diffusible ligand which may be capable of activating JAK/STAT signalling at greater distances from its source (Harrison et al., 1998; Agaisse et al., 2003; Hombría et al., 2005). Upd and Upd2 show overlapping expression patterns during embryogenesis and act semi-redundantly, with *upd2* mutants being completely rescued by Upd, whereas Upd2 only partially rescues phenotypes cause by *upd* loss-of-function mutants (Gilbert et al., 2005; Hombría et al., 2005).

Key differences in the dynamics and strength of JAK/STAT activation elicited by these three ligands were detected in cell culture. Upd is the most potent but short-lived, whereas activation via Upd2 is longer lasting (Wright et al., 2011). Upd2 appears to have very similar roles to mammalian leptin, being released in the *Drosophila* fat body in response to feeding and triggering insulin secretion (Rajan and Perrimon, 2012). Upd2 and Upd3, but not Upd, have crucial roles in activating the immune response against parasitic wasp infection (Yang et al., 2015). Therefore, the three *Drosophila* JAK/STAT ligands appear to have some overlapping roles, but also key different roles which to-date are not fully understood.

1.1.4.2 The receptor, Domeless

The *Drosophila* JAK/STAT pathway receptor was the final component of the pathway to be identified, and was done so by two independent laboratories. Dome mutants produced similar abnormalities in larval posterior spiracles to that of STAT92E mutants (Brown et al., 2001) and are capable of rescuing *Drosophila* eye overgrowth phenotype caused by ligand overexpression. Genetic analysis demonstrated the Dome acted downstream of Upd, but upstream of Hop (H. Chen et al., 2002). Sequencing of Dome cDNA indicated that it encodes a transmembrane protein and that, although sequence

similarity to vertebrate cytokine receptors is low, Dome contains many similar features such as fibronectin-type-III domains which resemble CBMs in vertebrate cytokine class I receptors (Brown et al., 2001). Dome has similarities to the IL-6R family including, LIFR, CNTFR, gp130, as well as the IL-3Rα and IL-2Rβ. This suggests that the variety of vertebrate receptors seen today evolved from a single ancestral receptor that gave rise to Dome (Brown et al., 2001; H. Chen et al., 2002). These early studies demonstrated that Upd binds Dome to activate JAK/STAT signalling, that STAT92E directly interacts with Dome (H. Chen et al., 2002) and that the intracellular domain of Dome is important for its function in signal transduction (Brown et al., 2001). In line with data from multiple vertebrate cytokine receptors, Dome homo-dimerisation occurs in a ligand-independent manner but is required for signal transduction, suggesting a role for pre-dimerised Dome in signalling (Brown, 2003).

In recent years a number of Dome regulators that alter JAK/STAT signalling have been identified. In the *Drosophila* genome, *dome* is located adjacent to *cg14225* (*latran*), which encodes a structurally similar transmembrane protein (Hombría and Brown, 2002). Latran negatively regulated multiple JAK/STAT luciferase reporters, and its knockdown caused hyperphosphorylation of STAT92E and hyperactivation of target genes *in vivo* (Kallio et al., 2010). A direct interaction with Dome (Kallio et al., 2010) were shown to result in signalling incompetent Dome:Latran dimers, whereby increased Latran concentrations reduced JAK/STAT signalling (Fisher et al., 2016). Interestingly, Latran also associates with Hop and STAT92E, but not with extracellular Upd ligands, suggesting that Latran may act by sequestering signalling components into an inactive complex. This negative regulation is thought to be key in controlling tissue specific immune responses. Unlike Domeless, Latran is not expressed in the developing embryo and therefore transcriptional control of the two related proteins must be different. This uncoupled transcriptional regulation allows Dome and Latran ratios to change dependent on the context. For example upon parasitic wasp infestation, haemocyte precursors increase Latran expression and reduce Dome expression, enabling lamellocyte differentiation (Makki et al., 2010). Therefore, expression of Latran may present a mode of fine tuning JAK/STAT signalling in a tissue-specific and context-dependent manner.

Using an RNAi screen Fisher et al., 2018, identified Multiple Ankyrin repeats and KH-domain containing protein (MASK) as a Dome interactor that promotes dimerisation and stability of the receptor, acting as a positive regulator for the ligand-induced JAK/STAT activation both in cells and *in vivo*. This protein is closely related to human

ANKHD1, whose knockdown caused a reduction JAK/STAT signalling in mammalian cells (Fisher et al., 2018). This study suggests that MASK is crucial in stabilising Dome, however its role Dome endocytosis has not been studied.

1.1.4.3 The JAK, Hopscotch

The JAK tyrosine kinase Hopscotch (Hop), was the first *Drosophila* JAK/STAT pathway component to be identified by the Perrimon group in 1994, and indicated a key role of the pathway in *Drosophila* development (Binari and Perrimon, 1994). Its sequence is similar to vertebrate JAK1 and JAK2, and includes kinase, pseudo-kinase, SH2-like and FERM domains. Hyperactivation mutations of *hop* results in proliferation and aggregation of *Drosophila* plasmatocytes resulting in tumour formation (Hou et al., 2002). Hop also appears to be a key protein involved in pathway crosstalk, as it activates D-Raf which is important for blood cell development (Luo et al., 2002).

1.1.4.4 The transcription factor, STAT92E

The Drosophila STAT protein, STAT92E, was discovered by two laboratories simultaneously due to its high sequence similarity to vertebrate STAT5 (37% identity) and as mutants displayed the same phenotypes as *hop* mutants (Hou et al., 1996; Yan et al., 1996). STAT92E acts downstream of Hop (Hou et al., 1996), and is phosphorylated at the highly conserved tyrosine 704 (Yan et al., 1996). This phosphorylation is key for STAT92E dimerisation, nuclear accumulation, DNA binding and expression of downstream targets both *in vitro* and *in vivo* (Karsten et al., 2006; Ekas et al., 2010). Dimerised STAT92E can bind to the conserved STAT DNA consensus sequence, TTC(n)GAA, specifically where there are 3 or 4 amino acids in the spacer (n) region. Binding to 3n sites occurs at higher affinity than to 4n sites, suggesting there is flexibility to STAT92E DNA binding that may influence transcriptional output and be dependent on other regulatory factors (Rivas et al., 2008).

In addition to the conserved tyrosine residue, further residues and modifications of STAT92E have been demonstrated to have essential roles in its function. A conserved arginine, at position 442, does not alter STAT92E activation or nuclear translocation, but is essential for DNA binding and gene transcription (Ekas et al., 2010). This residue is conserved in multiple vertebrate STATs but has not been studied. Mutation of methionine 647 to histidine (M647H) results in increased STAT92E nuclear accumulation, phosphorylation at tyrosine 704 and DNA binding in the absence of ligand. However, the M647H did not activate gene expression and acted as a

dominant-negative *in vivo*, suggesting that this mutant can bind DNA and may occupy the promotor region preventing wild-type STAT92E from binding. This study speculates that the M647H mutant may lack a post-translational modification that, in addition to phosphorylation of the conserved tyrosine residue, is required for transcriptional activity (Karsten et al., 2006). To date, the extra post-translational modifications that regulate the action of mammalian STATs have not been identified in STAT92E. However, SUMOylation of STAT92E at an non-conserved lysine 187 was demonstrated to have inhibitory roles on JAK/STAT signalling (Gronholm et al., 2012).

STAT92E has also been associated with non-canonical JAK/STAT functions. The regulation of heterochromatin formation is essential for genome stability and accurate gene expression. Unphosphorylated, 'transcriptionally inactive', STAT92E is involved in maintaining heterochromatin stability through association with, and subsequent distribution of, heterochromatin protein 1 (HP1). HP1 is required for gene silencing of heterochromatin. Phosphorylation of STAT92E results in relocalisation to euchromatin, causing displacement of HP1 and destabilisation of heterochromatin (Shi et al., 2008). Reduced levels of unphosphorylated STAT92E caused *Drosophila* to be more sensitive to DNA damage, highlighting the role of unphosphorylated STATs in maintaining genome stability (Yan et al., 2011).

1.1.4.5 Regulation of the JAK/STAT pathway in Drosophila

Key protein regulators of the JAK/STAT pathway in vertebrates are also conserved in *Drosophila*. There are three SOCS-like proteins in *Drosophila*, SOCS16D, SOCS44A, SOCS36E. SOCS36E has the closest homology to mammalian SOCS5 and is a direct target of the JAK/STAT pathway, therefore generating in a negative feedback system. SOCS36E regulates *Drosophila* JAK/STAT signalling via two mechanisms. Firstly, it associates with Elongin-B/C and Cullin-5, via its C-terminal SOCS box, creating a Elongin-Cullin-Socs (ECS) complex which acts as an E3-ubquitin ligase. This alters Dome stability, likely by targeting it for lysosomal degradation (Stec et al., 2013), and therefore reducing Dome-signalling. SOCS36E also regulates JAK/STAT signalling in a cullin-2 dependent manner *in vivo* (Monahan and Starz-gaiano, 2015), through its SH2 domain, which acts by preventing binding of STAT92E to the activated receptor (Stec et al., 2013; Monahan and Starz-gaiano, 2015). SOCS36E also negatively regulates EGFR signalling in the wing disc. In contrast, SOCS44A is not a target of JAK/STAT pathway activity and instead up-regulates EGFR signalling. The SOCS proteins also have distinct roles during oogenesis suggesting different functions (Jason S. Rawlings et al., 2004). Dephosphorylation of JAKs and STATs is key in regulating

signalling. Protein tyrosine phosphatase 61F (PTP16F) was identified as a negative regulator of *Drosophila* JAK/STAT signalling in an RNAi screen. PTP16F is the homolog of PTP-1B, and acts by dephosphorylating Hop *in vitro* and *in vivo*. *Ptp61f* expression is increased upon JAK/STAT activation, therefore creating a further negative feedback system (Baeg et al., 2005). *Drosophila* PIAS (dPIAS) also acts as a negative regulator, acting as a SUMO-ligase to STAT292E (Betz et al., 2001).

Ken and Barbie (Ken) is a negative regulator of the JAK/STAT pathway in *Drosophila*, acting as a transcriptional repressor of STAT92E and recognising a DNA consensus sequence which overlaps with STAT92E consensus sequences. Interestingly, not all STAT binding sites are Ken binding sites, meaning that Ken only alters the expression of a subset of target genes, and does not alter *socs36E* expression due to different DNA binding sites. Therefore Ken is a key regulator of the JAK/STAT pathway that regulates the expression of specific targets (Arbouzova et al., 2006).

Nuclear import is required for STAT92E function in signalling and heterochromatin regulation. Phosphorylated STAT92E translocates to the nucleus in response to ligand stimulation (Karsten et al., 2006). dsRNA against *Drosophila* Ran homologs resulted in increased nuclear phospho-STAT92E and signalling, suggesting their importance in the regulation of nuclear transport and JAK/STAT signalling (Baeg et al., 2005). This indicates an importin/exportin mediated nuclear translocation of STAT92E. As unphosphorylated STAT92E has roles in heterochromatin stability, it must be able to enter the nucleus independent of tyrosine phosphorylation, however this has not been studied.

As in mammals, relatively little is known about the spatial and temporal regulation of the *Drosophila* JAK/STAT signalling elicited by endocytosis.

1.2 Endocytosis and trafficking

Endocytosis is a ubiquitous process in eukaryotic cells and is crucial for multiple cellular events such as nutrient uptake, immune surveillance, modulation of membrane composition and signal regulation. Invagination of the plasma membrane and subsequent pinching off allows extracellular molecules to be internalised into intracellular vesicles (Damke, 1996; Mettlen et al., 2018). There are a plethora of mechanisms (Figure 1.3) that facilitate the internalisation of a diverse range of cargo molecules. Transmembrane cargo molecules, such as receptors, contain internalisation signals on their cytoplasmic tails, which are recognised by adaptor proteins and cause clustering of the cargo molecules into specialised areas of the plasma membrane. Uptake can occur via clathrin-dependent or clathrin-independent mechanisms, both creating vesicles that can fuse with the early endosome (EE) (Conner and Schmid, 2003; Mettlen et al., 2018). After fusion with the EE, cargo such as plasma membrane components, receptor-ligand complexes and cell debris, are sorted and trafficked through the endocytic pathway for recycling, degradation or storage (Huotari and Helenius, 2011). Endocytic uptake and trafficking is a process originally thought as distinct from signalling, yet it is being increasingly understood to have quantitative and qualitative roles in the regulation of signalling pathways.

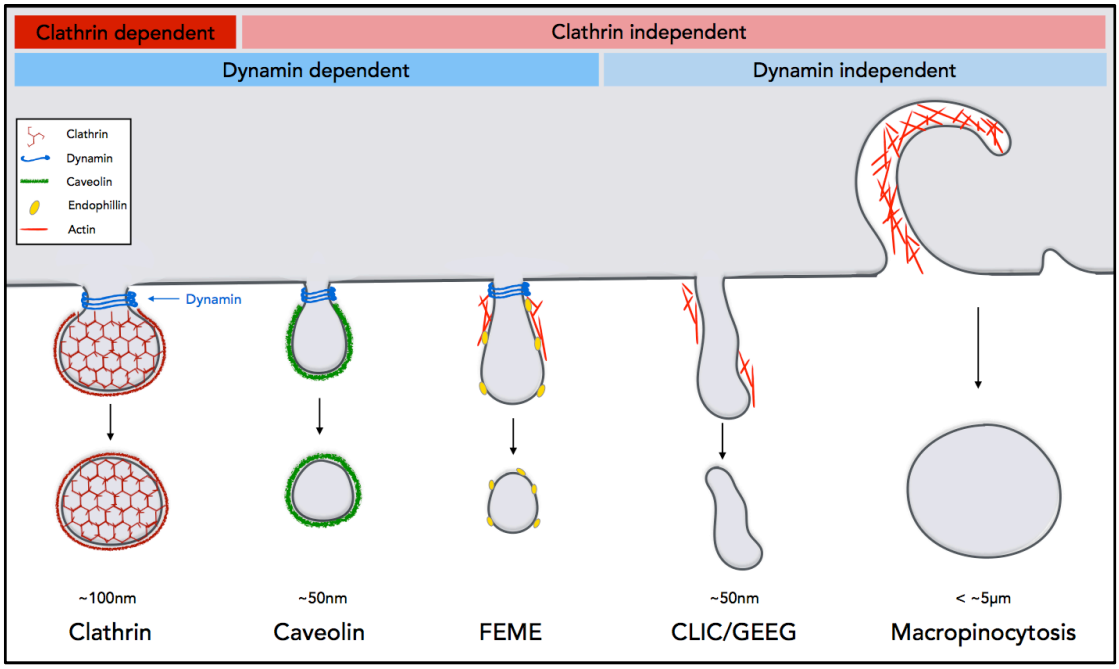


Figure 1.3: Mechanisms of endocytic uptake. Schematic demonstrates the major internalisation pathways, their recruitment for dynamin and clathrin, and size of vesicles produced (if known).

1.2.1. Mechanisms of cargo internalisation

1.2.1.1 Clathrin-mediated endocytosis

Clathrin-mediated endocytosis (CME) is the best characterised mechanism for internalisation of activated receptors and regulation of their cell-surface expression (Mettlen et al., 2018). Internalisation requires incorporation of cargo molecules into clathrin coated pits (CCP) at the plasma membrane. Clathrin, first isolated in 1975 by Barbara Pearse, is a highly evolutionarily conserved coat protein, consisting of clathrin heavy and clathrin light chains that form a three-legged trimer, or triskelia (Pearse, 1975). This clathrin triskelia is the major structural component of a CCP, assembling into a lattice of hexagons and pentagons that enables deformation of the membrane to form invaginated pits (Kanaseki and Kadota, 1969; Maib et al., 2017). The GTPase dynamin is recruited to the neck of a deeply invaginated pit where it self-assembles into a ring, causing membrane fission (Antonny et al., 2016) and the production of a clathrin-coated vesicle (CCV). CCVs are then uncoated within the cytoplasm by heat shock cognate factor 70 and auxilin, allowing fusion with the endosomal pathway (Conner and Schmid, 2003; McMahon and Boucrot, 2011).

Aside from the role of clathrin in internalisation of cargo from the plasma membrane, it is also involved in multiple other trafficking events within the cell. For example, clathrin forms flat lattices at restricted regions of early endosomes and is required for efficient sorting of ubiquitinated cargo (Raiborg et al., 2001; Wenzel et al., 2018). It is also crucial in recycling and intracellular transport, transporting cargo between trans-Golgi network, the endoplasmic reticulum and endosomes (Lauvrak et al., 2004; Robinson, 2015) and, unrelated to its functions in trafficking, clathrin plays a role in crosslinking microtubules at the mitotic spindle (Royle, 2012).

Cargo proteins do not interact directly with clathrin, but instead through the association with endocytic adaptor proteins. These adaptors bind to both the clathrin triskelia and cargo proteins, along with phosphoinositides at the plasma membrane (Traub and Bonifacino, 2013). Multiple endocytic adaptor proteins have been identified, and are specific to the intracellular location of CCPs (Robinson, 2015). The heterotetrameric adaptor protein 2 (AP2) is the major adaptor protein at the cell surface, and includes an α, β2, μ2 and σ2 subunit. Recruitment of AP2 to phosphatidylinositol 4,5-bisphosphate (PIP_2) at the cell surface results in a drastic conformational change in AP2 (Kelly et al., 2014). This allows AP2 to associate with clathrin through a clathrin-box-motif in its β2-subunit, and to bind cargo via distinct subunits. Hence, AP2 acts as

a bridge between PIP_2, cargo and clathrin (McMahon and Boucrot, 2011). AP2 also associates with membrane sculpting proteins, allowing for successful pit formation that is stabilised by the binding and polymerisation of clathrin (Schmid et al., 2014; Robinson, 2015). Removal of AP2 from the newly formed vesicle is mediated by dephosphorylation of the μ2 subunit (Semerdjieva et al., 2008). Multiple alternative adaptors, or clathrin-associated sorting proteins, facilitate cargo-specific uptake in AP2 dependent or independent manners. For example, arrestins enable the uptake of G-protein coupled receptors (Hamdan et al., 2007), Dab2 and Autosomal Recessive Hypercholesterolemia (ARH) are specific to the uptake of the LDL receptor (Maurer and Cooper, 2006), and clathrin assembly lymphoid myeloid leukaemia (CALM) is specific to the internalisation of SNAREs (Maritzen et al., 2012). Patients with AHR disorder have impaired LDLR-mediated uptake in liver (Garuti et al., 2005).

Selection of transmembrane cargo for CME is specific and dependent on the recognition of short, cytoplasmic, internalisation motifs or post-translational modifications. AP2 binds cargo with YxxΦ and acidic dileucine ([D/E]xxxL[L/I]) motifs, where Φ is any bulky hydrophobic side-chain (L, I, V, M or F), and x is any residue. The YxxΦ motif is recognised by the μ2 subunit of AP2 (Ohno, et al, 1995), whereas the acidic dileucine motif binding site is found in the α-σ2 hemicomplex of AP2 (Doray et al., 2007; Kelly et al., 2008). Alternative adaptors involved in CME recognise distinct internalisation motifs. Arrestins bind to phosphorylated serines and threonines on GPCRs, ARH recognise a FDNPVY motif (Mishra et al., 2002), and epsins/Eps15 recognise ubiquitinated transmembrane cargo (Polo et al., 2002). Ubiquitination of the erythropoietin receptor (EpoR), an RTK that requires JAK2 for downstream signalling, is required for its internalisation (Bulut et al., 2011). EpoR ubiquitination triggers subsequent ubiquitination of p85, which can then associate with phospho-tyrosines on EpoR and enables binding to the adaptor epsin-1, resulting in EpoR CME (Bulut et al., 2013).

1.2.1.2 Clathrin-independent endocytosis

Clathrin-independent endocytosis (CIE) encompasses a multitude of pathways that differ in their mechanism of internalisation and are defined by the molecular machinery used for uptake. Advances in biological tools, such as high-resolution microscopy and genetic manipulation, has facilitated the study of these pathways (Mayor et al., 2014). We now understand that a diverse range of internalisation pathways are responsible for the uptake of extracellular material, and that they can function in distinct

physiological processes. CIE pathways have also been linked to disease, including cancer, atherosclerosis and lysosomal storage disorders (Ferreira and Boucrot, 2018).

Caveolae are membrane invaginations enriched in cholesterol and characterised by the presence of caveolin and cavins. Caveolae-dependent endocytosis requires dynamin and appears to have cell type specific functions (Mayor et al., 2014). Caveolin and cavins have not only been linked to endocytosis, but with processes such as transcytosis in endothelial cells, cholesterol homeostasis and mechanotransduction, and their mutation has been associated with multiple diseases (Nassoy and Lamaze, 2012). The fast endophilin-mediated endocytosis (FEME) pathway also requires dynamin, along with endophilin and RhoA. Uptake via this method occurs at patches of membrane pre-enriched with endophilin, and is important for rapid internalisation of GPCRs and activated RTKs (Ferreira and Boucrot, 2018). The interleukin-2 receptor (IL-2R), an activator of STAT1, STAT3 and STAT5 in T, B and NK cells, is internalised via a distinct mechanism but also requires RhoA. The receptor itself is thought to concentrate upon activation and, via its cytoplasmic tails, stimulate a cascade of protein activity that causes Arp2/3-mediated actin polymerisation, ultimately resulting in the formation of an endocytic pit around the activated receptor (Mayor et al., 2014; Ferreira and Boucrot, 2018). The GPI-anchored protein-enriched endocytic compartments (CLIC/GEEC) pathway is a constitutive pathway that mediates fluid phase internalisation and uptake of lipid-anchored proteins. This pathway requires Arf1 and the RhoGTPase Cdc42, but not dynamin (Mayor et al., 2014). Arf6 defines a further CIE pathway important for the uptake of MHC class 1 molecules (Donaldson et al., 2016).

CIE also includes pathways that result in large regions of the plasma membrane being internalised into micrometre scale endocytic vacuoles. Macropinocytosis occurs when regions of the plasma membrane ruffle. These ruffles may subsequently fuse with the plasma membrane and allow for bulk uptake of extracellular material in a non-selective manner. This process can be stimulated by the presence of high concentrations of growth factors, and may represent a cellular response to high extracellular ligand concentrations. Macropinocytosis exhibits cell specific roles in immune cells, enabling capturing of antigens for presentation, and in neurons, regulating synaptic signalling (Buckley and King, 2017). Over recent years, research into the details of macropinocytosis has become substantial, as it is an attractive method for the delivery of therapeutics.

Signal sequences and adaptors for the uptake of cargo via CIE mechanisms are considerably less well characterised than for CME. Although no consensus sequences have been identified to date, some cargo specific sequences have been identified. For example, Lectin-like oxidized low-density lipoprotein receptor-1 contains a DDL motif in its cytoplasmic tail which is essential for internalisation via dynamin-dependent CIE (Murphy et al., 2008), whereas Arf6- and β-arrestin-dependent CIE of the M2 muscarinic acetylcholine receptor requires a KKKPPPS sequence (Reiner and Nathanson, 2009; Wan et al., 2015). However, neither of these studies identified interacting partners that bindqw these signalling sequences.

1.2.1.3 Mechanism of cargo internalisation can be context dependent

Cargo uptake is often not restricted to internalisation via a specific endocytic portal, and cargo can harbour multiple internalisation motifs. Uptake via multiple pathways may present a mechanism whereby endocytosis can respond to different physiological contexts and influence signalling outcomes. For example, EGFR can enter cells via both CME and CIE pathways, in a ligand concentration-dependent manner. After stimulation with low (~1ng/ml) EGF concentrations the receptor is internalised primarily through the CME route, whereas stimulation with high levels of EGF (>20ng/ml) results in receptor monoubiquitination and internalisation through CIE routes (Sigismund et al., 2005). Phosphorylation of EGFR occurs in a graded response to EGF levels, yet a threshold concentration of ligand causes ubiquitination, resulting in a molecular switch to CIE that controls EGFR fate (Sigismund et al., 2013). Monoubiquitination enables binding of epsin/eps15, demonstrating their multifunctional role in both CME and CIE (Sigismund et al., 2005). Unlike CME, internalisation via CIE does not appear to contribute to EGFR signalling and therefore may be a mechanism to reroute receptors and regulate signalling. Syp1 is a yeast homolog of the FCHo1 endocytic adaptor. Recently, Syp1 was shown to have roles in both CME and CIE, suggesting that the adaptors involved in internalisation routes may not be entirely distinct (Apel et al., 2017).

1.2.2. Endocytic trafficking and cargo sorting

Following internalisation and uncoating, endocytic vesicles fuse with the early endosome (EE), the major endosomal-sorting compartment (Figure 1.4). Here cargo proteins can be targeted for lysosomal degradation via the late endosome (LE), recycled to the cell surface or trafficked to further intracellular organelles such as the

trans-Golgi network. Comparable to internalisation, cargo destination is dependent on specific sorting signals and interacting proteins. Trafficking pathways are complex and interconnected, facilitating the delivery of diverse cargo molecules to specific membranous organelles in an accurate and context-appropriate manner. Movement through and between these pathways requires a specific set of components, yet the master regulator of progression through the endocytic pathway is the rab family of small GTPases (Hutagalung and Novick, 2011).

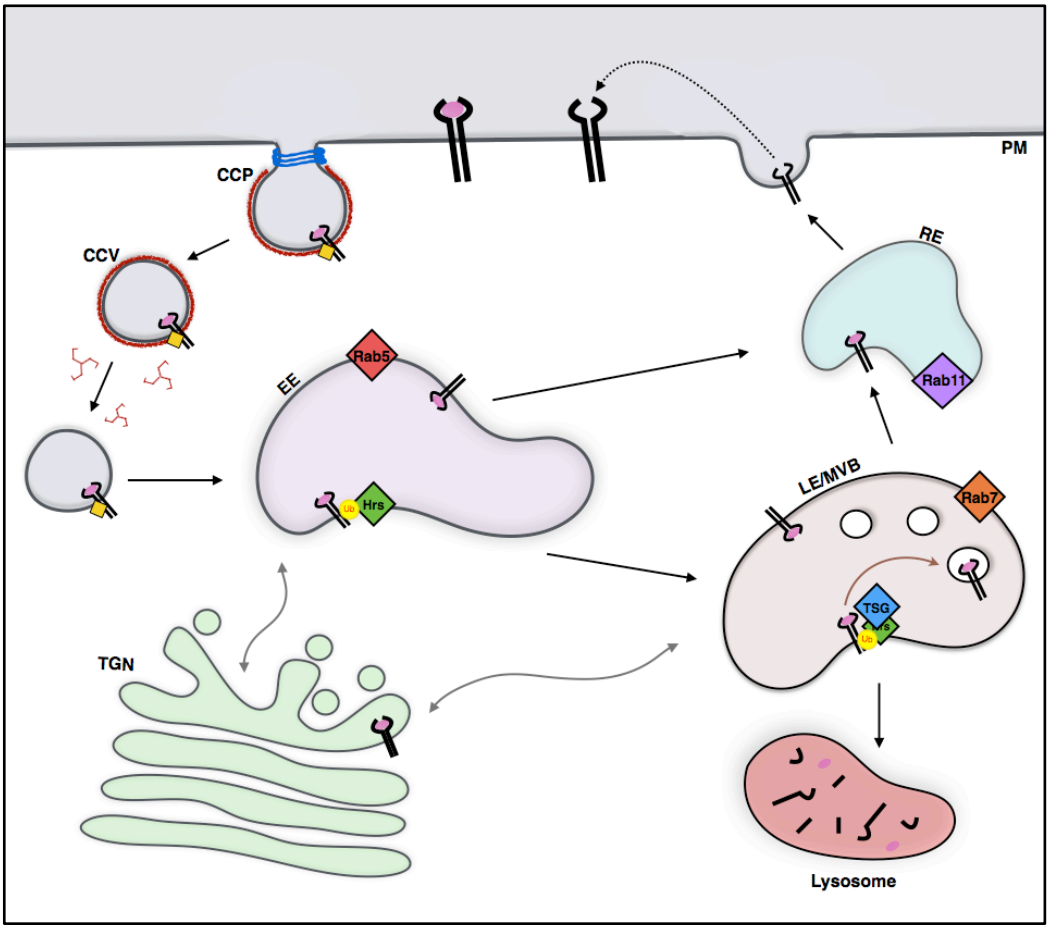


Figure 1.4: Schematic of endocytic trafficking. Transmembrane cargo such as receptors are localised to clathrin-coated pits (CCP) through associations with an endocytic adaptor protein (yellow square). Clathrin-coated vesicles (CCV) are pinched off and after uncoating vesicle fuse with the early endosome (EE). Here cargo can be sorted into the late endosome (LE) / multivesicular bodies (MVB), to the recycling endosome (RE) or the trans-Golgi network (TGN). At the LE cargo is incorporated into ILVs through the action of ESCRT complexes. The degradation of the cargo is then achieved in the lysosome.

Rabs are monomeric ras-like GTPases that cycle between an inactive, GDP-bound form in the cytoplasm and an active, GTP-bound form that is membrane-associated. The activation of rab proteins is catalysed by guanine nucleotide exchange factors

(GEFs), which mediate the exchange of GDP for GTP, whereas GTPase activating proteins (GAPs) cause GTP hydrolysis and hence inactivate the rabs. This cycling results in conformational changes of the rab GTPases, altering their function, localisation and interacting partners. Active rabs are master regulators of membrane trafficking, functioning in cargo selection, vesicle movement and fusion, cytoskeletal interactions, and signalling, which are mediated by interactions with cytoplasmic effector proteins (Hutagalung and Novick, 2011; Wandinger-Ness and Zerial, 2014). Approximately 70 human rab GTPases have been identified, and these specifically localise to distinct endosomal membranes, providing organelle identity and interaction with specific effectors (Wandinger-Ness and Zerial, 2014). Rab5 is the major regulator of the early endocytic pathway, modulating traffic between the plasma membrane and the EE. Through its effectors, such as, VPS34, EEA1, and APPL1/2, rab5 regulates PtdIns(3)P production, vesicle fusion, and signal transduction (Jovic et al., 2010). Rab7 is important in the control of LE trafficking, whereas rab4 and rab11 regulate endocytic recycling (Huotari and Helenius, 2011; Bhuin and Roy, 2014). Rab proteins recruit lipid kinases and phosphatases, causing endosomes to be enriched in specific phosphoinositides and hence contributing to organelle identify. Maturation of endosomes is marked by changes in the rab proteins that associate with the endosomal membrane, a process termed rab conversion. A transition from EE to LE occurs through a cascade of rab activity, whereby active rab5 recruits a GEF for rab7, which in turn recruits a GAP for rab5 (Rink et al., 2005; Hutagalung and Novick, 2011).

At the LE, ubiquitinated cargo destined for degradation is sorted into intraluminal vesicles (ILVs) that are produced from inward budding of the endosomal membrane, leading to the formation of multivesicular bodies (MVBs). This process attenuates signalling by encompassing the cargo in the endosome and preventing interactions with cytosolic signalling proteins. ILV formation occurs in the opposite topology to endocytosis at the cell surface, and requires the sequential action and assembly of endosomal sorting complexes required for transport (ESCRT) complexes (Figure 1.4). These are a family of four protein complexes that are highly evolutionarily conserved (Leung et al., 2008). ESCRT-0 is a heterodimer of HGF-regulated tyrosine kinase substrate (Hrs) and STAM and is crucial in initiating ILV production by binding ubiquitin-tagged proteins. Hrs localises to endosomal membranes through a FYVE domain that binds phosphatidylinositol 3-phosphate (PI3P). In Drosophila Hrs is essential for endosomal membrane invagination (Lloyd et al., 2002). The Hrs/STAM dimer recruits the ESCRT-I complex through interaction with the TSG101 protein. A distinct domain of ESCRT-I binds with ESCRT-II, which can associate with the membrane, cargo and

ESCRT-III. Transient assembly of ESCRT-III results in membrane scission and recruitment of Vps4, which results in termination of MVB formation and recycling of ESCRT complexes. MVBs then fuse with lysosomes in a unidirectional manner, where cargo is inactivated and degraded (Hurley and Emr, 2006; Huotari and Helenius, 2011; Schmidt and Teis, 2012; Schöneberg et al., 2016; Moore et al., 2018).

1.3 The role of Endocytosis in signal regulation

Traditionally endocytosis was viewed as a mechanism to attenuate signalling from activated receptors through their downregulation and degradation. However, now it is understood to have roles in maintaining, defining and even generating distinct signals. Research has demonstrated that endocytosis and signalling are bidirectionally linked, with signalling also regulating the endocytic pathway. However, here I will focus on the role of endocytosis in the regulation of signalling pathways.

1.3.1. Quantitative regulation of signalling

Early studies assumed that all signalling occurs at the cell surface and that internalisation plays housekeeping roles to remove available receptors. We now know that endocytosis has complex, cargo specific, effects on signalling. Quantitative regulation is the notion that endocytosis can regulate the level of signalling. Endocytosis is a mechanism that can physically reduce the number of cell-surface receptors and disrupt membrane signalling complexes. For example, G-protein signalling from activated GPCRs is attenuated due to an association with arrestin and subsequent endocytosis. Following sustained agonist stimulation of GPCRs arrestins are recruited and attenuate G-protein signalling by uncoupling it from the activated GPCR, whilst also targeting the receptor for endocytosis (Moore et al., 2007). Endocytosis also results in the physical separation of receptors from plasma membrane factors. For example, signalling of GPCRs through phospholipase Cγ (PLCγ) only occurs at the plasma membrane as the substrate of PLCγ, $PI(4,5)P_2$, is enriched here (Sorkin and Zastrow, 2009). Trafficking of receptors to the LE for lysosomal degradation can attenuate signalling through ligand dissociation, and also results in receptor degradation which can desensitise cells to ligand stimulation. In contrast, recycling of the receptor maintains a pool of cell surface receptors, causing resensitisation and signalling potential. An example of this is the association of

arrestins with GPCRs. Prolonged interaction of the β_2 adrenergic receptor with β-arrestin results in receptor degradation, whereas transient association causes receptor recycling, hence enabling further signalling (Tian et al., 2014; Takenouchi et al., 2018).

1.3.1.1 Endosomes as signalosomes

Endosomes also act as mobile signalling platforms, facilitating the assembly of signalling complexes. This can allow for signal maintenance, enabling signalling at the cell surface to persist after receptor internalisation. For example, some GPCRs can continue activating MAPK signalling pathways after endocytosis due to sustained association with arrestins at 'signalling endosomes' (Daaka et al., 1998). Spatial regulation of signalling is key for precise biological outcomes, and endocytosis provides a mechanism by which this can occur. For example, during *Drosophila* border cell migration, EGF and PDGF/VEGF receptors are localised at the leading edge to provide direction and guidance cues. Preventing endocytosis of these RTKs does not inhibit signalling, but it does result in cell migration defects. This suggests that spatially restricting signalling through key interactions and endocytosis is essential for appropriate signalling (Jékely et al., 2005).

1.3.1.2 Endocytic trafficking can modulate signalling

The method of endocytic uptake can modulate signalling from receptors, such as the TGFβR, which can be internalised via both CME and caveolin-1 dependent pathways (Di Guglielmo et al., 2003). Activation of TGFβR results in Smad signalling, whereby R-Smad transcription factors associate with Smad4 to regulate the transcriptional output. Interaction of the receptor with Smad6 or Smad7 negatively regulates signalling (Balogh et al., 2013). Entry via CME results in the delivery of TGFβRs to endosomes that are positive for EEA1. Here, the Smad anchor for receptor activation (SARA) provides a bridge between the activated receptor and R-Smads. Although SARA is also present at the cell surface, it is predominantly associated with the EE membrane due to the enrichment of PI3P, and hence delivery of TGFβR to the EE enhances Smad signalling (Di Guglielmo et al., 2003; Runyan et al., 2005). Originally it was shown that entry via caveolin-dependent endocytosis reroutes the receptor to avoid the EE. These receptors are associated with Smad7 and are hence targeted for lysosomal degradation (Di Guglielmo et al., 2003). Therefore, internalisation via CME and delivery to the EE was considered as crucial for TGFβR signalling, whereas uptake via the CIE regulated receptor turnover, suggesting subcellular location of TGFβR is critical in the modulation of Smad signalling. However, recent studies have

demonstrated that clathrin- and caveolae- derived vesicles can fuse, resulting in the delivery of caveolin-1, and TGFβR internalised via CIE, to EEA1 positive early endosomes (He et al., 2015). Smad7 and ubiquitin ligases were also associated with these endosomes, suggesting that the caveolin-1 positive EEs facilitates both TGFβR signalling and degradation. This demonstrates that the internalisation pathways are not functionally separated and that the EE may represent a multifunctional organelle for cargo sorting (He et al., 2015), whereby membrane microdomains may be crucial for specific functions (see Chapter 1.3.2.1).

1.3.2. Qualitative regulation of signalling

Endocytosis of cell-surface receptors does not only affect the level of signalling, but also the specific transcriptional output. The location of the receptor-ligand complex within the endocytic pathway may produce a unique response (Figure 1.5), which is regulated spatially and temporally. *In vivo* signal transduction and live-cell FRET studies demonstrated that uptake of EGFR, in response to EGF stimulation, results in hyperphosphorylation of specific EGFR tyrosine residues and maximal association with a signalling complex including SHC, GRB2 and mSOS at endosomes (Di Guglielmo et al., 1994; Jiang and Sorkin, 2002). These signalling components are required for downstream Ras activation and MAPK/ERK signalling, and subsequent studies showed that EGFR internalisation is required for full MAPK activation, whereas other EGF-dependent signalling events are not altered by inhibition of endocytosis (Vieira et al., 1996). Trafficking of the EGFR to the LE is important in the specific activation of ERK1. P14 is localised to the cytoplasmic face of the LE, where it acts as an endosomal adaptor protein and specifically interacts with MEK partner 1 (MP1), which is a scaffold for ERK1 activation (Teis et al., 2002). Therefore the localisation of the signalling complex spatially within the cell is important for specific cellular interactions and distinct signals. In budding yeast, GPCR signalling has also been shown to be qualitatively regulated by endocytosis. Although MAPK signalling occurs at the cell surface, the Gα subunit can become activated at the endosome causing activation of Cdc42 signalling (Slessareva et al., 2006). The tumour necrosis factor receptor (TNFR) associates with multiple components at the cell surface allowing for NF-κB signalling and upregulation of prosurvival proteins. However, internalisation of TNFR is required for the activation of caspase 8 and subsequent apoptotic signalling (Schneider-Brachert et al., 2004). Therefore compartmentalisation of the receptor causes differential signalling and phenotypic outputs.

Cell-surface and endosomal membranes differ in their enrichment of phospholipids, facilitating the assembly of unique protein complexes. Therefore, endosomes can serve as signalling platforms to generate distinct signals. PI(4,5)P_2 at the plasma membrane facilitates association of toll-like receptor 4 (TLR4) with an adaptor complex that activates p38 and inhibitor of NF-κB. After internalisation and depletion of PI(4,5)P2 there is dissociation of the adaptor from the receptor, enabling association with other interacting proteins that activate IRFs (Kagan et al, 2008).

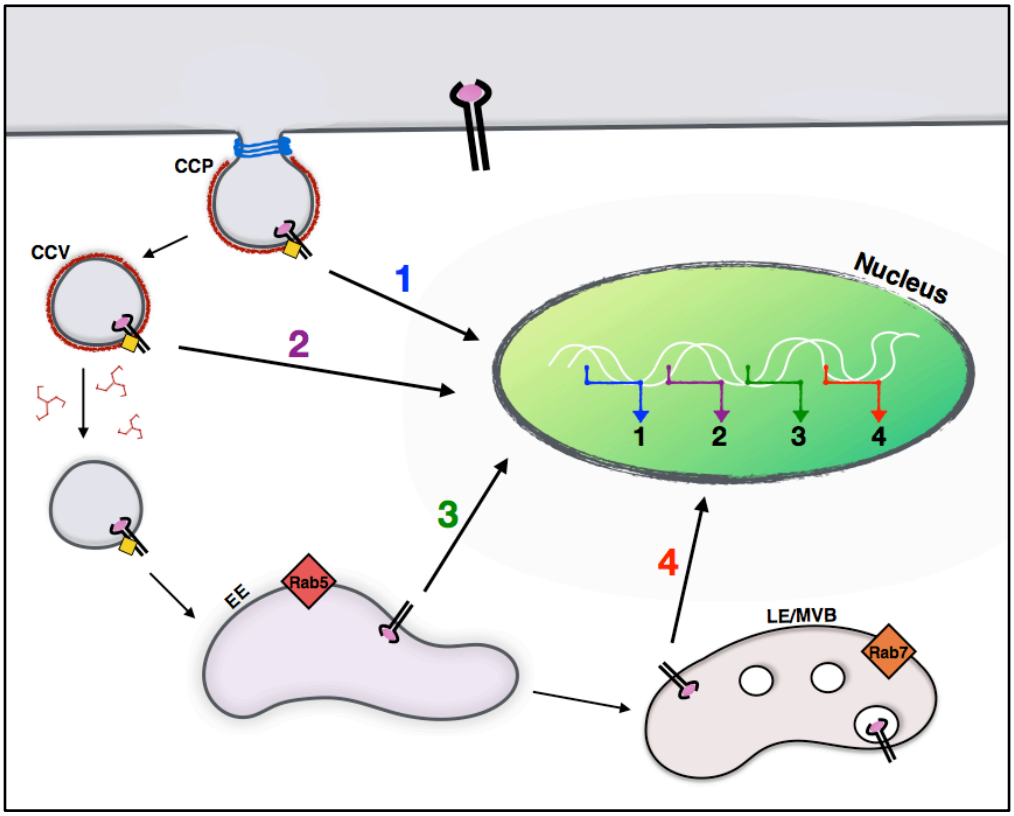


Figure 1.5: Qualitative signal regulation during endocytosis allowing for differential signalling. Schematic demonstrating how endocytic trafficking can facilitate the expression of subsets of transcriptional targets from the same receptor.

1.3.2.1 Endosomal subpopulations and microdomains

For endocytosis to provide qualitative regulation to signalling pathways, receptors must be located within a specific membrane environment which is competent for a subset of signalling. These different membrane environments may be formed by different subpopulations of endosomes, or microdomains within the same endosomal limiting membrane, that recruit specific scaffold proteins, adaptors and enzymes (Moore et al., 2018).

Although (Early Endosome Antigen 1) EEA1-positive endosomes are considered as the canonical early endosomal compartment, there are other early endosomal populations involved in receptor trafficking. Adaptor protein, phosphotyrosine interacting with PH domain and leucine zipper 1 (APPL1) is a rab5 effector that interacts with the endosomal membrane and labels a subpopulation of early endosomes that are important in EGFR signalling. Endosomes positive for APPL1 were original believed to be absent of EEA1, however more recent studies have identified endosomes that are positive for both APPL1 and EEA1 (Kalaidzidis et al., 2015). APPL1 endosomes are positioned close to the cell surface and are suggested to play cargo specific roles; EGFR traffics to APPL1 endosomes quickly after stimulation, yet there is limited transferrin labelling (Miaczynska et al., 2004). Following EGF stimulation, EGFR physically interacts with APPL1 which is required for EGF-induced Akt activation (Wang et al., 2002; Zhou et al., 2016). APPL1 endosomes are also key for specifying Akt signalling in zebrafish, with signalling of Akt through Gsk-3β being dependent on APPL1, whereas signalling via Tsc2 is APPL1-independent (Schenck et al., 2008). Active, endocytosed, EGFR was shown to transit though an APPL1-independent route that is positive for SNX15 (Danson et al., 2013). Subpopulation of endosomes have also been discovered for later endocytic compartments. Internalisation of EGFR by ligand-independent stress responses and the chemotherapeutic cisplatin, results in the delivery of EGFRs to a subset of MVBs. In contrast to MVBs containing EGF activated receptors, these appear to be a distinct non-degradative subset of MVBs, that result in sustained signalling, delayed apoptosis and therefore may contribute to chemotherapy resistance (Thomas et al., 2015).

Microdomains within plasma membranes may also act as signalosomes. At the cell surface, localisation of the IFNGR into cholesterol and glycosphingolipids rich subdomains is required for IFN-γ activated signalling (Marchetti et al., 2006), and active signalling domains have been demonstrated for IL-2R (Blouin et al., 2016). Through association with flat clathrin lattices at the EE, Hrs is clustered into microdomains on the endosomal limiting membrane that are absent of EEA1. This clustering facilitates Hrs recognition of ubiquitinated cargo and is required for efficient targeting for degradation (Raiborg et al., 2002, 2008; Wenzel et al., 2018). In the same limiting membrane, cargo destined for trafficking to the trans-Golgi network are separated in microdomains distinct from Hrs (Norris et al., 2017). Therefore, there is heterogeneity in endosomal compartments and dynamic organisation of

microdomains, enabling cargo to be spatially segregated or associated with signalling molecules.

1.3.2.2 Endocytosis provides a platform for signalling cross-talk

Endosomes may also act as platforms for regulating physical interactions between various signalling pathways and therefore providing a spatial location for modulating signalling crosstalk. For example, endosomal proteins such as APPL1, EEA1 and Hrs have been seen to regulate various signalling pathways. APPL1 enables crosstalk between insulin and WNT signalling, associating with components from both pathways (Pálfy et al., 2012), whereas EEA1 mediates angiotensin II induced Akt activation (Nazarewicz et al., 2011). Hrs plays a role in crosstalk between BMP and MAPK signalling pathways, being required for activation of SMADs and TAK1/p38 in response to BMP signalling (Miura and Mishina, 2011). Through *in silico* analysis other endosomal scaffolds have been proposed that could be involved in crosstalk (Pálfy et al., 2012).

1.3.3. Endocytic regulation of signalling in *Drosophila*

Endocytosis has also been identified to play key roles in specifying signalling outcomes in lower complexity organisms such as *Drosophila*. For example, *hrs* mutant flies were demonstrated to have impaired ILV formation, which prevented degradation of EGF and Torso receptors and resulted in prolonged signalling (Lloyd et al., 2002).

In *Drosophila* endocytosis also provides mechanisms for consistent signalling in varying contexts. Internalisation of the Notch receptor via different endocytic routes enables robust signalling throughout a variety of temperatures. Notch signalling is activated in response to membrane bound ligands such as Delta and Serrate, which initiate proteolytic cleavage events that ultimately result in the release of the Notch intracellular domain (NCID). NCID then translocates to the nucleus to regulate transcription. Activation of Notch can also occur via the intracellular E3 ubiquitin ligase, Deltex, which promotes Notch ubiquitination, endocytosis and lysosomal degradation independently of ligand activation, with release of NCID occurring at the late endosomal membranes. Suppressor of Deltex causes Notch to enter MVBs, therefore preventing its Deltex-induced Notch activation (Wilkin et al., 2008; Baron, 2012). Interestingly, the activity of Deltex and Suppressor of Deltex are temperature sensitive

and a network of competing endocytic routes have different temperature sensitivities (Shimizu et al., 2014). This enables signalling from Notch receptors to be stable at a variety of different temperatures, which is essential for normal development and homeostasis in this ectothermic organism.

A further role for endocytosis in the *Drosophila* system is the use of multiple internalisation mechanisms to spatially regulate signalling. Wingless (Wg) is the *Drosophila* Wnt protein and interacts with the receptor DFrizzled2 (DFz2). Wg is secreted apically, whereas DFz2 is located basally on the signal-receiving cell. In the wing disc, Wg is endocytosed from the apical surface by the CLIC/GLEEC pathway. In contrast, DFz2 is internalised via CME at the basal surface into distinct endosomes. Fusion of these endosomes allows for interactions of the receptor-ligand and mediates Wg signalling (Hemalatha et al., 2016). Therefore, endocytosis appears to provides cells with a mechanism to fine-tune ligand/receptor interactions and downstream signalling.

1.3.4. Endocytic regulation of the JAK/STAT signalling pathway

Studies investigating the internalisation of nRTKs that activate the JAK/STAT pathway have also demonstrated a crucial roles for endocytosis in the regulation of the signalling.

1.3.4.1 Regulation of the mammalian JAK/STAT pathway by endocytosis

The most well studied JAK/STAT pathway in relation to endocytic regulation is that activated in response to the mammalian IFNs. IFN-α or IFN-γ bind to distinct receptors, the IFN-α/β receptor (IFNAR) and IFN-γ receptor (IFNGR), respectively, yet both stimulate STAT1 activity (see Chapter 1.1.2). Therefore, a major question surrounding IFNs is how different signals and receptors can activate the same effector molecules whilst maintaining signal specificity and distinct transcriptional responses. Marchetti et al., 2006, demonstrated that both receptor complexes are internalised via clathrin and dynamin-dependent endocytosis, yet the role on signalling is different. Inhibition of CME prevents the induction of STAT1 by IFN-α, but has no effect on the phosphorylation of STAT1 by IFN-γ. CME is therefore crucial for IFNAR signalling, but not for IFNGR (Marchetti et al., 2006). In contrast, signalling from the IFNGR is modulated by lipid nanodomain compartmentalisation. The glycosylation of IFNGR

subunits and galectin binding plays a crucial role in the localisation of IFNGR into membrane nanodomains. Excess galectin binding results in IFNGR confinement to actin-rich nanodomains, where IFNGR cannot activate STAT phosphorylation. In contrast, depletion of galectin binding facilitates IFNGR portioning into lipid nanodomains, resulting in a conformation of IFNGR that can activate signalling. Thus, lipid nanodomains can control transmembrane receptor signalling at the cell surface (Blouin et al., 2016). Therefore, the same effector molecule, STAT1, is activated in two distinct mechanisms through CME and endocytosis, or by lipid microdomain compartmentalisation, providing a mechanism by which specificity in the JAK/STAT pathway is achieved.

Endocytic trafficking of the IFNAR receptor mediates finetuning of the JAK/STAT pathway in a spatial and temporal manner. Delivery to the EE enables association between the IFNAR2 subunits and the retromer subunit, VPS35. This causes dissociation of the IFNAR complex, differential sorting of the subunits and the termination of JAK/STAT signalling (Chmiest et al., 2016). Therefore, internalisation and trafficking causes STAT1 to be activated at precise times following IFNAR activation.

The regulation of endocytosis on STAT3 activation also appears to be ligand/receptor-specific. The cytokine oncostain-M (OSM) acts via JAK2 (Fossey et al., 2011) to elicit strong and rapid phosphorylation of STAT3, which is independent of receptor endocytosis (Kermorgant and Parker, 2008). STAT3 is also activated by mesenchymal-epithelial transition factor (c-Met), the hepatocyte growth factor (HGF) receptor (Boccaccio et al., 1998). Yet in contrast to OSM, HGF causes a low and delayed level of STAT3 phosphorylation. STAT3 nuclear translocation in response to HGF requires endocytosis of c-Met and microtubule dependent delivery to a perinuclear endosome enriched in rab7 (Kermorgant and Parker, 2008). Endocytosis is therefore considered a mechanism to overcome a weak signal by trafficking the receptor to an endosomal subpopulation close to the nucleus, and enabling effective STAT3 phosphorylation, which is hypothesised to be due to limited phosphatase exposure in the cytoplasm (Kermorgant and Parker, 2008).

Internalisation into endosomes appears to be key in concentrating JAK/STAT components, facilitating interactions and signalling. Subunits of the IL-4R complex were shown to require uptake into distinct cortical signalling endosomes before signal transduction occurred. Heterodimerisation of the IL-4 bound interleukin-4Ra (IL-4Ra)

subunit and interleukin-2Ry (IL-2Ry) forms the type 1 IL-4R complex. Interestingly, the affinities of these two subunits for one another at plasma membrane concentrations are too low for efficient dimerization. Kurgonaite et al., 2015, demonstrated that constitutive uptake of these receptor subunits into endosomes tightly associated with the cell cortex increases their concentration, facilitating ligand-induced dimerization and downstream signalling (Kurgonaite et al., 2015). This emphasises the role of endocytosis in spatially arranging molecules for appropriate signalling. In the case of the type 1 IL-4R complex, found primarily in hematopoietic cells and hence important in immune responses, this may allow for the integration of weak signals or buffer against sudden fluctuations (Kurgonaite et al., 2015).

Although IL-6R internalisation is constitutive and ligand independent (Thiel et al., 1998), endocytosis is required for IL-6 induced activation STAT3 and MAPK activation (German et al., 2011). STAT3 was seen in cytoplasmic puncta and interacts with the early endosomes (Shah et al., 2006; German et al., 2011). Interestingly, STAT3 appears to directly interact with clathrin heavy chain (Shah et al., 2006), which may mediate STAT3's membrane association. STAT3 is thought to be transcriptionally competent at the EE, but phosphorylation of serine 727 which is required for maximal transcriptional output occurs at the LE, providing a platform for local signal activation. S727 phosphorylation is also dependent on MAPK activation, and therefore endocytosis may provide a mechanism to carefully regulate pathway cross-talk in a location specific manner (German et al., 2011).

As discussed previously, the endocytic regulation of the EGFR has been well studied, yet the EGFR is capable of activating STATs in JAK dependent and independent mechanisms. Bild et al., 2002, demonstrated that CME of EGFR is crucial for EGF activated STAT3 nuclear translocation, dimerisation and transcriptional activity. STAT3 was seen localised to endocytic vesicles following EGF stimulation. Interestingly, STAT3 is still phosphorylated on the conserved tyrosine residue (705) when endocytosis of EGFR is blocked, suggesting tyrosine phosphorylation is not sufficient for nuclear translocation (Bild et al., 2002). Although this endocytic regulation may be independent of nRTKs and JAK activity, it demonstrates the importance of STAT endocytic regulation and also how cross-talk between signalling pathways may be important.

1.3.4.2 Endocytic regulation of the Drosophila JAK/STAT pathway

Studies using *Drosophila* to investigate the role of endocytosis on the JAK/STAT pathway suggested both positive and negative regulation. The first report of Dome internalisation was by Ghiglione, 2002, where the receptor was seen to accumulate in intracellular vesicles in follicle cells. Ligand expression caused an increase in Dome containing vesicles, indicative of receptor-mediated endocytosis (Ghiglione, 2002), and subsequent studies have demonstrated the importance of Dome endocytosis during development (Silver et al., 2005). Inhibiting endocytosis through the expression of a dominant-negative form of Shibire (the *Drosophila* homolog of dynamin), prevented border cells from migrating towards the oocyte and increased cell surface levels of Dome, suggesting that endocytosis is crucial in regulating Dome cell surface levels and signalling. Cytoplasmic levels of STAT92E were also increased, indicating that endocytosis is required for STAT92E degradation (Silver et al., 2005).

In 2007, Devergne et al., suggested that ligand-dependent, CME is required for Dome signalling. Colocalisation of Dome with Rab5, Rab7 and an MVB marker, but not Rab11, indicated that the Upd/Dome complex is trafficked via the LE for lysosomal degradation. Mutation of Clathrin heavy chain (CHC) prevented Dome internalisation, whilst decreasing STAT92E expression and nuclear translocation in follicle cells, suggesting CME of the receptor is required for JAK/STAT signalling (Devergne et al., 2007), In contrast, Vidal et al., 2010, argued that endocytosis negatively regulates JAK/STAT signalling. Pulse-chase experiments in the Kc_{167} *Drosophila* cell line confirmed that uptake of the Upd2/Dome complex occurred via CME. However, knockdown of CHC, AP2 and Hrs increased STAT92E activity in a luciferase assay, suggesting that internalisation and trafficking negatively regulates Dome signalling (Vidal et al., 2010). Vidal et al, confirmed that inhibiting endocytosis increases JAK/STAT signalling *in vivo*, in cells of the imaginal disc and in migrating border cells (Vidal et al., 2010). Dissimilarities between the outcomes of these groups may be due to differences in the context (*in vitro* vs *in vivo*), ligand stimulation and assays used to study and modulate endocytosis.

A recent study by Ren et al., 2015, identified Windpipe (Wdp) as a JAK/STAT target and pathway modulator that is important in the regulation of intestinal stem cell homeostasis and tissue regeneration in the *Drosophila* midgut. Wdp is a transmembrane protein, whose expression is positively regulated by JAK/STAT signalling. Interestingly, Wdp acts to negatively regulate the JAK/STAT pathway, creating a negative-feedback loop to tightly control JAK/STAT pathway activity. Wdp

was shown to interact with Dome, causing Dome aggregation at the cell surface, receptor internalisation and lysosomal degradation in an ligand-independent mechanism (Ren et al., 2015). Inducing Dome internalisation may be a key mechanism to regulate a cells responsiveness to JAK/STAT pathway ligands in a context-specific manner. This study highlights a novel modulator of the *Drosophila* JAK/STAT signalling whose mechanism of action involves endocytic trafficking.

Data regarding the role of endocytosis on *Drosophila* JAK/STAT signalling appears to be contradictory. However, all studies suggest that endocytosis plays a role in regulating downstream signalling, whether this is positive or negative. Increased understanding of the role endocytosis in regulating multiple signalling pathways, in various organisms, has demonstrated that the mechanisms can be complex and context specific. Therefore, endocytosis can provide a method to finetune signalling to enable appropriate responses from pathways, such as the JAK/STAT pathway, which are involved in a range of physiological outcomes.

1.4 Project aims

Previous data from the Smythe lab suggested a role for specific Rab5 guanine nucleotide exchange factors (GEFs) in JAK/STAT signalling in the developing *Drosophila* wing disc. This suggested that endocytosis may regulate the JAK/STAT pathway in a more complex manner than described by Devergne et al., 2007, and Vidal et al., 2010. This study also established signalling assays in the *Drosophila* S2R+ cell line and demonstrated biochemically that internalisation of the a GFP tagged version of Upd2 occurs in receptor dependent manner and primarily via clathrin-mediated endocytosis (Figure 1.6) (Vogt and Smythe, unpublished), verifying previous immunofluorescence data (Vidal et al., 2010). Therefore, I aimed to enhance our understanding of the molecular mechanisms that enable endocytosis to regulate signalling.

In the present study, I initially focused on investigating how perturbing endocytic trafficking altered the transcriptional output of JAK/STAT signalling in *Drosophila* S2R+ cells. These results revealed an essential role for endocytosis in qualitatively regulating signalling, with subsets of JAK/STAT targets being expressed as the receptor is trafficked through the endosomal pathway. These results caused me to carry out a

comprehensive study of STAT92E activity during endocytic trafficking, and led to the identification of a novel phosphorylation site that appears to be crucial in transcription factor activity.

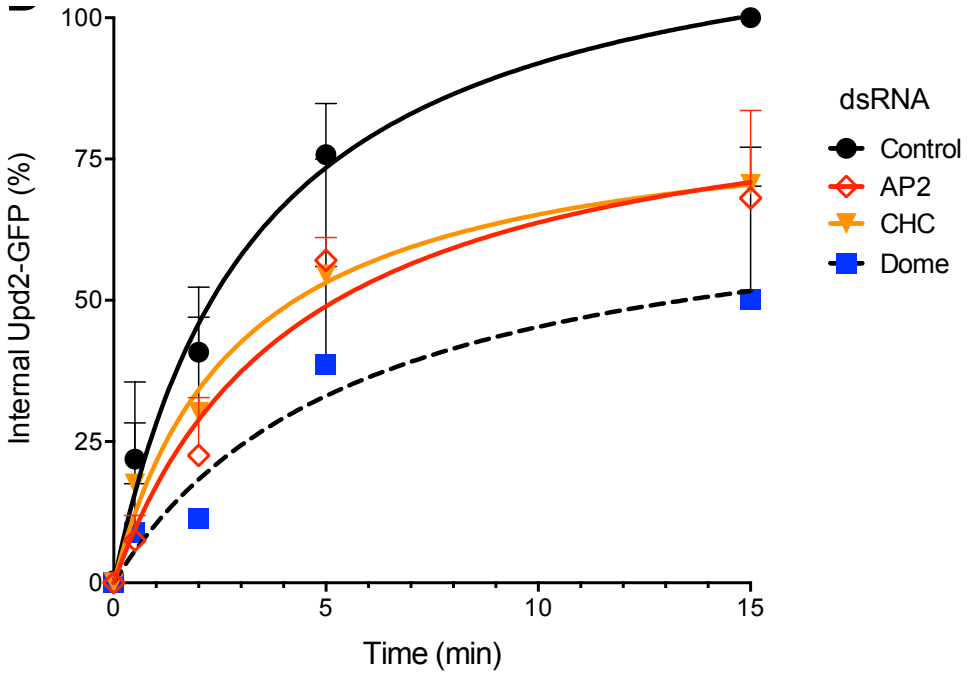


Figure 1.6: Uptake of Upd2-GFP is due to Clathrin-mediated endocytosis. S2R^{+} cells were treated for 5days with dsRNA. Cells were then incubated with Upd2-GFP (20nM) for indicated time points at 25ºC, after acid washes cell-lysates were analysed with the anti-GFP ELISA. The internal Upd2-GFP amount is represented as %, whereby the longest time point (15min) of the control was set to 100% in each individual experiment, to allow comparison. Graph represents at least 3 independent experiments and error bars show standard error of the mean. In a two-way ANOVA statistical test, the Dome dsRNA has an effect on the uptake rate into cells, which is considered significant with a P value of 0.0216 (Vogt and Smythe, unpublished).

Chapter 2. Materials & Methods

2.1 Common Buffers

- *PBS:* Fisher BioReagents® tablets (#BPE9739-1). 1x Solution = 137mM NaCl, 10mM phosphate buffer 2.7mM KCl.
- *TAE (50x):* 50mM Na_2EDTA, 2M Tris, 1M acetic acid.
- *TBS:* 1x solution: 20mM Tris-HCl, pH7.4, 150mM NaCl
- *TBST:* TBS + 0.1% tween-20
- *SDS-PAGE electrophoresis running buffer (1x):* 25mM Tris, 192mM glycine, 0.1% (w/v) SDS
- *Western blot transfer buffer (1x):* 25mM Tris, 192mM glycine, 20% methanol
- *Lysis buffer:* 20mM Tris pH7.5, 150mM NaCl, 1mM EDTA, 1mM EGTA, 1% Triton X-100, 1mM β-Glycerophosphate, 25mM Na-Pyrophosphate, 1mM Na_3VO_4, 1µg/mL microcystin, 25mM N-ethylmaleimide supplemented with cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche #11836170001).

2.2 *Drosophila* cell culture and manipulation

2.2.1. *Drosophila* cell culture

S2R+ cells were cultured at 25°C in Schneiders Insect Tissue Culture media (Gibco #21720024), supplemented with 10% heat inactivated FBS (Sigma #F4135), penicillin-streptomycin (100x stock, Gibco) and 2mM L-Glutamate (Gibco). Cells were grown to confluency in T75cm^2 flasks and routinely passaged at a 1:3 dilution every 3-4 days. Cells were dislodged by pipetting media across the flask surface. For long term storage cells were resuspended in 90% HiFBS and 10% DMSO before storage in liquid nitrogen. Reactivation of cells was accomplished by thawing briefly at 37°C and spinning down in 10mL fresh media. Cells were used for no more than 25 passages.

2.2.2. DNA transfection of *Drosophila* cells

Cells were seeded at 1x10^6 cells per well of a 6 well plate 1 day prior to transfection. In general, a total of 2µg DNA was transfected per well using Effectene Transfection

Reagent (Qiagen) following manufacturer's instructions, with the DNA (µg) to Enhancer (µL) ratio kept at 1:8. Briefly, 20µL DNA was incubated with 180µL EC Buffer and 16µL Enhancer for 3mins, before addition of 20µL Effectene for 7mins. The DNA-reagent mix was then diluted in 1mL fresh media and added in a dropwise fashion to PBS washed cells. Cells were incubated at 25°C for 6hrs-5days. Relevant ratios were used for transfection in different-sized wells/plates (Table 2.1).

2.2.2.1 Production of Upd2-GFP condition media

Cells were transfected, as described above, with 2µg *pAct-Upd2-GFP*. After 2 days 3 wells of transfected cells from a 6 well plate were transferred to a T75cm^2 flask, 20mL fresh media added and incubated for 4 days. Conditioned media was removed and cells pelleted (at 1000xg for 3mins), and supernatant was filtered, aliquoted and snap-frozen in liquid N2 and stored at -80ºC. Concentration of Upd2-GFP in media was determined using anti-GFP ELISA (Chapter 2.3.1.1).

2.2.2.2 Production of CRIPSR/Cas9 cell lines.

To calculate the lowest concentration of puromycin required to kill non-transfected S2R+ cells, cells were plated at 5x10^{5}cells/well in a 12 well plate and incubated overnight in standard culture media (Chapter 2.2.1). The following day media was replaced with fresh media containing 1.25, 2.5, 5, 10 or 50µg/mL puromycin and cells were examined for 7days, with media being replaced every 3 days. After 6 days cells grown in 50, 10 and 5µg/mL puromycin containing media were dead, and therefore 5µg/mL puromycin was chosen for antibiotic selection of transfected cells.
S2R$^+$ cells were plated at 5x10^{5}cells/well in a 12 well plate and transfected with 1µg *pAc-sgRNA-Cas9* constructs using Effectene. After 3 days antibiotic selection was performed for 4 days before subsequent analysis (Bassett et al., 2014).

2.2.3. dsRNA knockdown in *Drosophila* cells

dsRNA knockdown in S2R+ cells was carried out an using a RNAi bathing protocol. Cells were split a day prior to knockdown and resuspended in serum free media on the day of knockdown. The correct volume of dsRNA, prepared in water at ~1µg/µL (Chapter 2.6.1), was added to a tissue culture plate (Table 1.1). The desired number of cells (maintaining a ratio of 1x10^6 cells to 15µg dsRNA), were diluted in serum free media and added to the wells containing dsRNA. After a 1hr incubation at 25°C, an

equal volume of fresh media containing 10% FBS was added. Cells were incubated at 25°C for a total of 5 days before subsequent experiments.

Plate/Flask size	Cell number	dsRNA (µg)	Typical [DNA] (µg) for transfection
150mm = 152cm^2	1.6x10^7 cells/well	240	32
100mm = 55cm^2	5.8x10^6 cells/well	86	12
60mm = 21cm^2	2x10^6 cells/well	33	4.5
6 well = 9.5cm^2	1x10^6 cells/well	15	2
12 well = 3.8cm^2	4x10^5 cells/well	6	0.8
24 well = 1.9cm^2	2x10^5 cells/well	3	0.4
48 well = 0.95cm^2	1x10^5 cells/well	1.5	0.2
96 well = 0.32cm^2	4x10^4 cells/well	0.5	0.07

Table 2.1: Concentrations of cells, dsRNA and DNA used in various cell culture plates.

2.3 Protein analysis

2.3.1. Enzyme Linked ImmunoSorbent Assay (ELISA) Buffers:

- *ELISA lysis buffer*: 1mM $MgCL_2$, 0.1% (w/v) BSA, 0.5% Triton-X 100 supplemented with cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche #11836170001).
- *ELISA wash buffer*: 0.2% (w/v) BSA, 0.5% Triton-X 100 in PBS.
- *HRP assay buffer*: 51mM Na_2HPO_4, 27mM citric acid, pH 5.0 filtered (0.2µm).
- *HRP developing solution*: 0.012% H_2O_2, 0.4mg/ml o-phenylenediamine in HRP assay buffer (prepared fresh for each assay).

2.3.1.1 Anti-GFP ELISA

The concentration of Upd2-GFP in condition media was measured using the anti-GFP ELISA described by Wright et al., 2011. A 96-well high-binding EIA plate, from Costar, was coated with 0.0625µg/mL goat anti-GFP antibody in 100mM Sodium Bicarbonate ON at 4ºC. The plate was washed 3x with wash buffer and then blocked in wash buffer for 1h at RT or O/N at 4ºC. Condition media was serially diluted across the plate and incubated for 3h at 37ºC. Recombinant GFP (Cellbiolabs STA-201) was also serially

diluted across the plate, starting at 5ng/mL, for reference. The plate was then washed and incubated with rabbit anti-GFP (Ab290) at 1:20,000 for 2h at RT or O/N at 4ºC. After a further wash the plate was incubated with the secondary HRP-linked anti-rabbit antibody (sc-2004) at 1:5000 for 1h at RT. The plate was then washed three times in wash buffer, followed by 3 washes in PBS. 200µL/well of freshly prepared HRP developing solution was added to the plate and colour change was observed. To stop the reaction 50µL/well 2M H_2SO_4 was added and the absorbance read at 492nm. A standard curve for recombinant GFP was made to allow calculation of Upd2-GFP concentration.

2.3.1.2 Endocytosis assay

Cells were seeded in a 24 well plate ($2x10^5$ cells/well) a day prior to experiment. Media was replaced with media conditioned with Upd2-GFP and incubated in a 25ºC incubator for desired times. Endocytosis was stopped by placing cells on ice and washing twice with ice-cold PBS. Cell-surface ligand was removed by 2x acid washed with 0.2M glycine, 0.15M NaCl pH2.5 for 2mins. Cells were then washed again in PBS before lysis in ELISA lysis buffer.

2.3.2. Protein biotinylation

2.3.2.1 Labelling of cell-surface Dome-flag

All reagents and protocol stages were carried out on ice unless specified. S2R+ cells were plated at $1x10^6$ cells in a 6 well plate and transfected with 2µg *pAc-Dome-flag* for 48hrs. Growth media was aspirated and cells were washed 2x with ice-cold PBS. EZ-link™ Sulfo-NHS-SS-Biotin (Thermo Scientific™ 21331) was made up immediately prior to use at 0.25mg/mL in PBS. Cells were incubated for 1hr on ice before biotin was quenched by washing twice with PBS containing 100mM glycine.

2.3.2.2 Internalisation of cell-surface receptor

Biotinylated cell surface proteins were allowed to internalise for various time-points by adding pre-warmed Upd2-GFP and incubating at 25ºC. Cells were then placed back on ice and washed 2x PBS. Cell surface biotin was cleaved by washing cells 3x for 20mins in MesNa buffer (50mM Tris-HCL pH8.6, 100mM NaCl, 1mM EDTA, 0.2% (w/v) BSA and 100mM 2-mercaptoethanesulfonate, which is added fresh to the buffer for each incubation). Cells were then washed 3x in PBS. Reduced disulphide bonds

were alkylated for 10mins with 500mM Iodoacetamide (IAA) made up in PBS, before 2 final PBS washes. Cells were then lysed for 30mins in lysis buffer (Chapter 2.1). Lysates were centrifuged at 13,000rpm for 10mins to remove insoluble protein and concentration was measured using Bradford assay.

2.3.2.3 Streptavidin-agarose pulldown

15µL streptavidin-agarose was washed 3x with lysis buffer and incubated with cell lysate (10-30µg) overnight at 4ºC with rotation. Beads were then washed 3x with lysis buffer and boiled for 5mins at 95ºC in 20µL 4x Laemmli SDS-PAGE buffer. For western blot analysis, an equal concentration of input/cell lysate is loaded in a separate lane to allow calculation of the proportion of internalised protein.

2.3.3. Luciferase assay

Cells were set up in a 12 well plate at $5x10^5$ cells/well a day prior to transfection. Cells were transfected with 0.5µg *10xSTAT92E-luciferase* and 0.5µg *pAct-Renilla* for 1 day then transferred to a 96well plate at $5x10^4$ cells/well. Cells were treated with conditioned media for 18hrs. The assay was carried out using the Dual-Glo® Luciferase Assay System (#E2920, Promega), following manufacturer's instructions, but at a 1:5 dilution in distilled water. Briefly, the Dual-Glo® Luciferase Assay reagent was added to the plate at an equal volume to the culture media in the wells, and incubated for at least 10mins. The Luciferase firefly signal was then measured using the Thermo Scientific Varioskan Flash Luminometer. An equal volume of Dual-Glo® Stop & Glo® Reagent is then added and incubated for a further least 10mins. The Renilla firefly signal was then measured.

2.3.4. Bradford assay

Protein concentration was determined using Bio-Rad Protein Assay Reagent (#500-0006), which is based on the protein dye binding method of Bradford, 1976. When Coomassie Brilliant Blue G-250 dye binds to protein there is a shift in its absorbance from 465 nm to 595 nm, allowing measurement using a spectrophotometer. Bio-Rad Protein Assay Reagent concentrate was diluted 1:5 in MilliQ H_20. 2-5µL of protein sample was mixed with 1mL of dilute protein assay reagent in 1mL cuvettes and incubated for 5mins at room temperature before measuring the absorbance at 595nm.

BSA standards, serially diluted from 2-0.0625mg/mL, were used to create a standard curve from which protein concentration was calculated.

2.3.5. SDS-PAGE (Poly Acrylamide Gel Electrophoresis)

Gels were cast and run using Bio-Rad Mini-PROTEAN Tetra apparatus. The separating gel was created by mixing 8 to 12% Bis-Acrylamide, 375mM Tris-HCl pH8.8, 0.1% (w/v) SDS, 0.05% (w/v) Ammonium persulfate (APS) and 0.05% (v/v) Tetramethylethylenediamine (TEMED), which was covered with a layer of isopropanol to allow for polymerisation. After washing the polymerised gel in H_2O, the stacking gel made with 4% Bis-Acylamide, 125mM Tris-HCl, pH6.8, 0.1% (w/v) SDS, 0.05% (w/v) APS and 0.05% (v/v) TEMED. This was layered on top of the separating gel, followed by a plastic comb to produce wells. Samples were boiled for 5mins at 95˚C in SDS gel loading buffer, prior to loading. Electrophoresis was carried out at 100-130V in SDS-PAGE running buffer (Chapter 2.1).

2.3.6. Western blotting

Protein samples separated by SDS-PAGE were transferred to 100% nitrocellulose blotting membrane (pore size: 0.45 µm, GE Healthcare #10600007) at 100V for 75mins in transfer buffer (Chapter 2.1). Following transfer, membranes were blocked for 1hr at room temp, or overnight at 4˚C, in 5% (w/v) skimmed milk powder in TBST. Membranes were then incubated with desired primary antibodies (Chapter 2.3.2) diluted in blocking buffer, overnight at 4˚C. After 3x TBST washes membranes were incubated for 1hr at room temp with appropriate LI-COR secondary antibody at 1:20,000. Following a further 3x TBST washes, membranes were washed 1x MilliQ H_20 and dried between sheets of Whatman blotting paper. Membranes were imaged using the LI-COR Odyssey® Sa and analysed using Image Studio™ Lite.

2.1.1 Calf intestinal alkaline phosphatase (CIAP) treatment

To remove phosphate groups cells were lysed in lysis buffer, without sodium orthovandate and EDTA, as these inhibit CIAP. Lysates were immunoprecipitated with STAT92E antibody as described previously. After washes with lysis buffer 1unit CIAP (#M0290S NEB) per 1µg protein was incubated in final 1x buffer.

2.4 Mass spectrometry

2.4.1. Preparation of cell lysates and immunoprecipitation for mass spec analysis

Cells were washed 2x in PBS and lysed for 30mins on ice in lysis buffer (Chapter 2.1). Lysates were cleared of insoluble material by spinning at 13,000rpm for 10mins and collecting the supernatant.
For immunoprecipitation of overexpressed STAT92E-GFP, cell lysates (2mg) were incubated with 15µL pre-washed GFP-Trap® agarose beads (chromotrek) for 2hrs at 4°C with rotation. Beads were then washed 3x with TBS. For analysis by western blot the sample was boiled for 5mins at 95°C in 15µL Laemmli SDS-PAGE buffer. For analysis by mass spectrometry beads were resuspended in 30µL 100mM ammonium bicarbonate and stored at -20°C.

2.4.2. Preparation of peptides

To reduce the proteins 1µL of 50mM tris(2-carboxyethyl)phosphine (TCEP) was added directly to the beads for 10mins at 70˚C. The reaction was then allowed to cool before the proteins were alkylated with 2µL of 50mM IAA for 30mins at RT, in the dark with shaking. An on-bead digest was then performed for 4hrs or O/N with 20µL of 0.1µg/µL trypsin or O/N with 20 µL of 0.1 µg/µL GluC at 37˚C. Digestion was stopped by lowering the pH to ~3 with 10% formic acid.

Peptides were purified and concentrated using Pierce C18 spin columns (#89870), following manufacturer's instructions. Briefly, spin columns were washed twice with 200µL of 50% acetonitrile, then primed for binding with two washed with 200µL of 0.5% formic acid. Samples (pH3) were then loaded to the spin columns, retained after the spin and reloaded. Unbound material was removed with 2 washes with 200µL of 0.5% formic acid. Peptides were eluted with 2x 20µL of 70% acetonitrile. All steps were centrifuged at 1500*xg* for 1min. Samples were then vacuum dried.

2.4.3. Phosphoenrichment

Phosphorylated peptides were enriched using Thermo Scientific™ HyperSep™ SpinTip Microscale Titanium Dioxide Tips (60109-422). Packed porous titanium dioxide was removed from tips and resuspended 1:1 in 0.5% formic acid. 5µl of this suspension was then placed back in the tip. By decreasing the volume of packed TiO2 we aimed to reduce background binding of unphosphorylated proteins. Tips were then washed 3x in 50µl 0.5% formic acid with a 1min centrifugation at 1500*xg*. Predigested, desalted and dried samples were resuspended in 20µl 0.5% formic acid and sonicated in a water bath. Samples were then loaded into the tips and centrifuged at 1500*xg* for 1min. Flow-through was reloaded to the tip twice to increase binding of phosphopeptides. Tips were then washed 3x in 50µl 0.5% formic acid with a 1min centrifugation at 1500*xg*. Bound peptides were then eluted using 50µl 5% NH4OH. This was repeated and elution fractions were combined prior to drying in the speed vacuum.

2.4.4. LC-MS/MS Analysis

Vacuum dried samples were resuspended in 40µL 0.5% FA, vortexed, sonicated for 2 minutes in a water bath and centrifuged for 5 minutes at 13000*xg*. 20 uL was placed in an auto-sampler vial, and 18uL was used for LC-MS/MS analysis with a Dionex Ultimate 3000 uHPLC system hyphenated to a Thermo Scientific Orbitrap Elite mass spectrometer. Peptides were injected onto a C18 trap-column (#164535, Thermo Fisher) and then loaded onto a 15cm EASY-Spray LC column (#ES801, Thermo Fisher) for analysis using a 60 minute gradient of increasing ACN(2.4 – 28% ACN/0.1% FA). Eluted peptides were ionised using the EASY-Spray™ Ion Source, precursor ion scans (375-1600 m/z) were acquired in the Orbitrap and the top 20 most abundant precursor ions (2+ and higher charge states) were fragmented by collision-induced dissociation (CID) and fragments were detected in the linear ion trap.

2.4.5. Data processing

Raw data was processed using MaxQuant (version 1.5.5.1, Max Plank Institute of Biochemistry, Martinsried, Germany) using default settings, unless specified. Protein sequences of the Drosophila melanogaster proteome downloaded from Uniprot (25-08-2016) along with the sequence of STAT-GFP were utilised for protein identification.

Carbamidomethylation of cysteines was set as a fixed modification due to reduction via IAA, oxidation of methionine and phosphorylation of serine, threonine and tyrosine were set as variable modifications. iBAQ (intensity-based absolute quantification) and LFQ quantification was performed using default settings. Identification of phosphorylation sites was verified through manual inspection of annotated spectra using pLabel (version 2.4). Files were converted from .raw into .mgf using MSConvertGUI.

2.4.6. Statistical analysis

Raw and LFQ intensities were filtered and analysed using Perseus (version 1.5.5.3, Max Plank Institute of Biochemistry, Martinsried, Germany). Proteins “only identified by site” and reverse matched were removed. Only proteins that were identified in 3 repeats of at least 1 group were included in further analysis. Values were transformed by log2(x) and missing intensity values (NaN) were imputed using values from the normal distribution (width = 0.3, down shift = 1.8). Groups were compared using volcano plots and unpaired Students t-test. A fold change of 1.5 (log2) was utilised for identification of significant protein abundance changes.

2.5 Microscopy

2.5.1. Immunostaining of S2R^{+} cells

Poly-L-lysine was diluted 1:10 in sterile Milli Q H_20 and incubated with coverslips for 30mins, before 3x wash with sterile Milli Q H_20. Cells were plated on sterile Poly-L-lysine coverslips for at least overnight. Cells were fixed in 4% Paraformaldehyde for 20mins, then quenched with 2x 5mins washes of 50mM Ammonium Chloride in PBS. Cells were then permeabilised in 0.2% Triton X-100 for 5mins, washed 3x in PBS and blocked in 0.2% Fish Skin Gelatine (FSG) for 1hr. The primary antibody was diluted in 0.2% FSG to desired concentration and coverslips were incubated for 1hr at RT. Coverslips were then washed 3x in 0.2% FSG and incubated with secondary antibody for 45mins in the dark. To stain cell nuclei, a final concentration of 1µg/mL DAPI was added for 5mins. Coverslips were then washed a further 3x in 0.2% FSG, followed by 3x washed in Milli Q H_20 before mounting onto slides with ProLong™ Gold Antifade Mountant (ThermoFisher Scientific #P10144).

2.5.2. Primary antibodies

Antibody	Antigen	Species	Source	Application
Anti-GFP	Recombinant full length GFP	Rabbit polyclonal	Abcam (ab290)	**ELISA** 1:20,000 **WB** 1:5000
Anti-GFP	Full length GFP	Goat polyclonal	Abnova (PAB10341)	**ELISA:** 0.0625µg/mL **IF:** 1:2000
Anti-Hrs (27-4)	Drosophila Hrs (N-term, amino acids 2-647 of NP_722831.1)	Mouse monoclonal IgG2a	DSHB*	**IF:** 1:50 **WB:** 1:1000
Anti-Rab5	Drosophila Rab5 (C-terminal)	Rabbit polyclonal	Abcam (ab31261)	**IF:** 1:1000
Anti-Rab7	Drosophila Rab7 (amino acids 184-200 of NP_524472.1)	Mouse monoclonal IgG1	DSHB*	**IF:** 1:50 **WB:** 1:1000
Anti-STAT92E (dN-17)	STAT92E (N-terminal)	Goat polyclonal IgG	Santa Cruz Biotechnology (sc-15708)	**WB** 1:1000 **IP** 0.5µg per 20µL protein G beads and 250µg lysate
Anti-β-actin	β-actin	Mouse monoclonal	Sigma (A1978)	**WB** 1:1000
Anti-Flag		Mouse	Sigma	**IF** 1:2000 **WB** 1:2000

Table 2.2: List of antibodies, application and dilutions used during this study. *DSHB: Developmental Studies Hybridoma bank, University of Iowa. Antibodies deposited by Sean Munro (MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus) (Riedel et al, 2016).

2.5.3. Secondary antibodies

Alexaflour secondary antibodies used for immunofluorescence were used at 1:1000 dilution. Licor secondary antibodies for western blots were used at 1:10,000.

2.5.4. Widefield imaging

A DeltaVision/GE Healthcare OMX optical microscope (version 4) with oil-immersion objective (60x NA 1.42, PlanApochromat Olympus) was used for widefield and SIM immunofluorescence image acquisition. Deconvolution and image registration (for alignment of SIM images) was carried out using the DeltaVision OMX softWoRx 6.0 software.

2.5.5. ImageJ analysis

Analysis of microscopy images was carried out using ImageJ.

2.5.5.1 STAT-GFP nuclear accumulation

Four regions of interest (ROI) of equal size were drawn within each transfected cell; two within the nucleus and two within the cytoplasm. Intensity measures were averaged for the nucleus and divided by the average intensity for the cytoplasm.

2.5.5.2 Colocalisation analysis

To quantify colocalisation ROIs were drawn around individual cells and Pearson's correlation coefficient was generated using the script found in Appendix 1. Further colocalisation analysis was carried out using Segmentation and quantification of subcellular shapes (Squashh), as described by Rizk et al, 2014. This is an ImageJ plugin produced by the MOSAIC group (Center for Systems Biology, Dresden)

2.6 RNA manipulation and analysis techniques

2.6.1. dsRNA design and amplification

PCR products were used as templates for dsRNA production. They were obtained from the Sheffield RNAi Screening Facility whose dsRNA database is based on the HD2.0 generated with Next-RNAi (Horn et al., 2010). These PCR templates (Table 2.3) are designed to contain a tag at either end in order to amplify with common primers, TU, TS1, TS2, TS3 and TS4 (see Chapter 2.6). The templates were amplified using ThermoPrime 2x Reddymix PCR mastermix (ThermoPrime AB0575DCLD).

dsRNA target	BKN ID	Size	Reverse primer
Alpha-Adaptin	BKN20148	764	TS4
Dome	BKN25660	351	TS2
Hrs	BKN27923	470	TS3
TSG101	BKN28961	333	TS1
Rab5	BKN22991	331	TS1
Rab7	BKN28849	161	TS3
STAT92E	BKN20615	402	TS1

Table 2.3: Details of dsRNA used during this study.

The MEGAscript® RNAi Kit (Life Technologies) was used to amplify dsRNA from appropriate PCR products, following manufacturer's instructions. The following reaction mix was prepared on ice:

PCR product	3µL
each dNTP	2µL
10x T7 Reaction Buffer	2µL
T7 Enzyme Mix	2µL
Nuclease-free Water	10µL

As products were generally <500bp, the reaction was incubated ON at 37ºC. 1µl of DNAseI was then added to digest the remaining PCR product for 30min at 37ºC. dsRNA incubated with 56µL 96% ethanol and 2µL 3M sodium acetate for 30mins at -80ºC, before ethanol precipitation and resuspension in sterile, nuclease and RNase free water.

2.6.2. RNA extraction

RNA extraction was carried out using TRI Reagent (Sigma #T9424), which lyses cells whilst inhibiting RNases. 1mL of TRI Reagent was added to cells plated in a 6-well plate for 10mins at room temperature. The TRI Reagent was then transferred into RNase free Eppendorf's, 200µL of chloroform was added and samples were vortexed for 20secs. After incubation for 5mins at room temperature samples were centrifuged at 13,000rpm for 15mins at 4˚C. Centrifugation separates the mixture into 3 phases: a red organic phase (containing protein), an interphase (containing DNA), and a colourless upper aqueous phase (containing RNA). The aqueous phase was transferred to fresh RNase free Eppendorfs and 500µL of isopropanol was added. After a 20secs vortex and incubation at room temperature for 10mins, samples were spun at 13,000rpm for 10mins at 4˚C. Supernatant was removed from the precipitated RNA pellet, which was washed with 1mL of 75% ethanol with a 13,000rpm spin for 5mins at 4˚C. The RNA pellet was then air dried for 10mins, before resuspension in 10µL RNase free H_2O.

2.6.3. Reverse transcription

Reverse transcription of RNA was carried out using the High-Capacity RNA-to-cDNA™ Kit (Applied Biosystems #4387406). 1.5µg of RNA (measured on nanodrop) was converted to cDNA following manufacturer's instructions. Briefly, the following reaction was incubates at 37˚C for 1hr, then heated to 95˚C for 5mins.

2x Buffer mix	10µL
20x RT Enzyme mix	1µL
RNA (1.5µg)	upto 9µL
Nuclease-free Water	upto 20µl

For downstream qPCR cDNA was diluted 1:10 in H_2O (to ~7.5ng/µL), as components from the RT reactions, such as salts, can interfere with efficient qPCR amplification.

2.6.4. Quantitative PCR (qPCR)

Relative mRNA levels were quantified using qPCR. This was performed in the BioRad CFX96 Real time system, C100 Touch™ thermal cycler or the Applied Biosystem® QuantStudio 12K Flex. 10µL/well total reaction was performed in a High Profile Semi-Skirted 96well plate (BIO RAD #HSS-9645) or Applied Biosystems® MicroAmp® Optical 384-Well Reaction Plate (#4343370), respectively. A standard curve of 4x 1:10 dilutions to cDNA was added to each plate to identify linear range and ensure efficient amplication.

1-2µL Template cDNA (dependent on target)
1µM Forward primer
1µM Reverse primer
5µL SYBR® Green JumpStart™ Taq ReadyMix™ (Sigma #S4438)

Cycling parameters:

Segment	**Cycles**	**Temperature**	**Time**
1	1	95˚C	3min
2	42	90˚C	30secs
		62˚C	30secs
			Plate read
3	1	68˚C	10mins

2.6.5. Microarray analysis

2.6.5.1 RNA preparation and chip hybridisation

RNA was prepared for hybridisation to the GeneChip Drosophila Genome 2.0 array by the Sheffield Institute for Translational Neuroscience (SiTRaN). The procedure was carried out using Affymetrix GeneChip 3' IVT PLUS Reagent Kit, and followed the manufacturer's instructions.

Briefly, 200ng of RNA is mixed with Poly-A RNA control samples containing *B. subtilise* genes that are absent from eukaryotes, allowing for monitoring of target preparation. Samples undergo two rounds of synthesis in order to produce single stranded cRNA of the correct orientation. First-strand cDNA synthesis is first carried out to produce ss-cDNA containing a T7-promotor sequence at its 5'end, which is then used to as a template and ds-cDNA is produced. Using the T7 RNA polymerase, which catalyses the RNA formation 5' to 3', biotinylated cRNA is synthesised from the second-stranded cDNA template via *in vitro* translation. cRNA is then purified with magnetic beads and fragmented with divalent cations and elevated temperature. Fragmented, biotin-labeled cRNA is mixed with a hybridisation cocktail, loaded onto the microarray and incubated for 16hrs at 45˚C with rotation at 60rpm. Arrays are then washed and stained prior to scanning with an Illumina HiScanHQ system.

2.6.5.2 Microarray analysis using R

The R software suite (version 3.4.2) was used to quantify and analyse microarray expression data. Bioconductor packages affy, gplots, drosophila2cdf, oligo, puma and lima were utilised. Full scripts can be found in Appendix 2.

2.7 Molecular Biology

2.7.1. Transformation of competent E. *coli*

In-house chemically competent DH5α E. coli were transformed as follows: Cells were thawed on ice and 20-50µL per reaction aliquoted into 14mL bug tube. ~10ng of plasmid DNA (see Chapter 2.7) was added to cells and incubated on ice for 20mins. Cells were then heat shocked at 42˚C for 45secs and returned to ice for 2mins before addition of 500µL pre-warmed LB containing no antibiotics. Cells were then shaken at

200rpm for 1hr at 37˚C. Cells were pelleted at 4000x*g* for 5mins, resuspended in 50µL LB and spread onto LB agar plates supplemented with appropriate antibiotic. The plate was incubated overnight at 37˚C and single colonies were selected.

2.7.2. Site-directed mutagenesis

Site-directed mutagenesis was carried out using the QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies) according to manufacturer's instructions. Primers used to create mutations were designed using PrimerX (www.bioinformatics.org/primerx/) and are listed in Table 2.3. The mutant strand synthesis reaction was carried out at a volume of 50µL and contained 5µL 10x Pfu ultra buffer, 10ng DNA template, 125ng forward primer, 125ng reverse primer, 1µL dNTP mix (40mM) and 3µL QuickSolution. 1µL of Pfu ultra HF DNA polymerase was added to the reaction, and all reactions were carried out with a minus polymerase control.

Cycling parameters:

Segment	**Cycles**	**Temperature**	**Time**
1	1	95˚C	1min
2	18	95˚C	50secs
		60˚C	50secs
		68˚C	90secs/Kb
3	1	68˚C	10mins

Parental, methylated DNA was digested with 2µL of DpnI at 37˚C for 30mins. The remaining mutated DNA was transformed into XL10-Gold® Ultracompetent Cells (Agilent Technologies #200314), following manufacturer's instructions. Cells were spread on LB agar plates supplemented with appropriate antibiotic. The plate was incubated overnight at 37˚C and single colonies were selected.

Gene	**DNA template**	**Mutation**	**Primer**	**Sequence**
STAT	*pAc-STAT-GFP*	Tyr 704 to Phe	Y704F-F	GATCCTGTGACCGGTTTCGTGAAGAGCACATTG
			Y704F-R	CAATGTGCTCTTCACGAAACCGGTCACAGGATC
		Lys 187 to Arg	K187R-F	GTATGGTCACACCCCGAGTGGAGCTGTACGAGGTC
			K187R-R	CACCTCGTACAGCTCCACTCGGGGTGTGACCATAC
		Thr 42 to Val	T47V-F	GATAATGTCCGAACAAATAGTGCCCAACACTACCGATCAG
			T47V-R	CTGATCGGTAGTGTTGGGCACTATTTGTTCGGACATTATC

		Ser 227 to Ala	S227A-F	CTAAACTCCACATCCGCGCCGAACGCAGAAG
			S227A-R	CTTCTGCGTTCGGCGCGGATGTGGAGTTTAG
		Thr 702 to Val	T702V-F	CGTCAAGATCCTGTGGTCGGTTATGTGAAGAGC
			T702V-R	GCTCTTCACATAACCGACCACAGGATCTTGACG
		Thr 702 to Glu	T702E-F	CGTCAAGATCCTGTGGACGGTTATGTGAAGAGC
			T702E-R	GCTCTTCACATAACCGTCCACAGGATCTTGACG
		Thr 702 to Asp	T702D-F	GCGTCAAGATCCTGTGGAGGGTTATGTGAAGAGCAC
			T702D-R	GTGCTCTTCACATAACCCTCCACAGGATCTTGACGC
Dome	*pAc-Dome-Flag*	diLeu 985 to diAla	LL985AA-F	GCTGTGCGCAGCGCCGCTTTCGAGCTCTCACTGCCGCAGC
			LL985AA-R	GCTGCGGCAGTGAGAGCTCGAAAGCGGCGCTGCGCACAGC
		Ser 980 to Ala	S980A-F	CTTCAGTAGCTGCGGCGCTGAGAGCTCGAAACTG
			S980A-R	CAGTTTCGAGCTCTCAGCGCCGCAGCTACTGAAG
		Ser 982 to Ala	S982A-F	GTAGCTGCGGCAGTGAGGCTTCGAAACTGCTG
			S982A-R	CAGCAGTTTCGAAGCCTCACTGCCGCAGCTAC
		Tyr 966 to Ala, Gln 969 to Ala	Y966A/ Q969A-F	GGCAGATCCCGCGGCGTGCGTGGCGGCGAGTC
			Y966A/ Q969A-R	GACTCGCCGCCACGCACGCCGCGGGATCTGCC
		Glu 981 to Ala	E981A-F	CTTCAGTAGCTGCGGCAGTGCGAGCTCGAAAC
			E981A-R	GTTTCGAGCTCGCACTGCCGCAGCTACTGAAG
	pAc-Dome^{LL}-Flag	Ser 980 to Ala	LLS980A-F	CTTCAGTAGCTGCGGCGCTGAGAGCTCGAAAG
			LLS980A-R	CTTTCGAGCTCTCAGCGCCGCAGCTACTGAAG
		Ser 982 to Ala	LLS982A-F	TTCAGTAGCTGCGGCAGTGAGGCTTCGAAAGCGGC
			LLS982A-R	GCCGCTTTCGAAGCCTCACTGCCGCAGCTACTGAA
		Glu 980 to Gly	E981ALL-F	CTTCAGTAGCTGCGGCAGTGGGAGCTCGAAAG
			E981ALL-R	CTTTCGAGCTCCCACTGCCGCAGCTACTGAAG
	pAc-Dome^{S1LL}-Flag	Ser 980 to Ala, Glu 981 to Ala **(all A)**	SE980AA-F	GCCGCTTTCGAAGCCGCAGCGCCGCAGCTACTGAA
			SE980AA-R	TTCAGTAGCTGCGGCGCTGCGGCTTCGAAAGCGGC

Table 2.4: Primers used for mutagenesis

2.7.3. Glycerol stocks of transformed E. *coli*

500µL of overnight culture was added to 500µL sterile 30% (w/v) glycerol and stored at -80˚C. To recover strain a sterile pipette tip was used to scrape off ice from the top of the vial and placed in LB or spread on agar plates containing appropriate antibiotic.

2.7.4. Plasmid purification from E. *coli*

Small scale plasmid purification (up to 20µg DNA) was carried out using the QIAprep Spin Miniprep Kit (Qiagen #27104), following manufacturer's instructions. 3mL of overnight bacterial culture was spun at 7000*xg* for 3mins.
Large scale plasmid purification (up to 1mg DNA) was carried out using the PureLink® HiPure Plasmid Filter Maxiprep Kit (Invitogen #K210017). 300mL of overnight bacterial culture was pelleted at 7000*xg* for 8mins, and purification was carried out following manufacturer's instructions.

2.7.5. Genomic DNA isolation

Isolation of genomic DNA from cell lines was carried out using DNAzol™ (Invitrogen #10503027). 1mL of resuspended cell culture was centrifuged at 1000*xg* for 3mins, supernatant removed and pellet resuspended in 1mL DNAzol™. 0.5mL of 100% ethanol was then added, tube mixed via inversion and incubated at RT for 3mins. Precipitated DNA was swirled onto a pipette tip and transferred to a new Eppendorf. DNA was then washed twice with 0.8mL 75% ethanol and allowed to air dry before resuspension in 8mM NaOH. After solubilisation 6.4µL 1M HEPES (free acid) was added to neutralise. DNA was stored at -20˚C.

2.7.6. Agarose gel electrophesis

Agarose gels were prepared by melting required mass of agarose in 50-100mL TAE buffer in microwave. To visualise the DNA 10mg/mL of ethidium bromide or 1:10,000 SYBR® Safe DNA gel stain (Invitrogen #S33102) was added to molten agarose before pouring into casting apparatus. Gels were run at 100v in a Bio-Rad Sub-Cell® GT containing TAE.

2.7.7. DNA sequencing

Sequencing of plasmid DNA was carried out at the University of Sheffield's Core Genomic Facility using BigDye v3.1 and results were analysed using ApE.

2.7.8. Production of sgRNA and ligation into pAc-sgRNA-Cas9 for creation of CRISPR S2R$^+$ cell lines

sgRNA for gene knockouts were designed to target the N-terminal coding region. For *STAT92E* the mRNA sequence of all 10 isoforms were aligned and common sequences were used in order to target all splice variants. mRNA sequences were loaded into crispr.mit.edu, created by the Zhang laboratory at MIT. Here, the sequence is scanned for potential sgRNA, 20bps upstream of NGG, whilst also comparing off-target mismatches in the chosen genome (fly, dm6). The output provides a score which relates to "the faithfullness of on-target activity computed as 100% minus a weighted sum of off-target hit-scores in the target genome". sgRNA sequences used in this study all had scores of 99, and therefore were unlikely to target other genomic regions. Sequences were also verified using NCBI blast to confirm potential off-targets. The NGG sequence was then removed, and a G was added to the 5' end of the sgRNA sequence to allow transcription from the U6 promoter in *pAc-sgRNA-Cas9* vector.

sgRNAs were cloned into the *pAc-sgRNA-Cas9* expression vector according to Bassett et al., 2013. Briefly, 10µL sgRNA oligos (100mM) were annealed in 20µL 20mM Tris, 2mM EDTA, 100mM NaCl, pH 8.0 by adding to a prewarmed 98˚C heat block, turning off the head block and allowing it to cool to room temperature. 1µL of annealed oligos were added to 8µL DNA ligase buffer and 1µL T4 polynucleotide kinase (NEB), incubated at 37˚C for 30mins, then diluted 10x in H_2O. 2µg of pAc-sgRNA-Cas9 was digested with 20 U *Bsp* QI (NEB) for 2hrs at 50˚C. The reaction was then treated with 10 U calf intestinal alkaline phosphate for 2hrs at 37˚C followed by purification using the QIAquick PCR Purification Kit (Qiagen #28104) following manufacturer's instructions. 50ng vector and 2µL diluted oligos were ligated with T4 DNA ligase (NEB) for 2hrs at 18˚C and transformed into chemically competent DH5α E. coli.

Positive colonies were selected with colony PCR. Single colonies were resuspended in 30µL sterile H_2O and a 20µL PCR reaction was set up as follows: 10µL resuspended colony, 4µL 5x One*Taq*® reaction buffer, 0.4µL dNTP mix (40mM), 0.4µL 10µM U6F primer, 0.4µL 10µM specific sgRNA reverse primer, 0.1µL One*Taq*® DNA polymerase.

Cycling parameters:

Segment	Cycles	Temperature	Time
1	1	94˚C	1min
2	30	94˚C	30secs
		65˚C	60secs
		68˚C	30secs
3	1	68˚C	5mins

PCR products were analysed with on a 1.5% agarose gel. Colonies showing a band at 418bp, demonstrating successful sgRNA insertion, were selected for DNA sequencing.

2.7.9. T7 endonuclease I reaction

To detect Cas9 induced mutations within the genomic DNA of S2R$^+$ CRISPR cell lines a T7 endonuclease assay was carried out to identify mismatched, heteroduplex, DNA. Firstly, PCR products were produced by amplifying a ~1kb region around the Cas9 cut site with a 50µL PCR reaction: 10µL 5x Q5 reaction buffer, 10µL Q5 GC enhancer, 1µL dNTP mix (40mM), 2.5µL 10µM forward primer, 2.5µL 10µM reverse primer, 100ng genomic DNA and 0.5µL Q5 high fidelity DNA polymerase (NEB #M0491S).

Cycling parameters:

Segment	Cycles	Temperature	Time
1	1	98˚C	30secs
2	35	98˚C	5secs
		68˚C	10secs
		72˚C	20secs
3	1	72˚C	2mins

5µL of PCR product was ran on a 2% agarose gel to verify band size. PCR products were then denatured and annealed to form heteroduplexes in the following reaction: 5-10µL PCR products, 2µL NEBuffer 2 made up to 19µL with nuclease free water. The reaction was held in a 95˚C heat block for 10mins and allowed to cool to room temperature by turning off the heat block. 1µL of T7 endonuclease was then added to

reactions and incubated at 37˚C for 15mins. The reaction was stopped by addition of 1.5µL of 0.25M EDTA before running on an agarose gel.

2.8 Primers

Primer name	Sequence (5'-3')	Source
dsRNA production		
TU	TAATACGACTCACTATAGGGTGGCGCCCCTAGATG	
TS1	TAATACGACTCACTATAGGGCGACGCCCGCTGATA	
TS2	TAATACGACTCACTATAGGGTAGGTCTAGCCCCGC	
TS3	TAATACGACTCACTATAGGGCGCATGTAGCCTGCC	
TS4	TAATACGACTCACTATAGGGTAGCCTCCCTAGCGC	
Quantitative PCR		
Rpl-F	GACGCTTCAAGGGACAGTATCTG	Vidal et al.,
Rpl-R	AAACGCGGTTCTGCATGAG	Vidal et al.,
SOCS36E-F	AGTGCTTTACTGCTGCGACT	Vidal et al.,
SOCS36E-R	TCGTCGAGTATTGCGAAGT	Vidal et al.,
Dome-F	ACTTTCGGTACTCCATCAGC	Vidal et al.,
Dome-R	TGGACTCCACCTTGATGAG	Vidal et al.,
AP2-F	AGCAGGCTCAGATGTTCCG	This study
AP2-R	CCCCGTCCATGATTGTTGTG	This study
Hrs-F	AGCAGTTTCCTCGAGTCGAC	This study
Hrs-R	CAGGATGGTCATGGTGTCCTTG	This study
TSG101-F	CGCCACTTTACCGACAGTTAC	This study
TSG101-R	CGGCCTTTTGTCTGCACTTC	This study
p53-F	CCAAGCTAGAGAATCACAACATCG	Zhang et al., 2014
p53-R	TCGAGTACATCCAAAGAGACTTGG	Zhang et al., 2014
SP555-F	GGATCAGATTGATGAGCCCCT	This study
SP555-R	GGTTGACGAACTCCCTCCTG	This study
CG5246-F	CCATTGCAGTCACGACAAGC	This study
CG5246-R	TGGAGTACCTTCCCCAGGTC	This study
Lama-F	TGATATTGCTGCTTCCTGGAC	Flaherty et al., 2009
Lama-R	TGGTTTGGCGATGGTTTTAT	Flaherty et al., 2009
Chinmo-F		
Chinmo-R		
Oligonucleotides for sgRNA production		
STAT1.1	TTCGACAACACGCCCATGGTTACC	
STAT1.2	AACGGTAACCATGGGCGTGTTGTC	
STAT2.1	TTCGACCATGTACCCGGTAACCAT	
STAT2.2	AACATGGTTACCGGGTACATGGTC	

Dome.1	TTCGCTGCTGCTCATGCTGCTTGC
Dome.2	AACGCAAGCAGCATGAGCAGCAGC

2.9 Plasmids

Plasmid	Backbone	Promoter	Insert	Source (#addgene)
pAc5-Upd2-GFP	*pAc5.1*	Actin-5c	Upd2-GFP	Hombria et al, 2005
pAct-RL	*pPac5c-PL*	Actin-5c	Renilla luciferase	Muller et al, 2005
10xSTAT92E-luciferase	*pUAST*	hsp70	10x SOCS36E enhancer upstream of firefly luciferase	Baeg et al, 2005
pAc-STAT-GFP	*pPac5c-PL*	Actin-5c	STAT92E isoform C cDNA + eGFP	Karsten et al 2006
pAc-Dome-Flag	*pAFW*	Actin-5C	Domeless	Stec et al, 2013
pAc5.1B-EGFP	*pAc5.1B*	Actin-5c	eGFP	Elisa Izaurralde (#21181)
pAc-sgRNA-Cas9	*pAc-STABLE1-Puro*	Actin-5C	Human codon optimised Cas9	Bassett et al, 2013 (#49330
pAc-STAT1-CRISPR	*pAc-STABLE1-Puro*	Actin-5C	Annealed STAT1 oligos	This study
pAc-STAT2-CRISPR	*pAc-STABLE1-Puro*	Actin-5C	Annealed STAT2 oligos	This study
pAc-Dome-CRISPR	*pAc-STABLE1-Puro*	Actin-5C	Annealed Dome oligos	This study

Chapter 3. Expression of JAK/STAT pathway transcriptional targets during endocytosis

Studies using *Drosophila melanogaster* to investigate JAK/STAT pathway signalling have suggested both positive and negative roles for endocytic regulation. Preliminary data from the Smythe lab indicated that only specific Rab5 GEFs influence JAK/STAT signalling (Vogt and Smythe, unpublished). Together these studies suggest a complex relationship between endocytosis and the JAK/STAT signalling pathway. Consequently, I wished to understand how JAK/STAT transcriptional targets are regulated during endocytic trafficking.

3.1 Differential expression of *10xSTATluciferase* reporter and *SOCS36E*

Drosophila cell lines are widely used for RNA interference (RNAi) mediated gene silencing due to ease of delivery that does not require transfection reagents. The Sheffield RNAi screening facility specialises in dsRNA-mediated knockdown in *Drosophila* cells, and routinely utilises a pre-designed dsRNA library. The Heidelberg 2 (BKN) library was designed using NEXT-RNAi software (Horn and Boutros, 2010), and excludes motifs that induce off-target effects. I therefore utilised dsRNA from this library to efficiently knockdown key components of the endocytic pathway (Figure 3.1). When these experiments were performed there were no antibodies available for these *Drosophila* proteins and therefore knockdown was measured using qPCR to investigate changes to mRNA levels of targets. Subsequently, an antibody specific for Hrs was produced (Riedel et al., 2016), enabling analysis of efficient knockdown at the protein level (Figure 4.2).

The *Drosophila* JAK/STAT receptor, Domeless, has previously been shown to undergo clathrin-mediated endocytosis (CME), in response to both Upd1 (Devergne et al., 2007) and Upd2-GFP (Vidal et al., 2010). Receptor-ligand complexes then traffic through early and late endosomes for lysosomal degradation (Devergne et al., 2007; Vidal et al., 2010; Stec et al., 2013). Therefore, dsRNA mediated knockdown of the AP2, Hrs and TSG101 (Figure 3.1) disrupts this trafficking pathway. Knockdown of

alpha-adaptin (AP2) prevents CME, and therefore will trap Dome at the cell surface. Hrs depletion prevents trafficking of Dome from early to late endosomes (Tognon et al., 2014), whereas TSG101 knockdown prevents multivesicular body maturation and therefore Dome will not be trafficked to the lysosome for degradation.

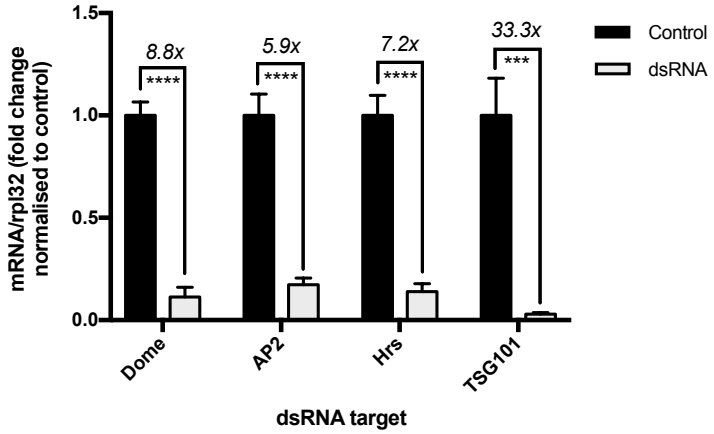


Figure 3.1: mRNA levels of dsRNA targets after knockdown. S2R+ cells were treated with dsRNA five days prior to TRIzol RNA extraction. mRNA levels were analysed using qPCR, with levels of target mRNA normalised to rpl32 mRNA. Ratios are plotted as fold change compared to control dsRNA for each target mRNA. Graph represents the mean of triplicates ± SEM for at least 2 independent experiments. Parametric, unpaired students t-test was performed to compare control knockdown with targeted dsRNA knockdown, with ***p≤0.001, ****p≤0.0001.

Confident that levels of dsRNA knockdown were significant (Figure 3.1), I investigated the effect of AP2, Hrs and TSG101 knockdown on Upd2-GFP induced JAK/STAT signalling. The *10xSTATluciferase* reporter (Baeg et al., 2005) is a well-characterised reporter with 10 potential STAT92E binding sites upstream of a Firefly luciferase, and is a particularly sensitive reporter of JAK/STAT activity. In this assay, an actin-driven *Renilla* luciferase reporter is co-transfected to allow for normalisation to the number of transfected cells. Upd2-GFP stimulated *10xSTATluciferase* reporter expression in a concentration-dependent manner that plateaued at ~10nM (Figure 3.2A). However, due to consistently low yields in Upd2-GFP preperation, (see Methods 2.2.2.1, Wright et al., 2011) 3nM Upd2-GFP was used throughout this study, unless otherwise stated. It was previously confirmed that the increase in luciferase expression is Upd2-dependent, and not triggered by the GFP tag (Vogt and Smythe, unpublished).

Utilising the luciferase reporter assay I confirmed that use of control dsRNA targeting *C. elegans* genes had no off-target effect on the ligand-induced *10xSTATluciferase* expression (Figure 3.2B). Importantly, depletion of Dome prevented ligand induced

luciferase activation, demonstrating that the response is receptor dependent. Dome depletion also reduced JAK/STAT signalling in mock treated cells. This may be because in culture S2R+ cells secrete all 3 Upd ligands (Cherbas et al., 2011), resulting in activation of the pathway prior to addition of Upd2-GFP. Using this system, I demonstrated that activation of the *10xSTATluciferase* reporter was altered only when AP2 was knocked down (Figure 3.2B), suggesting the receptor requires internalisation, or an association with the clathrin coat, for JAK/STAT signalling and luciferase expression. In contrast, knockdown of later endocytic proteins did not affect luciferase expression, suggesting signalling occurs at an early stage of the endocytic pathway.

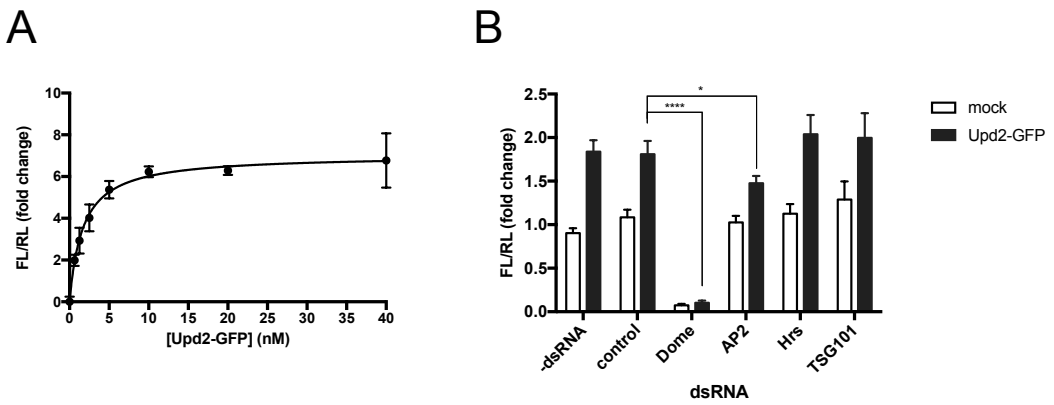


Figure 3.2: Expression of 10xSTATluciferase reporter is Upd2-GFP dependent and endocytically regulated. A) S2R+ cells were transfected with an actin driven Renilla Luciferase (RL) and 10xSTATLuc (FL) reporter construct for 6hrs before transferring to 96well plate. Cells were then treated with varying concentrations Upd2-GFP for 30mins, then incubated for 18hrs in fresh media before bioluminescence was measured. Graph represents mean of triplicates + SEM for 2 experiments; B) S2R+ cells were transfected with RL and FL for 6hrs prior to treatment with dsRNA and incubated for five days. Cells were treated with Upd2-GFP for 18hrs and then bioluminescence was measured. Data is expressed as the ratio of FL to RL and normalised to control, mock treated cells. Graph represents mean of triplicates+ SEM for 4 experiments. Parametric, unpaired students t-test carried out to compare Upd2-GFP stimulated samples only, with *p≤0.05, ****p≤0.0001.

As the *10xSTATluciferase* reporter is an artificial target of JAK/STAT pathway activation, I wanted to confirm that AP2 is also required for endogenous pathway activity. SOCS36E is negative regulator of the JAK/STAT pathway (Discussed in Chapter 1.1.3.1), yet is also a well-characterised downstream target of pathway activation, and has been shown to be highly sensitive to JAK/STAT signalling in *Drosophila* (Karsten et al., 2002; Bina et al., 2010). I first investigated the concentration dependence of Upd2-GFP induction on *SOCS36E* mRNA expression utilising qPCR

(Figure 3.3A). *SOCS36E* appears to be maximally expressed at 3nM Upd2-GFP whereas *Domeless*, alos a transcriptional target of the pathway (Hombría et al., 2005), seems to require higher concentrations of ligand for maximal expression (Figure 3.3B). Although this data is only from a single individual experiment, the result is plausible as *SOCS36E* and *Domeless* genes have different STAT92E DNA-binding sites in their regulatory regions. *Domeless* has two upstream STAT92E binding sites, containing a sequence with a N_4 spacer (Rivas et al., 2008). In contrast the *SOCS36E* enhancer contains sequences with a N_3 spacer region (Baeg et al., 2005; Müller et al., 2005; Bina et al., 2010), which STAT92E binds to with higher affinity (Rivas et al., 2008).

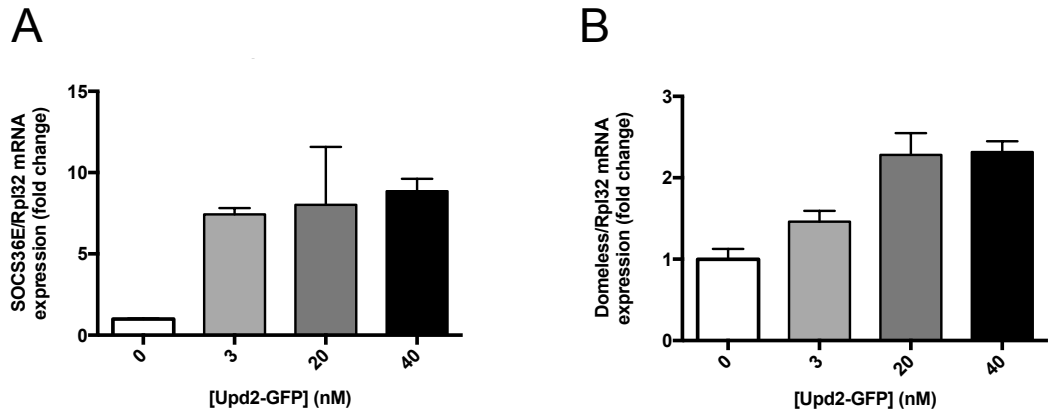


Figure 3.3: SOCS36E expression requires lower concentrations of Upd2-GFP compared to Domeless expression. S2R$^+$cells were stimulated with varying concentrations of Upd2-GFP for 2.5hrs at 25˚C. mRNA was harvested using TRIzol and analysed by qPCR. A) SOCS36E mRNA and B) Domeless mRNA levels where measured. Data is normalised to reference gene rpl32, and compared to 0nM Upd2-GFP. Data represents mean of triplicates + SD for one experiment.

I then investigated changes to *SOCS36E* expression under conditions where endocytosis had been perturbed. As with the exogenous reporter, knockdown of AP2 inhibited ligand induced-expression of *SOCS36E*. However, in contrast to the activation of the *10xSTATluciferase* reporter, knockdown of Hrs also inhibited *SOCS36E* expression (Figure 3.4). This suggests that the two targets are differentially regulated by endocytosis. dsRNA against TSG101 does not alter SOCS36E expression. Thus, for SOCS36E expression, the receptor needs to traffic through a Hrs positive endosome, but does not need to reach a TSG101 positive compartment.

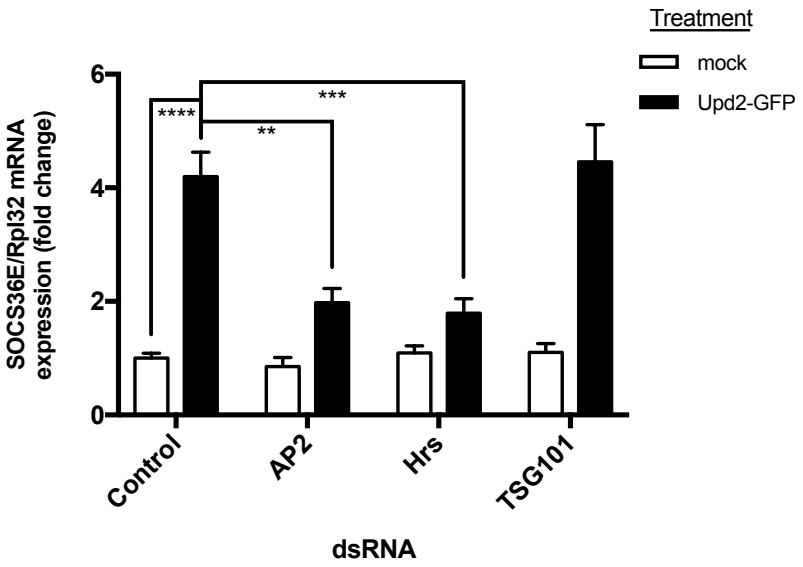


Figure 3.4: Endocytic trafficking regulates the expression of JAK/STAT pathway target SOCS36E. dsRNA knockdown of endocytic components in S2R+ cells was performed for 5 days. Cells were incubated with 3nM Upd2-GFP for 2.5hrs prior to RNA extraction. SOCS36E mRNA levels were normalised to that of reference gene Rpl32, and presented as fold change compared to mock-treated control samples. Results are expressed as means of triplicates + SEM for 3 independent experiments. Parametric, unpaired students T-test was carried out to compare Upd2-GFP stimulated samples only. **p≤0.01, ***p≤0.001, ****p≤0.0001 (P-values not stated in the figure were not significantly different).

Together this data suggests that endocytosis differentially regulates the expression of JAK/STAT pathways targets. The *10xSTATluciferase* reporter is expressed early after Upd/Dome complex internalisation requiring only internalisation or an association with AP2, whereas *SOCS36E* is expressed after trafficking through a Hrs-positive compartment (Figure 3.5). This is interesting as activation of the JAK/STAT pathway results in a broad range of cellular outcomes, and therefore suggests that endocytosis may provide a mechanism by which these broad outcomes can be fine-tuned in a specific context.

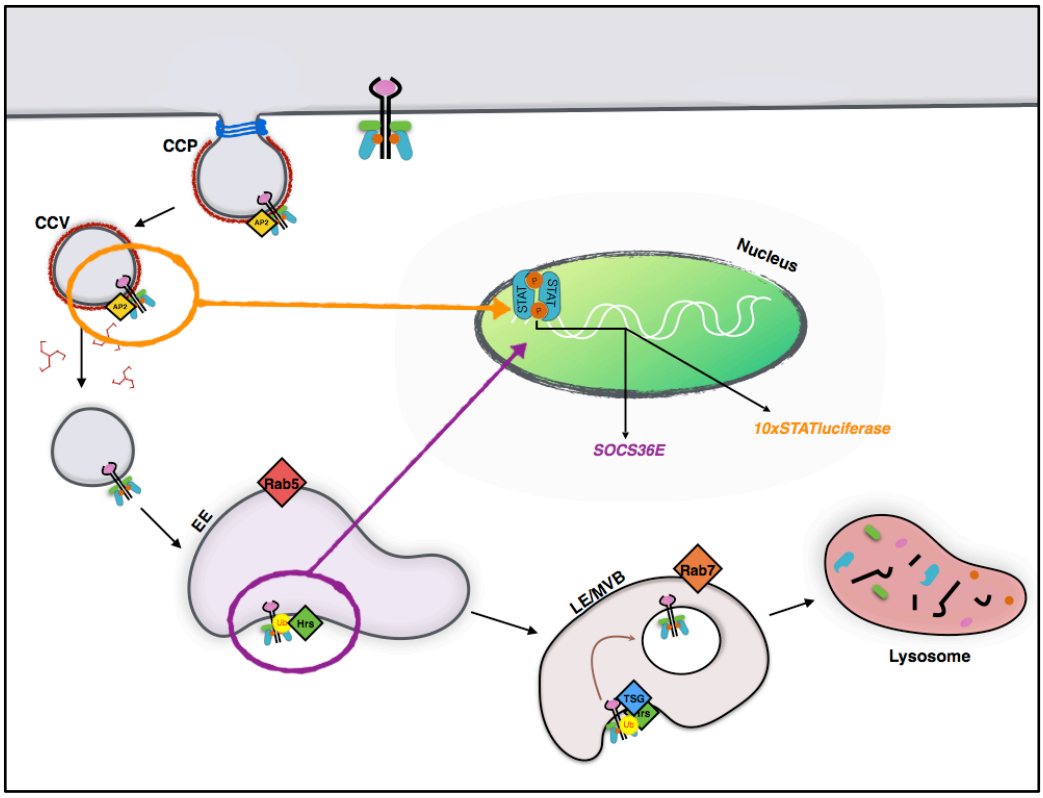


Figure 3.5: Schematic of differential JAK/STAT target expression during endocytic trafficking. 10xSTATluciferase expression occurs at an earlier stage of Upd2/Dome endocytosis than expression of SOCS36E.

3.2 Transcript profiling of endocytically regulated JAK/STAT targets

Having demonstrated that endocytosis qualitatively regulates two JAK/STAT pathway targets, I aimed to investigate how endocytosis regulates the expression of further endogenous JAK/STAT targets. To investigate changes in the transcriptome of S2R+ cells I used Affymetrix Genechip™ microarrays. These chips enable measurement of the expression level of thousands of genes simultaneously, and have been an important tool in understanding transcriptional changes in various biological systems. Unlike other available microarray array platforms which compare different signals on the same array, Affymetrix arrays allow generation of absolute expression values and therefore comparisons can be made between multiple arrays and conditions (Dalma-Weiszhausz et al., 2006).

When preparing samples for microarray analysis, I initially stimulated cells with 40nM Upd2-GFP to increase the likelihood of altering the expression of less sensitive

JAK/STAT targets. However, this caused *SOCS36E* expression to no longer be regulated by endocytosis (Figure 3.6A). This appears to be a concentration dependent effect (Figure 3.6B), which suggests that at high concentrations of ligand Dome is taken up via a different endocytic route (Discussed in Chapter 4). Therefore, I used 3nM Upd2-GFP for transcriptome analysis.

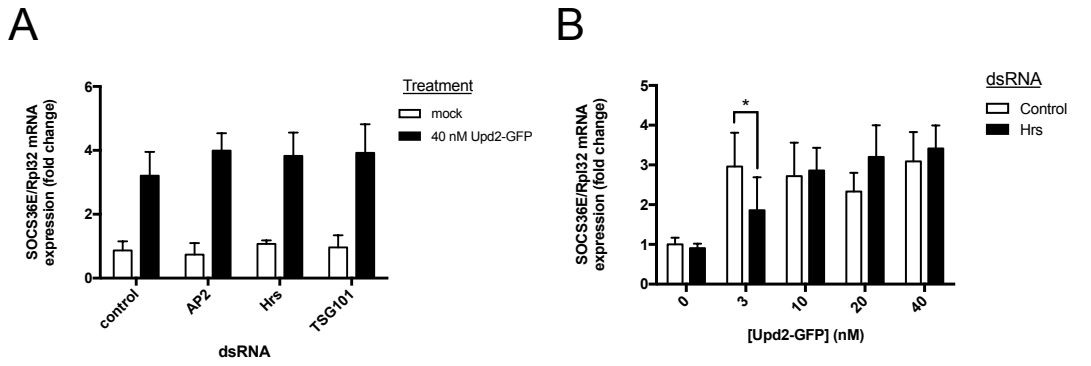


Figure 3.6: SOCS36E expression is only regulated by endocytosis at low ligand concentrations. A) S2R$^+$ cells were treated with targeting dsRNA for 5 days prior to stimulation with 40nM Upd2-GFP for 2.5hrs at 25˚C. B) S2R$^+$ cells were treated with dsRNA targeting Hrs for 5 days prior to stimulation with varying concentration of Upd2-GFP. mRNA was extracted with TRIzol and analysed with qPCR. Data is normalised to reference gene rpl32, and compared to mock or 0nM Upd2-GFP. Data is presented as mean of triplicates + SD from at least 2 independent experiments. Parametric, unpaired students t-test was carried out to compare samples, with *p≤0.05.

For transcript profiling S2R+ cells were treated, as for *SOCS36E* expression experiments (Figure 3.4), with dsRNA for 5 days prior to ligand stimulation and RNA extraction. Figure 3.7A represents the experimental procedure and quality control points used to ensure that the RNA was suitable for the downstream application. For the microarray analysis, I used 5 conditions: control dsRNA with or without Upd2-GFP treatment, AP2 dsRNA with treatment, Hrs dsRNA with treatment and TSG101 dsRNA with treatment. These conditions would allow me to search for targets whose expression changes upon addition of ligand in control cells, then compare these to changes in cells treated with dsRNA targeting endocytic components. Following extraction, extensive examination of RNA quality was carried out to ensure downstream results were trustworthy. Samples were measured on the Nanodrop and electrophoresed on a bleach agarose gel (Aranda et al., 2012) to check for contaminants and integrity, respectively. Due to *Drosophila* 28s rRNA being processed as two fragments, RNA appears as one main band on an agarose gel and not the two seen in other species (Winnebeck et al., 2010) (Figure 3.7B). Samples were finally

checked for efficient knockdown and for the phenotypic *SOCS36E* expression. I selected RNA from three independent experiments to take forward for microarray analysis (Figure 3.7C), to ensure true biological changes were investigated.

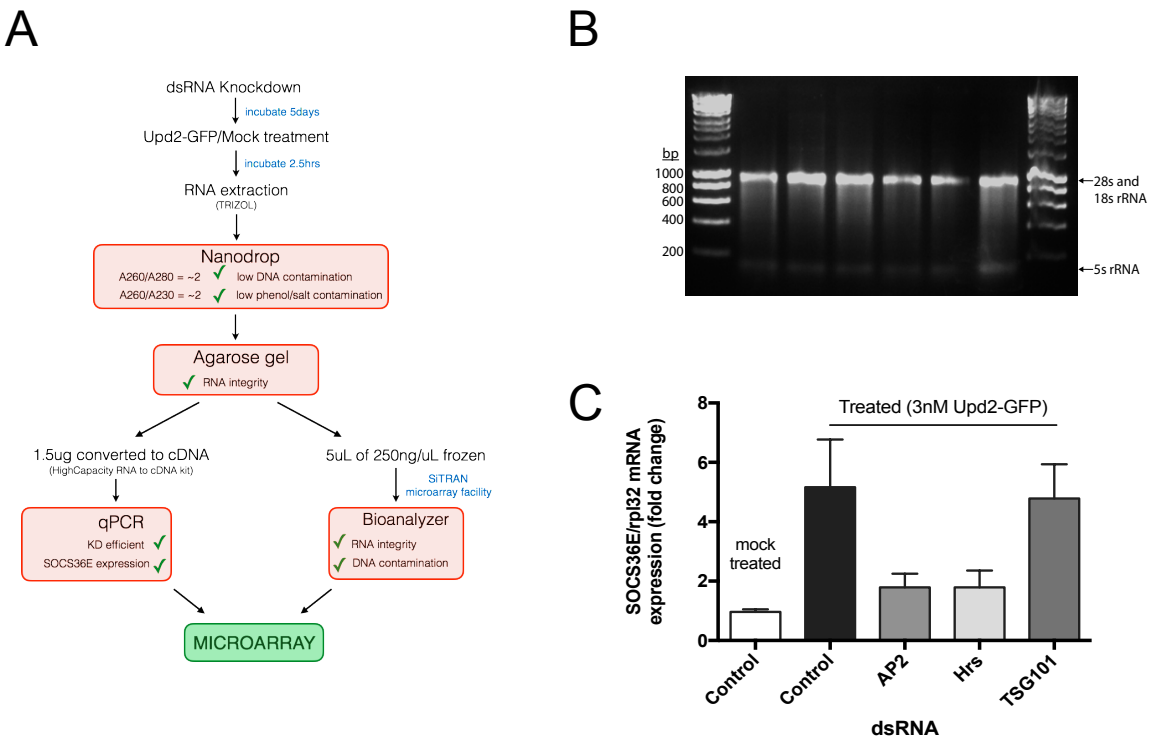


Figure 3.7: Preparation of samples for transcriptome analysis. A) Flowchart demonstrating procedure of RNA preparation and quality control points (in red boxes). B) Example agarose gel of extracted mRNA showing specific 28s and 18s rRNA band with minimal degradation. C) SOCS36E expression of RNA samples used for microarray. S2R^{+} cells were incubated with non-targeting or dsRNA targeting endocytic components for 5 days. 3nM Upd2-GFP stimulation was carried out for 2.5hrs prior to RNA extraction with TRIzol. SOCS36E mRNA levels were normalised to that of reference gene rpl32, and presented as fold change compared to mock-treated control samples. Data is presented as mean + SD of the 3 biological replicates taken for microarray analysis (each from a separate individual experiment).

At the Sheffield Microarray Core Facility RNA samples underwent further quality control for integrity and contamination using the Agilent BioAnalyzer 2100. The BioAnalyzer uses a micro-capillary electrophoretic cell that produces electropherograms and enables analysis of RNA integrity. An RNA integrity number (RIN) is calculated using the 28s and 18s rRNA. Figure 3.8A is an example of mammalian RNA with a high RIN. However, since Drosophila 28s rRNA is processed as two fragments, a RIN is not produced. Nonetheless, the electropherograms produced allowed confirmation that all samples used for the microarray were high-quality *Drosophila* RNA (Figure 3.8B and 3.8C). The RNA was then processed by the

Sheffield facility and hybridised onto GeneChip™ *Drosophil*a Genome 2.0 array (described in Chapter 2.6.5).

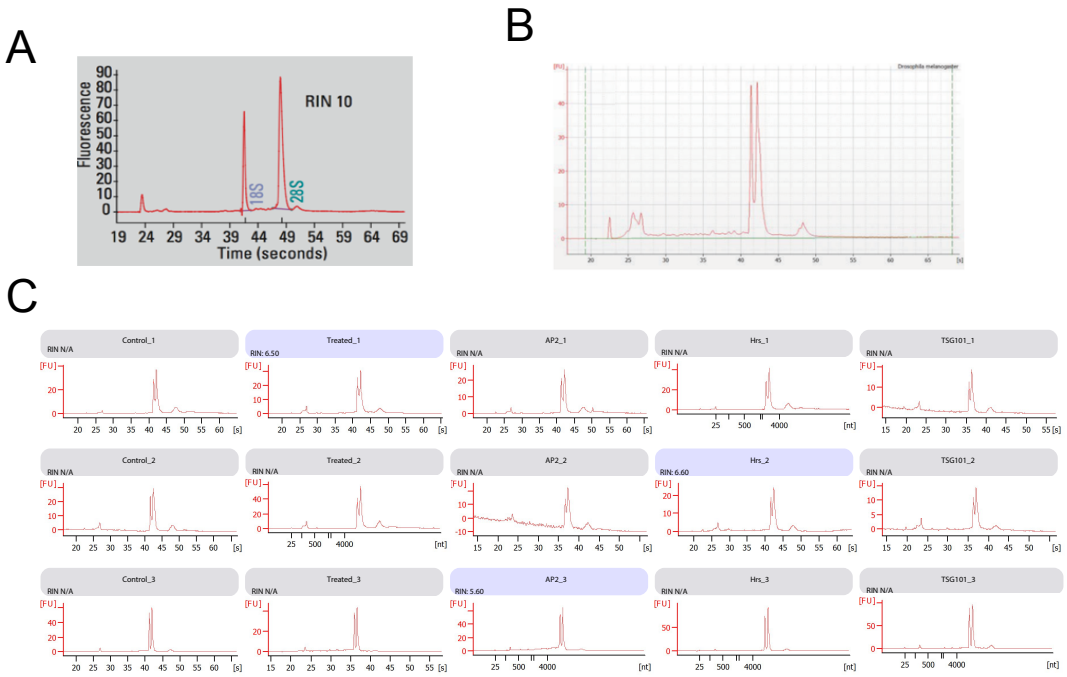


Figure 3.8: BioAnalyzer electropherograms show high RNA quality of samples. A) Example bioanalyzer result from mammalian RNA with RIN of 10. Bioanalyzer identifies 18S and 28S rRNA peaks in the electropherograms (from https://www.agilent.com/cs/library/applications/5989-1165EN.pdf). B) Example bioanalyzer results from high quality Drosophila Melanogaster RNA, where 18S and 28S rRNA peaks cannot be identified (from http://www.agilent.com/cs /library/applications/5991-7903EN.pdf). C) Electropherograms from RNA samples used for the microarray. All show high integrity although RIN could not be calculated.

3.1.1. Microarray analysis with PUMA package

The GeneChip™ *Drosophil*a Genome 2.0 array by Affymetrix contains probe sets to measure the expression of 18,500 transcripts. These probe sets are based on the Flybase 3.1 *Drosophila melanogaster* genome and consist of 14 pairs of oligonucleotide probes to measure the abundance of each sequence within a sample. Probe pairs consist of a perfect match (PM) and mismatch (MM) probe, both of which are 25 oligonucleotides in length. PM probes are the exact complement to the sequence of interest, whereas MM probes have a single base pair change at position 13 and are therefore assumed to account for non-specific hybridisation to that probe (Dalma-Weiszhausz et al., 2006; Wang et al., 2007). Subsequent analysis is then required to summarise the information from these 14 probe pairs into a single

expression level and, due to the complexity of this task, there have been multiple methods developed. However, many of these methods often produce a single value and the uncertainty associated with each probe set is rarely retained, hence information regarding the confidence of the expression value is lost. Because of this, and expertise available at the University of Sheffield, I chose to utilise the PUMA (Propagating Uncertainty in Microarray Analysis) package within the R suite, which retains confidence levels during downstream analysis and therefore adds statistical power to the expression data (X. Liu et al., 2005; Pearson et al., 2009; Liu et al., 2013). All scripts used for microarray analysis can be found in Appendix 1.

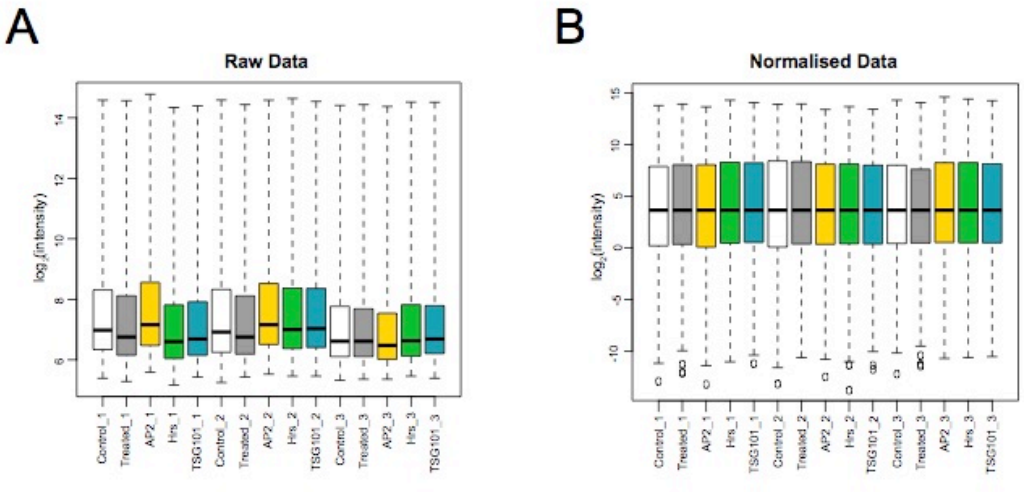


Figure 3.9: Boxplots of chip data produced in this study. A) Boxplots expressing spread of raw probe data within each individual array. B) Spread of data within each array after normalisation using the PUMA's mmgmos package for probe-level analysis. Boxes are labelled with sample and biological replicate (eg- Control_1 = control sample, replicate 1) and colour coded for samples. Control- white, treated- grey, AP2- yellow, Hrs- green, TSG101- blue.

Due to the complexity of the multistep microarray procedure, and the technical variability that may be introduced at each stage, differences in background signals across individual arrays can occur. The variability across the chips in this study can be visualised by a boxplot (Figure 3.9A), which demonstrates the spread of raw signal intensities (log2). Therefore, it is important to normalise signal intensities when considering differentially expressed genes so as to decrease the possibility of false-positive hits. The multi-chip modified gamma Model for Oligonucleotide Signal (mmgMOS) (X. Liu et al., 2005) command globally normalises across all chips, so that the median distribution is similar across all arrays (Figure 3.9B). One of the major advantages of this method is that it considers the nucleotide sequence of a specific

probe, and hence the different binding affinities associated with it, which should be similar for the same probe across all chips (X. Liu et al., 2005).

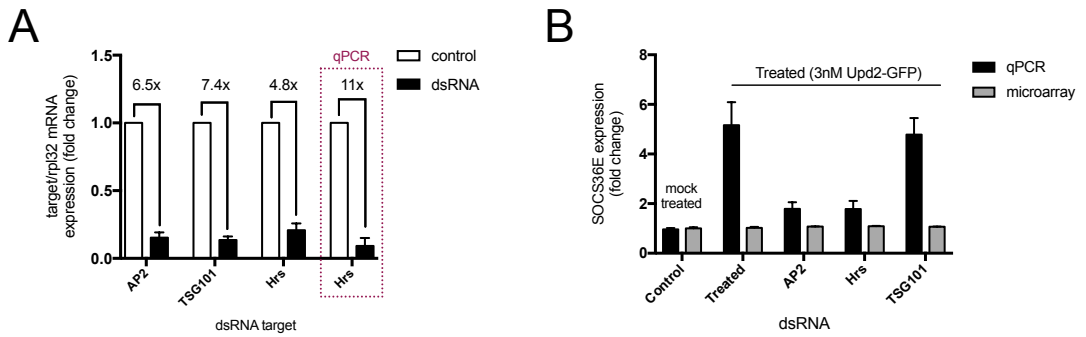


Figure 3.10: Analysis of genes whose expression is known to change. A) mRNA expression in control dsRNA or AP2, TSG101 and Hrs dsRNA treated cells. Hrs was measured using both microarray and qPCR (highlighted). Data has been normalised to expression of reference gene rpl32 in the same sample, and is presented as fold change compared to control dsRNA treated cells. B) SOCS36E mRNA expression in microarray samples, as measured by qPCR or microarray.

Following puma analysis, I examined the expression of targets that were predicted to change. As anticipated, there was a significant decrease in the expression of endocytic components in cells treated with targeting dsRNA, compared to control (Figure 3.10A). However, the fold change in Hrs expression level in the microarray data is noticeably less than when the same samples were measured by qPCR. This may be due to the reduced dynamic range of microarrays in comparison to qPCR. In qPCR, the PCR product is amplified exponentially, and therefore allows for detection of low RNA concentrations. In contrast, microarrays inherently have high background noise, due to the hybridisation procedures, and therefore transcripts with low expression levels are often not detected above the background (Allanach et al., 2008). In this context, when *SOCS36E* expression was measured via microarray analysis no changes in expression levels were observed which contradicts the qPCR results (Figure 3.10B). This may also relate to the dynamic range of the microarray. The chips have a limited number of probes to which the transcript can bind, therefore if the probe is oversaturated with oligonucleotides, information about changes in expression may be lost. *SOCS36E* has an average expression of 12.40 (log_2) across the 15 chips, which is above the determined 11.81 (log_2) cut off determined by Zhao et al, 2014. Although this value was determined from a different Affymetrix GeneChip (Zhao et al., 2014), *SOCS36E* has one of the highest values on the chips in this study. Probe sets do not

all saturate at the same level due to difference in probe composition so it is difficult to generate a saturation limit across all probes sets (Skvortsov et al., 2007), and therefore I cannot easily determine whether SOCS36E was saturated in this experiment. However, due to the small dynamic range microarrays may only provide limited information about differentially expressed genes, and this is dependent on their level of expression.

To identify differentially expressed (DE) genes I used the function *pumaComb* to summarise the data from the 3 repeats. Unlike other methods of analysing DE genes, the advantage of *pumaComb* is that it combines the uncertainty values calculated previously to produce a Probability of Positive Log Ratio (PPLR) which gives statistical power to expression values (X. Liu et al., 2005). Following this I used the *pumaDE* function to compare expression data from the different conditions and produce $\log_2$ fold changes. This data was then filtered to include only targets whose expression changed by more/less than ±0.2 (log2) fold, which is 1.5x the standard deviation of fold changes across all chips from the Treated (plus Upd2-GFP) vs Control comparison. Targets were also selected on the basis of having PPLR values above 0.8 or below 0.2. PPLRs close to 1 gives high confidence that target expression has increased, whereas those with PPLRs close to 0 give a high confidence of decreased expression. As seen in the summary in Figure 3.11, the expression of 961 genes are altered upon ligand stimulation, in the absence of dsRNA. Of the 606 where expression increased, 80 of these were affected by AP2 knockdown alone, and 113 by AP2 and Hrs knockdown. 37 targets were altered upon knockdown of AP2 and TSG101, but not Hrs. However, these targets are unlikely to be altered by endocytosis as their expression is unaltered upon knockdown of Hrs, and therefore this group was not investigated further.

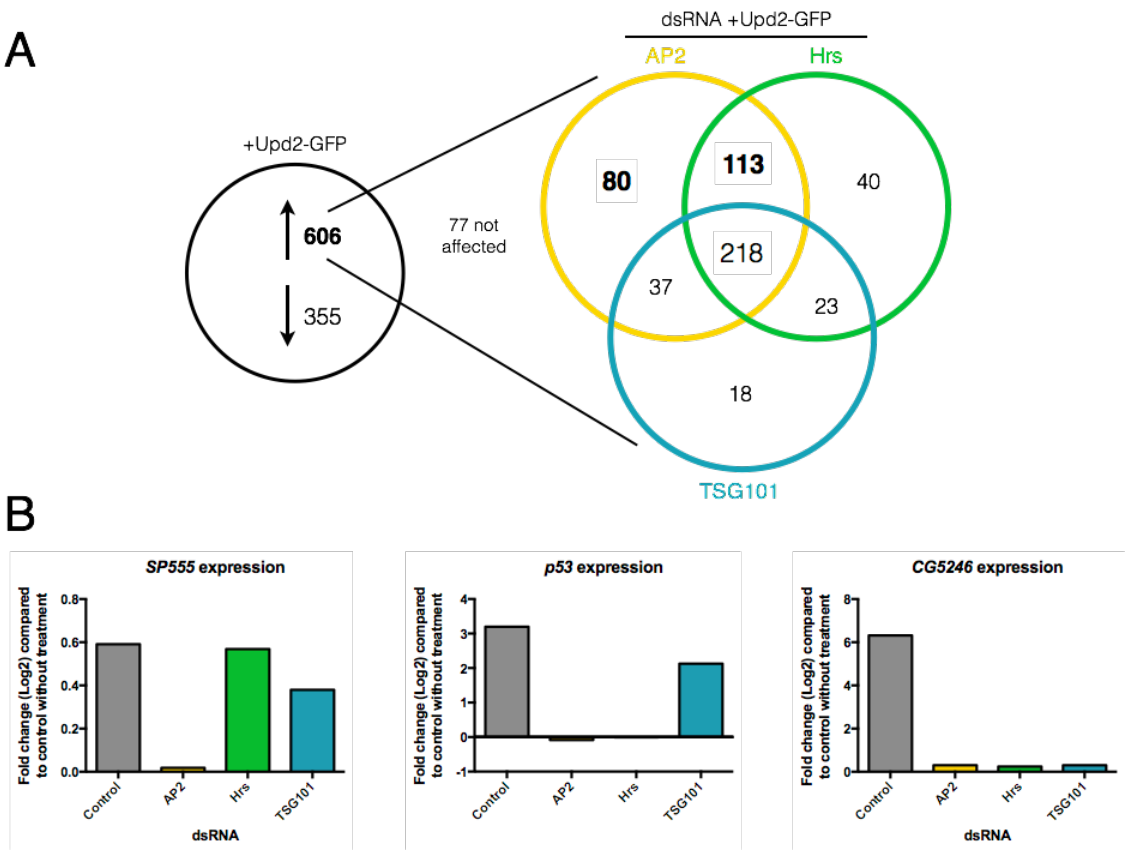
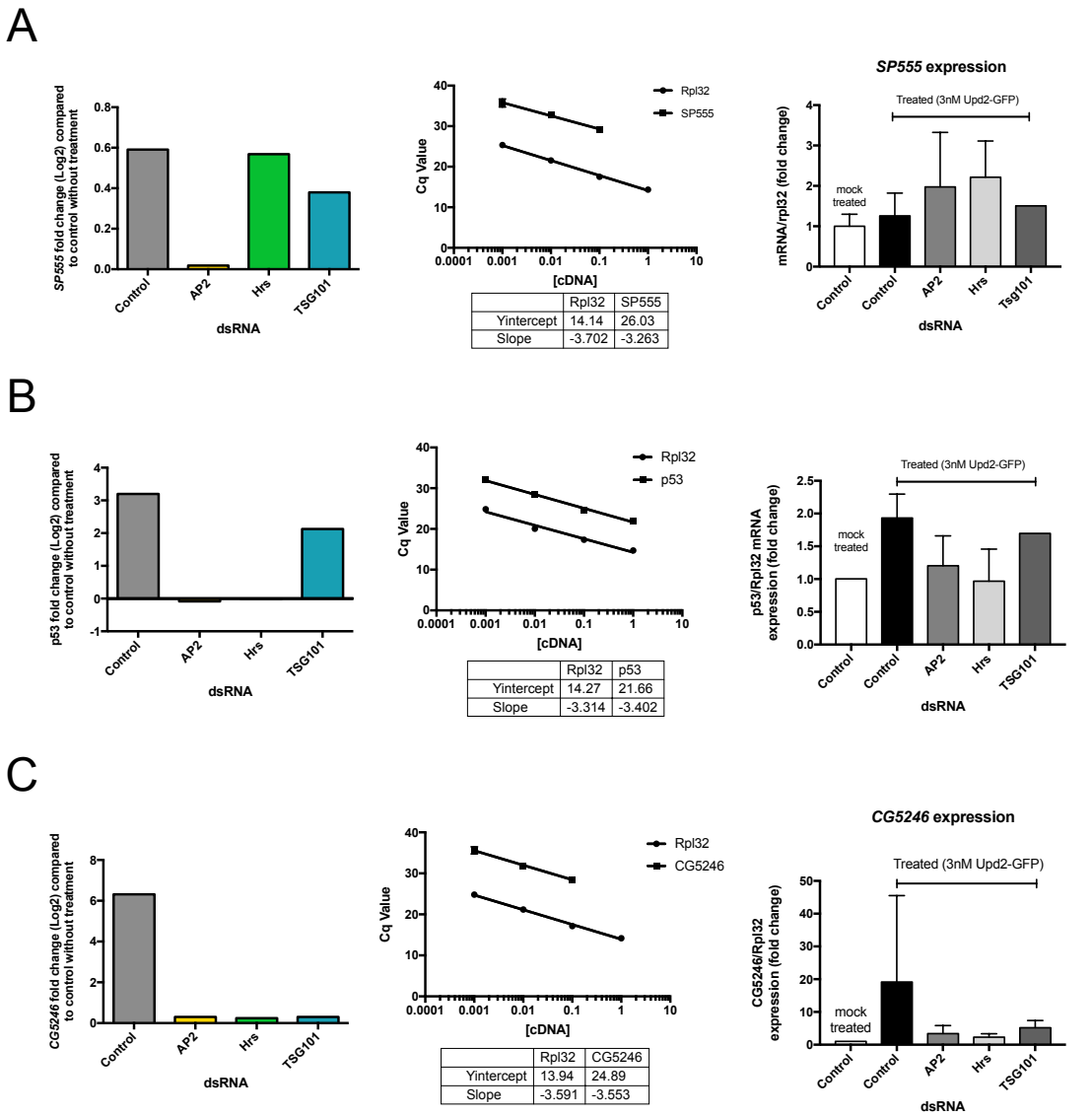


Figure 3.11: Differentially expressed targets regulated by endocytosis. A) Venn diagram illustrating genes whose expression changes by a fold change ≥0.2 and PPLR ≥ 0.8 between conditions. 606 genes were upregulated upon stimulation with Upd2-GFP, with 529 of these being altered when endocytic components were knocked down. B) Example of targets whose expression was affected by AP2 knockdown, AP2 and Hrs knockdown or all knockdown conditions, respectively. Graphs show outcome of PUMA-comb analysis, which combines microarray repeats, and represents fold change compared to control without treatment.

3.2.1. Validation of identified targets

Due to the multistage microarray protocol, complexity of downstream analysis and non-specific hybridisation that can occur (Discussed in Chapter 3.2.4), it is feasible that changes in the expression of targets may be artefacts. Therefore, it is paramount that potential hits are validated before biological conclusions can be made. In the first instance, several promising candidates (SP555, p53 and CG5246) which showed positive results by microarray, were selected (Figure 3.11B). SP555 is a SPRY domain and SOCS box containing protein. Although little is known about this protein, SOCS boxes are key in JAK/STAT pathway regulation (Kile et al., 2002) and SPRY domains bind to a variety of proteins, such as the RTK c-met (D. Wang et al., 2005). p53 is key for cell cycle regulation, stress response and tumour suppression, and mammalian STAT1 directly interacts with p53 resulting in increased expression of pro-apoptotic

genes (Townsend et al., 2004). CG5246 is an uncharacterised protein that has serine protease domains. Multiple serine endopeptidases have been identified as targets of Upd induced JAK/STAT activity (Bina et al., 2010). qPCR primers were designed to target the same region of the transcript as the microarray probes. Surprisingly qPCR analysis, using exactly the same RNA as in the microarray, produced very different results for the selected targets with the exception of *p53* (Figure 3.12). I therefore carried out experiments to further validate p53 by preforming qPCR on freshly prepared RNA samples, but saw no Upd2-GFP dependent activation or endocytic regulation (Figure 3.13).



	Rpl32	SP555
Yintercept	14.14	26.03
Slope	-3.702	-3.263

	Rpl32	p53
Yintercept	14.27	21.66
Slope	-3.314	-3.402

	Rpl32	CG5246
Yintercept	13.94	24.89
Slope	-3.591	-3.553

Figure 3.12: Validation of targets regulated by endocytosis using same RNA as microarray. A) Validation of *SP555* expression, B) Validation of *p53* expression, C) Validation of *CG5246* expression. The first panel is a graphical representation of the data from the microarray. The 2nd panel demonstrates validation of the qPCR primers. Panel three shows the mean of triplicates ± SEM from 3 independent qPCR experiments using cDNA produced from the identical RNA used for microarray analysis.

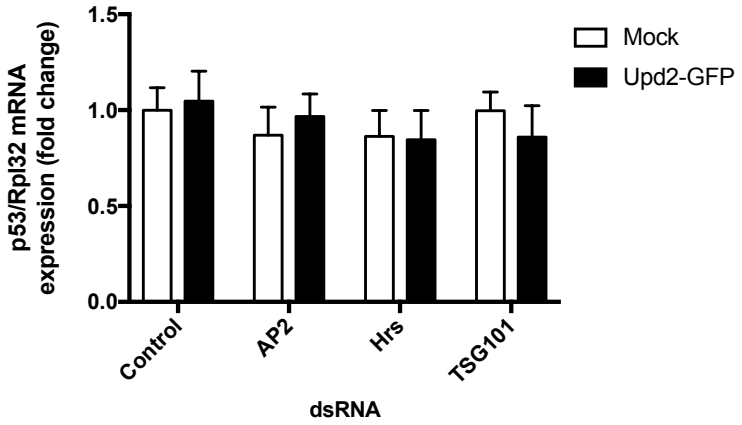


Figure 3.13: Validation of p53 in fresh RNA samples. dsRNA knockdown in S2R+ cells was performed for 5 days. Cells were incubated with 3nM Upd2-GFP for 2.5hrs prior to RNA extraction. P53 mRNA levels were normalised to that of reference gene Rpl32, and presented as fold change compared to mock-treated control samples. Results are expressed as means of triplicates + SEM for 3 independent experiments. Parametric, unpaired students T-test was carried but all were not significant.

Examination of microarray expression levels from individual experimental repeats, revealed that for both p53 and CG5246, a single repeat with a high expression level was capable of skewing the combined result (Figure 3.14). Higher expression levels appear to be accompanied by a reduced standard error. Therefore, when repeats are combined using the *pumaComb* function, which retains and utilises uncertainty measures (X. Liu et al., 2005), the high value has greater influence on the result, as the associated standard error is smaller and so the measurement has a greater probability of being correct. These data also demonstrate that large fold changes within the data set may arise from the inclusion of data with negative Log_2 values (see Figure 3.9). Because of this, I also manually investigated differentially expressed genes. Firstly, I filtered targets to include those whose expression was between 2 and 10, thereby excluding low and negative values that exhibit large standard errors. I then calculated the fold change between Treated and Control samples, and averaged this across the repeats. Using this method, 401 targets were altered by ±0.2 (log2) fold and a standard error of <0.2. Of these 21 appeared to be affected by AP2 knockdown, however these were not followed up as none appeared to have a robust expression pattern or upstream STAT92E DNA binding sites.

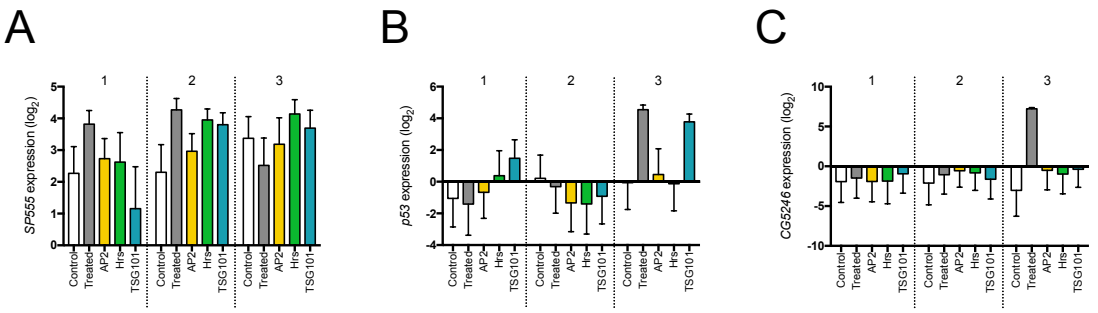


Figure 3.14: Expression of targets across each array, following mmgmos normalisation. Expression of A) SP555, B) p53 and C) CG5246 across all 15 chips. Chips are grouped into biological replicate (1, 2 or 3) and colour coded for samples. Control- white, treated- grey, AP2- yellow, Hrs- green, TSG101- blue. Data was produced using the PUMA's mmgmos package and is log_2 expression + SE.

3.2.2. Quality control of microarray

Due to unsuccessful validation of identified microarray targets, I firstly decided to examine array quality in more detail, as this can have a detrimental effect on downstream analysis. Although RNA quality was checked before microarray analysis, degradation may have occurred during the sample preparation. The Affymetrix Genechips™ include controls to determine whether RNA degradation may have occurred during the labelling and hybridisation. This analysis is crucial as hybridisation of degraded RNA to the arrays provides no useful information. *Drosophila* Actin, eIF-4a and GAPDH probes are included as 'housekeeping' controls for RNA degradation, as their expression is likely to remain the same across experimental conditions. Three different probe sets are included for each gene, probing the 5' region of the transcript, the middle region and the 3' region. If RNA is fully intact the probe intensities at all positions should be similar. In contrast, if RNA degradation has occurred, intensities at the 5' probe would be lower due to reverse transcription occurring from the 3' end of the mRNA. Therefore, calculating the 3'/5' ratio allows us to examine RNA degradation (Jaksik et al., 2015). Figure 3.15A demonstrates there is little RNA degradation as all values are ~1. Probe sets contain multiple probes that target various regions across the transcript and it is thus possible to examine the signal vs probe position. Although the signal is reduced at the 5' probes (Figure 3.15B), the important feature is that the trend is the same across all chips. The data has been shifted and scaled so that each individual array can be visualised. Therefore, there appears to be no obvious problems with RNA degradation.

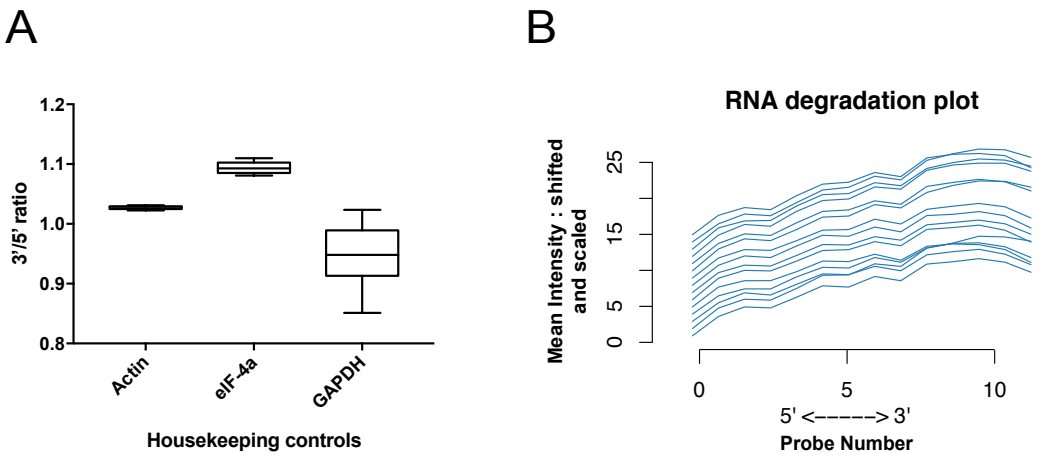


Figure 3.15: RNA degradation after preparation for hybridisation to microarray chip. A) Ratio of 3' probe set to 5' probe set for each housekeeping control set across the 15 arrays. B) Graph of 5' to 3' slope of probe intensities across entire array, for all 15 arrays, using the function plotAffyRNAdeg.

Efficient amplification, labelling and hybridisation are key factors required for trustworthy microarray data and downstream biological conclusions. During the Affymetrix microarray procedure external RNA controls are used to validate effective cDNA synthesis and hybridisation to the chip. Before cDNA synthesis, Poly-A spike in controls are added. These are poly-adenylated transcripts for *B.subtilis* genes which are added at varying concentrations, to monitor in-vitro translation and labelling, independent of RNA quality (Jaksik et al., 2015). RNA for the *lys* (*diaminopimelate decarboxylase*) gene is added at the lowest concentration, near to the sensitivity limit of the microarray, whereas *dap (dehydrodipicolinate reductase)* RNA is added at the highest concentration and near the saturation level of the microarray. Analysis of expression confirms that *dap* gives the highest intensity (Figure 3.16A), followed by *thr (homoserine kinase)*, *phe* (*threonine repressor*) and then *lys*, suggesting that the amplification and labelling process has not favoured either high or low expression targets. However, the range of signal identified by the microarray is low, only ~10-fold change between Dap and Lys when the true value is 15-fold, suggested a reduced dynamic range.

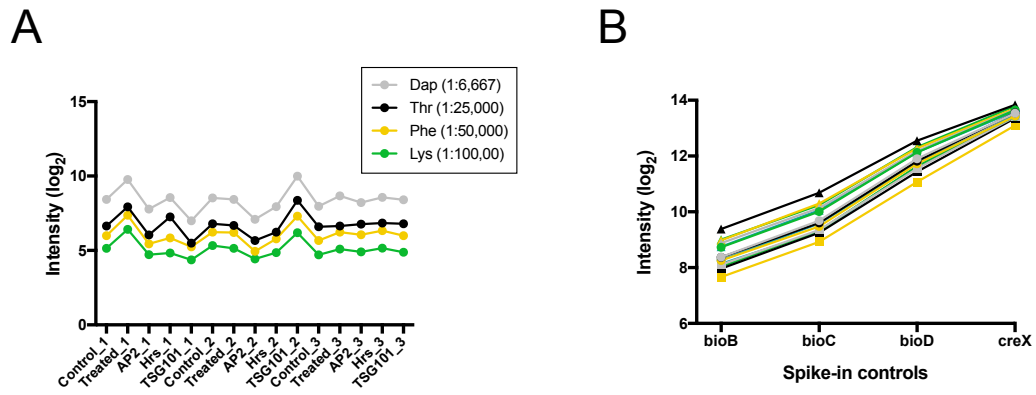


Figure 3.16: Quality control plots for RNA amplification, labelling and hybridisation. A) log2 intensities of Poly-A spike in controls. B) log2 intensities of spike-in hybridasation controls. Each line represents a different array.

Prior to hybridisation of labelled, fragmented antisense RNA to the array, spike-in hybridisation controls are added to the sample. These are pre-fragmented and labelled transcripts and are again added at varying concentrations, allowing for analysis of the hybridisation step. The intensities of these (Figure 3.16B) increased as the concentration increases, *bioB*, *bioC*, *bioD* (biotin synthesis genes from *E.coli*) and *creX* (recombinase gene from P1 bacteriophage), suggesting hybridisation has occurred successfully.

Therefore, using the various Affymetrix controls included on the array or in the RNA preparation, there does not appear to be any obvious issues with quality RNA degradation, amplification, labelling or hybridisation.

3.2.3. Discussion of microarray analysis process

There did not appear to be any obvious issues with the RNA preparation or microarray hybridisation, and therefore I decided to examine the method of analysis used. As discussed previously, PUMA analysis considers the binding affinities for different probe sequences, and therefore allows for a level of uncertainty to be associated with each probe set, overcoming some of the issues associated with non-specific binding of some MM probes (X. Liu et al., 2005). However, multiple studies discuss the validity of using MM probes as an estimate of background binding at all (Naef et al., 2002; Irizarry et al., 2003), with Wang et al, 2007 demonstrating that ~20% of MM probes

have a higher signal intensity compared to the equivalent PM probe (Wang et al., 2007).

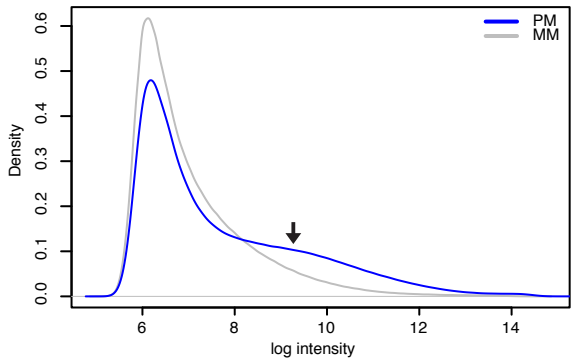


Figure 3.17: Distribution of probe intensities across arrays. Histogram represents average density distribution across all 15 arrays. Arrow highlights region in which PM probes are higher than MM probes.

Global array histograms (Figure 3.17) demonstrate that for my data MM probes have a similar density distribution of $\log_2$ intensities to that of PM probes. There are a high number of MM probes with low signal intensities, and therefore this area is likely to represent background hybridisation. If PM probes hybridise with higher affinity, we would expect this curve to shift to the right of the curve for MM probes. However, the only region whereby more target is bound to PM probes is at highly-expressed targets (shown by arrow, Figure 3.17), a level which may coincide with oversaturation. As the density distribution looks similar for both probes, removing MM from PM during normalisation results in an increase in signal intensity variation across the chip due to the presence of negative values (Figure 3.18C). This may result in false positives when identifying DE genes, that may not represent true biological differences. Hence, I decided to look at normalisation methods that do not use MM probes. Robust multi-array average (RMA) (Irizarry et al., 2003) is a normalisation technique routinely used during microarray analysis that considers PM probes only. PUMA also introduced a PM only analysis, called PMmmgmos (Liu et al., 2013), which enables the same preservation of uncertainty measurements described previously, without subtracting MM probes. I therefore used these two normalisation methods and compared their density distributions. Normalisation that excludes MM probes produces data with a distribution similar to that of the raw data, with all values positive (Figure 3.18).

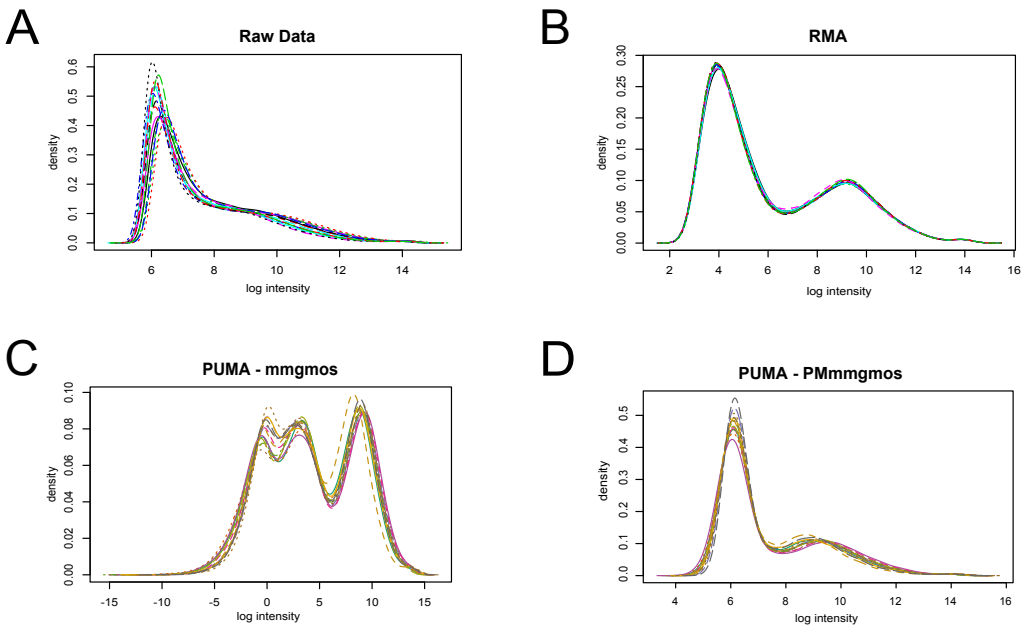


Figure 3.18: Intensity distribution histograms after various normalisation methods. A) Distribution of raw data in Affybatch object. Histogram representing distribution of data after using B) RMA function, C) mmgmos function or D) PMmmgmos function.

Using the data produced by these normalisation methods I investigated the top 20 genes that were differentially regulated between Treated and Control samples. No targets are top hits in all methods, and only one target is seen more than once (Table 3.1, Figure 3.19). This is MYPT-75D, one of *Drosophila's* two non-muscle myosin targeting subunits, which has not been identified as a JAK/STAT target before. Interestingly, ligand-dependent activation appears to be altered when endocytic components are knocked down (Figure 3.19A). However, STAT92E binding sites were not identified upstream of this gene, suggesting that its expression is not directly regulated by STAT92E and therefore I chose not to follow up this target. None of these normalisation method used show SOCS36E expression to have the same trend as when measured via qPCR (Figure 3.19B).

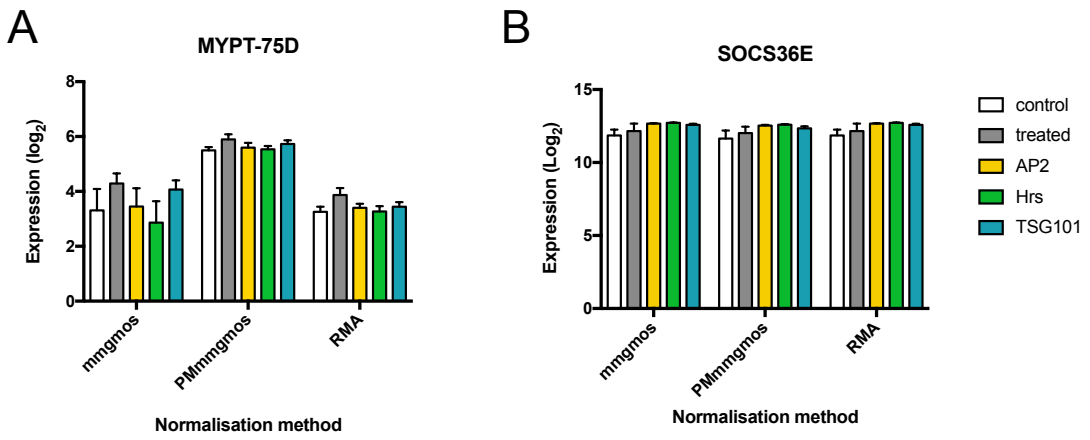


Figure 3.19: Expression values for targets after normalisation with either mmgmos, PMmmgmos or rma. Graphs represent the mean log2 expression values for A) MYPT-75D B) SOCS36E, across the 3 array repeats +SD.

Hence, although there appears to be no obvious issues with the microarray, and I used multiple analysis methods, unfortunately no usable data was obtained from these transcriptomics experiment. The small dynamic range of the microarray and the validity of MM probes as an estimate for background hybridisation are potential reasons that have been discussed above. Another factor that could affect the microarray results is the model used. S2R+ cells are a particularly heterogenic *Drosophila* cell line, showing obvious phenotypic differences in cell culture and as identified in this study (Figure 4.3), where only 80% take up ligand, and even this seems to be variable in quantity. However, previous published microarray experiments have been successful using a variety of *Drosophila* cell lines.

	mmgmos				PMmmgmos				RMA			
	Gene	**Probe ID**	**FC**	**PPLR**	**Gene**	**Probe ID**	**FC**	**PPLR**	**Gene**	**Probe ID**	**FC**	**P value**
1	CG5246	1640355_at	6.302	1.000	CG12896	1631628_s_at	0.326	0.858	CG31381	1628813_at	0.747	0.004
2	CG31041	1624805_at	4.265	1.000	CG17124	1633795_a_at	0.310	0.814	Ef1100E	1624989_s_at	0.636	0.003
3	Esp	1637292_at	3.712	1.000	CG34033	1634750_at	0.307	0.762	**MYPT-75D**	**1640082_at**	**0.614**	**0.001**
4	CG30502	1639923_at	2.881	0.998	vir-1	1625012_s_at	0.303	0.862	CG7888	1627964_s_at	0.611	0.012
5	CG6356	1624448_at	2.851	1.000	His1:CG33837	1629740_at	0.288	0.841	CG17549	1634507_s_at	0.517	0.011
6	Tace	1639144_a_at	2.834	0.999	CG15678	1632964_at	0.281	0.877	Atet	1629559_s_at	0.490	0.010
7	CG13482	1636798_at	2.815	1.000	Ugt86Da	1624156_at	0.263	0.928	charybde	1627511_at	0.442	0.016
8	CG10824	1631523_at	2.763	1.000	**MYPT-75D**	**1640082_at**	**0.262**	**0.997**	Cdep	1636023_at	0.428	0.002
9	Dsk	1637290_at	2.628	1.000	CG40100	1632652_s_at	0.237	0.983	CG3775	1640399_at	0.404	0.015
10	CG9875	1631174_at	2.573	1.000	CG32245	1626837_a_at	0.217	0.999	CG9813	1634855_s_at	0.365	0.010
11	Tdc1	1635665_at	2.505	1.000	Cyp6a9	1628345_at	0.214	0.978	Est10	1626350_at	0.346	0.010
12	CG9691	1628187_s_at	2.397	1.000	sens	1632294_at	0.203	0.998	CG15601	1629501_at	0.342	0.008
13	CG31217	1633893_at	2.270	0.999	CG1600	1638498_s_at	0.203	0.849	CG33347	1635574_at	0.341	0.004
14	CG7191	1627069_at	2.186	0.999	CG6569	1639974_a_at	0.203	0.883	CG31856	1637345_at	0.329	0.005
15	CG11672	1630159_at	2.091	0.998	CG4267	1631783_at	0.201	0.877	CG6560	1638592_at	0.323	0.013
16	Fas3	1628543_a_at	2.018	1.000	CG2064	1634019_at	0.200	0.958	CG9117	1632406_at	0.318	0.015
17	inx7	1638225_a_at	1.993	0.998	CG30403	1641573_at	0.200	0.891	CG13126	1632049_at	0.301	0.015
18	CG34002	1626699_at	1.950	1.000	GstE7	1640065_at	0.198	0.927	CG11248	1638930_s_at	0.295	0.015
19	CG14448	1636898_at	1.926	0.998	c(2)M	1624856_at	0.198	0.843	CG5755	1641138_at	0.269	0.008
20	Sip1	1636440_at	1.831	1.000	Punch	1639469_a_at	0.191	0.708	Indy	1630894_s_at	0.259	0.005

Table 3.1: Top 20 DE genes, between control and treated, as measured by different analysis processes

The Smythe lab previously investigated JAK/STAT targets that were identified in a screen by Bina et al, 2010 when Kc_{167} cells were stimulated with Upd. However, these targets were not reproduced in S2R+ cell treated with Upd2 (Vogt and Smythe, unpublished). Although Upd and Upd2 are semi-redundant ligands (Hombría et al., 2005), distinct roles are being discovered with Upd2, but not Upd, being important in response to parasitic wasp infection (Yang et al., 2015). Therefore, it is likely different targets are activated, and hence why previously identified targets (Bina et al., 2010) may not be activated when S2R+ cells are exposed to Upd2-GFP. As Upd2 is also a freely diffusible ligand (Hombría et al., 2005) and key in immune responses, it may not be as essential in cell culture as it is in a tissue scenario and therefore interesting DE genes may not be identified. Interestingly, unpublished data by Bina et al., 2009 demonstrated that addition of Upd2 to Kc_{167} cells gave a very different transcriptional response to the addition of Upd (Bina, 2009).

Differing cellular contexts may also alter the transcriptional output of JAK/STAT pathway activation. In 2011, Cherbas et al., investigated the transcriptome of 25 different well-used *Drosophila* cell lines and demonstrated variability in gene expression. Out of ~15,000 genes investigated, only 21% were identified in all cell lines. A subset of genes was expressed at much higher levels in all cell lines than in tissue experiments, suggesting an adaptation to cell culture that has altered their transcriptome. Therefore, cell lines may not be physiologically representative for transcriptomic analysis.

3.3 *Lama* and *Chinmo* are not regulated by endocytosis

With the aim of identifying further JAK/STAT targets regulated by endocytosis, I investigated the expression of published *Drosophila* targets. I chose targets that had been previously identified *in vivo*, and therefore likely to be functionally-relevant, by methods that also identified *SOCS36E* as a target. A microarray study of JAK/STAT hyperactivation in the *Drosophila* eye disc identified *Domeless*, *SOCS36E* and *wingless* as upregulated targets, as well as novel targets including *lamina ancestor* (*lama*) and *chronically inappropriate morphogenesis* (*chinmo*) (Flaherty et al., 2009). *lama* is a gene with two upstream clusters of STAT92E binding sites and encodes a phospholipase B protein. *chinmo*, a zinc finger transcription factor, also has one cluster of STAT92E binding sites in the regulatory region, suggesting both targets are directly regulated by STAT92E (Flaherty et al., 2009). Further study of *chinmo* determined a

role in STAT92E-dependent stem cell renewal, tumour formation and eye disc development (Flaherty et al., 2010).

Therefore, I utilised the previously published qPCR primers to investigate the regulation of these two targets by endocytosis. As these targets have been studied in response to addition of Upd *in vivo*, I first demonstrated that their expression is upregulated by the addition of Upd2-GFP in S2R^{+} cells (Figure 3.20).

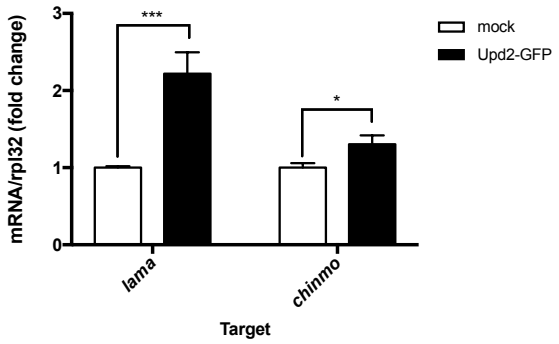


Figure 3.20: Expression of lama and chinmo in S2R+ cells upon addition of Upd2-GFP. S2R^{+} were treated with 3nM Upd2-GFP for 2.5hrs prior to RNA extraction Data represents mean of triplicates ±SEM from 3 independent experiments. *p≤0.05, ***p≤0.001.

Expression of both targets was not altered when endocytic components were knocked down (Figure 3.21), suggesting Dome does not require internalisation for expression of these targets. This data complements results discussed later in this thesis, whereby STAT92E, the transcription factor, appears to still be transcriptionally competent when endocytosis of Dome is prevented (Chapter 5).

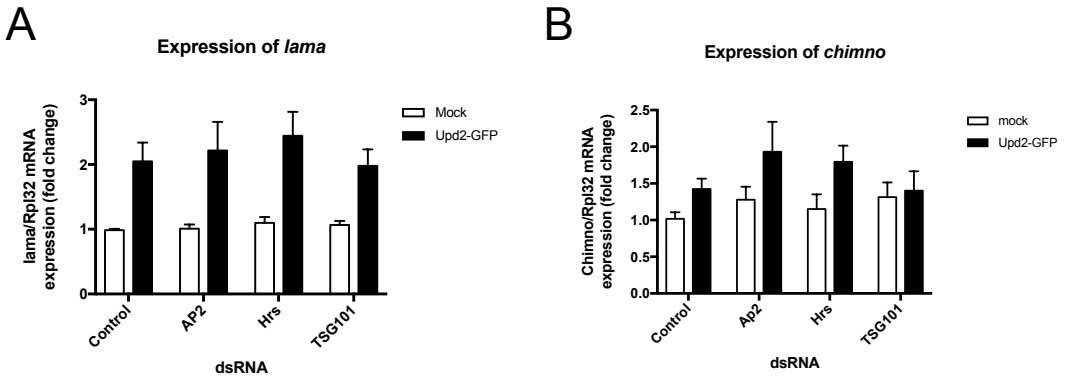


Figure 3.21: Endocytosis does not regulate expression of JAK/STAT targets lama and chinmo. dsRNA knockdown of endocytic components in S2R+ cells was performed for 5 days. 3nM Upd2-GFP stimulation was carried out for 2.5hrs prior to RNA extraction. A) lama mRNA levels B) chinmo mRNA levels. mRNA expression of targets was normalised to that of reference gene Rpl32, and presented as fold change compared to mock-treated control samples. Results are expressed as mean of triplicates + SEM for 3 independent experiments.

3.4 Conclusions

3.4.1. Conclusions from microarray analysis

There are multiple stages during the microarray procedure where errors can be introduced. Although these were analysed, there may be subtle issues that have not been identified due to lack of extensive expertise in *Drosophila* RNA analysis in the Sheffield community, therefore making it difficult to identify problems. For example, when analysing the density distribution of intensities across the arrays (Figure 3.17), there appears to be a group of genes with low expression, and a group of highly expressed genes. With mammalian samples a unimodal distribution is typical, with the peak at low intensities and a tail to the high intensities (Huber et al., 2003; Altirriba et al., 2009). As the extra peak at high intensities seen in my experiments is present in all arrays it is unlikely to represent an artefact on the array. Therefore, this extra peak at high intensities may pose a concern when considering normalisation methods, and the small dynamic range of the microarray. As the microarray can only provide data on differential expression for targets whose expression falls within a defined range, it is possible highly-expressed genes are outside of this range and therefore are not identified, for example SOCS36E.

Due to the heterogeneous cell line used, and the variability in ligand uptake (Chapter 4.1.2) the transcriptional changes may be subtle. Although dsRNA treatment was demonstrated to reduce AP2, Hrs and TSG101 mRNA effectively, this may not relate to protein levels, and it is unlikely that all cells were knocked down to the same extent, therefore adding to the variability. However, few targets were identified to change in response to Upd2-GFP treatment in control cells, where level of knockdown plays no role in the outcome.

If I were to continue with transcriptomic analysis it may be beneficial to use a different cell line or tissue, whereby ligand addition may have more relevant responses. *In vivo* experiments however, would also prevent the ability to add a specific concentration of Upd-ligand, which may alter the mechanism of endocytosis and signalling (Discussed in Chapter 4.2). It may also be beneficial to hybridise varying concentrations of RNA, with the aim of reducing the SOCS36E signal. This may help in understanding why the transcriptomic analysis in this project identified no targets, for example if the microarray had been oversaturated. If changes to SOCS36E were not observed at lower RNA concentration then RNA-sequencing (RNA-seq) may be more useful. This method has a significantly larger dynamic range, enabling detection of low abundance transcripts that are otherwise missed using microarrays and the identification of more subtle changes to gene expression. Overcoming problems with non-specific probe hybridisation, and probe hybridisation bias reduces technical variation. RNA-seq does not require an annotated genome and therefore allows for the detection of novel transcripts, and specific isoforms. However, RNA-seq is more expensive and analysis is very complex, therefore microarrays are still regularly used (Zhao et al., 2014).

3.4.2. Summary

This chapter demonstrates that the expression of JAK/STAT transcriptional targets is dependent upon trafficking of Dome and where in the endocytic pathway signalling is occurring. Firstly, I showed that the *10xSTATluciferase* reporter is expressed at an early stage of endocytosis, requiring only Dome internalisation or an event/interaction requiring AP2. In contrast, *SOCS36E*, a negative regulatory and sensitive downstream target of the JAK/STAT pathway, is expressed only when Dome reaches a later stage of the endocytic pathway. I aimed to identify further JAK/STAT targets that are endocytically regulated using microarray transcriptomic analysis. Unfortunately, and for unknown reasons, no targets were identified. Although quality control analysis

revealed no obvious problems, even SOCS36E did not appear to be expressed upon addition of Upd2-GFP, though this was confirmed prior to microarray analysis using the same RNA samples. This may be related to the oversaturation of the probe set, but this is hard to verify. Therefore, I chose to examine two known JAK/STAT targets from the literature, *Lama* and *Chinmo*. Although these were activated in response to ligand, perturbation of endocytosis did not alter their expression. This data suggests that *Lama* and *Chinmo* are both expressed when Dome is at the cell surface, whereas *10xSTATluciferase* expression requires dome internalisation and *SOCS36E* expression requires further Dome trafficking (Figure 3.22).

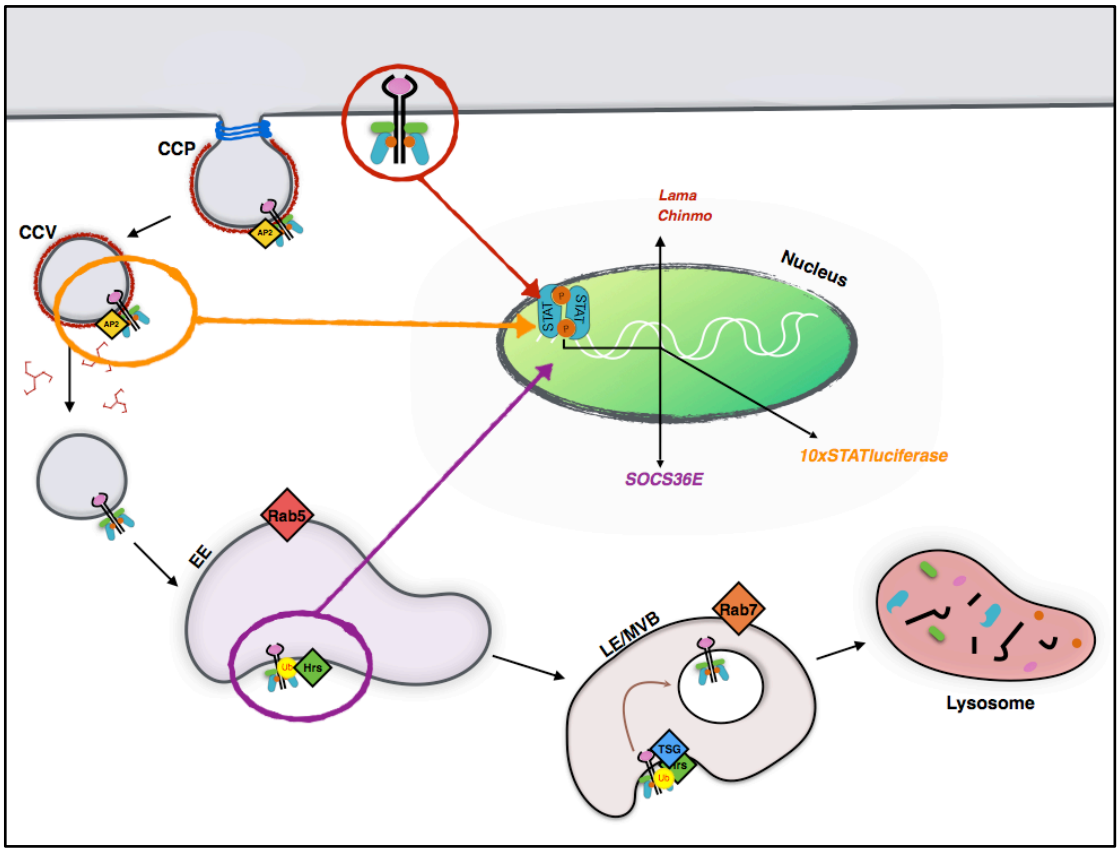


Figure 3.22: Schematic of differential JAK/STAT target expression during endocytic trafficking. *lama* and *chinmo* appear to be expressed at the cell surface. Expression of *10xSTATluciferase* requires Upd2/Dome internalisation, whereas expression of SOCS36E requires trafficking to a Hrs positive endosome.

Chapter 4. Defining Domeless endocytic uptake and trafficking

Previous reports have demonstrated that Dome, the *Drosophila* JAK/STAT pathway receptor, and its ligands, are internalised via clathrin-mediated endocytosis (CME) and trafficked for lysosomal degradation via Rab5-positive and late Rab7-positive endosomes (Devergne et al., 2007; Stec et al., 2013). During this study I have manipulated CME and trafficking, with the aim of trapping Dome in specific endosomal compartments and preventing further trafficking. This resulted in changes to the transcriptional outputs of the JAK/STAT pathway (Chapter 3.1), suggesting that uptake and localisation of Dome to different endosomal compartments is key for signalling. Therefore, I aimed to characterise the molecular environment required for regulation of a subset of JAK/STAT targets by investigating how knockdown of Hrs alters Dome trafficking.

4.1 Clathrin mediated uptake and endocytic trafficking of Domeless

To understand at a molecular level how endocytosis regulates the transcriptional output of the JAK/STAT pathway in *Drosophila*, I first aimed to define the endocytic pathway of the Upd2/Dome complex in further detail. My previous results showed that whilst treating cells with dsRNA against Hrs did not alter expression of a luciferase reporter, it did prevent ligand-induced *SOCS36E* expression (Chapter 3). Therefore, it appears that knockdown of Hrs prevents the Upd2/Dome complex from reaching an endosomal environment conducive to STAT92E–mediated *SOCS36E* transcription. Hrs depletion *in-vivo* has been shown to cause accumulation of Dome in endosomal puncta, suggesting reduced lysosomal degradation (Tognon et al., 2014), however, these studies did not determine the compartment or molecular environment in which Dome is accumulated. Vidal et al., 2010, demonstrated that in KC_{167} cells colocalisation of Upd2-GFP with Rab5 was seen after 5mins pulse-chase, and with Rab7 and lysosomal markers after 40mins. Upd2-GFP did not colocalise with Rab11, however, and after a 90mins chase ligand puncta were mostly undetectable (Vidal et al., 2010). However, this colocalisation was scored manually, so I wished to identify a method of quantitatively measuring subtle changes in trafficking, such as accumulation of ligand in various different endosomal populations.

4.1.1. Characterisation of antibodies for endocytic markers

I first defined the time-course of Upd2/Dome endocytosis. Due to the lack of antibodies available against Dome I investigated the uptake of the ligand, Upd2-GFP. I was unable to detect Upd2-GFP without amplification of the GFP (data not show). After testing a range of antibodies (Figure 4.1), I continued using the Abnova anti-GFP antibody, as this is commercially available and can be used at low concentrations. This antibody enabled the visualisation of puncta, and similar results to the in house antibody (KV, first panel) of which there was only a limited volume remaining. In contrast, a sheep anti-GFP and a GFP-booster from chemotrek, a nanobody produced in alpacas and attached to a fluorescent dye, gave high background staining and did not produce punctate staining upon the addition of Upd2-GFP. Vidal et al., 2010, demonstrated that Dome and Upd2-GFP puncta colocalise during endocytosis for at least 40mins, but that by 90mins the majority of ligand was no longer detectable. Therefore, for at least 40 mins after endocytosis Upd2-GFP location is an accurate readout for Dome endocytosis.

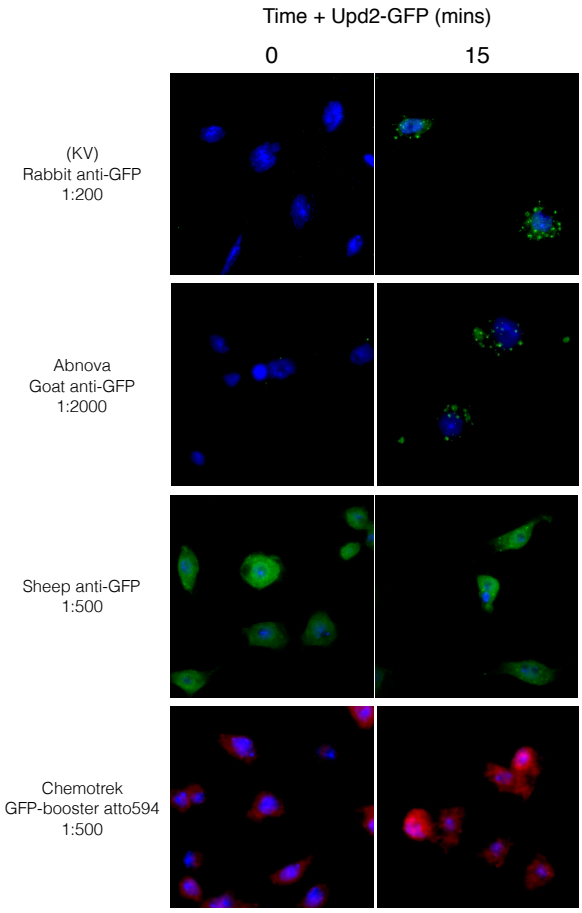


Figure 4.1: Antibody optimisation for Upd2-GFP uptake in S2R+ cells. Cells were incubated with 3nM Upd2-GFP for 0 or 15mins at 25˚C prior to fixation and staining with indicated antibodies and dilutions. Cells were imaged using the Nikon Dual-Cam.

Previous studies investigating endocytosis in *Drosophila* have commonly used overexpression of tagged endocytic proteins. However, due to the variable level of transfection in cell lines and the role of Rabs in endocytosis (discussed in 1.2.2), overexpression can sometimes alter endocytosis, which may have a drastic effect on signalling. For example, Rab5 overexpression results in abnormally large early endosomes which are associated with microtubules, increased rate of endocytosis and transferrin receptor recycling (Stenmark et al., 1994; Nielsen et al., 1999). Although this would be an interesting approach to manipulate endocytosis and investigate changes to JAK/STAT signalling, here I aimed to understand how knockdown of Hrs alters Dome trafficking and JAK/STAT signalling. Moreover, current constructs for endocytic proteins are largely tagged with GFP derivates, which would bind the anti-GFP antibody required to amplify Upd2-GFP. During this project, Sean Munro's group, at the MRC Laboratory of Molecular Biology, produced antibodies to label endocytic organelles in *Drosophila* (Riedel et al., 2016). Therefore, I could use these antibodies

under conditions where endogenous protein levels were maintained and therefore avoid artefacts due to protein overexpression. DsRNA knockdown of the endogenous proteins enabled successful characterise Rab5, Hrs and Rab7 antibodies and visualisation of an endosomal staining patterns (Figure 4.2).

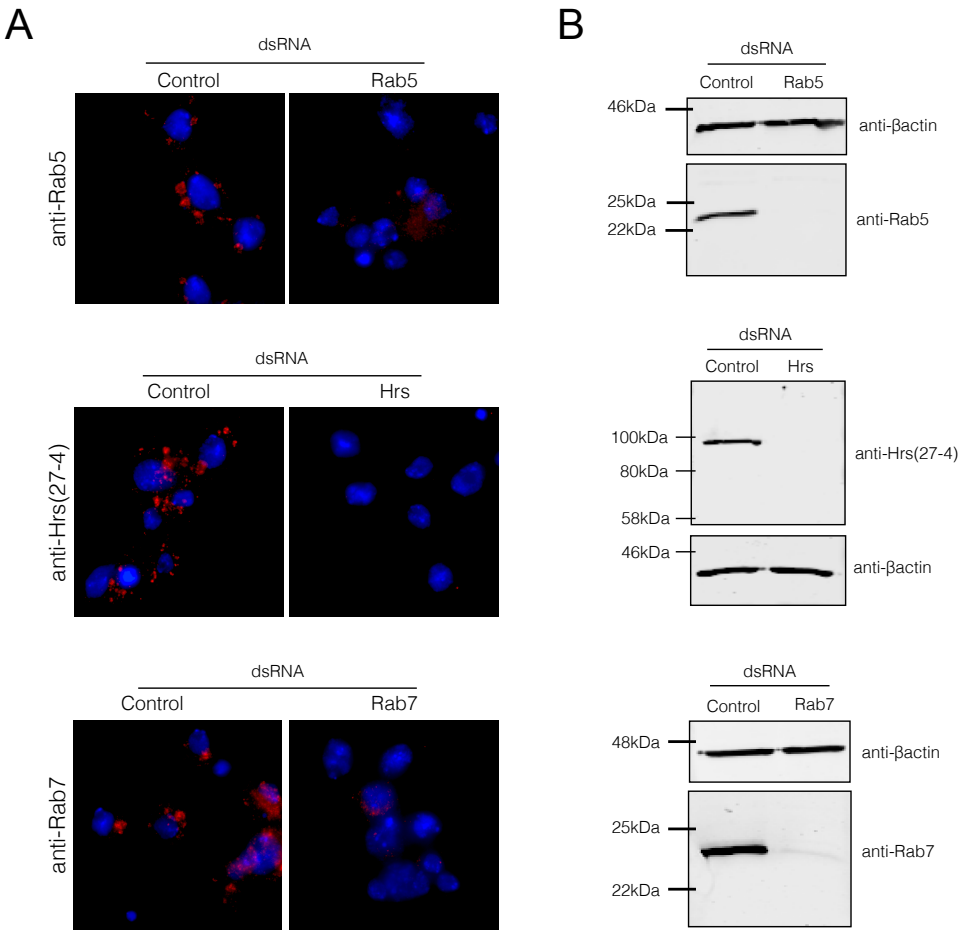


Figure 4.2: Characterisation of antibodies against endocytic markers. S2R+ cells were treated with dsRNA against Rab5, Hrs or Rab7 for 5 days prior to fixation and staining. A) Representative images of antibody staining for immunofluorescence, taken on the Nikon Dual-Cam. B) Immunoblots showing effective knockdown of proteins. Antibody concentrations for both uses were optimised and are documented in Methods.

4.1.2. Time-course for delivery of Upd2-GFP to different endocytic compartments

Utilising these antibodies, I investigated colocalisation of Upd2-GFP with the various endocytic markers. Interestingly, due to the heterogeneous nature of the S2R+ cell line, ~20% cells do not internalise Upd2-GFP (Figure 4.3). This was calculated

approximately, by counting the number of cells that did, or did not, take up ligand. No variability in amount of ligand internalised was considered. When carrying out the following experiments I was careful to image random populations of cells and not only those which were positive for Upd2-GFP uptake. Therefore, data includes up to 20% of cells that do not internalise ligand, and hence changes to endocytosis may be underestimated.

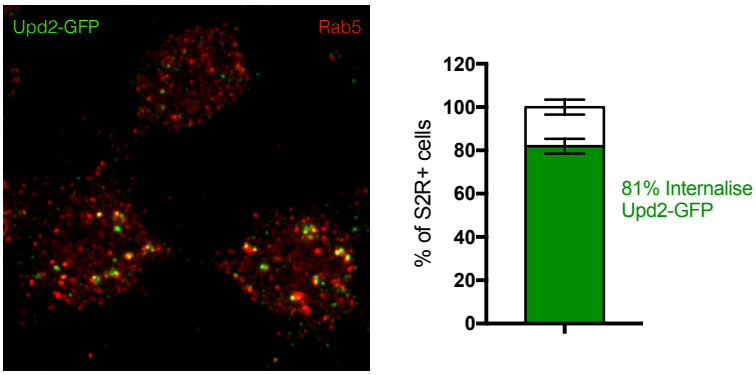


Figure 4.3: Variable uptake of Upd2-GFP by S2R+ cells. Image demonstrates 3 cells after 15min treatment with 3nM Upd2-GFP, imaged on the OMX with widefield settings. 2 cells show Upd2-GFP puncta, whereas the top cell does not internalise ligand. Data is mean ± SD from 2 independent experiments where at least 100 cells were counted for each.

Colocalisation of Upd2-GFP with Rab5, Hrs and Rab7 was investigated over a time-course of 60mins. Upd2-GFP is found in Rab5 and Hrs endosomes after 15mins, yet after 60mins some ligand is in endosomes negative for these endocytic markers (Figure 4.4 and 4.5). In contrast, Upd2-GFP does not enter Rab7 endosomes until 30mins after treatment with ligand (Figure 4.6). Pearson's correlation coefficient, a pixel-based method of colocalisation (Pearson, 1895; Dunn et al., 2011), was calculated to quantitatively measure the degree of colocalisation between these endocytic markers and Upd2-GFP (Figure 5.7). Quantification of the data demonstrates that ligand first enters Rab5 endosomes, before trafficking through Hrs positive endosomes to Rab7 endosomes, over a time-course which is similar to that seen by Vidal et al. for KC_{167} cells.

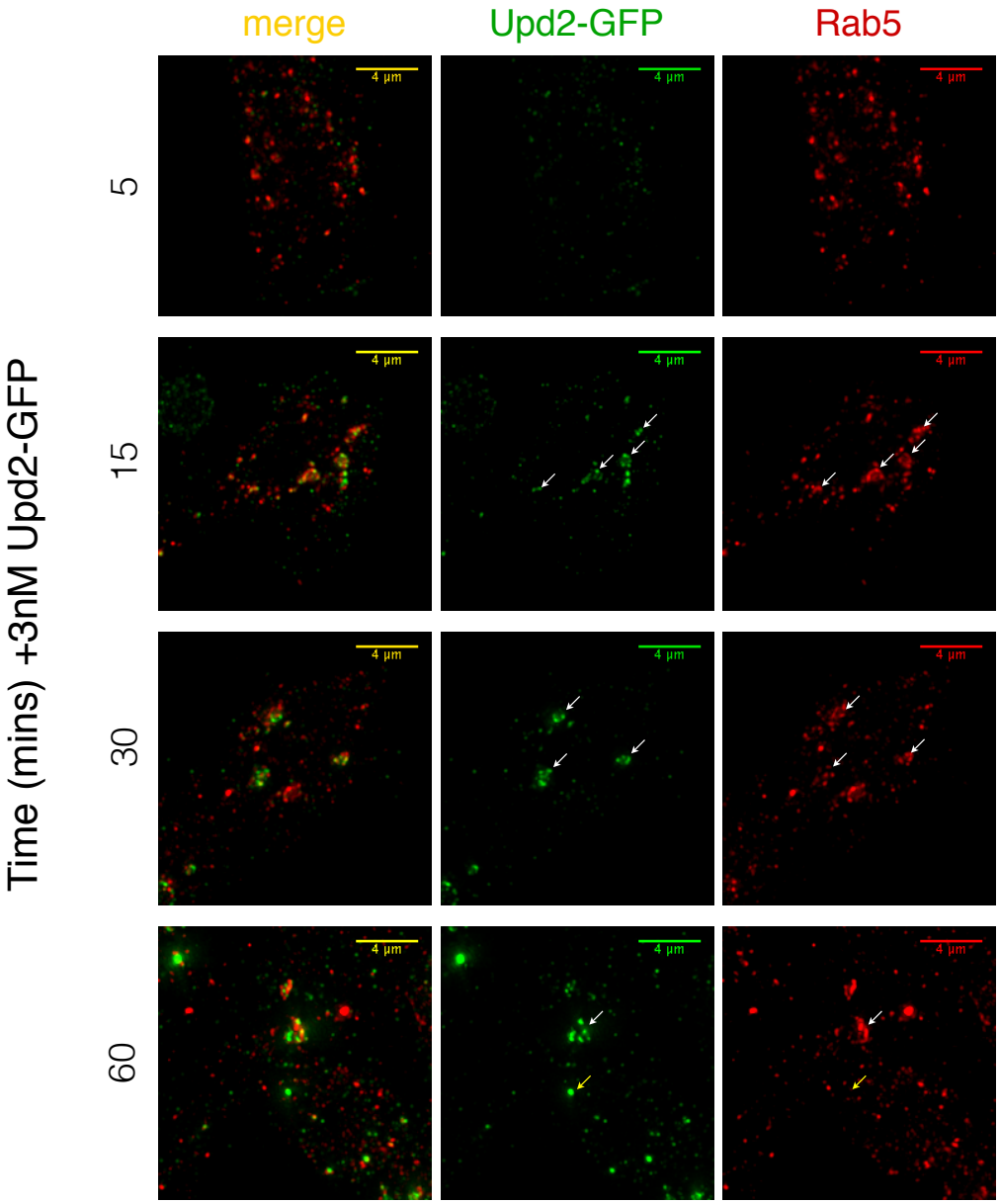


Figure 4.4: Upd2-GFP colocalises with Rab5 after 15mins. Representative images of Upd2-GFP colocalisation with Rab5 after incubation with Upd2-GFP for various timepoints. White arrows highlight Upd2-GFP puncta positive for Rab5, whereas yellow arrows highlight Upd2-GFP puncta devoid of Rab5. Scale bar is 4µM.

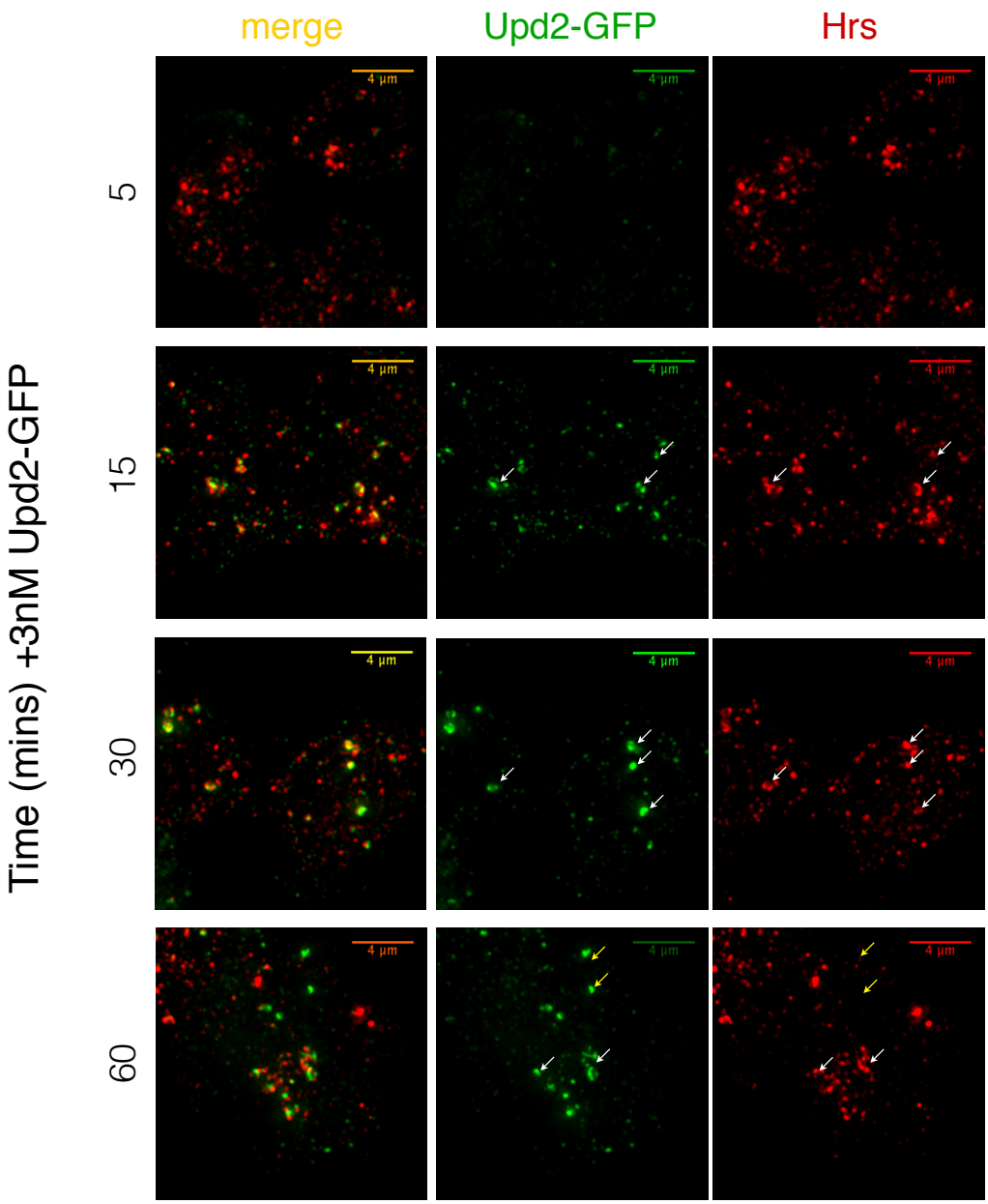


Figure 4.5: Colocalisation of Upd2-GFP with Hrs. Representative images of Upd2-GFP colocalisation with Hrs after incubation with Upd2-GFP for various times. White arrows highlight Upd2-GFP puncta positive for Hrs, whereas yellow arrows highlight Upd2-GFP puncta devoid of Hrs. Scale bar is 4µM.

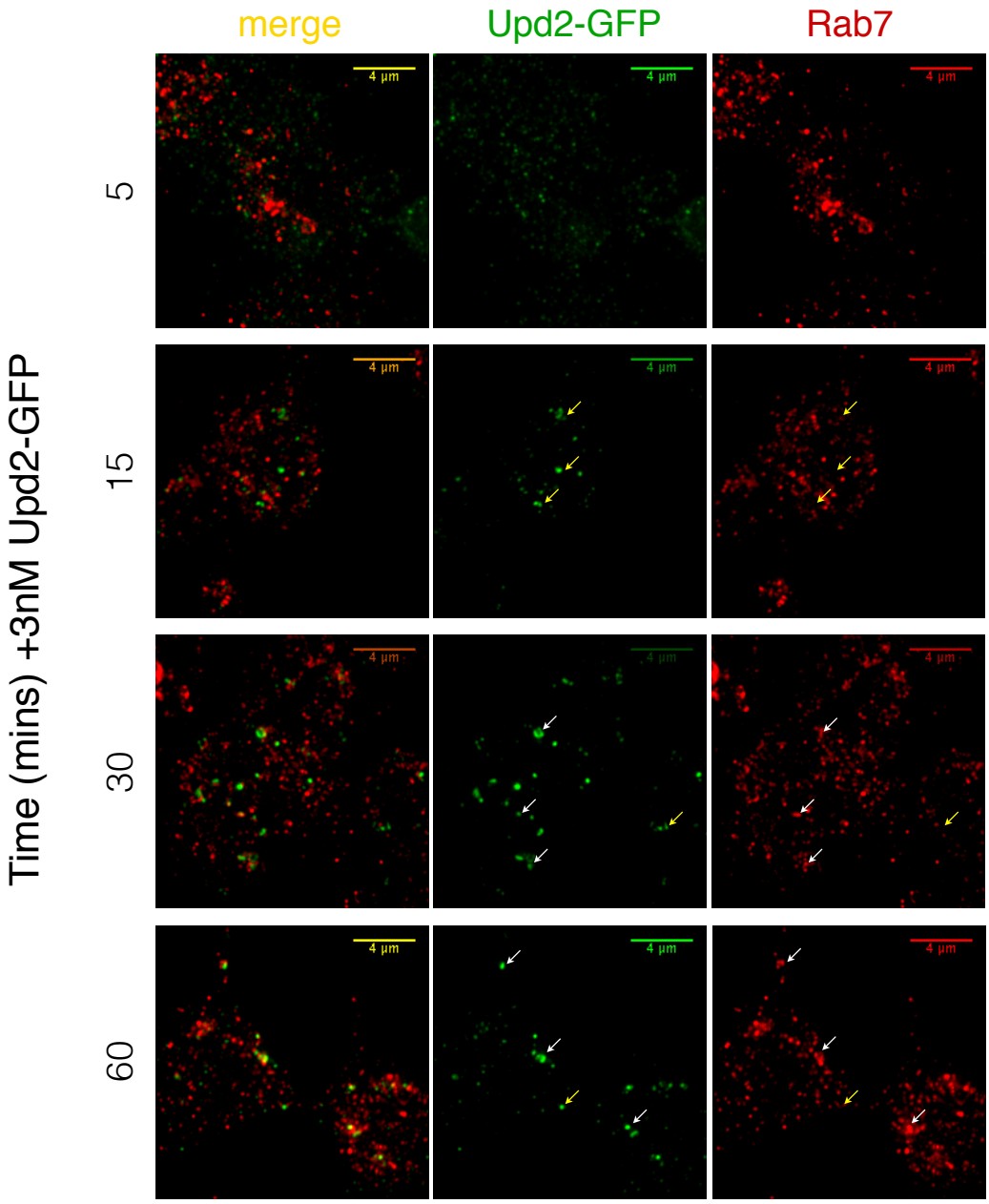


Figure 4.6: Upd2-GFP colocalises with Rab7 at later time-points. Representative images of Upd2-GFP colocalisation with Rab7 after incubation with 3nM Upd2-GFP for various times. White arrows highlight Upd2-GFP puncta positive for Rab7, whereas yellow arrows highlight Upd2-GFP puncta devoid of Rab7. Scale bar is 4µM.

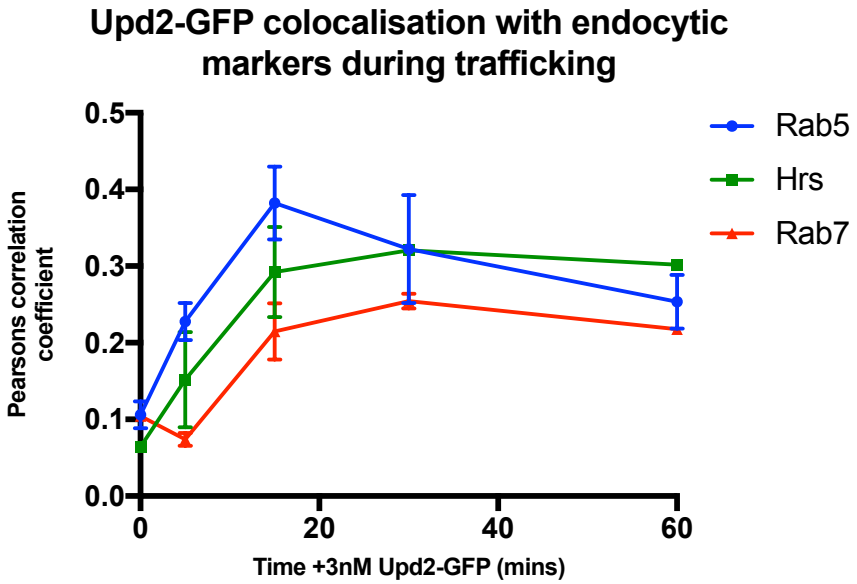


Figure 4.7: Timecourse of Upd2-GFP colocalisation with endocytic markers. Cells were treated with 3nM Upd2-GFP for variable times prior to fixation and antibody staining. Correlation was calculated using Pearsons correlation coefficient. Data is mean ± SEM for at least 2 independent experiments where ~30 cells were images per condition.

Subpopulations of endosomes and/or microdomains within the limiting membrane of an individual endosome could allow for qualitative differences in signalling, by creating an environment in which specific protein interactors could be recruited. Interestingly, Upd2-GFP does not appear to colocalise with all endosomes positive for Rab5 (Figure 4.4 and 4.8), suggesting trafficking is restricted to a subpopulation of early endosomes. Upd2-GFP also seems to be restricted to microdomain within Rab5 endosomes, which may indicate localisation to a specific signalling scaffold.

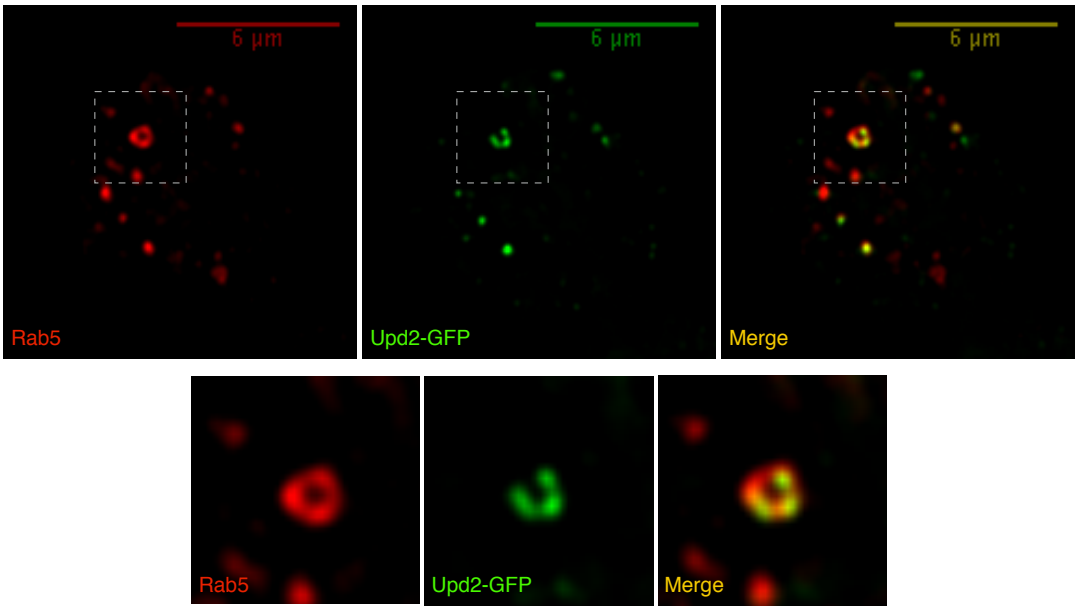


Figure 4.8: Upd2-GFP localises to subpopulations and subdomains of Rab5 endosomes. Image of cell treated with 3nM ligand for 45mins prior to fixation and staining. Insert demonstrates that Upd2-GFP is localised to defined areas of the limiting membrane.

4.1.3. Knockdown of Dome and clathrin reduces Upd2-GFP colocalisation with Rab5

To validate the use of Pearsons correlation coefficient for colocalisation analysis, I assessed the effect of dsRNA against Dome and Clathrin heavy chain (CHC), on the colocalisation of Upd2-GFP with Rab5. As expected, the Pearsons correlation coefficient is decreased after Dome knockdown (Figure 4.9).

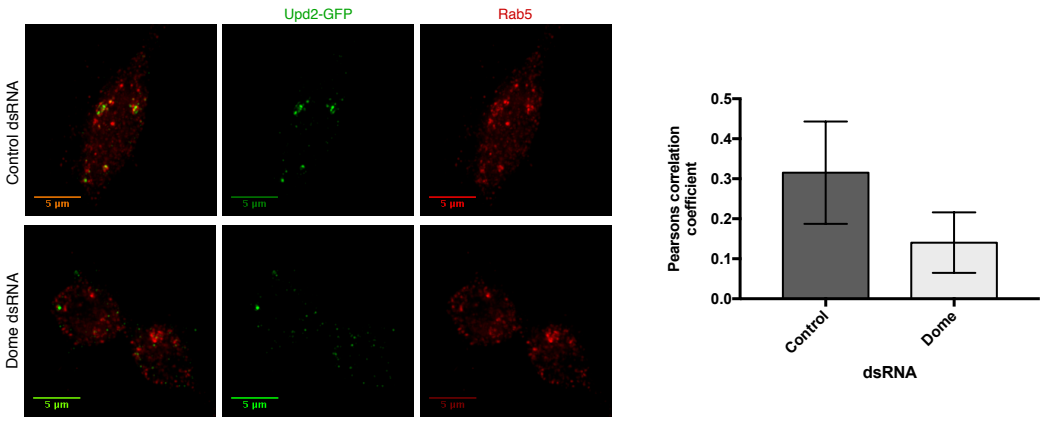


Figure 4.9 dsRNA against Dome abolishes uptake of Upd2-GFP and colocalisation with Rab5. Representative images of cells treated with either control or dome dsRNA 5 days prior to stimulation with 3nM Upd2-GFP for 30mins. Graph represents mean ± SD for one individual experiment.

DsRNA against CHC reduced uptake of Upd2-GFP (Figure 4.10A), however some Upd2-GFP puncta were still visible. This could suggest inefficient CHC knockdown, or that the receptor utilises a different endocytic route when CME is not available. Although these puncta do not appear to be positive for Rab5, Pearsons correlation still increases after ligand treatment (Figure 4.10B). To get a more accurate measure I utilised an ImageJ plugin, called Squassh (segmentation and quantification of subcellular shapes), developed by the MOSAIC group at the Center for Systems Biology Dresden (Rizk et al., 2014). The Squassh software employs masks and segmentation to identify subcellular shapes, and only uses information within these shapes for downstream analysis. The point-spread function of the microscope is accounted for, allowing for complex shapes to be identified. After identifying the correct parameters to allow for detection of specific shapes, Squassh can analyse multiple images at once, and an R script produces statistical, non-biased data. From this software, I can examine size and signal intensity of objects as well as colocalisation data including number of overlapping objects, size of overlapping objects, and the intensity of signal in overlapping regions (Figure 4.11). The use of Squassh to investigate the signal intensity of Upd2-GFP within objects identified in the Rab5 channel, revealed a difference between control and Clathrin dsRNA (Figure 4.10C).

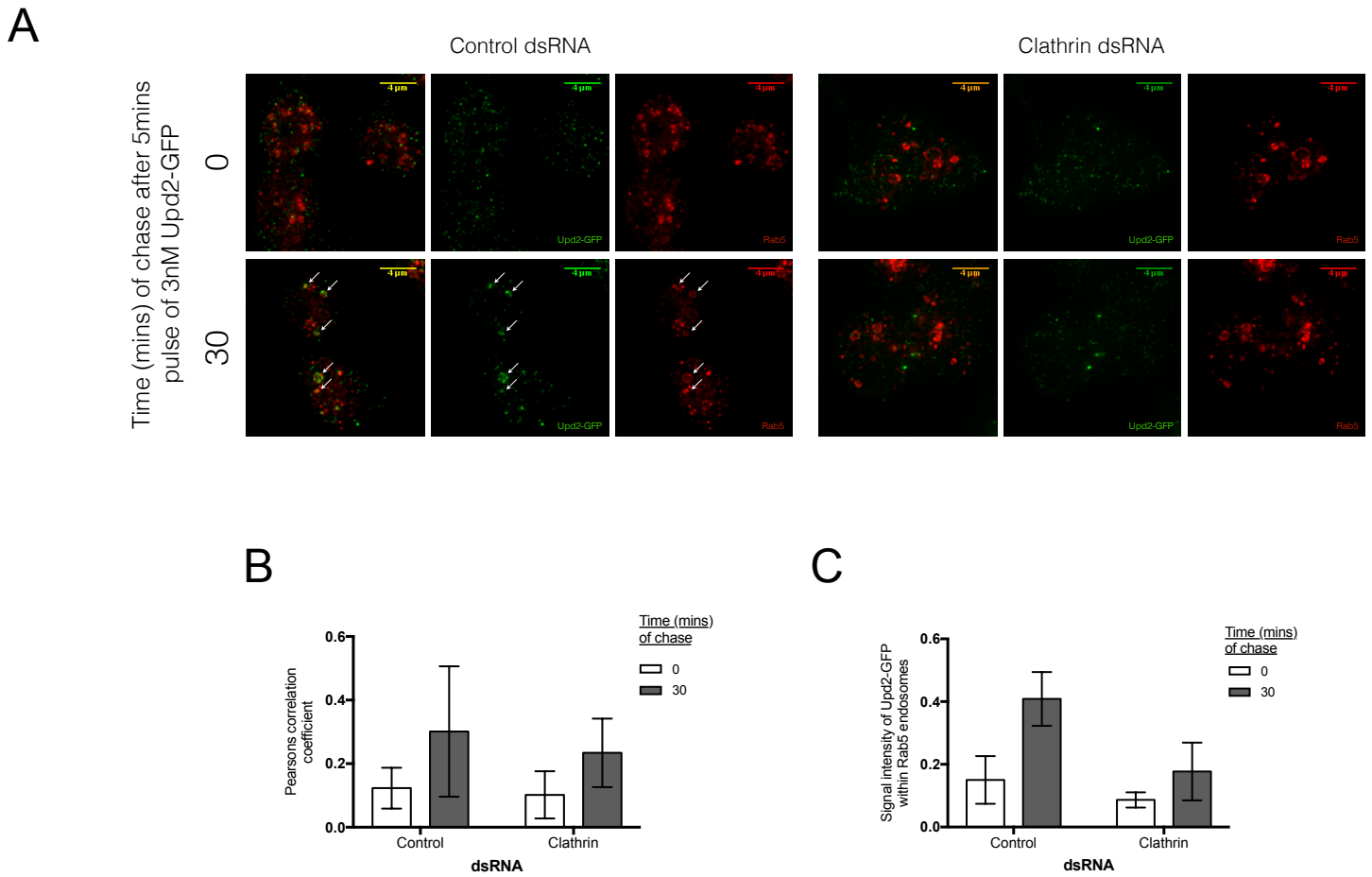


Figure 4.10: dsRNA against clathrin reduces colocalisation of Upd2-GFP with Rab5. A) Representative images of ligand uptake in cells treated with control or clathrin dsRNA 5 days prior to a 30min treatment with Upd2-GFP. Rab5 is shown in red, Upd2-GFP in green B) Pearson's correlation coefficient shows little difference between control and clathrin dsRNA treated cells. C) Using Squassh software the signal intensity of Upd2-GFP within the Rab5 objects is decreased in the clathrin knockdown. Data is mean ±SD from one experiment, n > 75 cells per condition (15 images).

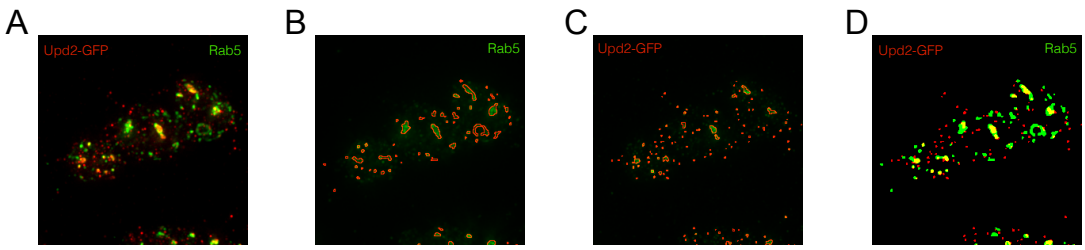


Figure 4.11: Example of Squassh analysis output. Cells were treated with 3nM Upd2-GFP for 5mins prior to 15min chase in fresh media. A) Original, merged, image prior to squassh segmentation. Upd2-GFP is shown in red and Rab5 in green to match software output. B) Object outlines for Rab5 channel using following parameters: background removal = 10 pixels, regularisation = 0.075, minimum object intensity = 0.1, local intensity estimation = low. C) Object outline for Upd2-GFP channel using following parameters: background removal = 10 pixels, regularisation = 0.04, minimum object intensity = 0.1, local intensity estimation = low. D) Overlay of object outlines to demonstrate colocalisation, shown in yellow.

4.1.4. Knockdown of Hrs results in enlarged Upd2-GFP endosomes

Having identified Squassh as a method to detect changes in colocalisation between Rab5 and Upd2-GFP, I asked where Upd2-GFP accumulates when Hrs is knocked down. As Hrs binds to ubiquitinated cargo, labelled for degradation at the early endosome (Chapter 1.2.2), I expect knockdown of Hrs to prevent progression to a late endosome. However, there have also been reports that Hrs is enriched at the plasma membrane of specific cell types (Welsch et al., 2006; Raiborg et al., 2008) and that Hrs prevents endosome fusion by inhibiting SNARE complex formation (Sun et al., 2003). Therefore, knockdown of Hrs may also effect earlier phases of Dome endocytosis.

Hrs knockdown was confirmed using microscopy (Figure 4.12A). After a 5-minute pulse and a 15-minute chase, Upd2-GFP was found in Rab5 positive structures. After 90-minutes the majority of Upd2-GFP puncta were no longer detectable in the control cells, however, in cells treated with Hrs dsRNA, Upd2-GFP was present in large endosomes positive for Rab5 (Figure 4.12B). These time-points were chosen as they are consistent with times used in the study by Vidal et al, 2010. Yet, analysis using both Pearsons colocalisation (Figure 4.12C) and Squassh signal intensity (Figure 4.12D) did not reflect these findings. This may be due to two reasons. Firstly, there were fewer individual Upd2-GFP structures at 90mins than at 15mins, likely due to endosome fusion, and therefore there are fewer Rab5 endosomes containing Upd2-

GFP. Hence, across the whole cell there is less Upd2-GFP signal in Rab5 positive structures. Secondly, the Rab5 signal labels the perimeter of these vesicles, whereas the ligand appears more luminal within the endosomes (Figure 4.12B). When the mask settings in Squassh selects Rab5 structures it does not fill the area within a shape. Therefore, the Upd2-GFP signal within a structure may not fall in the Rab5 mask. This is less of an issue in small endosomes, prior to Hrs knockdown. It is also important to remember that not all cells take up ligand and therefore apparent differences may be underestimated.

Therefore, I decided to investigate the size of the Upd2-GFP objects, independent of Rab5 staining. Although this does not provide quantitative information about the endosomal compartment in which the Upd2-GFP is trapped, it demonstrates that Upd2-GFP accumulates in cell treated with Hrs dsRNA (Figure 4.12E).

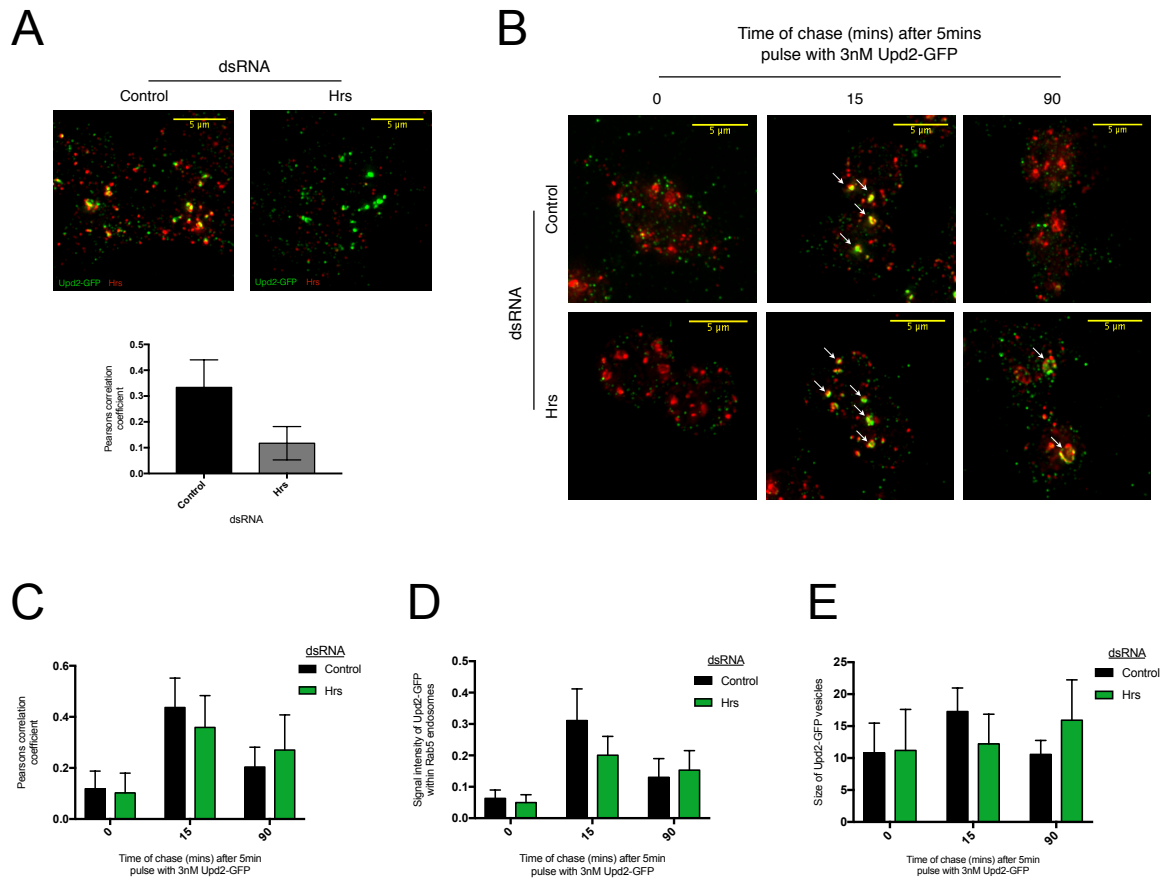


Figure 4.12: Knockdown of Hrs and Colocalisation with Rab5. A) Confirmation of Hrs knockdown, images from 15mins + 3nM Upd2-GFP. Control 63 cells, Hrs 50 cells. Mean ±SD. B) Representative images of Upd2-GFP localisation with Rab5, following Hrs knockdown. Upd2-GFP was pulsed for 5min, then chased for 0, 15 and 90 mins. C) Pearsons correlation coefficient between Upd2 and Rab5 signal. D) Upd2-GFP signal within Rab5 structures, as calculated with Squassh. E) Size of Upd2-GFP endosomes, irrespective of Rab5 signal. Graphs are mean ± SD of at least 60 cells from one independent experiment.

Repetition of this experiment demonstrated that Upd2-GFP objects are smaller at 15mins in the Hrs knockdown condition, compared to control (Figure 4.13). This may suggest a delay in the endosome fusion or maturation. What is also evident is that after 90mins, the size of Upd2-GFP endosomes in control dsRNA treated cells has decreased, whereas in the Hrs knockdown condition Upd2-GFP endosomes have continued to increase in size. This suggests that cargo is continuing to be delivered to the endosome, but may not be able to progress further through the trafficking pathway. This is likely as knockdown of Hrs has previously been shown to reduce ILV formation, and hence delivery to lysosomes, in *Drosophila* (Lloyd et al., 2002). These endosomes appear to be labelled with Rab5, although I have not been able to quantify this.

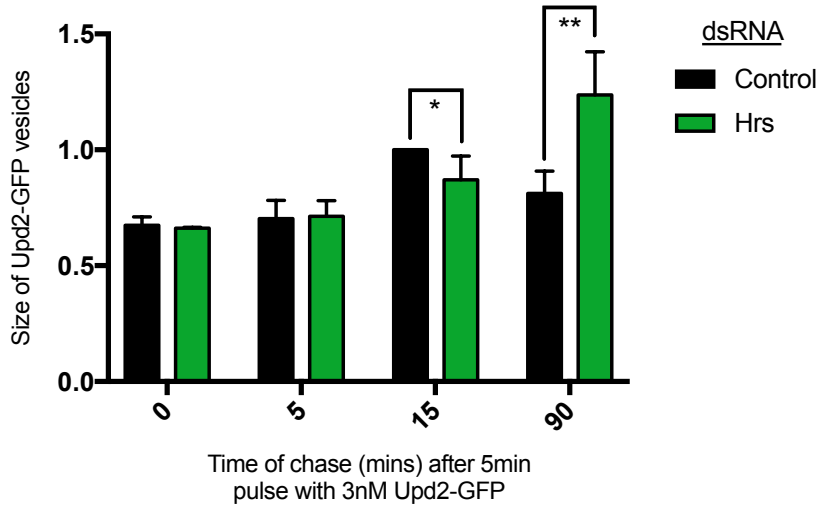


Figure 4.13: Knockdown of Hrs causes internalised Upd2-GFP to accumulate in large endosomes. Cells were treated with control of Hrs dsRNA 5 days prior to internalisation assay. Cells were treated for 5mins with 3nM Upd2-GFP, followed by a chase in fresh media for varying times. The 0 and 5minute time points are from 2 independent experiments, whereas 15 and 90minutes are from 5 and 4 independent experiments, respectively. Each experiment included at least 15 images, with Upd2-GFP object size being calculated by Squassh using the following parameters: background removal = 10 pixels, regularisation = 0.04, minimum object intensity = 0.1, local intensity estimation = low. Data is normalised to control dsRNA chased for 15mins and graph represents mean ± SD. Parametric, unpaired student t-test was carried out on data with 3 or more repeats, with *p≤0.05, **p≤0.01.

As I was unable to provide quantitative data about where Upd2-GFP is trapped using the Rab5 antibody, I decided to investigate colocalisation of Upd2-GFP with Rab7. Preliminary data suggests that at 30mins Upd2-GFP does not reach Rab7 compartments effectively. However, the 30min timepoint for Hrs knockdown is only

from 1 experiment, and therefore further experiments would be needed to understand delivery of Upd2-GFP to Rab7 endosomes following Hrs knockdown (Figure 4.14).

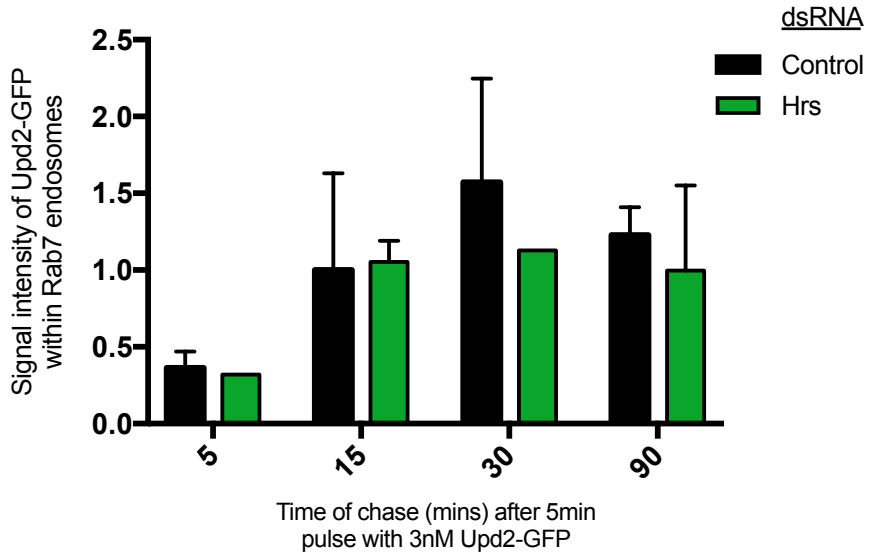


Figure 4.14: Preliminary data suggests Upd2-GFP does not efficiently traffic to Rab7 endosomes when Hrs is absent. Cells were treated with control of Hrs dsRNA 5 days prior to internalisation assay. Cells were treated for 5mins with 3nM Upd2-GFP, followed by a chase in fresh media for varying times. Data for 15 and 90-minute time-points are from 3 independent experiments, whereas others are from 1 or 2. Data is normalised to 15mins chase of cells treated with Hrs dsRNA and graph represents mean ± SD of at least 20 cells per experiment.

In summary, when Hrs is knocked down Upd2-GFP accumulates in large endosomes positive for Rab5. Preliminary data also suggests that Upd2-GFP cannot traffic through to Rab7 endosomes. It is important to note that I have not looked at markers of the recycling pathways which Upd2-GFP may be trafficked through if it is not reaching the lysosome for degradation.

4.2 Mechanism of endocytic uptake is dependent on ligand concentration

During this project, there have been hints that ligand concentration may determine which endocytic pathway is used for uptake. This has been documented for other receptors, such as EGFR (Sigismund et al., 2005). Although I have not studied this in much detail, mainly due to problems generating ligand at a high concentrations, it may suggest an important context-specific role for endocytosis in the *Drosophila* JAK/STAT

pathway. The first hint of this was when investigating the role of Hrs knockdown on *SOCS36E* expression (Chapter 3.1). At low ligand concentrations dsRNA against Hrs results in reduced *SOCS36E*, whereas at high ligand concentrations this effect is no longer seen (Figure 3.6). This may suggest, in the presence of high concentrations of Upd2, Dome can be trafficked through a different, signalling competent route, that is independent of Hrs.

Interestingly, when I looked at Upd2-GFP uptake via ELISA, I saw that at 3nM Upd2-GFP Clathrin knockdown decreases internalisation to the level of Dome knockdown. This suggests that receptor-mediated Upd2-GFP uptake is via CME (Figure 4.15). However, previous data using 25nM ligand results in ~50% decrease in uptake after clathrin knockdown, compared to Dome knockdown (Figure 1.6, Vogt and Smythe, unpublished). Therefore at higher ligand concentrations clathrin knockdown does not alter uptake to the same extent, suggesting Dome is also internalised via a CME independent route.

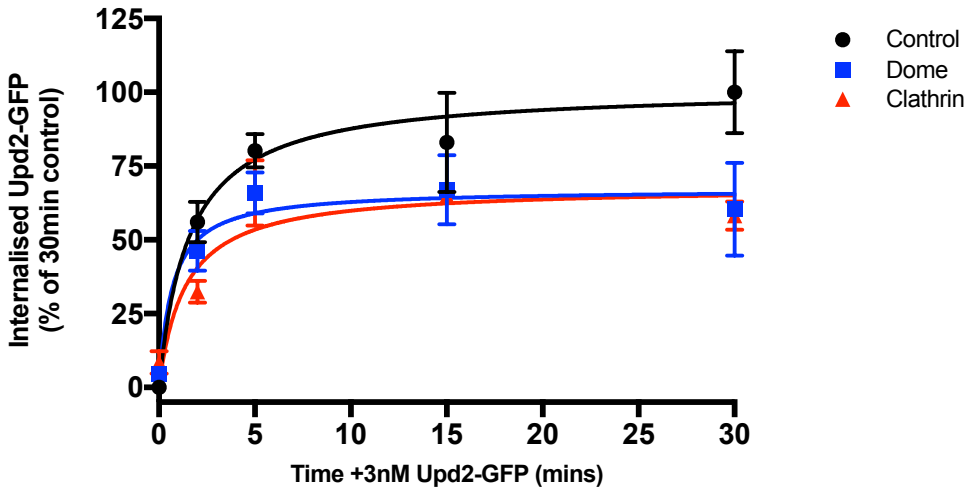


Figure 4.15: Uptake of Upd2-GFP is receptor- and Clathrin-dependent. S2R+ cells were treated for 5 days with control, CHC or Dome dsRNA. Cells were incubated with 3nM Upd2-GFP for indicated time points at 25ºC, after acid washes cell-lysates were analysed with the anti-GFP ELISA. The internal Upd2-GFP amount is represented as %, whereby 30min control was set to 100%. Graph represents mean of triplicates ± SD for one experiment.

4.3 Mechanism of clathrin-mediated endocytic uptake of the *Drosophila* JAK/STAT receptor Domeless

Transmembrane proteins are selected as cargo for CME through short, cytoplasmic peptide motifs that are bound by specific adaptor proteins. The motif responsible for CME of Dome has not been identified, however due to its evolutionary conservation with the mammalian IL-6/gp130 complex in mammals, I investigate the role of a conserved dileucine motif.

4.3.1. Mutation of potential AP2 binding motifs in the C-terminal of Dome.

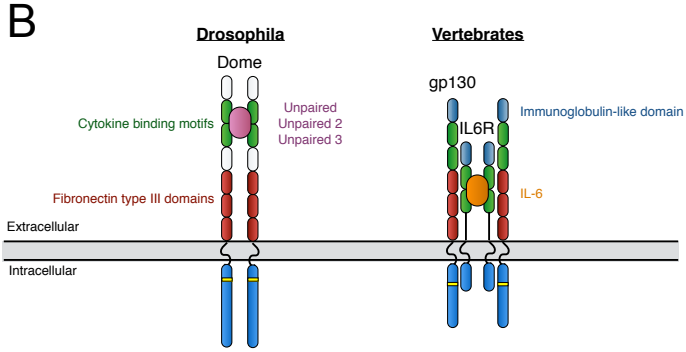
AP2, the major clathrin adaptor protein at the plasma membrane, binds cargo with YxxΦ and dileucine ([D/E]xxxL[L/I]) motifs, where Φ is any bulky hydrophobic side-chain (L, I, V, M or F), and x is any residue. The YxxΦ motif is recognised by the μ2 subunit of AP2 (Ohno et al., 1995), whereas the dileucine motif binding site is found in the α-σ2 hemicomplex of AP2 (Doray et al., 2007; Kelly et al., 2008). Other peptide motifs, such as NPXP motifs, are also crucial for cargo internalisation into clathrin coated pits, however accessory clathrin adaptor proteins are often required for their recognition (Pandey, 2009). Endocytic motifs are also key sorting signals for subcellular localisation of the cargo, i.e. whether it is targeted for degradation, recycling or storage (Pandey, 2010). However, here I focussed on identifying the residues responsible for Dome internalisation.

DsRNA against AP2 causes reduced Upd2-GFP internalisation (Vogt and Smythe, unpublished; Vidal et al., 2010), demonstrating a role for AP2 in the CME of Dome. Therefore, I searched the cytoplasmic sequence of Dome for YxxΦ and dileucine motifs that have the potential to be AP2-dependent internalisation motifs. I identified 3 potential YxxΦ motifs, and a single dileucine (Figure 4.16A).

A

YRKMSDIGLVLPQGIMETMKKPIDMGGLGLGLGPDSSVSGGIVCTRVDDSPPYTPQD
LPHDFSSCGSESSKLLLRTASSSGGGGCVDRDGYDDNHETGPISAVGPPTSYLAMR
HGLLVQNDRERERERDREQERERERQQQQQRESEMDREQSCTNGYIKPTQMKSWG
GNGPSDNDHTFSVPSTAMTAPMSQPLSQIPLSGYVPVPIPQSRFNPAPVQPFGSPAV
PSAATAAAASTFFPPAHLLNMDNYVQASDLHKLKPLVAAPLSQTGGPAFAGSSPATSP
PLQLPPVHAASPAAATPKMADIGYTTMEQLQLTGLIKPPLAATVGSPTHAAGGAPGG
GNQHSRLQPQINGYVTPQDLNAMAHNRHVL

B



C

	[D/E]XXXL[L/I]
	*
gp130	SESTQPLL
Dome	SESSKLLL
Dome^{LL985AA}	SESSKAAL
Dome$^{S979A/LL985AA}$	AESSKAAL
Dome$^{S981A/LL985AA}$	SEASKAAL
Dome^{AAASKAAL}	AAASKAAL
Dome^{E980A}	SASSKLLL
Dome$^{E980G/LL985AA}$	SGSSKAAL

Figure 4.16: Potential Dome internalisation motifs. A) Amino acid sequence of the cytoplasmic tail of Dome. Potential YxxΦ motifs are highlighted in red, the dileucine motif in blue, and the conserved STAT binding site in purple. B) Schematic of Drosophila Dome and the vertebrate gp130/IL6-R complex. C) Dileucine motif in the cytoplasmic tail of gp130 and Dome, and mutants produced to investigate internalisation.

Dome is most similar in sequence and structure to the vertebrate IL-6R (interleukin-6 receptor) complex (Brown et al., 2001; H.-W. Chen et al., 2002), which requires a dileucine motif for internalisation (Dittrich et al., 1994). The IL-6R complex is composed of two type I membrane proteins; the 80kDa IL-6R and the 130kDa gp130. IL-6R is the only IL-6 binding protein, yet is not responsible for IL-6 induced signalling (Taga et al., 1989). Instead, ligand binding causes IL-6R to interact with two gp130 molecules, whose homodimerisation results in kinase activation and downstream signalling (Murakami et al., 1993). This is different from the situation with Dome, whereby ligand binding and STAT activation occur from the same molecule, and therefore Dome may present an evolutionary hybrid of the IL-6R complex (Figure 4.16B).

The gp130 protein is also key for the efficient uptake of IL-6, with deletion of the cytoplasmic region inhibiting internalisation (Dittrich et al., 1994). Dittirich et al., 1996, later identified a dileucine motif and upstream serine (**S**ESTQP**LL**) as the crucial residues in IL-6 uptake. Mutation of the dileucine residues caused a decrease in ligand uptake and ultimately lessened receptor degradation. The rate of internalisation after

12mins is reduced by 50% upon mutation of an upstream serine to alanine, suggesting a role in rapid endocytosis (Dittrich et al., 1996). Although this serine was not shown to be specifically phosphorylated, phosphoserines have been identified previously as a requirement for the recognition of specific dileucine motifs (Pitcher et al., 1999). Based on this data, I therefore chose to prioritise study of the dileuicine box (SESSSKLL) on internalisation of Dome (Figure 4.16C).

I created various point mutations within the dileucine box (residues 979-987) of *pAc-Dome-flag* (Stec et al., 2013) (Figure 4.16), and investigated their effect on Dome internalisation using cell-surface biotinylation and subsequent endocytosis. Originally, I aimed to examine the uptake of Upd2-GFP and localisation of Dome-flag mutants using microscopy. Although Dome-flag colocalised with Upd2-GFP and Rab5 (Figure 4.17A), I saw little cell surface staining and multiple Dome-flag puncta devoid of Rab5 or Upd2-GFP (Figure 4.17B, red channel). This could be due to multiple reasons. Firstly, transfected Dome-flag may be inefficiently trafficked to the cell-surface and instead remain in the trans-golgi network. Secondly, only a low level of transfected Dome-flag may be at the cell surface at any one time, and therefore the signal from 'bright' intracellular structures prevents visualisation of Dome at the cell surface. A similar punctate staining with lack of visible cell surface dome was also seen with transfection of Dome-V5 in a study by Makki et al., 2010. Therefore, investigating changes to levels of cell-surface receptor was not possible using immunofluorescence. Although I had previously investigated Upd2-GFP uptake using microscopy and an anti-GFP ELISA, Upd2-GFP may still be internalised via bulk endocytosis such as macropinocytosis even if the receptor is not internalised and therefore may not be an appropriate readout. Hence, cell-surface biotinylation allowed me to specifically label cell-surface receptor and investigate its internalisation. For this I utilised EZ-link Sulfo-NHS-SS-biotin, a membrane impermeable form of biotin that reacts with primary amines in proteins to specifically label cell surface proteins (Figure 4.19A).

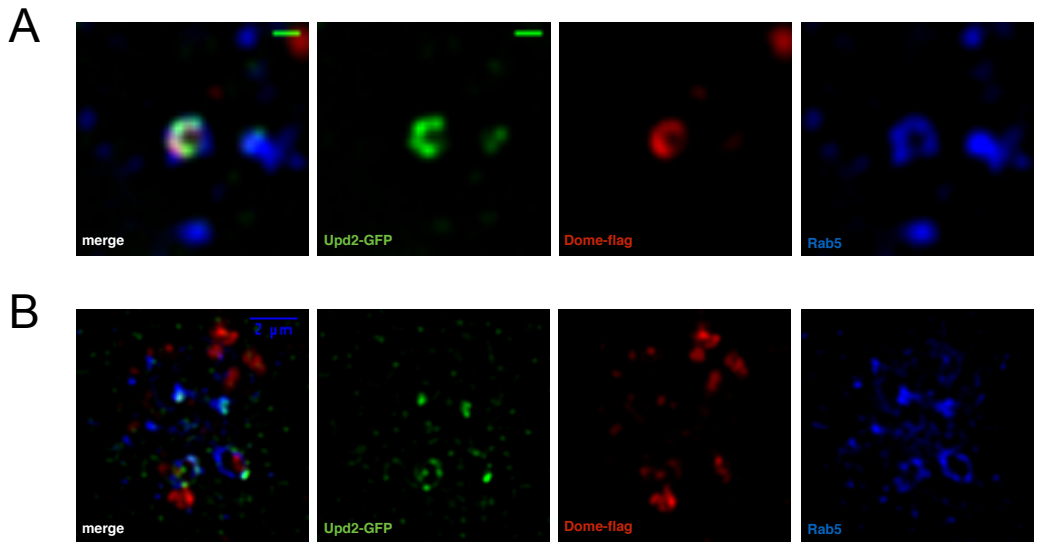


Figure 4.17: Images of Dome-flag expression in S2R+ cells. Cells were transfected with pAc-Dome-Flag for 48hrs prior to treatment with 3nM Upd2-GFP. Following fixation cells were stained for GFP (green), flag (red) and Rab5 (blue). A) scale bar = 0.5µm. B) scale bar = 2µm. Representative images taken on the OMX with widefield settings.

As a first experiment I wished to verify that Dome-flag was efficiently trafficked to the cell surface. Through optimisation of transfection conditions, biotinylated cell-surface Dome-flag was successfully detected after 48hrs (Figure 4.18). In subsequent experiments, I consistently saw ~30% of total Dome^{WT}-flag at the cell surface.

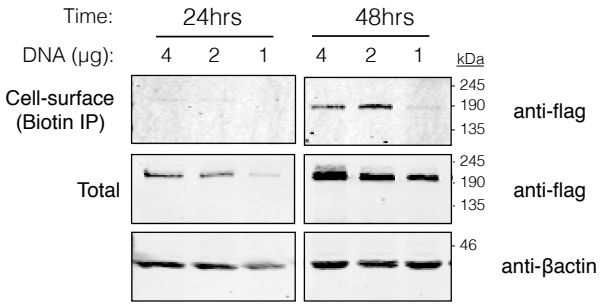


Figure 4.18: Example immunoblot after cell surface biotinylation. S2R+ cells were transfected for varying times with different concentrations of pAc-Dome-flag. Cell surface proteins were biotinylated for 1hr on ice, prior to lysis and incubation with streptavidin-agarose (Biotin IP). Sample were analysed via SDS-PAGE and western blot using the indicated antibodies.

Following this I carried out cell-surface biotinylation and subsequent endocytosis, followed by MESNa treatment. MESNa is a cell impermeable reducing agent that cleaves the disulphide bond in the linker of EZ-link Sulfo-NHS-SS-biotin. Therefore, cell surface biotin is cleaved whereas the internalised molecules remain biotinylated (Figure 4.19A) and can be isolated using streptavidin agarose. Utilising this procedure I demonstrated that ~50% of cell-surface Dome^{WT}-flag is internalised within 30mins (Figure 4.19B,C).

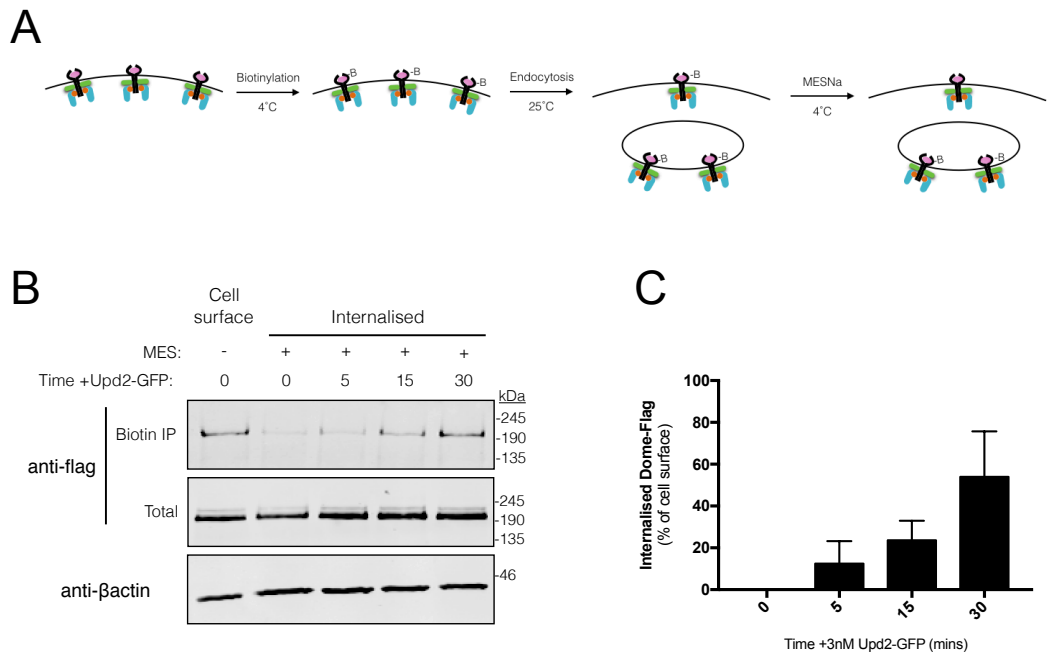


Figure 4.19: Internalisation of Dome-flag in S2R+ cells. A) Schematic of biotinylation and MESNa stripping process. Cells were transfected 48hrs prior to cell-surface biotinylation with EZ-link Sulfo-NHS-SS-Biotin and subsequent endocytosis in the presence of Upd2-GFP for varying times. MESNa was then used to strip cell surface biotin and therefore leave endocytosed Dome-Flag biotinylated. B) Example immunoblot of Dome-flag internalisation time-course. Biotinylated cell surface proteins were immunoprecipitated with streptavidin-agarose. Sample were analysed via SDS-PAGE and western blot using the indicated antibodies. C) Graph shows internalised Dome-flag as a percentage of total cell surface Dome at time zero. Signal from biotin IP blot was normalised to the total Dome-flag in the sample. This was then normalised by removing the 0 minute + MES sample, and calculating as a percentage of cell surface (no MES) sample. Data represents mean + SD for at least 4 independent experiments.

To examine the effect of the different mutants on receptor internalisation, I chose to investigate uptake after 15mins of Upd2-GFP stimulation. This allows study of uptake before the rate of internalisation has plateaued, and therefore may enable identification of mutants that slow the rate of endocytosis, that would not be evident at later timepoints. Mutation of the dileucine box to alanines (SESSKLL to AAASKAA)

prevented receptor internalisation after 15mins. Interestingly, the receptor is still internalised to a similar extent regardless of whether ligand is present or not (Figure 4.20). S2R+ cells secrete ligand into the media, and therefore Dome-flag may be occupied with ligand prior to addition of Upd2-GFP. With the aim of limiting this possibility, I added fresh media to the cells for 1hr prior to biotinylation. This should allow for occupied receptors to be internalised. However, this did not change the ligand independent uptake of Dome (data not shown).

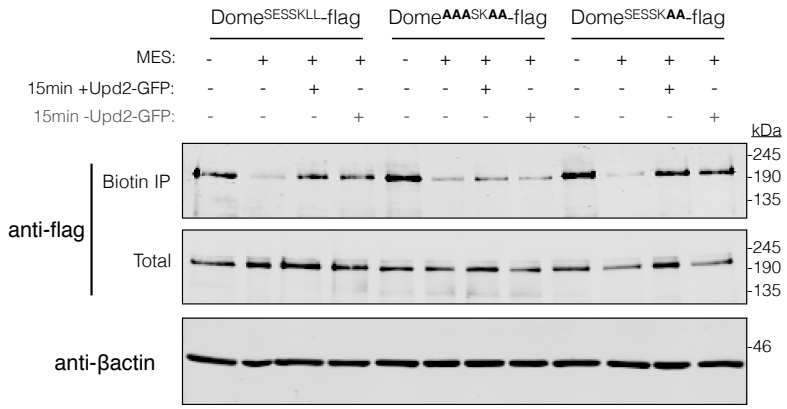


Figure 4.20: Dome internalisation is not dependent on ligand, but requires the SESSKLL box. Example immunoblot of cells transfected with Dome^{WT}-flag, Dome^{AAASKAA}-flag or Dome^{SESSKAA}-flag for 48hrs prior to cell surface biotinylation and endocytosis for 15-minutes +/- Upd2-GFP.

Interestingly, in the vertebrate system, gp130 association is key for internalisation of IL-6R, however it is independent of ligand binding. In unstimulated COS-7 cells, basal, constitutive endocytosis of gp130 occurs at similar rates to cell stimulated with IL-6. Pulldown experiments also demonstrated that AP2 is constantly associated with gp130, and that this interaction is not increased on stimulation (Thiel et al., 1998). Other groups have also demonstrated ligand independent uptake of Domeless (Fisher et al., 2016). Considering this, Dome may constantly associate with AP2 and therefore internalisation in the absence of ligand still occurs. Unfortunately, there is no antibody available for AP2 in *Drosophila*, and due to time constraints I did not produce a tagged version of AP2 to test this hypothesis.

Irrespective of whether stimulated or unstimulated Dome-flag is being internalised, the SESSKLL box within the cytoplasmic tail of Dome appears to be essential for receptor endocytosis (Figure 4.20, 4.21A and D). The expression of cell-surface Dome-flag

mutants is similar to that of WT (Figure 4.21C), suggesting the mutant receptors are being efficiently trafficked to the cell surface. Interestingly, unlike gp130, mutation of the two leucine residues alone does not alter internalisation (Figure 4.21A and D). Because of the role of the upstream serine on gp130 endocytosis, I chose to investigate the role of the two serines surrounding the glutamic acid. In combination with the dileucine mutation, the serine mutations appear to slightly inhibit internalisation by ~50%. Therefore, the serines appear important roles in internalisation, but other residues must also be key. Unfortunately, mutants where the serine had been altered, but the dileucine was not mutated, did not express within S2R+ cells.

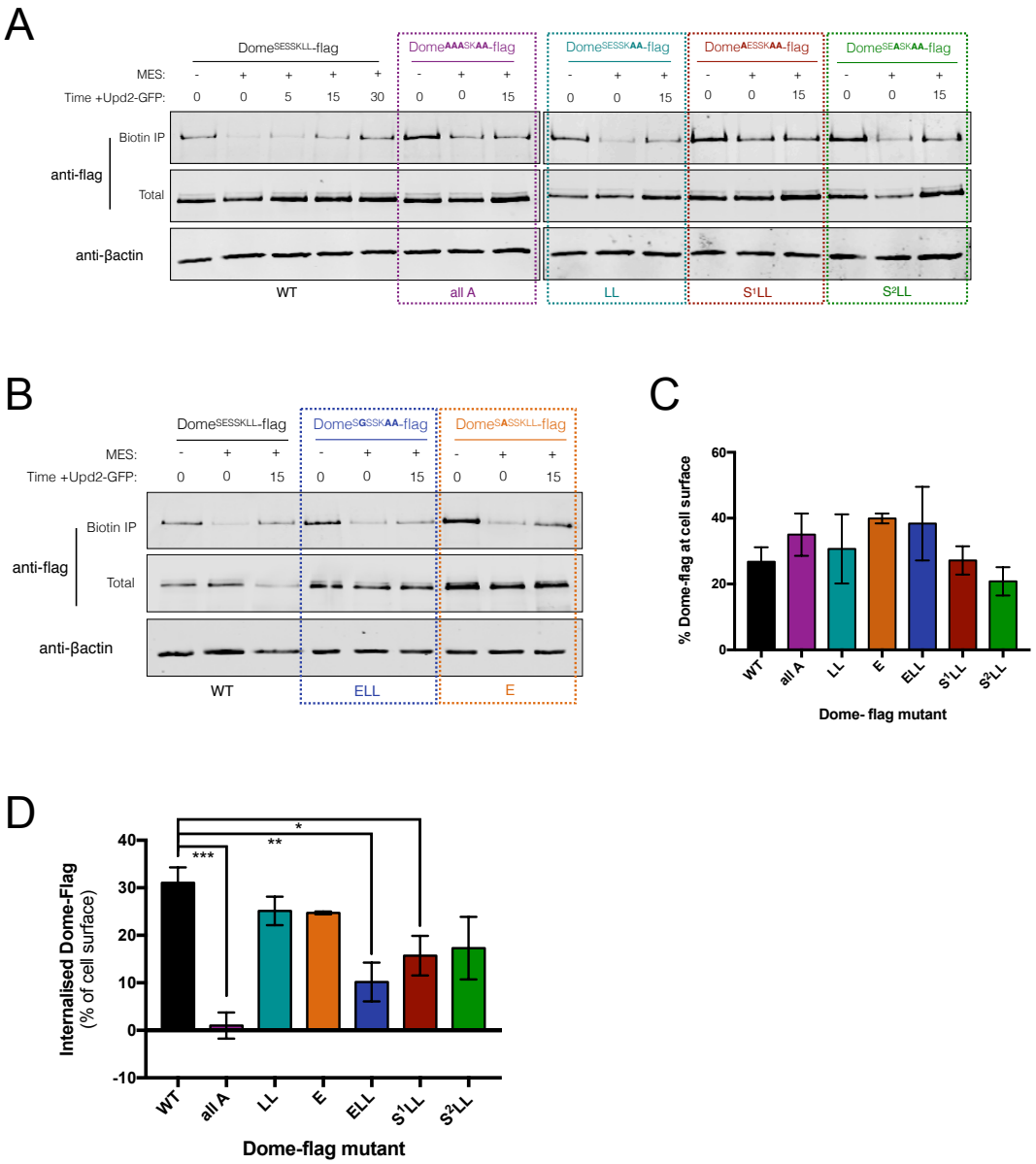


Figure 4.21: Uptake of Dome-flag internalisation motif mutants. A) Example immunoblot of cells transfected with Dome^{WT}-flag (WT), Dome$^{\mathbf{AAA}SK\mathbf{AA}}$-flag (all A), Dome$^{SESSK\mathbf{AA}}$-flag (LL), Dome$^{\mathbf{S}ESSK\mathbf{AA}}$-flag (S^{1}LL) or Dome$^{SE\mathbf{S}SK\mathbf{AA}}$-flag (S^{2}LL) for 48hrs prior to cell surface biotinylation and endocytosis for 15-minutes +/- Upd2-GFP. Membranes were probed with indicated antibodies. B) Example immunoblot of cells transfected with Dome^{WT}-flag (WT), Dome$^{\mathbf{S}G\mathbf{S}SK\mathbf{AA}}$-flag (ELL) or Dome$^{\mathbf{SA}SSKLL}$-flag (E) for 48hrs prior to cell surface biotinylation and endocytosis for 15-minutes +/- Upd2-GFP. Membranes were probed with indicated antibodies. C) Percentage of Dome-flag at cell surface compared to total levels of transfected Dome-flag. D) Percentage of cell-surface receptor that is internalised after 15mins at 25°C. Biotinylated cell surface Dome-flag after 0mins endocytosis and MESNa treatment was taken from internalised biotin, to remove background. Internalised Dome-flag was then calculated as a percentage total cell surface Dome-flag prior to MESNa treatment. Graphs represent mean ± SEM for at least 2 independent experiments (E =2repeats, ELL =3repeats, all other mutants ≥4 repeats).

An acidic residue (glutamate or aspartate) at -4 position is common in dileucine motifs, and its mutation has previously been shown to drastically decrease binding to the α-σ2 hemicomplex of AP2 (Doray et al., 2007). When the two leucine residues bind a hydrophobic pocket σ2 subunit, the acidic residue sits within a positively charged interface between α and σ2 subunits (Kelly et al., 2008). Therefore, both the dileucine and glutamate may be essential for AP2 binding. Hence, I mutated the glutamic acid to understand the requirement of this residue for Dome internalisation. Removal of this charged residue alone had no effect on receptor internalisation, yet mutation of both the glutamic acid and dileucine reduced internalisation by approximately 66% compared to WT (Figure 4.21B and D). This mutant did not completely abolish Dome uptake after 15mins, unlike the AAASKAA (all A) mutant. This suggests that the glutamic acid and dileucine are important but other residues, such as the surrounding serines, may also influence Dome internalisation.

In summary, I have identified a dileucine box within the cytoplasmic tail of Dome which is essential for internalisation. Mutation of this box prevents uptake after 15mins. I have not been able to identify the specific residues responsible for this result, yet mutation of the upstream glutamic acid or surrounding serines, in combination with mutation of the two leucines, reduces internalisation but is not sufficient for preventing endocytosis. Therefore, the residues surrounding the conserved [D/E]XXXL[L/I] motif may also play essential roles in enabling endocytosis. Considering that knockdown of AP2 inhibits JAK/STAT signalling (Chapter 3.1), I would expect mutation of the internalisation motif, and predicted abolishment of AP2 interaction, to have the same effects on signalling.

4.3.2. Production and characterisation of CRISPR/Cas9 engineered cell population lacking functional Dome.

In order to investigate the role of the internalisation deficient mutants on signalling, I chose to create a cell line lacking endogenous Dome. This will prevent WT endogenous Dome from signalling normally and masking any signalling changes due to mutation of Dome-flag. As transfection of S2R+ cells is low and variable, I created a cell line using CRISPR/Cas9 instead of using dsRNA so that levels of endogenous Dome remain constant across the experiments. This also allows me to use the Dome-flag constructs without the need for RNAi-resistant constructs.

In recent years, scientists have successfully adapted the type II CRISPR/Cas9 system (clustered regularly interspaced short palindromic repeats/CRISPR-associated) for genome engineering in cells lines and *in vivo*. This is a viral defence systems used by *Escherichia coli* that is essential in immunity. Bassett et al, 2014 adapted the CRISPR/Cas9 system from *Streptococcus pyogenes* for use in *Drosophila* cell lines. They created a vector expressing both the sgRNA and a puromycin resistant Cas9, under the control of the *Drosophila* U6 promoter and constitutive actin 5c promoter, respectively (pAc-sgRNA-Cas9). This allows for insertion of target sequences within the sgRNA and selection of cells expressing Cas9 with puromycin (Bassett et al., 2014). Utilising the detailed protocols documented in this paper, I created oligonucleotides to target exon 1 of the *dome* gene (Figure 4.22A), and inserted these into *pAc-sgRNA-Cas9*. These oligonucleotide sequences were carefully selected in order to limit any off-target effects (Chapter 2.7.8). Cells were transfected with the edited *pAc-sgRNA-Cas9* for 3 days, prior to a 7 day selection with puromycin. S2R+ cells cannot be single cell selected as they do not grow in isolation and therefore the cell-line created is a population of genetically modified cells and not a clonal cell-line where the mutation is constant.

As mentioned previously, there are no antibodies available for Dome. Therefore, I chose to investigate changes to the activation of downstream components of the JAK/STAT pathway as a readout of Dome functionality. Upon addition of Upd2-GFP STAT92E is phosphorylated, which results in a band-shift on an SDS-PAGE gel (discussed in Chapter 5.1). This is evident in WT cells, but absent in crDome cells (Dome CRISPR/Cas9 cells) (Figure 4.22B). This demonstrates that Upd2-GFP is now incapable of activating STAT92E, indicating Dome is no longer present in this cell-line. This result is maintained at later passages, suggesting the cell-line cannot recover. Addition of Upd2-GFP to the crDome cell line does not result in expression of the luciferase reporter or SOCS36E, yet this is rescued upon expression of Dome-flag (Figure 4.22C).

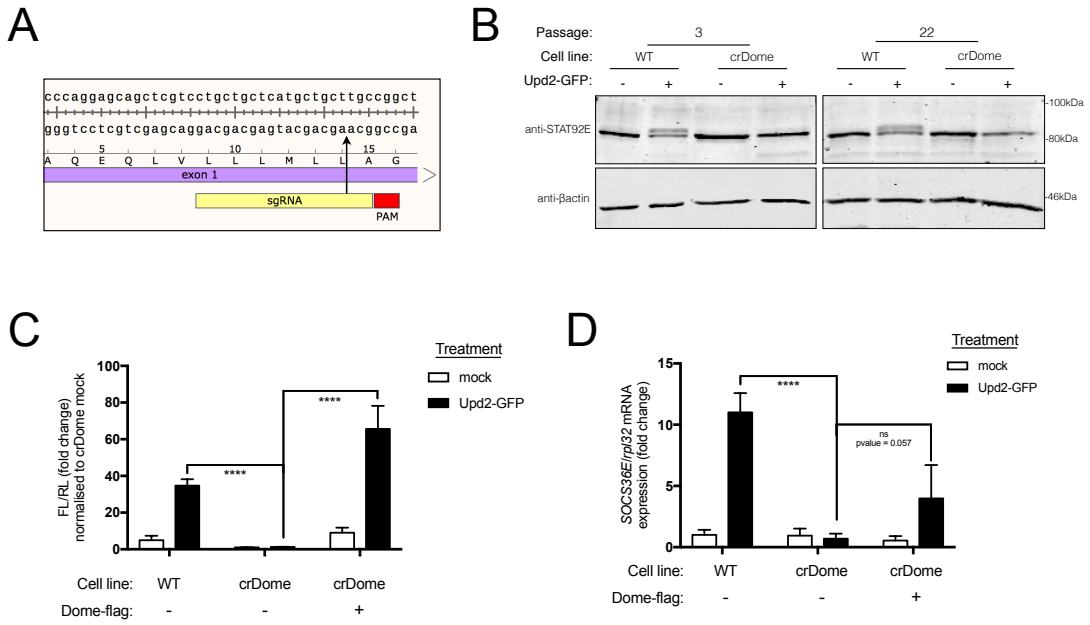


Figure 4.22: Generation and validation of Dome CRISPR knock out cell line. A) Location of sgRNA targeting within exon 1 of dome gene. B) Immunoblot demonstrating that in crDome STAT92E is not phosphorylated. 3nM Upd2-GFP was added to cells for 10mins prior to lysis. C) crDome cells cannot activate 10xSTATluciferase reporter, but expression of Dome-flag can rescue. Cells were transfected with Renilla luciferase (RL), 10xSTATluciferase (FL) and pAc5.1 or Dome-flag overnight before transferring to a 96well plate. Cells were treated for 30mins with 3nM Upd2-GFP, then incubated for 18hrs with fresh media before bioluminescence was measured. Graph represents the mean of triplicates + SEM for 3 independent experiments, and data is presented as FL/RL normalised to crDome mock treated cells. D) Upd2-GFP does not stimulate SOCS36E expression in crDome cells, but expression of Dome-flag rescues. Cells were transfected with pAc5.1 or Dome-flag for 2 days prior to incubation with 3nM Upd2-GFP for 2.5hrs and subsequent RNA extraction. SOCS36E mRNA levels were normalised to that of reference gene Rpl32, and presented as fold change compared to mock-treated WT cells. Results are expressed as means of duplicates + SEM for 2 independent experiments. Parametric, unpaired students t-test was carried out to compare Upd2-GFP stimulated samples, with ****p≤0.0001. For SOCS36E expression, the p-value of crDome compared to crDome cells transfected with Dome-flag was 0.057.

To confirm that transfection with *pAc-sgRNA-Cas9* caused mutations within the correct region of the genomic DNA I carried out a T7-endonuclease I assay, a method regularly used to investigate genomic mutations (Mashal et al, 1995, Guschin et al, 2010). The region surround the sgRNA target site is amplified from genomic DNA, then denatured and reannealed to cause the formation of mismatched heteroduplexes in DNA harbouring a mutation. These heteroduplexes are cut by the T7-endonuclease, and therefore when run on an agarose gel multiple bands are visible. Using this method, DNA amplified from WT cells was compared DNA amplified from crDome cells (Figure 4.23A). Addition of the T7 endonuclease to WT DNA did not produce any digested bands, whereas addition to crDome DNA resulted in multiple cleavage products

(Figure 4.23B). This suggests that the Cas9 has successfully cleaved and caused mutations within the *dome* gene.

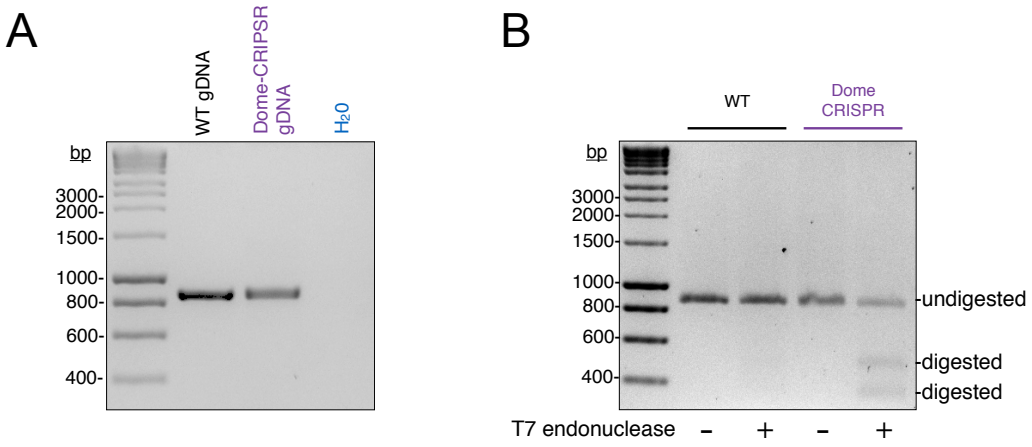


Figure 4.23: T7 endonuclease I assay demonstrates that Cas9 induced mutation in the *dome* gene. A) Genomic DNA was extracted from WT and crDome cell lines, and 871bp region around the sgRNA target site was amplified. B) Addition of T7 endonuclease to PCR product causes multiple bands for crDome cell line but not WT cells.

I further characterised this cell line by investigating internalisation of Upd2-GFP. Using both ELISA (Figure 4.24) and microscopy (Figure 4.25), I confirmed that the crDome cells were deficient in Upd2-GFP uptake. Although low levels of Upd2-GFP appear to be internalised in crDome cells using the ELISA, this result is similar to that of cells treated with Dome dsRNA (Figure 4.15) and is likely due to non-receptor mediated endocytosis, such as fluid-phase uptake.

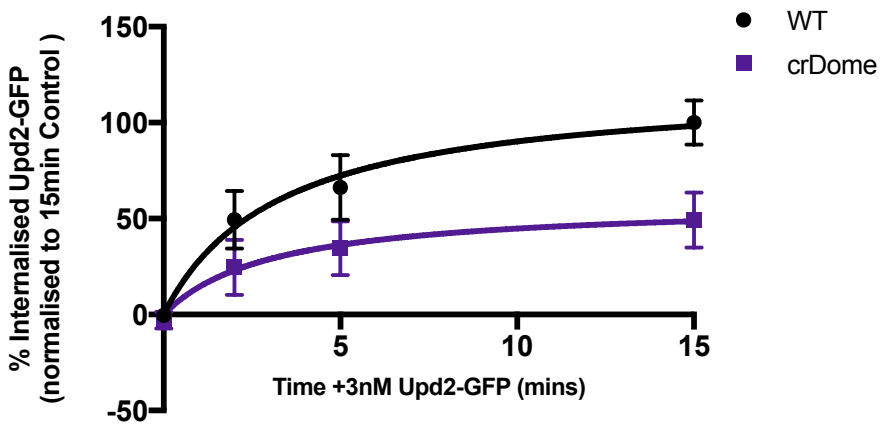


Figure 4.24: crDome cells are deficient in Upd2-GFP uptake. Cells were treated with 3nM Upd2-GFP for varying time-points, prior to acid wash and cell lysis. Internalised Upd2-GFP was measured via anti-GFP ELISA and is presented as %, whereby the 15-minute time-point was set to 100% in each experiment. Graph represents means ± SD for 4 independent experiments.

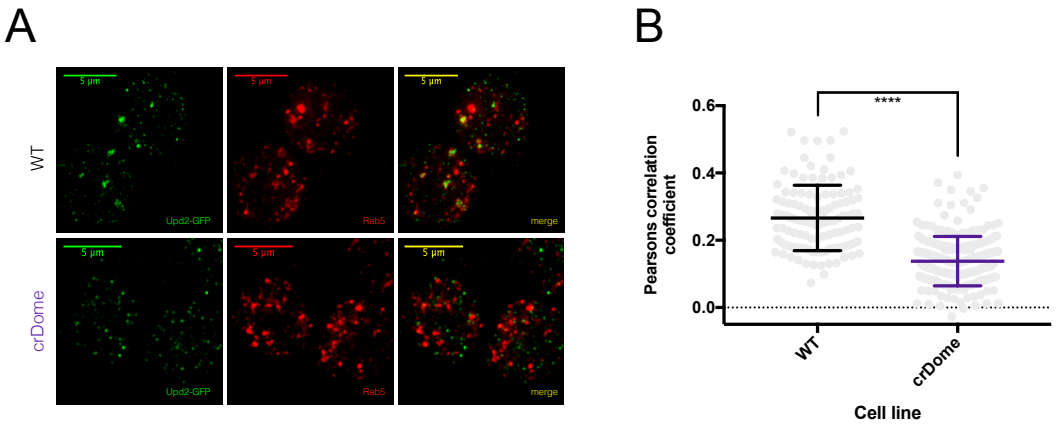


Figure 4.25: Upd2-GFP colocalisation with Rab5 is reduced in crDome cells. Cells were treated with 3nM Upd2-GFP for 15mins, prior to fixation and staining with anti-Rab and anti-GFP. A) Representative images of WT and crDome cells. B) Measure of colocalisation at 15mins via Pearson's correlation coefficient. Error bars represent mean ± SD for 1 experiment where over 100 cells were analysed per cell line (grey circles represent individual cells). Parametric, unpaired students t-test was carried out to compare Upd2-GFP stimulated samples, with ****p≤0.0001.

To investigate the role of the Dome internalisation motif on JAK/STAT signalling, I first wanted to confirm that the dynamics of Dome-flag internalisation in crDome cells is equivalent to WT cells. Uptake of Dome-flag appears similar in both cell lines (Figure 4.26A), though more variable at 30-uptake in crDome cells. This may be due to general reduction in transfection efficiency in this cell line (Chapter 4.27). Interestingly, during

biotinylation experiments I consistently saw a second anti-flag band that was not present in WT cells (Figure 4.26B). This band was the correct size for Cas9-flag from the *pAc-sgRNA-Cas9* plasmid. Comparing untransfected WT and crDome cells, along with crSTAT cells (Chapter 5), it is evident that this band is CRISPR dependent (Figure 4.26C) and hence likely to be due to integration of the Cas9 into the genome. Surprisingly, a portion of Cas9 is appears to be at the cell-surface. I therefore validated that EZ-link SS-NHS-biotin was not entering the cell during the biotinylation protocol by investigating if cytoplasmic β-actin was biotinylated. This was not the case (Figure 4.26B), suggesting that Cas9 is indeed at the cell-surface. This has not been documented before, and the reason for this expression remains unknown.

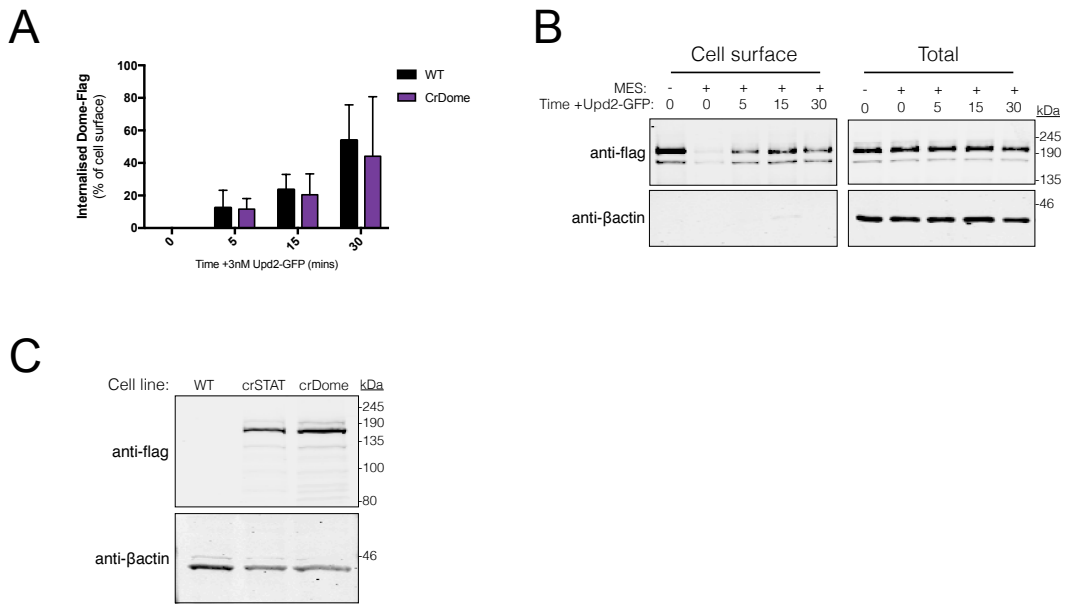


Figure 4.26: Dome-flag internalisation in crDome cells and extra flag band. A) Internalisation of cell-surface biotinylated Dome-flag in WT and crDome at varying time-points. Graphs represents mean + SD for at least 3 repeats for WT cells, and at least 2 repeats for crDome cells. B) Example immunoblot demonstrating Dome-flag uptake, extra flag band and lack of β-actin staining in biotinylated fraction. C) Immunoblot of untransfected WT, crSTAT and crDome cells reveals flag band is CRISPR specific.

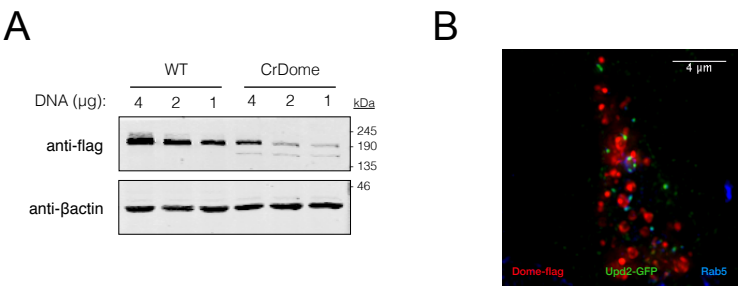


Figure 4.27: Transfection efficiency of WT cells vs Dome-flag cells. A) Example immunoblot of WT vs CrDome cells transfected with Dome-flag for 3 days shows reduced transfection in CrDome cells. B) Microscopy image demonstrating high levels of transfection in a single CrDome cell. Cells were transfected with Dome-flag 3 days prior to fixation with PFA and staining with anti-GFP, anti-flag and anti-Rab5. Cells were imaged on the OMX SIM, using widefield settings.

4.3.3. The role of the Dome internalisation motif in JAK/STAT signalling.

In Chapter 3.1, I demonstrated that knockdown of AP2 inhibits the expression of the *10xSTATluciferase* reporter and *SOCS36E* in response to Upd2-GFP. I have since identified a conserved SESSKLL motif required for Dome internalisation which, as it is a typical dileucine motif, I predict it to interact with AP2. Therefore, I expect to observe the same signalling defects with the internalisation motif as with AP2 knockdown. To study this I utilised the *10xSTATluciferase* reporter assay, as transfection of Dome-flag rescues expression of the reporter in crDome cells (Figure 4.22C). *SOCS36E* expression is also rescued upon Dome-flag transfection (Figure 4.22D), however this assay is more sensitive to difference in transfection efficiency between experiments (discussed in Chapter 5.5.3).

In this experiment I included two further Dome-flag mutants that are expected to prevent JAK/STAT signalling. STAT molecules have been shown to bind conserved YxxQ motifs in the receptor tail, and that this association is required for STAT activation (Stahl et al., 1994). A YTPQ sequence (residues 966-969) in the C-terminal of Dome is the only conserved YxxQ motif in the C-terminal of Dome. It is assumed to bind to STAT92E, however this has not been confirmed. The *Drosophila* kinase, Hop, binds to Dome a conserved proline residue (P925). Mutation of the residue results in decreased Hop binding, reduced Dome phosphorylation and decreased JAK/STAT luciferase reporter expression (Fisher et al., 2016). I therefore mutated both of these sequences in Dome-flag as positive controls for signal inhibition.

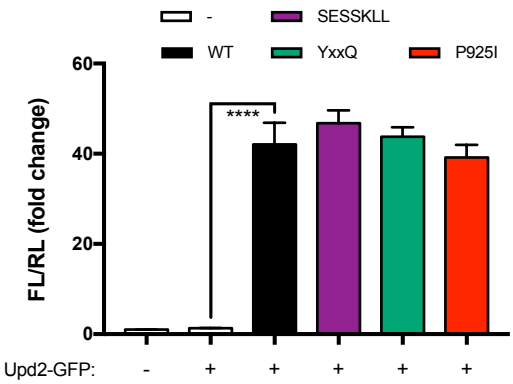


Figure 4.28: Dome-flag mutants do not alter 10xSTATluciferase expression in crDome cells. crDome cells were transfected with pAc-Ren (RL), 10xSTATluciferase (FL) reporter and pAc5.1 (-) or Dome-flag mutants. Cells were stimulated with 0.75 Upd2-GFP for 30mins, then incubated in fresh media for 18hrs. Graph represents mean of triplicates ± SEM for 2 independent experiments. Data is presented as FL/RL, normalised to unstimulated empty vector (-). Parametric, unpaired students T-test was performed, with ****p≤0.0001.

Unexpectedly, mutation of the Dome internalisation motif in crDome cells lead to a similar level of *10xSTATluciferase* activation as WT Dome (Figure 4.28). However, neither the YxxQ or P925I mutant reduced signalling either, suggesting issues with the luciferase assay and/or crDome cells. I previously observed that Dome-flag transfection levels were lower in crDome cells than in WT cells (Figure 27). Fewer crDome cells were transfected with Dome-flag than compared to WT cells, however the level of Dome-flag was higher in individual cells. Although the transfection was variable across the cell population, the image in Figure 4.27B demonstrates the high level of transfection seen in a single CrDome cell. This level of expression could force oligomerisation of the receptor and abnormal signalling. I therefore decided to examine if mutation of the internalisation motif alters signalling in WT cells.

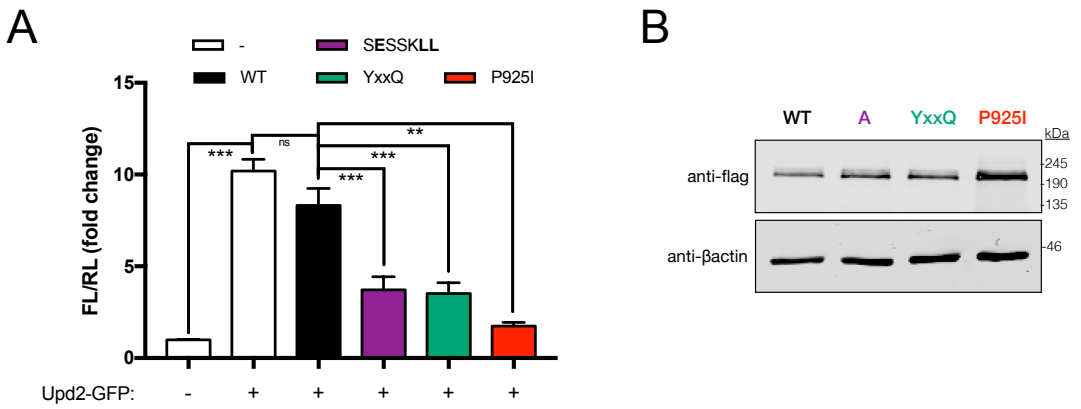


Figure 4.29: Mutation of Dome Internalisation motif inhibits ligand induced 10xSTATluciferase reporter activation in WT cells. A) WT cells were transfected with pAc-Ren (RL), 10xSTATluciferase (FL) reporter and pAc5.1 (-) or Dome-flag mutants. Cells were stimulated with 0.75 Upd2-GFP for 30mins, then incubated in fresh media for 18hrs. Graph represents mean of triplicates ± SEM for 4 independent experiments. Data is presented as FL/RL, normalised to unstimulated empty vector (-). Parametric, unpaired students T-test was performed, with **p≤0.01, ***p≤0.001. B) Example immunoblot of Dome-flag constructs and their relative transfection efficiencies.

In WT cells mutation of the STAT and Hop binding sites in Dome-flag resulted in a dominant-negative effect on Upd2-GFP induced *10xSTATLuciferase* expression (Figure 4.29). Mutation of the Dome internalisation motif also has a dominant-negative effect on JAK/STAT signalling, corroborating that endocytosis is required for *10xSTATLuciferase* activation. All constructs appear to be expressed at a similar level, apart from P925I which may account for its greater dominant negative effect.

4.4 Summary

This chapter builds upon our current understanding of Dome endocytic trafficking and identifies a motif required for Dome internalisation and *10xSTATLuciferase* activation. Firstly, I utilised recently developed antibodies to confirm that Upd2/Dome trafficks through Rab5, Hrs and Rab7 labelled endosomes. Knockdown of Hrs, which inhibits expression of *SOCS36E* but does not inhibit expression of *10xSTATluciferase,* causes accumulation of Upd2/Dome into large endosomes and appears to prevent trafficking to Rab7-positive endosomes. I then investigated motifs responsible for Dome internalisation. The mutation of a conserved [D/E]xxxL[L/I] in the cytoplasmic tail of Dome prevents internalisation in both the presence and absence of ligand. This suggests that Dome is constitutively endocytosed, independent of stimulation. Finally, I illustrated that this dileucine box and Dome internalisation is key for expression of the

10xSTATluciferase reporter. This confirms results from Chapter3.1, where AP2 knockdown also prevents *10xSTATluciferase* reporter expression.

Chapter 5. Characterisation of STAT92E as the target of endocytic regulation

The JAK/STAT pathway is a signal transduction pathway, whereby an extracellular cue is transferred into an intracellular signal to produce a defined cellular response. This occurs through a cascade of proteins that undergo successive biochemical changes to alter their activity. Consequently, regulation of any of these proteins can modulate downstream signalling. The JAK/STAT pathway in *Drosophila* is a straightforward example of a signalling pathway, with only four main components (ligand, receptor, kinase and transcription factor). My results have demonstrated that internalisation of Domeless, via a conserved [D/E]xxxL[L/I] motif (Chapter 4), and endocytic trafficking, is required for the transcription of a subset of *Drosophila* JAK/STAT targets (Chapter 3). Therefore, I aimed to identify the pathway component that undergoes endocytic regulation, and elucidate the molecular mechanism that results in the modulation of distinct JAK/STAT targets.

5.1 Phosphorylation of STAT92E is necessary but not sufficient for transcription of all JAK/STAT targets

Upon ligand activation of Dome, STAT92E is phosphorylated by Hopscotch at a conserved tyrosine residue (Yan et al., 1996). This residue is conserved across all vertebrate STATs, and its phosphorylation is essential for ligand-induced STAT dimerisation and target expression (Darnell, 1997). If STAT92E is phosphorylated as expected, the receptor and kinase are presumed to be behaving normally. I first addressed whether STAT92E phosphorylation was affected by endocytosis. The antibody previously used to investigate STAT92E phosphorylation (Baeg et al., 2005; Shi et al., 2008) was discontinued before this project, however phosphorylation of STAT92E results in an extra band that runs more slowly on an SDS-PAGE gel (Shi et al., 2008). The retarded mobility of phosphoproteins is a well-documented phenomenon (Smith et al., 1989) and for STAT92E is likely due to a change in charge, and the conformational change that occurs when STAT proteins are phosphorylated (Mao et al., 2005; Wenta et al., 2008). This method allows visualisation of STAT92E phosphorylation without the need to carry out immunoprecipitation and subsequent blotting with phosphotyrosine antibodies. I confirmed that the observed band-shift was

due to phosphorylation by treatment with phosphatases (Figure 5.1A). Using this method, I demonstrated that STAT92E phosphorylation is dependent on Upd2-GFP concentration (Figure 5.1B).

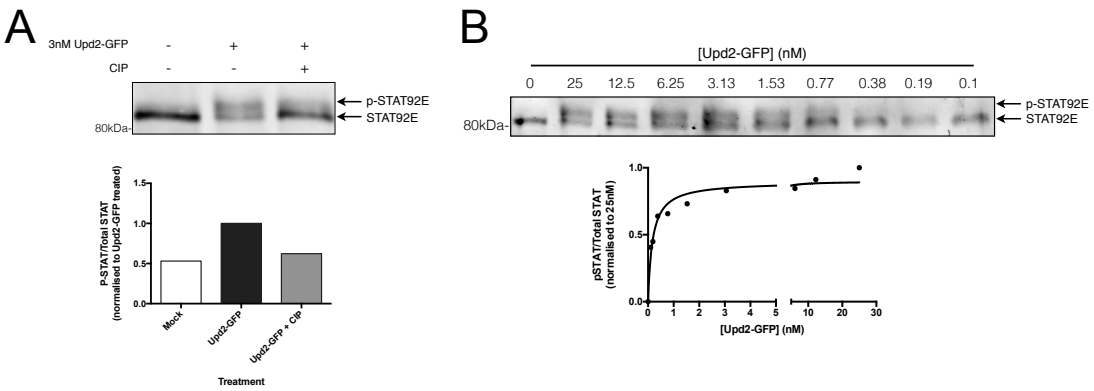


Figure 5.1: Upd2-GFP causes phosphorylation of STAT92E, which results in a species whose migration is slower during SDS-PAGE. A) Immunoblot of phosphatase treated STAT92E. S2R+ cells were treated with 3nM Upd2-GFP for 10mins and lysates incubated with anti-STAT92E antibodies. Immunoprecipitated protein was then treated with CIP, and analysed via SDS-PAGE and immunoblotting with anti-STAT92E. Graph represents proportion of p-STAT92E compared to total STAT92E. B) Immunoblot demonstrating the concentration-dependent effect of Upd2-GFP on STAT92E phosphorylation. Cells were treated with varying concentrations of Upd2-GFP for 10mins and lysates were analysed via SDS-PAGE and immunoblotting with anti-STAT92E.

Knockdown of AP2 prevents Upd2-GFP induced expression of both *10xSTATluciferase* reporter and *SOCS36E* (Chapter 3). Interestingly, the phosphorylation state of STAT92E is not altered when cells have been treated with dsRNA against AP2 (Figure 5.2A,B). This suggests two things: 1) both the receptor and kinase are functional when AP2 is knocked down, and 2) STAT92E is 'transcriptionally active' yet unable to regulate expression of specific targets. Phosphorylation of STAT92E was also maintained when Hrs and TSG101 were targeted with dsRNA (Figure 5.2C).

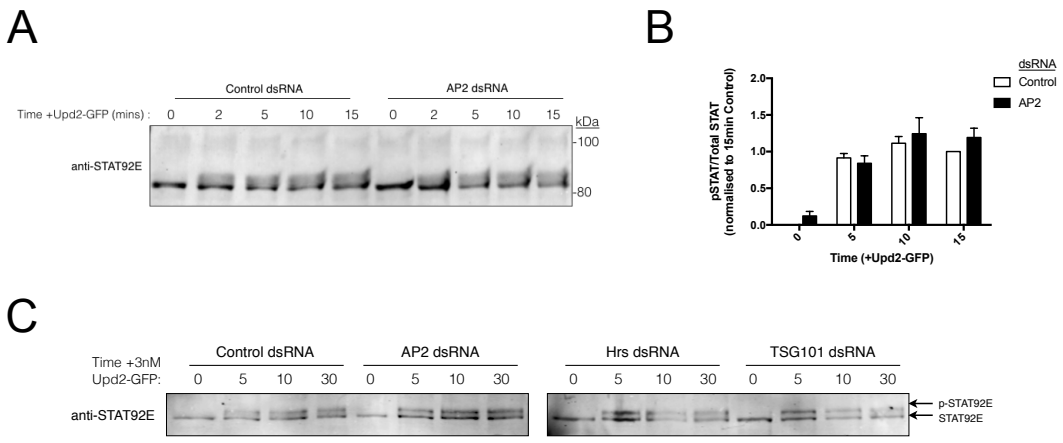


Figure 5.2: dsRNA knockdown of AP2, Hrs and TSG101 does not affect STAT92E phosphorylation. A) Representative immunoblot of control vs AP2 knockdown S2R+ cells treated with 3nM Upd2-GFP at 25°C for the indicated times. Cells were treated with targeting dsRNA cells and incubated for 5days at 25°C. Total protein extract was analysed by SDS-PAGE and immunoblotted with anti-STAT92E. B) Quantification of STAT92E phosphorylation after AP2 knockdown. pSTAT is normalised to total STAT, and compared to control sample treated for 15mins with Upd2-GFP. Results are expressed as mean +SEM from 4 independent experiments. Using students t-test there are no statistically significant differences between control and AP2 knockdown samples. C) Immunoblot of control vs AP2, Hrs or TSG101 knockdown S2R+ cells treated with 3nM Upd2-GFP at 25°C for the indicated times. Total protein extract was analysed by SDS-PAGE and immunoblotted with anti-STAT92E.

Interestingly, changing from HRP-based western blot techniques to the Licor system that uses secondary antibodies conjugated to infrared dyes, increased resolution of the band-shift and lead to visualisation of multiple STAT92E bands after treatment with ligand (Figure 5.3A). This is indicative of multiple phosphorylated protein states, and was confirmed using a PhosTAG gel (Figure 5.3B). Multiple phosphorylated species were seen for STAT92E isolated from untreated cells, but the level of these increased upon addition of Upd2-GFP. Unfortunately, I was unable to determine whether specific phosphorylated species were altered during endocytosis.

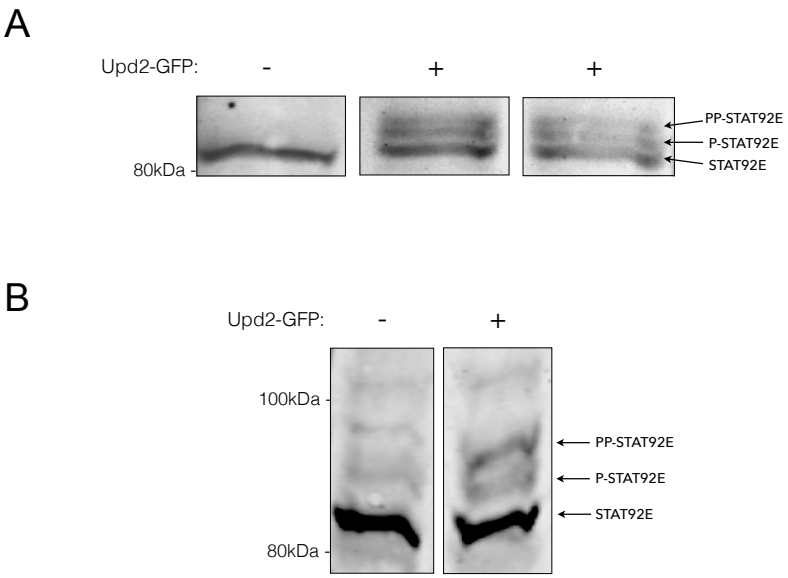


Figure 5.3: Multiple phosphorylated forms of STAT92E are visible by western blot. A) Standard immunoblot protocol imaged on the Licor Odyssey reveals 3 bands upon addition of 3nM Upd2-GFP. B) Lysate run on PhosTAG gel shows multiple phosphorylated bands.

As my data suggested that Dome can be internalised via different endocytic routes capable of modulating JAK/STAT signalling (Chapter 4.2; Figure 3.5B, Figure 4.15), I investigated whether AP2 knockdown altered STAT92E phosphorylation at different concentrations of Upd2-GFP. However, there were no obvious changes in the phosphorylation of STAT92E at different ligand concentrations when AP2 knockdown was compared to control knockdown (Figure 5.4).

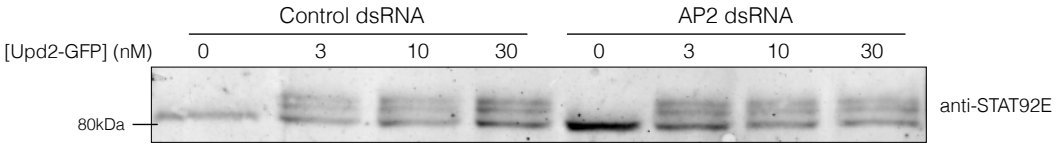


Figure 5.4: Ligand concentration does not alter the effect of AP2 dsRNA on STAT92E phosphorylation. S2R+ cells were treated with control or AP2 dsRNA for 5 days prior to treatment with indicated concentrations of Upd2-GFP for 15mins. Cell lysates were run on an 8% SDS-PAGE gel, transferred to nitrocellulose membrane and blotted with anti-STAT92E antibody.

Therefore, Dome and Hop appear to be functioning normally when endocytic components are knocked down or when the receptor enters via pathways other than CME, suggesting that the transcription factor itself is subject to endocytic regulation. Interestingly, although tyrosine phosphorylation of STAT92E is required for its transcriptional activity, it does not appear to be sufficient for expression of the *10xSTATluciferase* reporter. There also appears to be multiple phosphorylated STAT92E residues, however the exact residues involved and their role(s) are unknown.

5.2 STAT92E-GFP nuclear import is not affected by knockdown of endocytic components.

STAT92E is a transcription factor and therefore requires nuclear import to elicit its transcriptional functions. As STAT92E is phosphorylated at the canonical Y^{704} residue following AP2 knockdown (Figure 5.2), I investigated whether STAT92E still enters the nucleus in these conditions. In mammalian STATs, the N-terminal domain is essential for nuclear accumulation of phosphorylated STAT1 (Strehlow and Schindler, 1998; Meissner et al., 2004). Karsten et al, 2006, demonstrated that a methionine 647 to histidine mutation results in increased STAT92E nuclear accumulation, but has a dominant-negative effect on signalling even though efficiently phosphorylated. Therefore, endocytosis may alter transcription by preventing STAT92E nuclear accumulation through a mechanism independent of tyrosine phosphorylation.

Nuclear accumulation can be visualised in S2R+ cells transfected with *pAc-STAT-GFP* and treated with Upd2-GFP (Figure 5.5), with a maximum accumulation reached after 30 minutes stimulation. This is comparable to nuclear accumulation of mammalian STATs (McBride et al., 2002) and the time-point at which STAT92E phosphorylation is maximal (Vogt and Smythe, unpublished). The STAT-GFP construct contains STAT92E isoform C fused to a C-terminal eGFP, under the control of an Actin5c promotor (Karsten et al., 2006). Isoform C lacks 7 amino acids at residue 699, resulting in the conserved tyrosine being at position 704. This construct was characterised by Karsten, 2007, where it was demonstrated that JAK/STAT pathway activation resulted specifically in nuclear accumulation of STAT-GFP and not GFP alone, and that

transfection of STAT-GFP increased expressed of a luciferase reporter (Karsten, 2007). Enrichment of STAT-GFP in the nucleus prior to stimulation is likely due to a high level of protein overexpression as GFP inherently translocates to the nucleus (Seibel et al., 2007).

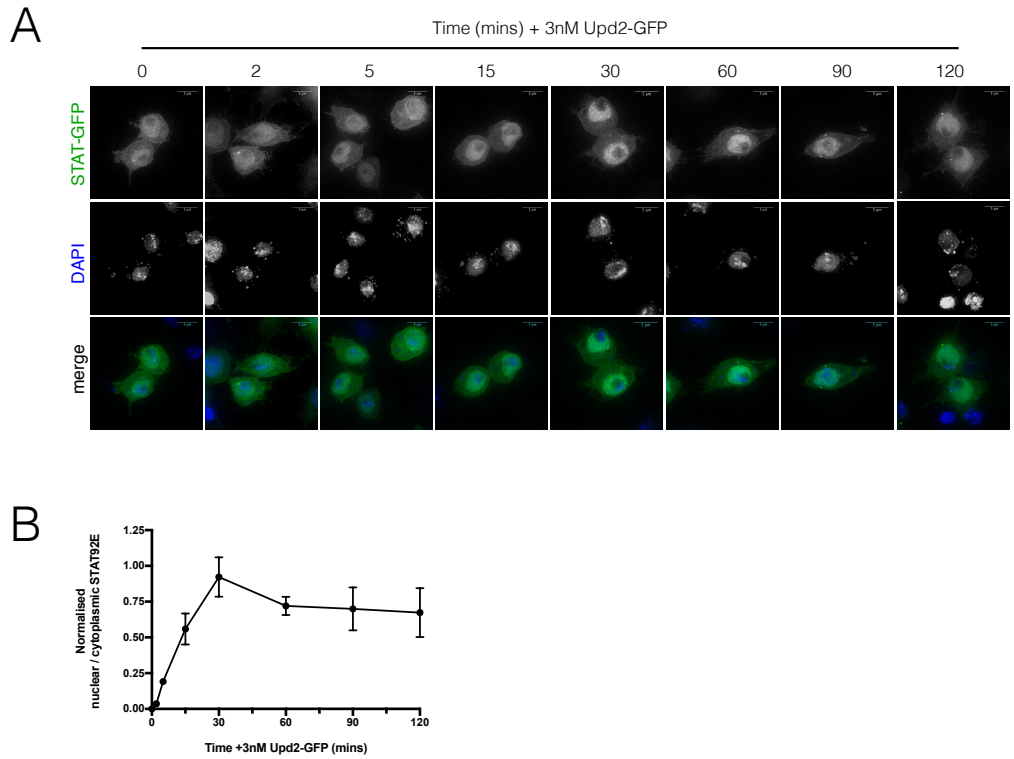


Figure 5.5: STAT-GFP translocates to the nucleus after addition of Upd2-GFP.
A) Representative images of S2R+ cells transfected with STAT-GFP and treated with 3nM Upd2-GFP for indicated times. B) STAT-GFP signal in the nucleus divided by the cytoplasmic signal, and normalised to 0mins. Data is presented as mean of at least 15 cells per time point ± SEM for 2 independent experiments.

Whilst knockdown of Dome abolishes ligand-induced nuclear accumulation of STAT92E, knockdown of AP2 and Hrs does not alter the ability of STAT92E to shuttle to the nucleus in response to ligand (Figure 5.6). Combined with efficient phosphorylation, it appears that STAT92E enters the nucleus and is transcriptionally active for some pathway targets, *lama* and *chinmo*, yet not transcriptionally competent for the expression of *10xSTATluciferase* and *SOCS36E* when endocytic trafficking is perturbed.

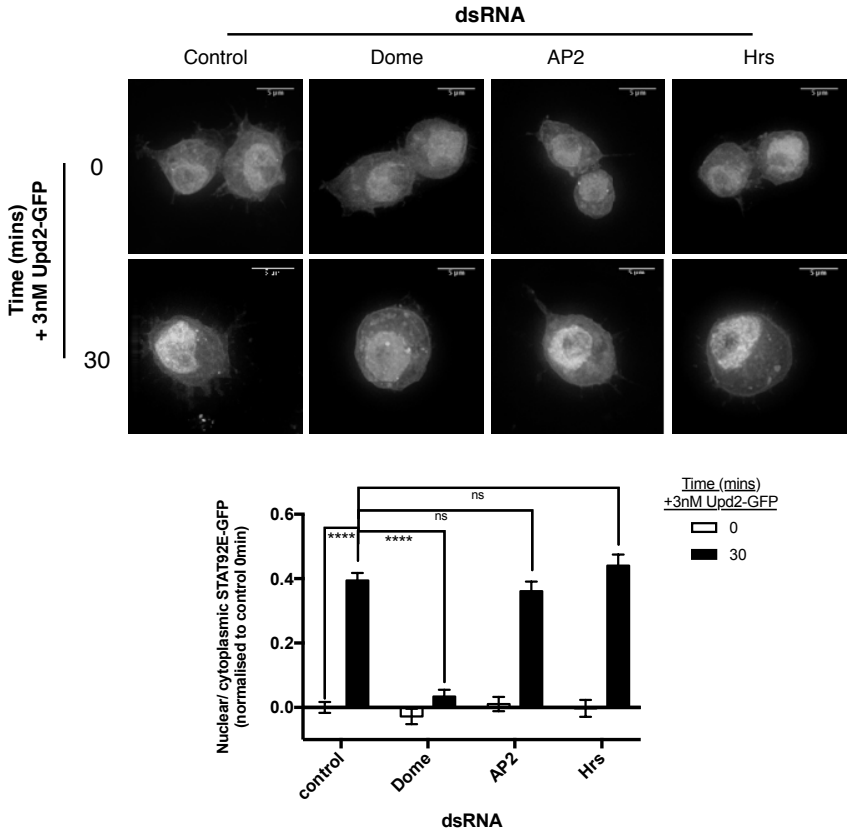


Figure 5.6: STAT92E-GFP can still enter the nucleus when endocytic components are knocked down. Representative images of cells treated with dsRNA for 3days prior to transfection of STAT-GFP. Cells were treated with 3nM Upd2-GFP for 0 or 30mins. Data is presented as mean ± SEM for 3 independent experiments. Nuclear signal was divided by cytoplasmic signal, and normalised to 0mins control cells. Data is presented as mean for at least 20 cells per condition ± SEM for 3 independent experiments, with parametric, unpaired students T-test being performed. ****p≤0.0001 (P-values not stated in the figure were not significantly different).

A phosphorylation deficient Y704F mutant was shown to prevent the nuclear accumulation in cells containing a constitutively active form of Hop (Karsten et al., 2006). Consistent with this observation, $STAT^{Y704F}$-GFP was unable to translocate into the nucleus upon ligand treatment, in contrast to $STAT^{WT}$-GFP (Figure 5.7). This further confirms that the phosphorylation state of Y^{704} is independent of Upd2/Dome endocytosis.

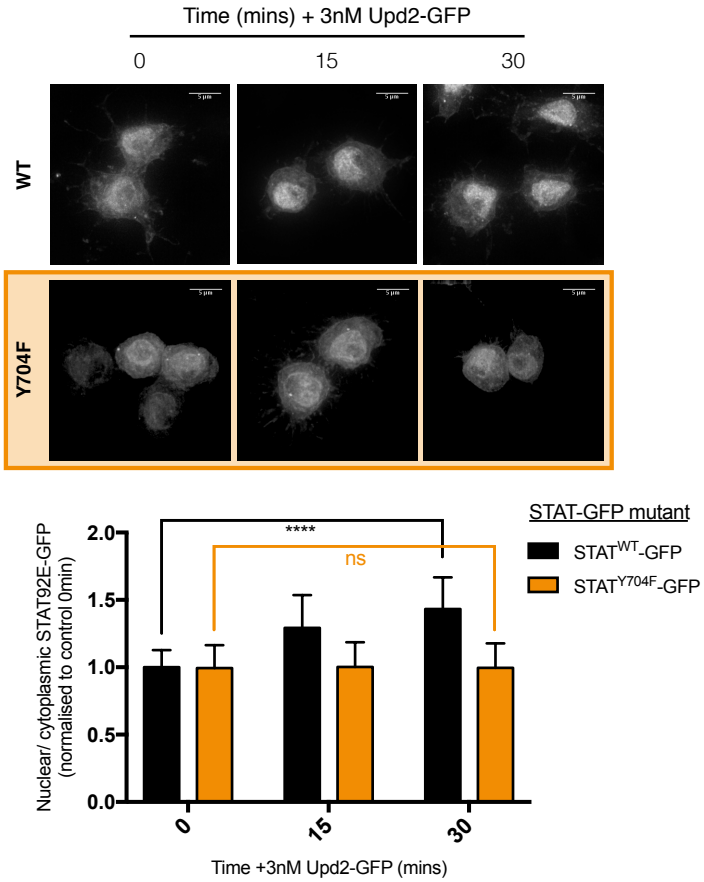


Figure 5.7: STAT-GFP nuclear translocation is dependent on phosphorylation of Y704. Representative images of cells transfected with either $STAT^{WT}$-GFP or $STAT^{Y704F}$-GFP, and treated with 3nM Upd2-GFP for 0, 15 or 30mins. Nuclear signal was divided by cytoplasmic signal, and normalised to 0mins control cells. Data is presented as mean ± SD for 3 independent experiments, where at least 20 cells were imaged per condition per experiment. Parametric, unpaired students T-test being performed. ****p≤0.0001 (P-values not stated in the figure were not significantly different).

Hence, STAT92E appears to be efficiently phosphorylated at Y^{704} and is able to enter the nucleus when AP2 and Hrs are knocked down, yet remains transcriptionally incompetent for expression of specific transcriptional targets. Therefore, other STAT92E post-translational modifications (PTMs) and/or interacting partners must be required for transcription of *10xSTATluciferase* and *SOCS36E*. These are likely to change as Dome is trafficked through distinct signalosomes to enable fine-tuned transcriptional outputs.

5.3 Mass-spectrometry to investigate STAT92E post-translational modifications and interacting partners during endocytosis

Post translational modifications (PTMs) and protein-protein interactions are key in the regulation of protein activity and signal transduction. In this study, I observed several phosphorylated species of STAT92E (Figure 5.3A) suggesting multiple phosphorylation sites. I used mass spectrometry (MS) with the aim of identifying further PTMs or interacting partners of STAT92E that are altered during endocytosis.

To carry out MS, I optimised immunoprecipitation to successfully isolate endogenous STAT92E from cell lysates along with an on-bead trypsin digest protocol. On-bead digestion eliminates the need to remove the isolated proteins from the beads prior to digestion, and does not require SDS-PAGE separation, therefore limiting the number of handling steps (Turriziani et al., 2014). Unfortunately, the STAT92E dN-17 antibody became discontinued and I was unable to continue investigating endogenous STAT92E. Therefore, I expressed STAT-GFP and utilised GFP-trap beads to isolate the GFP tagged protein. It is important to note that this means Upd2-GFP will also be immunoprecipitated by the GFP-trap beads, and therefore the investigation of interacting proteins will also include the Upd2-GFP interactome. Overexpression of STAT-GFP vastly increased the coverage of STAT92E, from ~35% when the endogenous protein was isolated, to ~70% after overexpression (Figure 5.8, red and green peptides respectively). This is likely due to the increased concentration of protein present and also potentially due improved accessibility of the enzyme to STAT92E, as it is the GFP tag that is interacting with the GFP-trap beads.

Immunoprecipitation of both endogenous STAT92E and STAT-GFP allowed for confirmation that the conserved tyrosine^{704} residue is phosphorylated upon addition of Upd2-GFP. To my knowledge this is one of the first demonstrations of phosphorylation at this site by MS, and is a proof-of-principle that phosphorylated peptides can be identified. When STAT-GFP was overexpressed and cells were treated with Upd2-GFP, mass spectrometry identified other PTMs. This is likely due to the increased abundance of phosphorylated STAT92E peptides in the sample compared to in the endogenous STAT92E sample. Three novel phosphorylation sites at threonine^{47}, serine^{227} and threonine^{702} were observed (Figure 5.8, Table 5.1), which have not been documented in *Drosophila* or mammalian STAT proteins before. MS spectra were

examined to ensure phosphorylation had been allocated to the correct amino acid within the peptide (representative spectra can be found in Appendix 2). For T702, it was difficult to determine whether the phosphorylation site was on T^{702} or Y^{704} in the tryptic peptide, although MaxQuant assigned it to T^{702}. The GluC peptide (discussed below) provided a spectrum that enabled confirmation of the phosphate group being attached to $T^{702.}$.

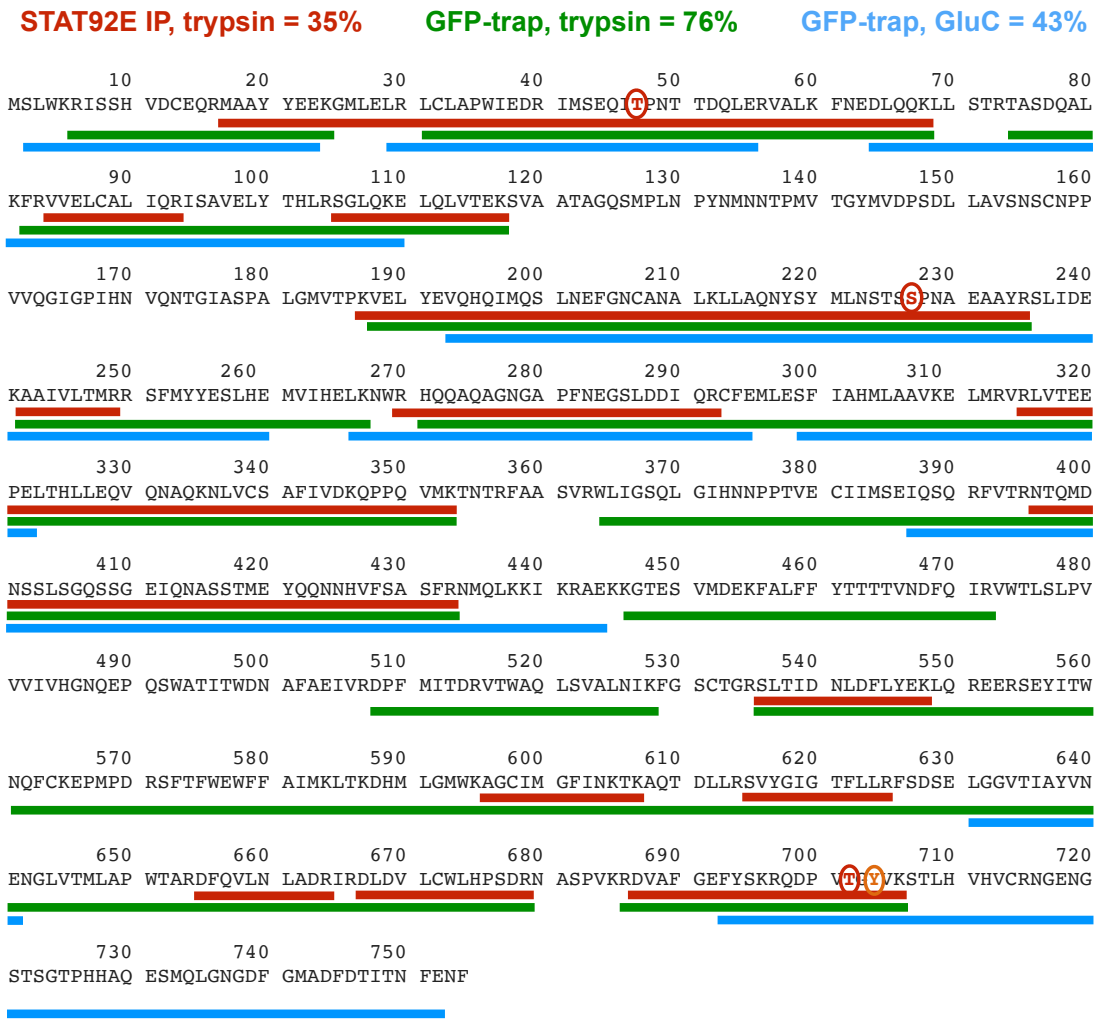


Figure 5.8: Coverage of endogenous STAT92E versus STAT-GFP, when digested with either trypsin of GluC. Maximal coverage of endogenous STAT after IP with anti-STAT92E and on-bead trypsin digest is highlighted in red. This coverage is doubled when cells are transfected with STAT-GFP (green) and an IP using GFP-trap beads is carried out. GluC digestion (blue) increases the coverage of the C-terminal. Tyrosine phosphorylation at 704 is circled in orange and novel phosphorylation sites identified whilst optimising the coverage of the STAT-GFP on-bead digest are highlighted in red.

Phosphorylation of serine residues in the transactivation domain of mammalian STAT1, STAT3 and STAT5 plays key roles in their transcriptional activity (Wen et al., 1995; Yamashita et al., 2001; Costa-Pereira, 2011). There are four serine residues in the transactivation domain of STAT92E, and hence I aimed to investigate whether these residues were phosphorylated. As trypsin digestion was not yielding coverage of residues 707-754 at the C-terminal of STAT92E, I utilised the enzyme GluC with the aim of achieving coverage of this region. GluC cleaves at the C-terminal of glutamic acid residues and therefore produces distinct peptides to those produced by trypsin. Digestion with GluC produced peptides covering 43% of STAT-GFP and peptides within the C-terminal region including residues 692-752 (Figure 5.8, blue peptides). However, no additional phosphorylation sites were identified within this region. Therefore, I continued to utilise trypsin digestion for further experiments due to the increased coverage of 76% and the identification of novel PTMs.

Enzyme	Peptide position		Sequence	Modification
	Start	End		
Trypsin	41	56	IMSEQI**T**PNTTDQLER	T7 (Phospho)
	213	235	LLAQ**N**YSYMLNSTS**S**P**N**AEAAYR	S15 (Phospho) N5,N17 (deamidated)
	697	706	RQDPVTG**Y**VK	Y8 (Phospho)
GluC	698	713	FYSKRQDPV**T**GYVKSTLHVHV**C**NGE	T10 (phospho) C22(Carbamidomethyl)

Table 5.1: STAT92E phosphorylated peptides identified by MS/MS.

To investigate if the phosphorylation of the newly identified sites (Figure 5.8, Table 5.1), or STAT-GFP interacting partners, were regulated during endocytosis, I prepared 3 biological repeats of 5 different conditions: Cntrl- (control dsRNA, mock treated), Cntrl+ (control dsRNA, Upd2-GFP treated), AP2+ (AP2 dsRNA, Upd2-GFP treated) and Hrs+ (Hrs dsRNA, Upd2-GFP treated) all transfected with *pAc-STAT-GFP*, and eGFP (control dsRNA, mock treated) transfected with *pAc-eGFP*. The eGFP sample was included as a control for STAT-GFP specific interacting partners. Cells were treated with control, AP2 or Hrs dsRNA 2 days prior to *pAc-STAT-GFP* or *pAc-eGFP* transfection. dsRNA treatment was confirmed to not alter the transfection efficiency of STAT-GFP (Figure 5.9A). After a 3 days transfection, cells were then treated with mock or 3nM Upd2-GFP for 75mins prior to cell lysis. All samples were examined for efficient

STAT-GFP transfection, immunoprecipitation (Figure 5.9B) and Hrs knockdown (Figure 5.9C). Figure 5.9B demonstrates that the GFP-trap beads efficiently immunoprecipitate STAT-GFP from all conditions (green asterisk at 112.5kDa), but also isolates Upd2-GFP (green asterisk at 68.3kDa) in samples treated with ligand.

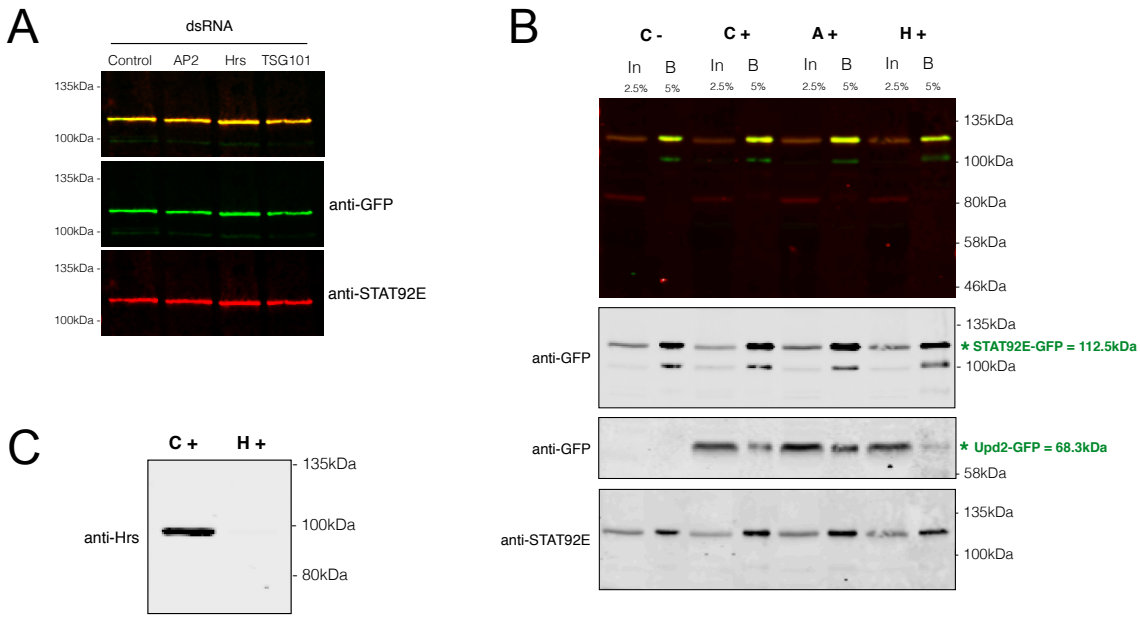


Figure 5.9: Immunoblots used as quality control for samples taken forward for MS experiments. A) dsRNA against control, AP2, Hrs and TSG101 does not alter STAT-GFP transfection. Cells were treated with dsRNA 2 days prior to STAT-GFP transfection. Cells were lysed after 3 days, ran on an SDS-PAGE gel and western blotted with indicated antibodies. B) Immunoblot for repeat 1 of final experiment, probed with indicated antibodies. Top panel represents full blot probed with both antibodies. C) Immunoblot showing successful Hrs knockdown in repeat 1 of final experiment.

Samples were prepared for MS, analysed with the Thermo Orbitrap Elite, and identified using MaxQuant, as discussed in Methods (Chapter 2.4.5). Unfortunately, no phosphorylated residues of STAT92E were identified in any of the STAT-GFP expressing samples. Examination of STAT92E coverage demonstrated that a total of 68% of the protein was identified across all samples, but that only 40% of the protein was identified in every sample (Figure 5.10). The peptide containing the conserved Tyrosine^{704} residue, which was regularly identified in previous MS experiments, was only identified in 1 sample.

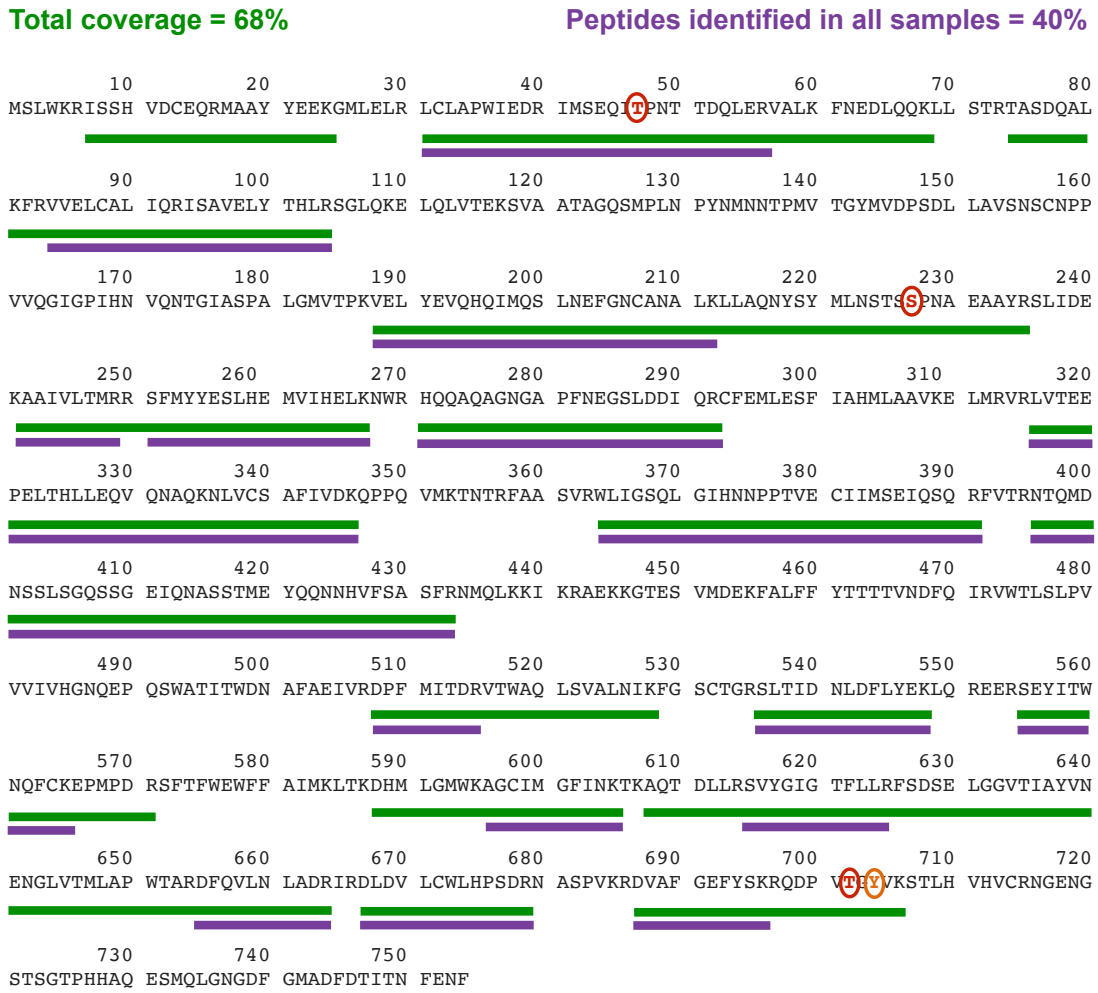


Figure 5.10: STAT-GFP coverage in final experiment where endocytic components were knocked down. Green peptides are those which appeared in at least one samples, whereas purple peptides were present in all samples. Phosphorylated residues highlighted were not identified in these samples.

Following further examination of the data it was evident that the intensity of STAT-GFP within these samples was significantly lower when compared to previous experiments (Figure 5.11). For example, in a previous experiment (experiment 1) STAT-GFP was 50x more abundant than in the final experiment (experiment 3), and this drastically increased the maximum coverage of STAT-GFP in a single sample, from 50% in experiment 3 to 76% in experiment 1. Whilst STAT-GFP it is still the most abundant protein detected in the samples of experiment 3, STAT-GFP is only 1.5x more abundant than the next protein, whereas in experiment 1 is it 11.5x more abundant. Although the concentration of cell lysate incubated with the GFP-trap beads was kept constant, S2R+ cells have variable levels of transfection efficiency, which may account for the difference in STAT-GFP levels. Within the MS run, the top 20 abundant peaks

are taken for further fragmentation. These precursor ions are fragmented into smaller product ions, which provides more detailed information about the peptide such as sequence composition. Therefore, if the concentration of STAT-GFP in the sample is reduced fewer peptides will fall within the top20 abundant precursor ions and therefore would not be selected for fragmentation. This would result in reduced coverage, and hence why in experiment 3 only 50% of STAT-GFP was identified.

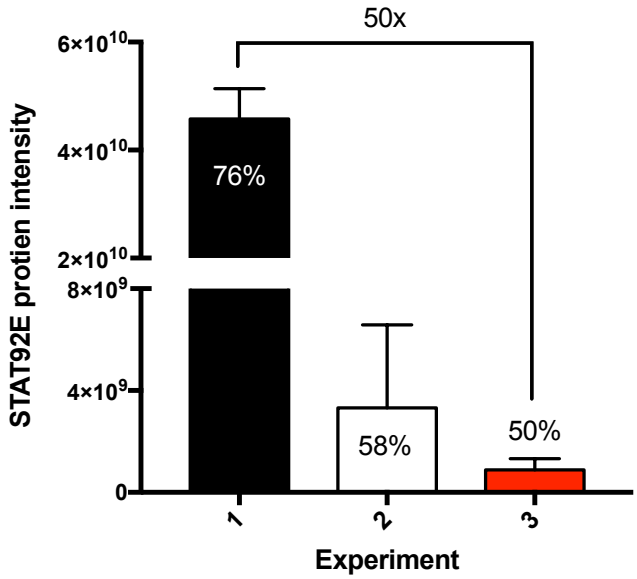


Figure 5.11: Intensity of STAT-GFP across different MS experiments alters protein coverage. Protein intensity for STAT92E over 3 different MS experiments. Percentage refers to the maximum protein coverage for a sample in that experiment. Experiment 3, highlighted in red, is the final experiment including AP2 and Hrs knockdown. Graphs represent mean ± SD of at least 2 samples from each experiment. STAT92E levels in experiment 1 are 50x more than in experiment 3**.**

5.3.1. Protein interactors

To investigate changes to the interacting partners of STAT-GFP during endocytosis, label free quantification (LFQ) intensities were produced to compare between samples. This is a function of MaxQuant that normalises samples on the assumption that the majority of proteins remain unchanged across varying samples and conditions. Therefore, protein intensity values are altered so that the majority of proteins have a 1:1 ratio across the 15 different samples. MaxQuant then quantifies the protein intensity across samples by comparing the intensity of common peptides and disregarding proteins which do not share common peptides across multiple samples (Cox et al, 2014). In a stringent IP, with only a few associated proteins, the

normalisation step may diminish real differences causing a loss of information, and therefore I also performed a LFQ without normalisation. Skipping the normalisation step prevents the MaxQuant software from stabilising large ratios, but still carries out the protein quantification. Due to the variable transfection efficiency of S2R+ cells, there are different raw intensities of STAT-GFP across the samples (Figure 5.12A). Normalisation by LFQ (Figure 5.12B) alters these differences dramatically, whereas LFQ with skipped normalisation maintains a similar trend to the raw intensities (Figure 5.12C), demonstrating the differences in these two methods.

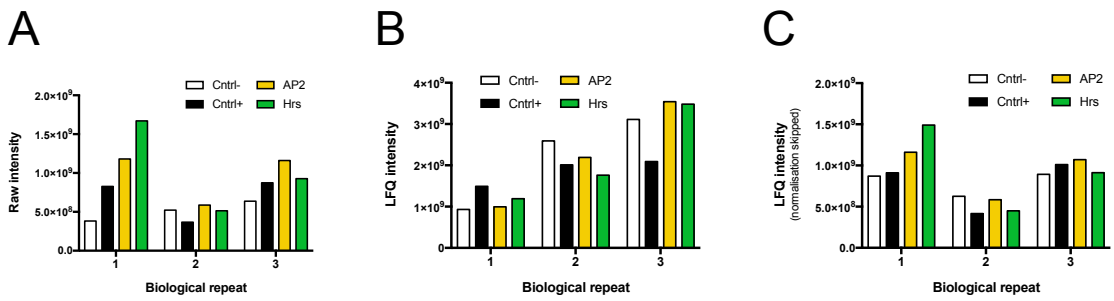


Figure 5.12: Intensity of STAT-GFP in different samples. A) Raw intensity values across the 3 biological repeats. B) LFQ intensity values with normalisation across repeats. C) LFQ intensity values when normalisation has been skipped**.**

Raw, LFQ and LFQ with skipped normalisation (LFQ_{SKIP}) protein intensities were loaded into Perseus, a software package produced by the Max Planck Institute of Biochemistry for the analysis of MS data. Proteins were filtered as described in Chapter 2.4.6, so that only proteins identified in all 3 repeats of at least 1 condition were retained. Intensity values were then log transformed and missing values, produced from transforming proteins with 0 intensity, were imputed using information from the normal distribution. This imputation is required for protein comparisons as valid values are required. I then visualised changes in protein intensity between conditions using volcano plots (Figures 5.13, 5.14 and 5.15). A two sample t-test was carried out on the average protein intensities from the conditions, and plotted on a scatter plot as fold change vs significance. I chose a cut-off of 1.5 $\log_2$ difference, which is equal to 2.83 fold change (demonstrated by the vertical dashed lines), as to categorise proteins whose abundance had altered. Using this method proteins expected to be enriched/depleted in specific samples were identified. For example, when comparing Cntrl- and eGFP, STAT-GFP was only present in the Cntrl- sample in which it had been transfected (A of Figures 5.13, 5.14 and 5.15). Moreover, Upd2-GFP was

enriched in treated control conditions, as was Dome and endogenous STAT92E (B of Figures 5.13, 5.14 and 5.15) due to the increased STAT dimerisation upon ligand treatment. AP2 and Hrs were depleted in cells where the corresponding mRNA had been targeted with dsRNA (C and D of Figures 5.13, 5.14 and 5.15). STAM, a ESCRT-0 component that dimerises with Hrs, was also reduced in samples where Hrs was knocked down.

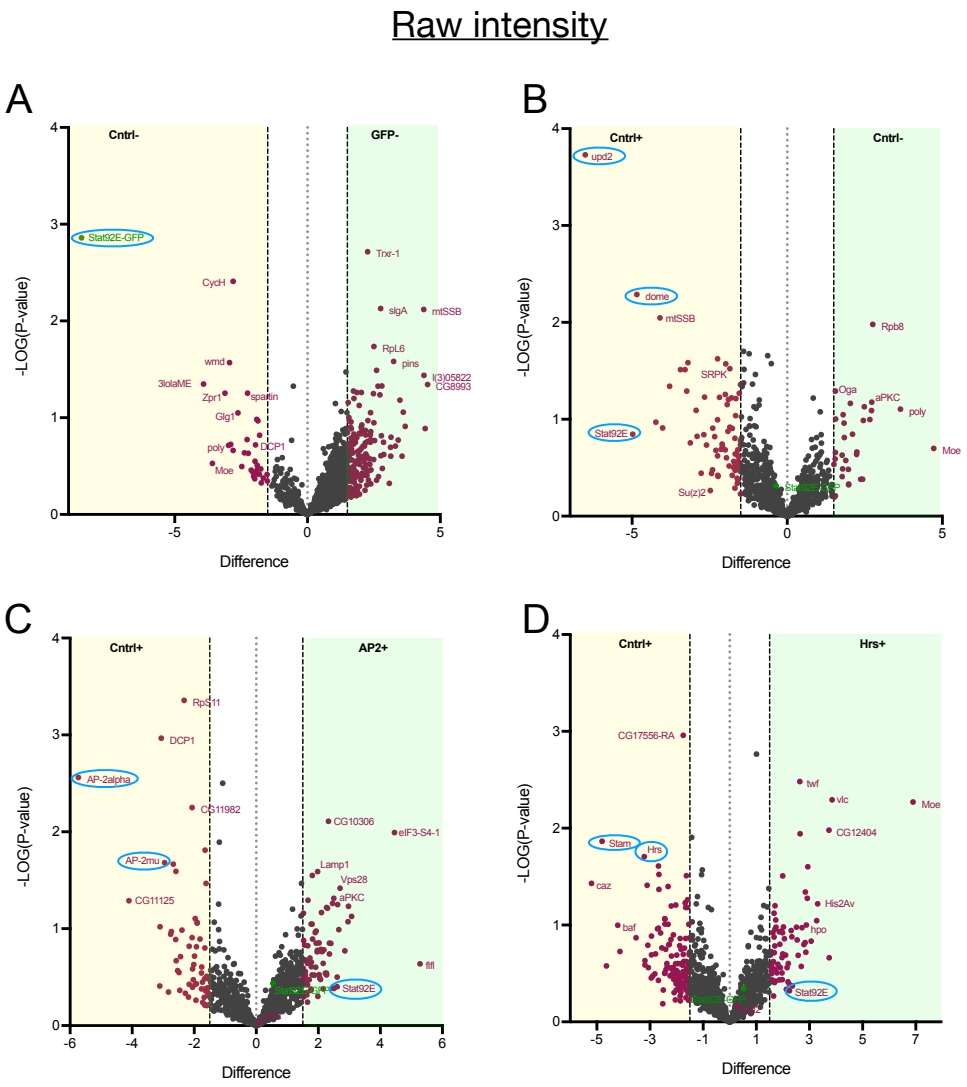


Figure 5.13: Volcano plots from raw intensity data. Volcano plot presenting comparison data from A) Cntrl- vs GFP B) Cntrl+ vs Cntrl- C) Cntrl+ vs AP2+ D) Cntrl+ vs Hrs+. Blue circles represent JAK/STAT pathway components or proteins expected to change.

As visualised by the volcano plots, there are fewer proteins on the plots for LFQ and LFQ_{SKIP} data sets. Whereas a raw intensity of 0 means no peptides were identified, for LFQ intensities 0 indicates that there was not enough information about that protein produce quantitative data. Therefore, proteins with little information are removed during the filtering process in Perseus and hence there are less data points on the LFQ plots. This is beneficial as it may reduce the detection of false positives identified for proteins with little information, but also may result in the loss of information. Therefore, I chose to use all three intensities to investigate differences and similarities between the methods.

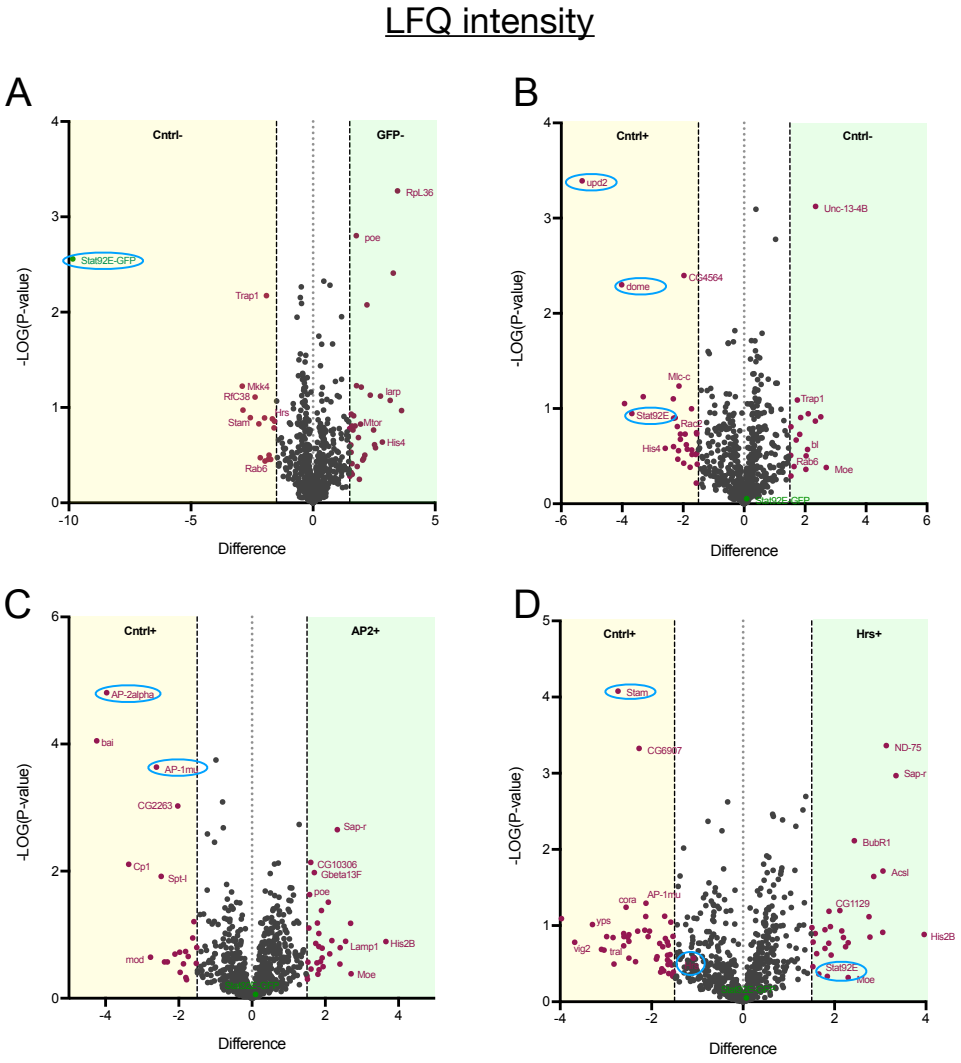


Figure 5.14: Volcano plots from LFQ intensity data. Volcano plot presenting comparison data from A) Cntrl- vs GFP B) Cntrl+ vs Cntrl- C) Cntrl+ vs AP2+ D) Cntrl+ vs Hrs+. Blue circles represent JAK/STAT pathway components or proteins expected to change.

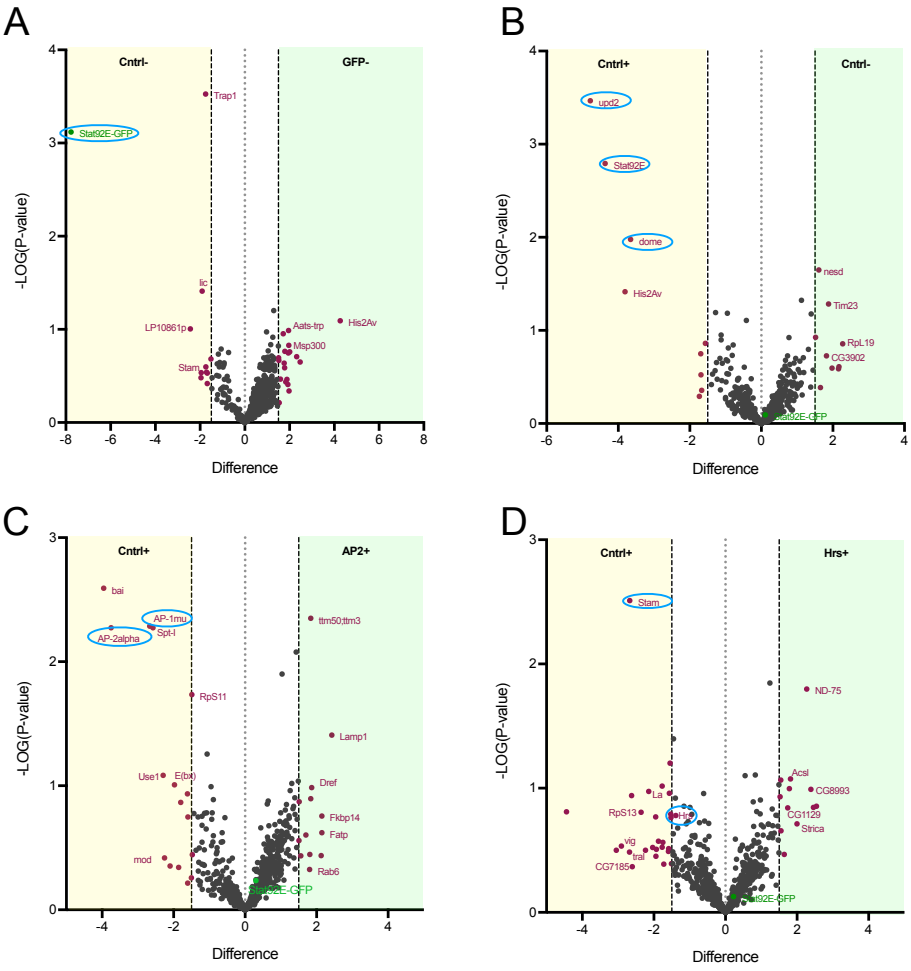


Figure 5.15: Volcano plots from LFQ$_{SKIP}$ intensity data. Volcano plot presenting comparison data from A) Cntrl- vs GFP B) Cntrl+ vs Cntrl- C) Cntrl+ vs AP2+ D) Cntrl+ vs Hrs+. Blue circles represent JAK/STAT pathway components or proteins expected to change.

The volcano plots only allow for the comparison of two conditions, yet to understand which STAT92E interactors are regulated by endocytosis I needed to consider all comparisons. Firstly, I aimed to determine the STAT-GFP interactors which are Upd2-GFP dependent. To do this I removed the proteins that are enriched in the eGFP-condition, when compared to Cntrl- and Cntrl+. I then selected proteins whose expression was altered in the comparison between the Cntrl- and Cntrl+ conditions (Figure 5.16). This revealed clear differences between comparisons made using raw, LFQ or LFQ$_{SKIP}$ intensities, with only Upd2, STAT92E and Dome being identified in all. I then took this list of proteins and examined their expression when AP2+ and Hrs+ were compared to Cntrl+ (Figure 5.16). This allows examination of Upd2-dependent STAT-GFP interactors whose association is altered when endocytic trafficking is perturbed.

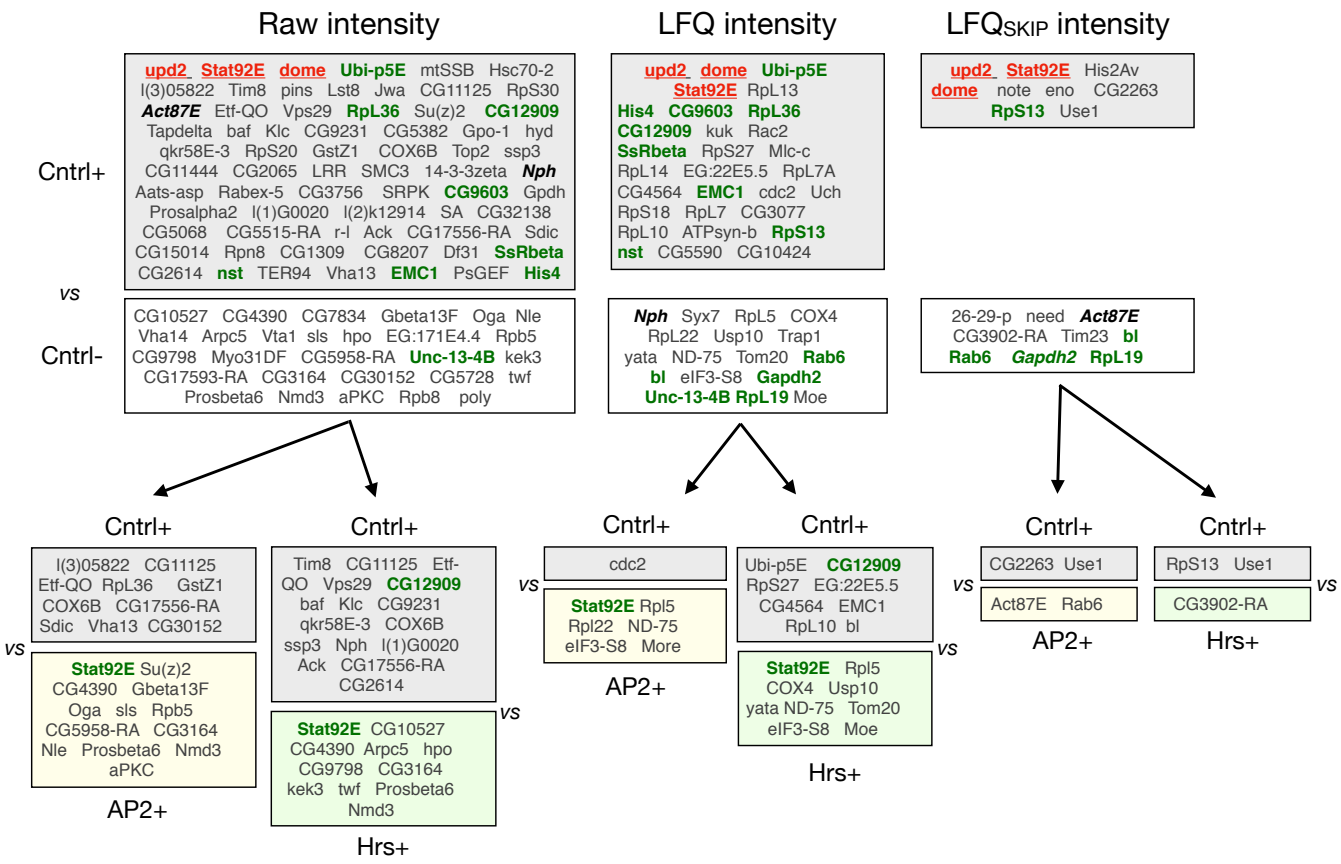


Figure 5.16: Comparison of significantly changed proteins between different analysis methods. Boxes in the top row represent proteins enriched in either Cntrl+ or Cntrl- conditions, and hence those that are Upd2-GFP dependent. These proteins were selected and their expression was examined when AP2 or Hrs,are knocked down (when compared to Cntrl+). These are represented in the bottom row of boxes. Grey boxes represents enrichment in Cntrl+, white boxes shows proteins enriched in Cntrl-, yellow is enriched in AP2+ and green enriched in Hrs+. Red proteins are those that appear in raw, LFQ and LFQ_{SKIP}, for that specific comparison. Green proteins appear in one other analysis. Bold italic proteins are those that appeared on other side of the comparison in a separate analysis (e.g. a protein enriched in cntrl+ when raw intensities are examined, yet enriched in cntrl- when LFQ values are compared).

As is evident in Figure 5.16, the expression of ~25 proteins were altered when raw intensity values were used, however, the majority of these are not consistent with outputs when comparisons were made with LFQ and LFQ_{SKIP} intensities. This may be due to the removal of proteins in the LFQ quantification due to lack of information, or proteins which fall just below the 1.5 log_2 difference. Only STAT92E and CG12909 changed in more than one comparison. CG12909 is an uncharacterised protein whose association with STAT92E has not been documented before, but according to this MS data may not associate when Hrs is knocked down. Interestingly, endogenous STAT92E appears to interact with STAT-GFP to a greater extent in conditions of endocytic knockdown. This suggests that when AP2 and Hrs are knocked down, STAT dimers are more prominent even if phosphorylation remains unchanged.

A comprehensive investigation of the proteins whose association changed upon AP2 and Hrs knockdown revealed that, when the raw intensities are examined, the majority were identified by 1 or 2 peptides. This means there is not enough information for LFQ values to be produced, and hence these would have been removed from the dataset and do not appear in the comparisons made with LFQ intensities. For example, atypical protein kinase C (aPKC) appears to be enriched when AP2 is knocked down when raw intensity values are considered, but not when LFQ values are (Figure 5.16). Study of the individual raw intensities demonstrates that in the Cntrl+ and Hrs+ samples aPKC levels are low or absent (Figure 5.17A). When this data is combined to look at differences between the conditions, Cntrl+ and Hrs+ therefore have a much lower expression level, accounting for the change between Cntrl+ and AP2+. However, when LFQ intensities are generated, the majority of the samples no longer contain information about the protein (Figure 5.17C), due to the lack of peptides identified. Therefore, as there are no conditions which contain values for all 3 biological repeats, this protein would have been filtered. This may account for the differences between the raw, LFQ and LFQ_{SKIP} intensities, but makes interpretation of the data and confidence in method choice more challenging.

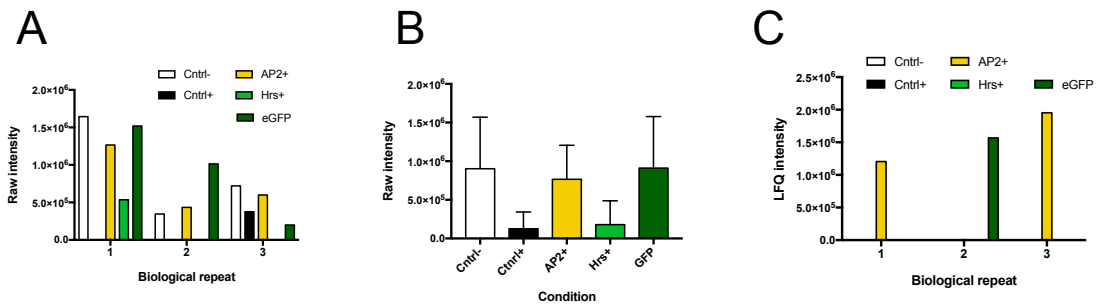


Figure 5.17: Intensity of aPKC across different. A) Raw intensity values for aPKC across all 15 samples. B) Raw intensity values for aPKC in each conditions. Data is mean +SD for the 3 biological replicates. C) LFQ intensity values of aPKC across the 15 samples**.**

In order to verify coimmunoprecipitation of STAT-GFP interactors and to investigate which intensity value (raw vs LFQ) is best to use, I investigated known STAT92E binding partners. There are 19 STAT92E binding partners documented on flybase, and 6 of these (Dsp1, His1, Jra, Su(var)205, Dome and Mor) were identified in the MS data of this study. Of these, there was only enough information about Su(var)205, Dome and Mor to create LFQ values. This suggests more optimisation may be required to enable detailed examination of interacting partners using LFQ values, such as more stringent washes to increase that abundance of specific interactors. However, one of the downfalls of on-bead digestion is that only low levels of detergent can be used in washes. High levels of detergent, for example triton x-100, cause polyethylene glycol (PEG) contamination. Triton x-100 is made up of 44Da PEG repeats, which are hydrophobic and suppress ionisation, whilst contaminating the MS spectra and preventing the identification of less abundant peptides (Scheerlinck et al., 2015). As mentioned above (Chapter 5.2), the abundance of STAT-GFP in the sample is significantly lower than in previous experiments. Therefore, a method to quantitively ensure high level of STAT-GFP prior to MS may be necessary for study of interacting partners. This would increase the chance of interactors being above the level of non-specific binders and hence identification via MS.

As no known STAT92E interactors were shown to be regulated by endocytosis and the variations between the different intensities (raw vs LFQ) I decided to focus on the STAT92E phosphorylation sites identified in this study instead of interacting partners. Downstream analysis of interacting proteins would also require a wide range of tools, such as dsRNA and antibodies against a number of proteins, which was not required for PTM studies.

5.3.2. Phosphoenrichment

The novel phosphorylation sites identified in previous MS experiments (Figure 5.8, Table 5.1) were not identified in the final experiment where endocytic components were knocked down. As this may be due to the reduced STAT-GFP protein intensity in comparison to previous MS experiments I aimed to increase the concentration of phosphorylated peptides within the sampled by carrying out a phosphoenrichment protocol. To further increase the probability of identifying phosphorylation sites repeats for each condition were pooled. For example, samples for repeat 1, 2 and 3 for AP2 were pooled to give a single AP2 sample. Therefore, there were only 4 samples; Cntrl-Cntrl+, AP2 and Hrs.

Columns containing titanium oxide (TiO_2) were used to enrich phosphorylated species. TiO_2 has a positively charged surface at low pH, allowing for selective binding of phosphorylated peptides (Dunn et al., 2010). Due to time constraints, the protocol used was not fully optimised. Thus, I maintained fractions from each stage of the protocol (flow through, wash and elute, described in Chapter 2.4.3) in order to analyse the efficiency of the TiO_2 enrichment. Following the MS run I calculated the total number of phosphorylated peptides in each sample, and determined the percentage of phosphorylated peptides that were present in each fraction. Figure 5.18A demonstrates that on average, across the 4 samples, ~66% of the phosphorylated peptides were in the eluted fraction after TiO_2 enrichment. Although 15% of the phosphorylated peptides did not bind to the TiO_2 (FT) and a further 19% were removed during wash steps, the majority of the phosphopeptides were retained in the correct fraction. There is, however, still a high proportion of unphosphorylated peptides found in the enriched fraction (Figure 5.18B). Therefore, the ratio of phosphorylated peptides vs unphosphorylated is only 1.5x more in the eluted fraction than in the total sample prior to enrichment (Figure 5.18C). Because of this, the number of phosphorylated peptides in the sample remains low (<10%). Key steps to improve this protocol for future use would involve increasing the wash stringency to without compromising on loss of phosphorylated peptides, in order to increase the ratio of phosphorylated to unphosphorylated peptides.

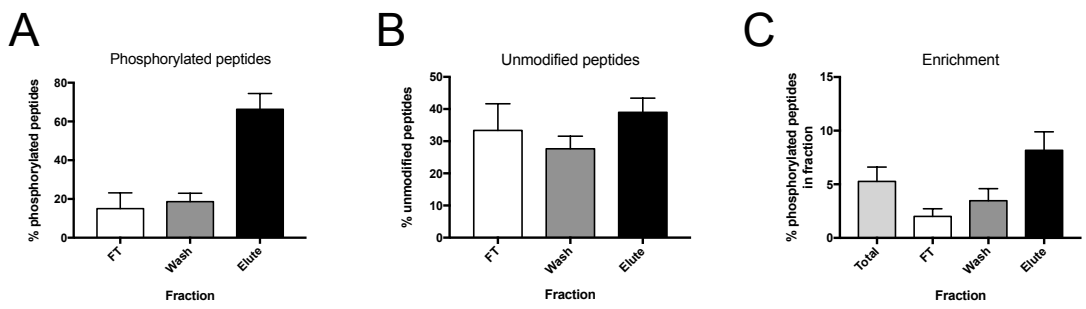


Figure 5.18: Analysis of fractions from phosphoenrichment with TiO_2. A) Number of phosphorylated peptides within each fraction as a percentage of total number of phosphorylated peptide in sample. B) Number of unphosphorylated peptides within each fraction as a percentage of unphosphorylated peptides in sample. C) Phosphoenrichment fractions. Number of phosphorylated peptides as a percentage of total number of peptides in each fraction. Graphs represent mean ± SD for the 4 samples (Cntrl-, Cntrl+, AP2 and Hrs). Total = starting material prior to phosphoenrichment, FT = flow through.

Unfortunately, none of the novel sites were identified following phosphoenrichment. However, the peptide containing the conserved tyrosine^{704} was identified in all samples treated with Upd2-GFP. This allowed for calculation of the intensity of phosphorylated peptides compared to unphosphorylated peptides (modification/base ratio) (Figure 5.19). This demonstrates that in all samples, Y^{704} is phosphorylated to the same amount, confirming the western blot data in Figure 5.2 and 5.3.

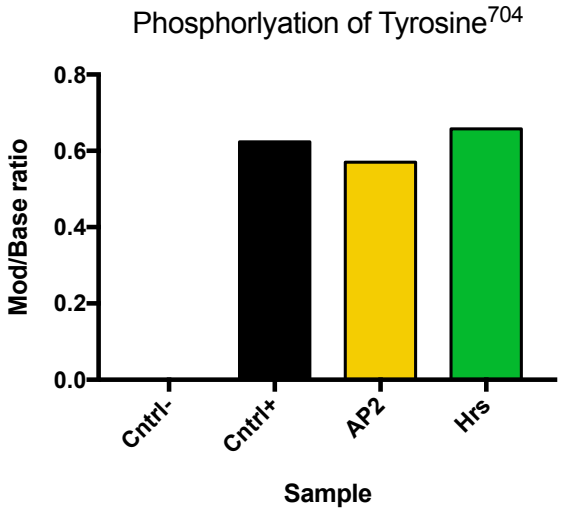


Figure 5.19: Phosphorylation of T^{702} is unchanged when AP2 and Hrs are knocked down. Phosphorylated/unphosphorylated peptide ratio for QDPVTGYVK peptide in samples from final experiment. Cntrl+ avg is the average of Mod/Base ratios from previous trypsin digested STAT-GFP MS experiments.

I used the on-bead digest protocol as it provided a high coverage of STAT-GFP and the identification of new phosphorylation sites, compared to experiments tested in-solution and in-gel. However, further optimisation of the in-gel digest may have provided greater information regarding STAT92E modifications, but removed any information regarding interacting proteins. This method allows for selection of proteins at the correct weight of STAT-GFP, by dissecting bands out of an SDS-PAGE gel. This therefore enriches STAT-GFP in the sample compared to other proteins, which are in turn depleted due to being the incorrect molecular weight. Hence, more STAT-GFP peptides are likely to fall within the top 20 abundant peaks and be selected for fragmentation, and may provide further information on STAT92E PTMs.

5.4 Functional characterisation of STAT92E modifications

Using proteomics I have identified novel STAT92E phosphorylation sites (Figure 5.8, Table 5.1). Although I could not determine if these newly identified modifications were Upd2-GFP dependent or endocytically regulated, I decided to investigate whether they are functionally relevant for JAK/STAT signalling.

5.4.1. Production and characterisation of CRISPR/Cas9 engineered cell population lacking functional STAT92E.

In order to understand the role of these novel phosphorylation sites on STAT92E function, I produced a cell line lacking endogenous STAT92E. This would prevent signalling changes caused by mutation of individual residues from being masked by the function of wild-type STAT92E. Using dsRNA to knockdown STAT92E persistently for 5 weeks, I observed that S2R+ cells could survive and grow at normal rates with minimal JAK/STAT activity (data not shown). I used the same method as described in Chapter 4.3.2 to create a STAT92E knock out CRISPR cell-line. I targeted exon 2 of the *STAT92E* gene, as this is common to all STAT92E isoforms (Figure 5.20A). Selection with puromycin enabled creation of cell lines (crSTAT) with low levels of STAT92E (Figure 5.20B). S2R+ cells cannot be single cell selected and hence the cell-line created is a population of genetically modified cells and not a clonal cell-line. A proportion of cells are likely to have successfully repaired the STAT92E gene and therefore contain functional STAT92E. This may explain why a low level of *10XSTATluciferase* activation is still observed upon addition of Upd2-GFP (Figure

5.20C). These cell lines maintain low levels of STAT92E through multiple passages, therefore suggesting the cells still expressing STAT92E do not have a growth advantage (Figure 5.21). Further experiments use the crSTAT2 line due to lower levels of STAT92E protein. An advantage of not producing a clonal CRISPR cell line is that signalling/phenotypic outcomes are unlikely to be due to artefacts of a specific mutation, as there are multiple Cas9 endonuclease generated mutations within the population of crSTAT cells.

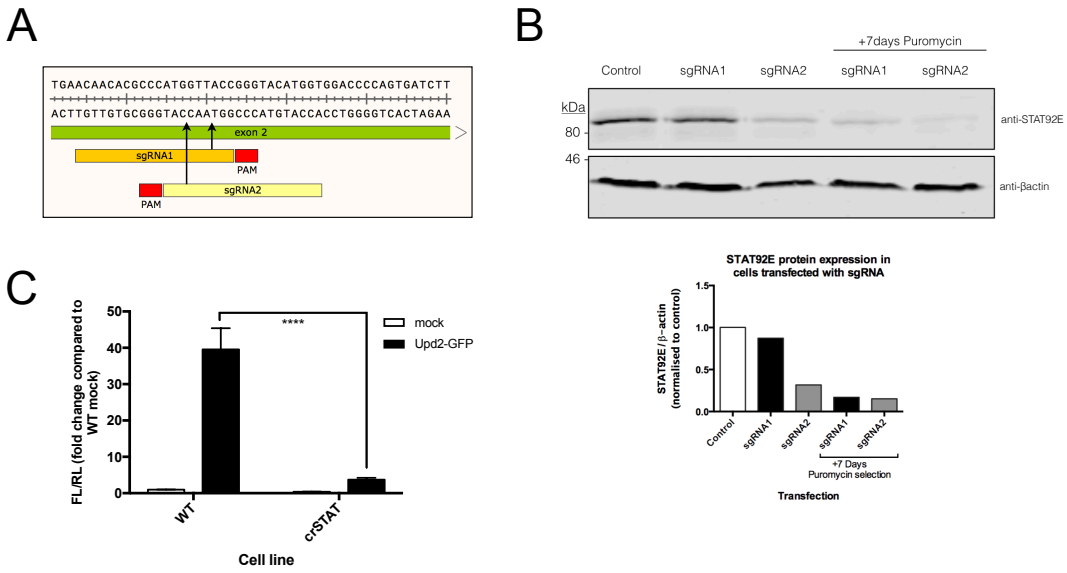


Figure 5.20: Generation of STAT92E CRISPR knock out cell lines. A) The region that sgRNA1 and sgRNA2 target the Cas9 to within exon 2 of STAT92E. B) Immunoblot (top panel) and quantification (lower panel) demonstrating levels of STAT92E protein in cells transfected with pAc-sgRNA-Cas9 targeting STAT92E for 3 days, with or without puromycin selection. C) Luciferase reporter assay demonstrating reduced activation of crSTAT cell lines in response to Upd2-GFP. Graph represents mean ± SEM of triplicates from 3 independent experiments. Parametric, unpaired students T-test was performed with ****p≤0.0001.

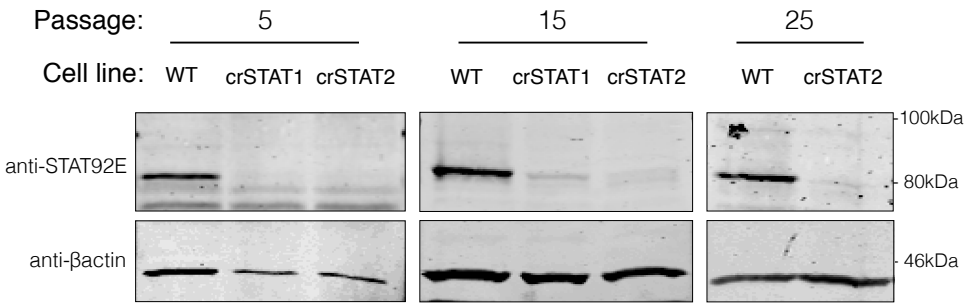


Figure 5.21: crSTAT cell lines have reduced STAT92E protein levels throughout passage Immunoblot of WT vs crSTAT cells after multiple passages, probed with indicated antibodies.

To confirm that the Cas9 endonuclease had cleaved and caused mutations within the genomic *STAT92E* locus, I performed a T7 endonuclease assay (as described in Chapter 4.3.2). Amplification of crSTAT2 genomic DNA results in a smeared band, in contrast to WT genomic DNA (Figure 5.22A). Denaturing and reannealing of DNA causes a band indicative of a heteroduplex for crSTAT2 DNA, and upon addition of the T7 enzyme multiple cleaved products are visible (Figure 5.22B). Therefore, I can be confident that the Cas9 has induced mutations within the *STAT92E* gene.

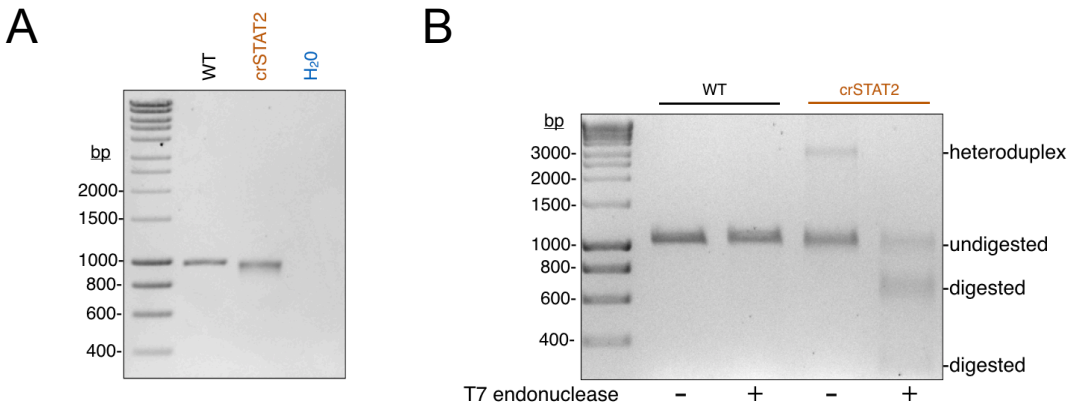


Figure 5.22: T7-endonuclease assay demonstrates Cas9 induced mutation in the stat92e gene. A) Genomic DNA was extracted from WT and crSTAT2 cell lines, and 989bp region around the sgRNA target site was amplified. B) Addition of T7 endonuclease to PCR product causes multiple bands for crSTAT2 cell line but not WT cells.

As a proof-of-principle that the crSTAT cells would be appropriate for investigating changes to signalling, I decided to verify a role for SUMOylation on STAT92E activity. Attachment of a Small Ubiquitin-like Modifier (SUMO) to a protein is a covalent modification whose conjugation, similar to ubiquitination, requires a E1-activating enzyme, E2-conjugating enzyme and E3-ligase. In mammalian STATs, SUMOylation negatively regulates STAT1 activity by reducing its DNA-binding activity (Grönholm et al., 2012) and STAT5a/b by preventing phosphorylation at the conserved tyrosine residue (Van Nguyen, 2012). In *Drosophila* there is a single SUMO gene, *smt3*, which encodes a protein of ~10kDa (Huang et al., 1998) and its conjugation to a non-conserved lysine (187) residue in STAT92E appears to inhibit signalling (Gronholm et al., 2012). Furthermore, the Smt3 activating enzyme, Uba2, has previously been identified as a negative regulator of the *Drosophila* JAK/STAT pathway (Vidal et al., 2010). Therefore, I investigated how mutation of the K187 residue to an arginine within STAT-GFP alters *10xSTATluciferase* expression in the crSTAT cell line. Consistent with the work of Groholm et al, 2012, in WT cells, this non-SUMOylatable mutant was able to increase luciferase expression greater than that of $STAT^{WT}$-GFP (Figure 5.23), supporting a negative role of Smt3 conjugation on STAT92E activity. This data confirms that the crSTAT cells, with their remaining low level of STAT92E activity, are an appropriate tool for studying changes to STAT92E signalling.

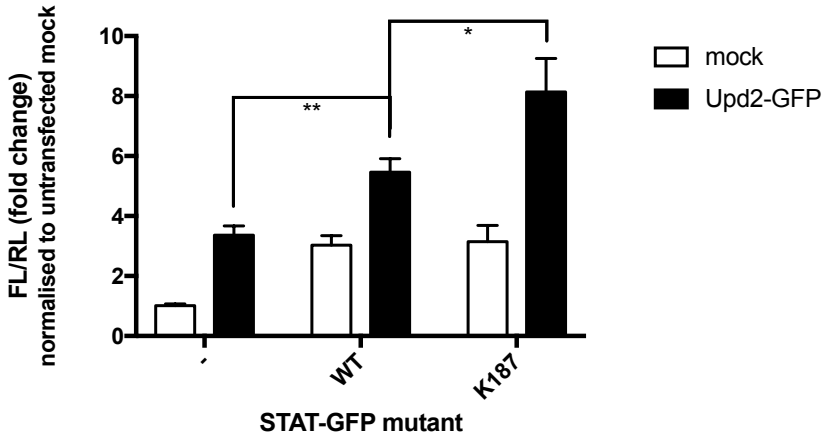


Figure 5.23: Mutation of lysine 187 increases STAT92E signalling. crSTAT cells were transfected with pAc-Ren, 10xSTATluciferase and pAc5.1 (-), $STAT^{WT}$-GFP or $STAT^{K187R}$-GFP. Cells were stimulation with 0.75nM Upd2-GFP for 30mins, then incubated in fresh media for 18hrs. Data is mean ± SEM from 3 independent experiments and normalised to cells transfected with pAc5.1 (-) and treated with 0nM Upd2-GFP.

5.4.2. STAT92E phospho-mutants affect signalling

It was unclear if the phosphorylated resides that I had identified by mass spectrometry were endocytically regulated. Since they were newly identified sites, I decided to explore their role(s) in STAT92E activity. Threonine 47 is within the N-terminal of STAT92E (Figure 5.24). In mammalian STAT proteins this region has roles in STAT4 phosphorylation (Chang et al., 2003), the formation of unphosphorylated STAT3 dimers (Vogt et al., 2011), cooperative binding of STAT dimers at low-affinity DNA binding sites (Xu et al., 1996) and the interaction with transcriptional co-activators (Zhang et al., 1996). The N-terminal is also important in *Drosophila* STAT92E activity, as a naturally occurring N-terminally truncated protein acts as a negative regulator of STAT92E activity (Henriksen et al., 2002). Serine S227 is within the coiled-coil domain of STAT92E, a region essential in STAT3 for receptor binding (Zhang et al., 2000). Finally, threonine 702 is only two residues upstream of the conserved tyrosine residue, whose phosphorylation is key for STAT92E activity. Therefore, all novel sites had the potential to alter STAT92E signalling.

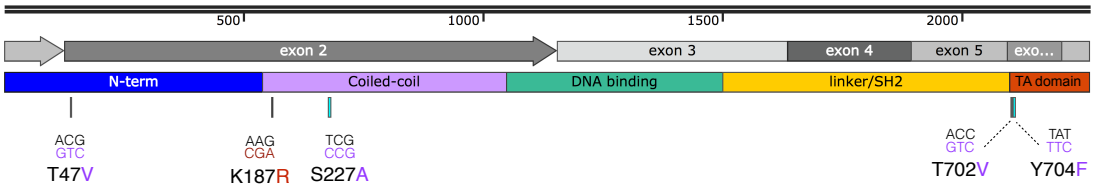


Figure 5.24: STAT92E domain structure and location of novel phosphorylation sites. Schematic showing STAT92E domains, exons, and changes to nucleotide sequence and amino acids following mutagenesis. TA domain = transactivation domain.

Using the crSTAT cell line characterised in Chapter 5.5.1, I studied changes in STAT92E signalling when the phosphorylation sites were mutated to non-phosphorylatable residues (Figure 5.24). For this, I utilised the *10xSTATluciferase* reporter assay. Interestingly, STAT-GFP rescued signalling in crSTAT cells, but acted as a dominant-negative when added to WT cells (Figure 5.25). The level of unstimulated signalling in both untreated WT and crSTAT cells was increased upon expression of STAT-GFP, suggesting a rise in pathway activity prior to the addition of Upd2-GFP. This is likely due to the overexpression of STAT-GFP allowing for heightened pathway activity. The dominant-negative effect of STAT-GFP in WT cells could be an assay specific phenomenon, due to the limited amount of

10xSTATluciferase in the system, as it is an exogenous reporter. High levels of STAT-GFP may also result in oligomeric structures at STAT DNA-binding sites that prevents initiation of transcription. STAT-GFP was shown to bind to consensus DNA using electro-mobility shift assay (Karsten et al., 2006), therefore the high levels of STAT-GFP may form oligomeric structures at STAT DNA-binding sites that prevents initiation of transcription.

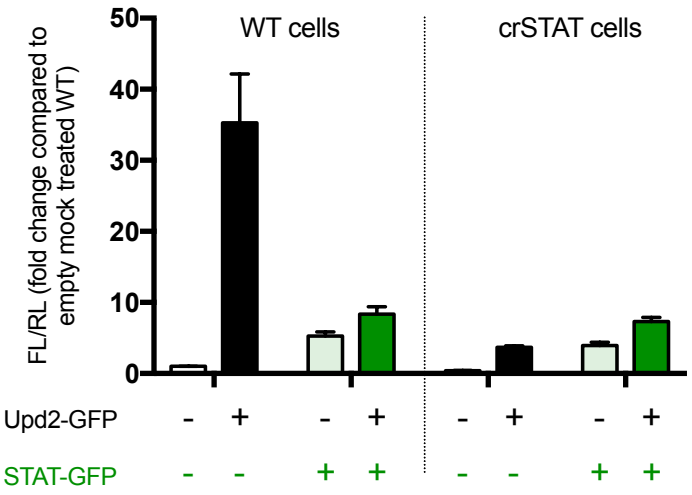


Figure 5.25: STAT-GFP rescues ligand-independent crSTAT signalling, but acts as a dominant-negative in WT cells. WT or crSTAT cells were transfected with pAc-Ren, 10xSTATluciferase and pAc5.1 (-) or STAT-GFP. Cells were stimulation with 3nM Upd2-GFP for 30mins, then incubated in fresh media for 18hrs. Data is mean ± SEM from 3 independent experiments and normalised to WT cells transfected with pAc5.1 (-) and treated with 0nM Upd2-GFP.

To understand greater the dominant-negative role of STAT-GFP on WT cells in the luciferase assay, I titrated the ligand in WT and crSTAT cells with or without STAT-GFP expression. In WT cells lacking STAT-GFP there is a ligand concentration dependence in the expression of luciferase (Figure 5.26A). Although WT cells expressing STAT-GFP do show a concentration dependence, this is severely reduced and appears to plateau at 0.75nM (Figure 5.26A). In contrast, crSTAT cells transfected with STAT-GFP produced a similar ligand dependent response as in untransfected cells (Figure 2.26B).

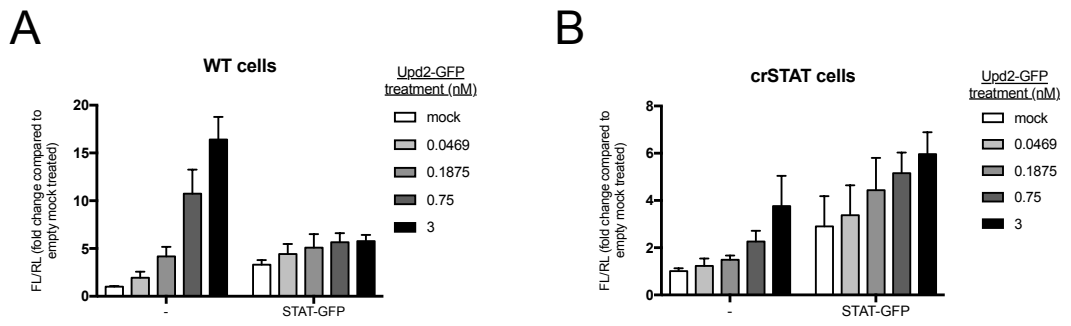


Figure 5.26: 10xSTATluciferase expression is not Upd2-GFP concentration-dependent in WT cells transfected with STAT-GFP. A) WT or b) crSTAT cells were transfected with pAc-Ren, 10xSTATluciferase and pAc5.1 (-) or STAT-GFP. Cells were stimulation with varying concentrations of Upd2-GFP for 30mins, then incubated in fresh media for 18hrs. Data is mean ± SEM from 2 independent experiments and normalised to cells transfected with pAc5.1 (-) and treated with 0nM Upd2-GFP.

As STAT-GFP rescued transcriptional activity in crSTAT cells, I investigated the effect of non-phosphorylatable mutants in this system. Transfection of eGFP was used as a negative control to ensure results were STAT92E specific. I treated cells with 0 and 0.75nM Upd2-GFP, as there was a marked difference between control and STAT-GFP transfected crSTAT cells (Figure 5.26B), and hence signalling changes were more likely to be identified. As expected, mutation of the conserved Y^{704} residue prevented luciferase expression in the presence and absence of Upd2-GFP (Figure 5.27). Mutation of either T27 or S227 did not affect STAT signalling, but the T702V mutant negatively regulated signalling in a ligand independent manner, similar to Y704F. Interestingly, this site is conserved in mammalian STAT1 and has been documented as being phosphorylated in tissue samples from the heart, lung and gastrointestinal-tract (Curated information from PhosphoSitePlus (Hornbeck et al., 2015), CST Curation Sets: 10453, 10501, 8327, 8196, 5513). This site is also a phosphomimetic in STAT5. However, no studies have investigated the role of this phosphorylation site on JAK/STAT signalling.

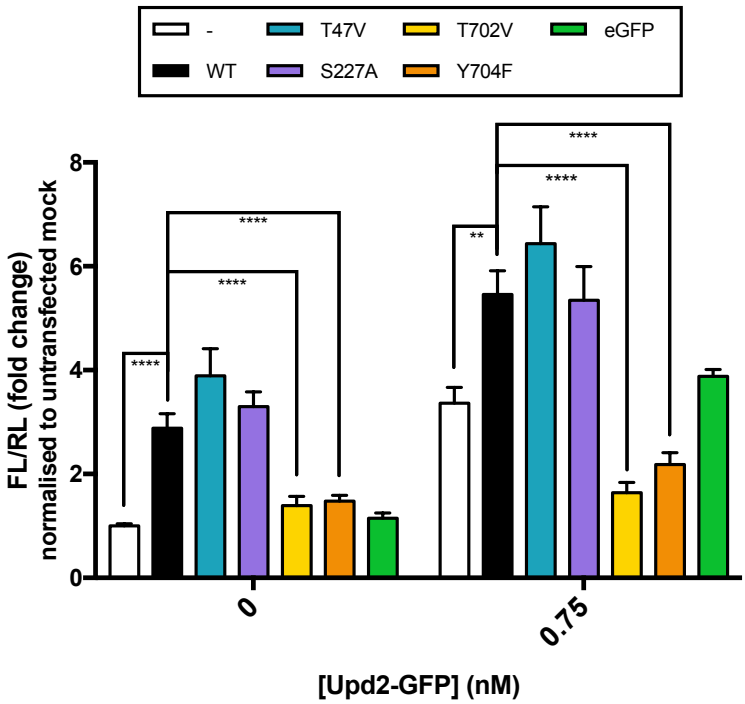


Figure 5.27: Non-phosphorylatable T702V and Y704F mutants inhibit STAT-GFP induced signalling. crSTAT cells were transfected with pAc-Ren (-), 10xSTATluciferase and pAc5.1, eGFP, or STAT-GFP mutants. Cells were stimulation with 0 or 0.75nM Upd2-GFP for 30mins, then incubated in fresh media for 18hrs. Data is mean ± SEM from 3 independent experiments each performed in triplicate and normalised to cells transfected with pAc5.1 (-) and treated with 0nM Upd2-GFP. Parametric, unpaired students T-test was performed with **p≤0.01, ****p≤0.0001.

When examined by western blot, all constructs are expressed at similar levels (Figure 5.28) and therefore I can be confident that signalling changes are due to the point mutation and not mutant-specific transfection levels. The T702V mutant appears to migrate slower through the gel compared to WT and the other point-mutants, indicative of a post-translational modification or a conformational change in STAT92E.

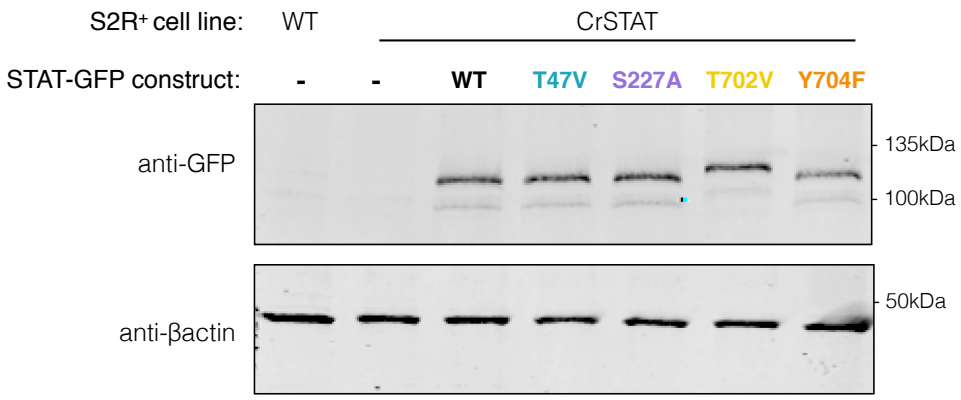


Figure 5.28: STAT92E^{T702V} migrates slower on an SDS-PAGE gel compared to STAT92E^{WT} and other phosphomutants. Cells were transfected with STAT-GFP mutants for 2days prior to cell lysis. Lysates were run on 10% SDS-PAGE gel and membrane was immunoblotted with anti-GFP and anti-βactin.

I then asked whether the T702V mutation inhibited expression of the endogenous JAK/STAT target, *SOCS36E*. crSTAT cells expressed similar levels of *SOCS36E* to WT when unstimulated, yet lacked the ligand-induced response seen in WT cells (Figure 5.29). Expression of STAT^{WT}-GFP increased the unstimulated level of *SOCS36E* in WT cells, but had little effect on Upd2-GFP induced expression. In crSTAT cells expression of STAT^{WT}-GFP marginally increased *SOCS36E* expression, with STAT^{T702V}-GFP and STAT^{Y704F}-GFP inhibiting this marginally. However, unlike the luciferase assay which exclusively measures transfected cells, this assay is very dependent on the expression levels of STAT-GFP as it is measures global mRNA levels across all cells. Therefore, these subtle changes between untransfected and transfected crSTAT cells may be background noise, with transfection levels being too low to influence the global expression of *SOCS36E* in the cell population. In order to overcome these issues and therefore understand changes to *SOCS36E* expression, I would first need to select transfected cells via FACS or use single cell transcriptomics.

However, due to time restrictions and lack of protocols for cell sorting of S2R+ cells I was unable to continue with this line of enquiry.

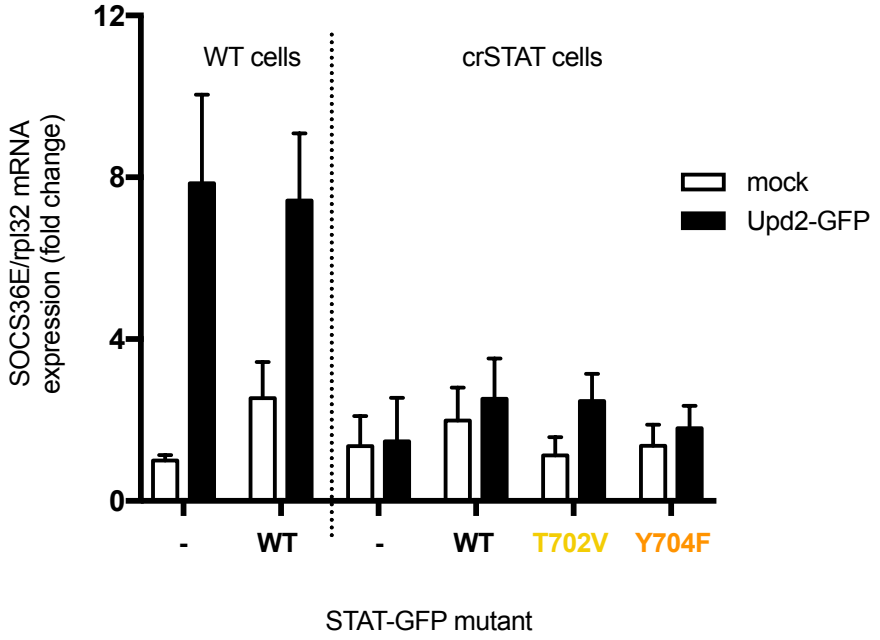


Figure 5.29: SOCS36E expression in WT and crSTAT cells. WT or crSTAT cells were transfected for 2 days prior to incubation with 3nM Upd2-GFP for 2.5hrs and subsequent RNA extraction. SOCS36E mRNA levels were normalised to that of reference gene Rpl32, and presented as fold change compared to mock-treated WT cells. Results are expressed as means of duplicates + SEM for 2 independent experiments.

5.4.3. Non-phosphorylatable T702V mutant prevents STAT nuclear accumulation, dimerisation and T704 phosphorylation

Since mutation of threonine at position 702 negatively affects JAK/STAT signalling in the luciferase assay, I aimed to investigate the mechanism by which this occurs. Firstly, I studied nuclear accumulation of $STAT^{T702V}$-GFP compared to $STAT^{WT}$-GFP in crSTAT cells. Similar to mutation of the conserved tyrosine residue, Y704F (Figure 5.7), the T702V mutant does not accumulate in the nucleus upon ligand addition (Figure 5.30). This result also suggests that $STAT92E^{T702V}$-GFP would not be capable of inducing the expression of any targets, including *SOCS36E*, in response to ligand.

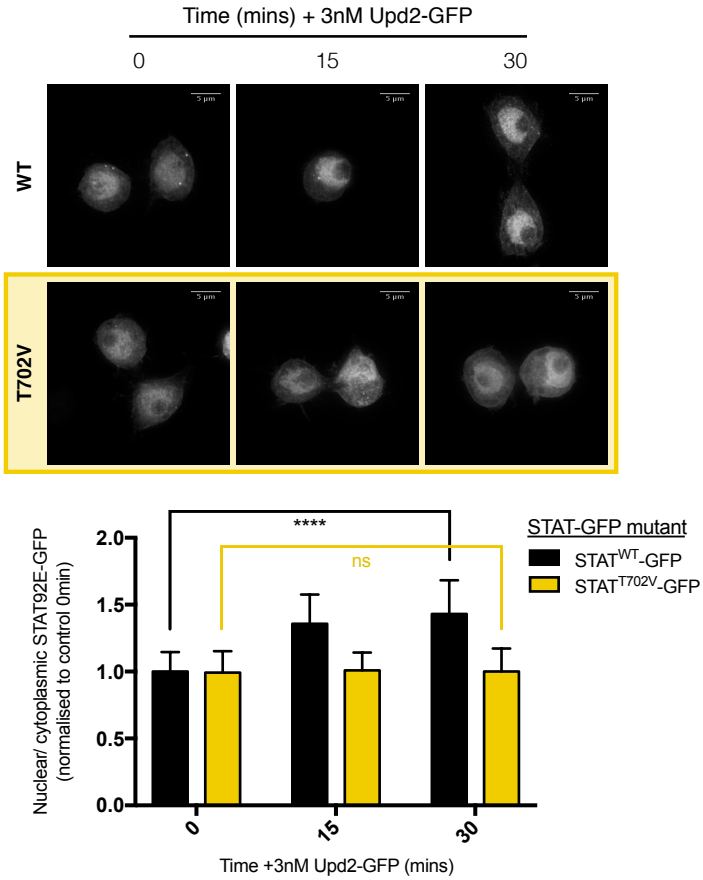


Figure 5.30: T702V mutation prevents STAT-GFP prevents nuclear translocation in response to ligand. Representative images of crSTAT cells transfected with either STAT^{WT}-GFP or STAT^{T702V}-GFP, and treated with 3nM Upd2-GFP for 0, 15 or 30mins. Nuclear signal was divided by cytoplasmic signal, and normalised to 0mins control cells. Data is presented as mean ± SD for 3 independent experiments, where at least 30 cells were imaged per condition per experiment. Parametric, unpaired students T-test being performed. ****p≤0.0001 (P-values not stated in the figure were not significantly different).

Due to the negative effect of the T702V mutation on signalling and its inability to accumulate in the nucleus, I examined whether STAT^{T702V} can still form a dimer. I investigated if expression of STAT-GFP phosphomutants could co-immunoprecipiate endogenous STAT92E in WT cells. Incubation of lysates from WT cells expressing STAT92E^{WT}-GFP with GFP-trap beads showed co-immunoprecipiataion of endogenous STAT92E (Figure 5.31A). Interestingly, endogenous STAT92E co-immunoprecipitates with STAT92E^{WT}-GFP in the absence of ligand treatment, but a bandshift is only seen upon Upd2-GFP, indicative of interaction with phosphorylated endogenous STAT92E. Studies in mammalian STATs have demonstrated that unphosphorylated STATs can form antiparallel dimers, but form parallel dimers upon tyrosine phosphorylation (Mao et al., 2005; Neculai et al., 2005; Wenta et al., 2008).

Hence unphosphorylated STAT-GFP may co-immunoprecipitate STAT92E in unstimulated conditions due to the formation of an antiparallel dimer.

In contrast to $STAT^{WT}$-GFP, co-immunoprecipitation of $STAT^{T702V}$-GFP is unable to pull down endogenous STAT92E (Figure 5.31B). Interestingly, $STAT^{Y704F}$-GFP appears to co-immunoprecipitate unphosphorylated STAT, further suggesting the formation of parallel dimers before tyrosine phosphorylation. Lysate from untransfected cells was used to enable identification of the endogenous STAT92E band, and eGFP was used as a control to ensure that the $STAT92E^{WT}$-GFP interaction was not influenced by the protein tag. This experiment was carried out in WT cells, due to requirement of endogenous STAT92E. Because of the dominant-negative effect of STAT-GFP on signalling in WT cells it would be interesting to investigate dimerisation via a further mechanism that used the crSTAT cells or did not require STAT-GFP overexpression. One method would be to utilise Native-PAGE gels, however initial efforts were unsuccessful and would require time-consuming optimisation.

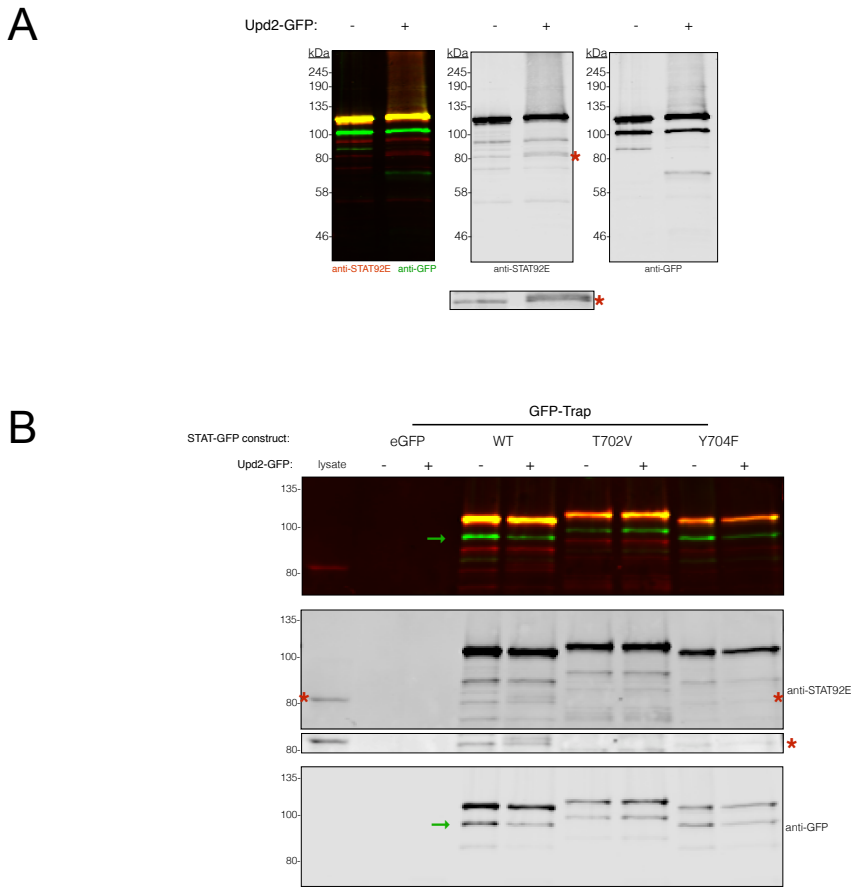


Figure 5.31: T702V mutant prevents STAT-GFP from co-immunoprecipitating endogenous STAT92. A) Example immunoblot demonstrating co-immunoprecipitation of endogenous STAT92E when $STAT^{WT}$-GFP is expressed in WT cells (red asterisk). Expanded panel highlights endogenous STAT92E and its phosphorylation upon addition of 3nM Upd2-GFP. B) Immunoblot of $STAT^{WT}$-GFP, $STAT^{T702V}$-GFP and $STAT^{Y704F}$-GFP immunoprecipitated from WT cell lysate. Coimmunoprecipitation of endogenous STAT92E is highlighted by red asterisk and in 3rd panel. Green arrow highlights a second anti-GFP band.

Interestingly, a band under STAT-GFP is routinely seen with the GFP antibody (Figure 5.31B, Green arrow). There is an N-terminally truncated version of STAT92E that is produced through an alternative promotor found between exon 1 and 2. However this is not present in the STAT-GFP construct, and so the extra band remains uncharacterised.

Due to the close proximity of T^{702} to the Y^{704} residue, I aimed to determine whether the T702V mutant inhibits signalling by preventing phosphorylation of the conserved tyrosine, or alters signalling via a separate, novel, mechanism. Unlike endogenous STAT92E, a band shift upon tyrosine phosphorylation is not visible with STAT-GFP. This is likely due to the increased size of the protein. Therefore, I required a different method to investigate phosphorylation. I tested gradient gels (Figure 5.32A), with the

aim of separating the higher molecular weight bands, however a band-shift after phosphorylation was still not visible. Gradient gels also prevented visualisation of the phosphorylation dependent band-shift of endogenous STAT92E. I also carried out immunoprecipitation of STAT-GFP followed by blotting with phospho-tyrosine antibodies (Figure 5.32B), yet anti-pY staining was present in the Y704F mutant and eGFP transfected cells. Therefore, I decide to use MS to investigate whether the Y^{704} residue was phosphorylated in the $STAT^{T702V}$-GFP when expressed in crSTAT cells.

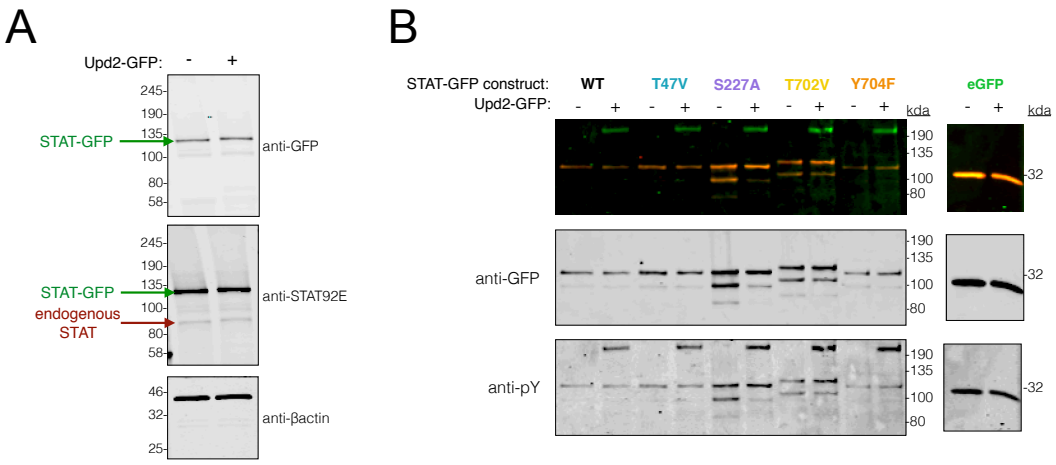


Figure 5.32: Phosphorylation of STAT-GFP cannot be visualised by blotting. A) S2R+ cells transfected with STAT-GFP for 2 days were lysed and run on a 4-15% Mini-PROTEAN® TGX™ precast gradient gel. B) S2R+ cells transfected with STAT-GFP mutants for 2 days were lysed and STAT-GFP was immunoprecipitated using GFP-trap beads. Immunoprecipitated proteins were separated in a 8% polyacrylamide gel and transferred onto nitrocellulose membrane. Membrane was immunoblotted for phosphorylated tyrosines using anti pY from cell signalling.

T^{702} and Y^{704} residues are within in the same tryptic peptide and, via MS, it was possible to identify the phosphorylation state specifically of the mutant T702V containing peptide. This was advantageous as, due to the low level of endogenous WT STAT92E remaining in crSTAT cells, a low intensity of WT peptide was identified in cells transfected with $STAT^{T702V}$-GFP. In the presence of Upd2-GFP, ~66% of the Y^{704} containing peptides were phosphorylated in $STAT^{WT}$-GFP (Figure 5.33). Although Y^{704} was also phosphorylated in some mutated peptides from cells transfected with $STAT^{T702V}$-GFP, the percentage of phosphorylated peptides decreased by ~10fold compared to WT. The difference in the proportion between unphosphorylated and phosphorylated peptides between these samples is >25 fold. Therefore, when threonine 702 is mutated to a valine, the Y^{704} residue cannot be efficiently phosphorylated, suggesting that T702 is required for phosphorylation of Y^{704}.

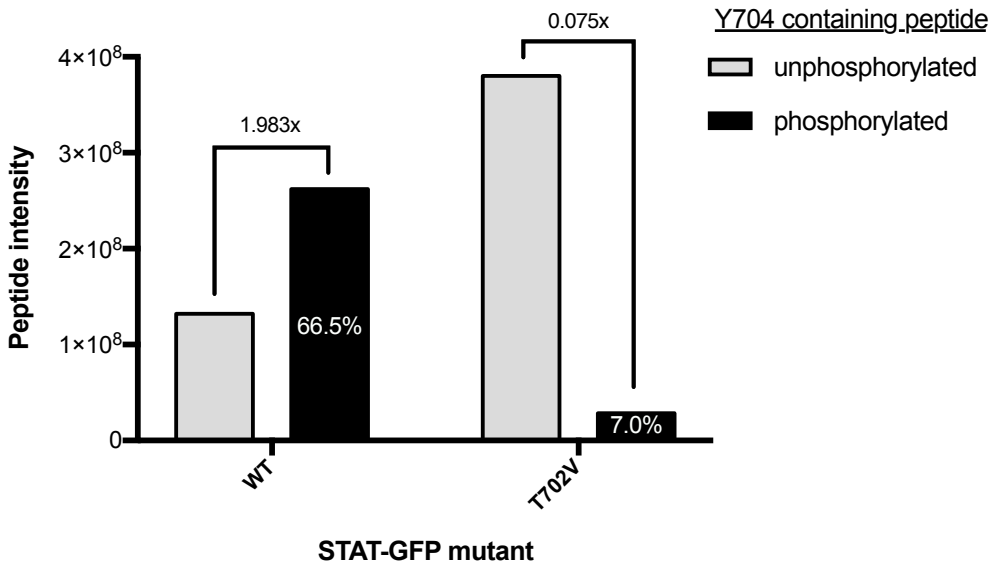


Figure 5.33. T702V reduces tyrosine phosphorylation. S2R+ cells were transfected with $STAT^{WT}$-GFP or $STAT^{T702V}$-GFP for 2 days prior to treament with 3nM Upd2-GFP for 75mins. Cells were lysed and incubated with GFP-trap beads. Samples then underwent reduction, alkylation, trypsin digestion and desalting, before injection into the Orbitrap. Data was anaylsed with MaxQuant and phosphorylation sites verified with plabel. Graph presents peptide intensity for unphosphorylated vs phosphorylated peptides in cells transfected with $STAT^{WT}$-GFP or $STAT^{T702V}$-GFP. Percentages are the intensity of phosphorlyated peptides compared to total, and the values above bars is the ratio of phosphorylated peptides in comparison to unphosphorylated.

5.4.4. T^{702} phosphomimetics rescue JAK/STAT signalling

To understand if the phosphorylation of T^{702} is key for STAT92E signalling, I asked whether mutation of the threonine to a phosphomimetic residue could rescue the signalling defect. Therefore, I replaced T^{702} with a glutamate and aspartate (Figure 5.34A), amino acids routinely used to mimic the charge change upon addition of the phosphate group. These point mutations rescued the inhibitory effects of the T702V mutant on *10xSTATluciferase* expression (Figure 5.34B), suggesting that the phosphorylation of the T^{702} is required for STAT92E signalling. Although the T702D mutant rescues the signalling defect to a greater extent than the T702E mutant, the levels of transfection were consistently low (Figure 5.35). This suggests that only a small amount of T702D is required for the rescue of signalling, and that a low level of expression may be closer to physiological levels.

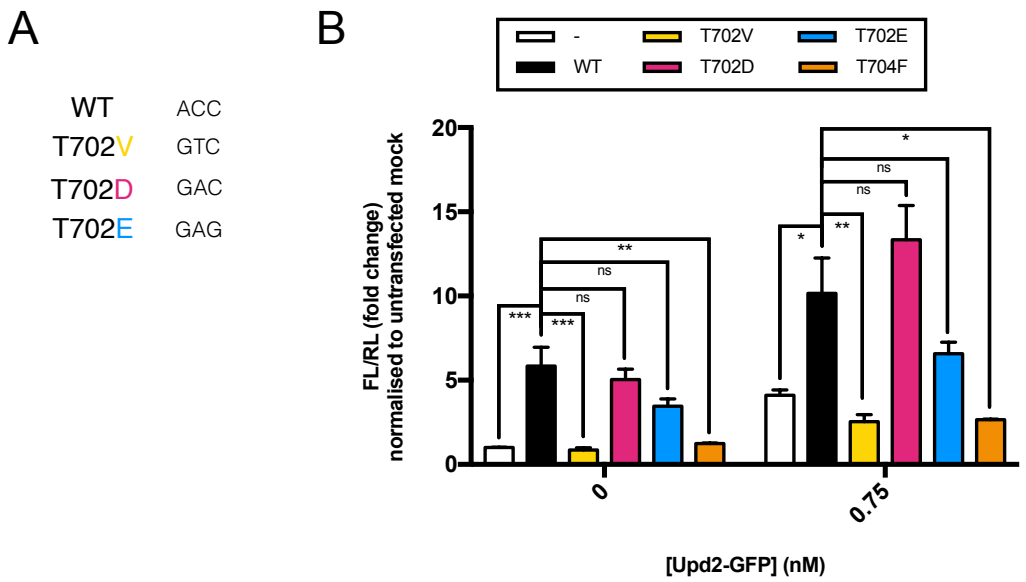


Figure 5.34: Phosphomimetics rescue inhibitory effects of T702V on STAT-GFP signalling. A) Nucleotide sequence of STAT-GFP T^{702} mutants. B) crSTAT cells were transfected with pAc-Ren, 10xSTATluciferase and pAc5.1 (-) or STAT-GFP mutants. Cells were stimulated with 0 or 0.75 Upd2-GFP for 30mins, then incubated in fresh media for 18hrs. Data is mean ± SEM from 3 independent experiments and normalised to cells transfected with pAc5.1 and treated with 0nM Upd2-GFP. Parametric, unpaired students T-test was performed, with *p≤0.05, **p≤0.01, ***p≤0.001.

In comparison to the T702V mutant, phosphomimetic mutants migrate at a similar speed on an SDS-PAGE gel to that of $STAT^{WT}$-GFP (Figure 5.35). This result indicates that their conformation may be similar to $STAT^{WT}$-GFP, and suggests that the negative charge of the phosphate or aspartate/glutamate may be required for a signalling-competent structure. Due to time limitations, I have not investigated the nuclear import or the Y^{704} phosphorylation state of the T^{702} phosphomimetic mutants. However, since they activate signalling, which requires STAT phosphorylation, dimerisation and nuclear import, it is reasonable to assume that these factors have also been rescued when compared to the T702V mutant.

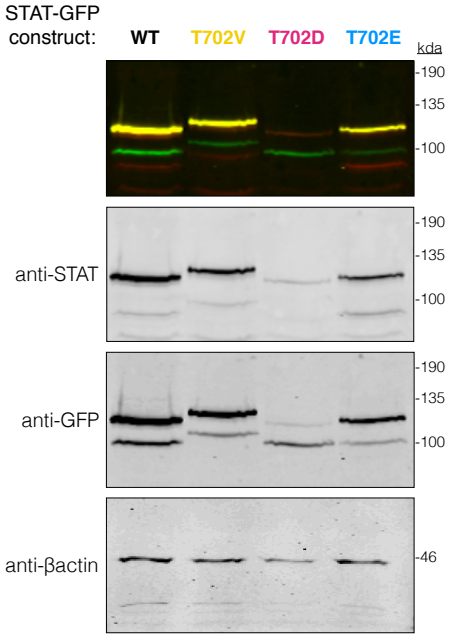


Figure 5.35: STAT T[702] phosphomimetics rescue mobility change on SDS-PAGE gel. Example immunoblot of S2R+ transfected with STAT-GFP mutants for 2days. Cells lysates were run on 10% polyacrylamide gel and membrane was immunoblotted with anti-STAT92E, anti-GFP and anti-βactin.

5.5 Summary

In this chapter I aimed to elucidate the mechanism by which endocytic trafficking regulates the transcriptional output of the JAK/STAT pathway. I have established that STAT92E is the component of the signalling pathway that is regulated by endocytosis. STAT92E is efficiently phosphorylated at the conserved Y^{704} residue when AP2 and Hrs are knocked down, yet is not transcriptionally competent for *10xSTATluciferase* and *SOCS36E* expression. This suggests that other factors, either post-translational modification or interacting partners, are required for specific STAT92E transcriptional responses. Therefore, I performed MS experiments with the aim of identifying novel PTMs and/or interacting partners key transcription of specific JAK/STAT targets.

Unfortunately, I was unable to identify endocytic-specific modifications or interactors of STAT92E (for reasons discussed in Chapter 5.3). However, I identified novel phosphorylation sites, whose phosphorylation state had the potential to regulate STAT92E activity. Creation of a CRISPR/Cas9 STAT92E cell-line, with minimal STAT92E activity, enabled investigation of the role of these phosphorylation sites on signalling. Mutation of a threonine at residue 702, two residues upstream of the

conserved Y704, to a valine prevented the rescue of *10xSTATluciferase* expression. Further study of the T702V mutant revealed that it ran slower on an SDS-PAGE gel, indicative of a conformational change, and was incapable of nuclear accumulation upon ligand stimulation. Due to the threonine's close proximity to the Y704 residue, I asked whether the T702V mutant inhibited signalling by preventing Y704 phosphorylation or by a novel mechanism. MS demonstrated that in crSTAT cells transfected with $STAT^{T702V}$-GFP, the phosphorylation of Y704 was reduced by ~10 times compared to cells expressing $STAT^{WT}$-GFP. This suggested that mutation of T702 was a requirement for Y704 phosphorylation. The signalling phenotype of T702V, and mobility shift during SDS-PAGE, was rescued upon the expression of T^{702} phosphomimetics.

To conclude, phosphorylation of a threonine two residues upstream from the conserved tyrosine residue appears to be required for efficient tyrosine phosphorylation, nuclear accumulation and signalling. This residue is conserved in human STAT1, and its phosphorylation has been identified by MS experiments. Therefore, future studies would investigate the role of this novel phosphorylation site on STAT1 tyrosine phosphorylation and transcription factor activity.

Chapter 6. Discussion

6.1 Summary of findings

The JAK/STAT pathway is used repeatedly during multiple developmental stages and in adult tissue, with its activation leading to a broad range of cellular outcomes. Therefore, tight regulation is required to provide the appropriate output in a specific context. Although multiple regulators of the JAK/STAT pathway have been characterised, the role of receptor internalisation and endosomal compartmentalisation remains less well-characterised. Endocytosis has traditionally been considered a mechanism to reduce signalling from cell-surface receptors, however it is increasingly becoming understood to play key roles in defining and generating signals. Therefore, this study focused on investigating endocytic regulation of JAK/STAT signalling in the low complexity *Drosophila* system, with the aim of increasing our understanding of how a signalling pathway with a broad range of potential outcomes, can produce defined context-specific responses.

Within this PhD project I have identified a [D/E]xxxL[L/I] motif within the cytoplasmic tail of the receptor, Dome, which is required for receptor internalisation and downstream expression of a *10xSTATluciferase* reporter. Moreover, I demonstrated that endocytosis differentially regulates the expression of the endogenous JAK/STAT target, *SOCS36E,* and the exogenous *10xSTATluciferase* reporter. In contrast, *lama* and *chinmo* expression is unaltered when endocytosis is perturbed. This data suggests that activated Dome is competent for signalling at the cell-surface, but for the expression of a specific subsets of transcriptional targets, internalisation and trafficking to specific endosomes is required. In line with this, STAT92E is efficiently phosphorylated and can enter the nucleus when endocytosis is inhibited, suggesting STAT92E activation can occur at the cell surface. Therefore, further STAT92E interacting partners and/or PTMs must to required to fine-tune its transcriptional competency during endocytic trafficking. Using Mass Spectrometry I identified a novel phosphorylation site, threonine 702 whose phosphorylation appears to be essential for STAT92E activity. I demonstrated that mutation of threonine 702 to a residue that cannot be phosphorylated, prevented phosphorylation at the conserved tyrosine 704, STAT-GFP nuclear accumulation and *10xSTATluciferase* expression. This threonine residue is conserved in human STAT1, and is a phosphomimetic in STAT5, therefore indicating a conserved role in STAT activity.

Thus, two main findings have come from this study. Firstly, internalisation and endocytic trafficking to specific endosomes regulates a subset of *Drosophila* JAK/STAT transcriptional targets through a mechanism that is independent of Y^{704} phosphorylation. Secondly, phosphorylation of a threonine two residues upstream of Y^{704} is essential in the activity of STAT92E.

6.2 The role of endocytosis in JAK/STAT signalling

Initially, I investigated the role of receptor internalisation in *Drosophila* JAK/STAT signalling as, to-date, there is contradictory evidence suggesting both positive and negative regulation (Devergne et al., 2007; Vidal et al., 2010). The receptor has been previously been shown to undergo CME, biochemically (Vogt and Smythe, unpublished) and via immunofluorescence (Devergne et al., 2007; Vidal et al., 2010). The ligand/Dome complex is then targeted for lysosomal degradation and trafficked through Rab5 and Rab7 positive endosomes (Devergne et al., 2007; Vidal et al., 2010; Stec et al., 2013), but is not observed in Rab11 recycling endosomes (Vidal et al., 2010). Firstly, I verified that localisation of Upd2-GFP in S2R+ cells using newly available antibody for Hrs and Rab7 (Figure 4.4, 4.5 and 4.6). Quantification of this data demonstrated that Upd2-GFP first colocalises with a Rab5, then Hrs and finally Rab7 (Figure 4.7). This confirmed that the Upd2/Dome complex traffics through to the late endosome via Rab5 and Hrs positive endosomes in S2R+ cells.

6.2.1. Trafficking differentially regulates subsets of JAK/STAT targets

To understand how trafficking through this endocytic pathway regulates JAK/STAT signalling, I knocked down key components of the trafficking pathway and investigated the expression of two well-characterised *Drosophila* JAK/STAT targets. This revealed that the Upd2-GFP induced expression of the *10xSTATluciferase* reporter and *SOCS36E* is dependent on AP2 (Figure 3.2 and 3.4). Knockdown of AP2, the major cell-surface adaptor protein, prevents AP2-dependent CME and hence traps Dome at the cell surface. Therefore receptor internalisation, and/or recruitment to a CCP via AP2, is required for the expression of both downstream targets. This was confirmed following identification of the Dome internalisation motif (Chapter 4.3). Mutation of this motif prevents Dome uptake (Figure 4.21) and also inhibits *10xSTATluciferase*

expression (Figure 4.28), validating that receptor endocytosis is important for expression of this target.

Lama and *chimno*, two *Drosophila* JAK/STAT targets first identified *in vivo* (Flaherty et al., 2009), are expressed in an Upd2-GFP dependent manner in S2R+ cells (Figure 3.20). In contrast to the expression of *10xSTATluciferase* and *SOCS36E*, the expression of *lama* and *chimno* is not dependent on endocytosis (Figure 3.21). suggesting a signalling-competent protein complex is formed at the cell-surface. This result was reiterated during a comprehensive study of the JAK/STAT pathway when endocytosis was perturbed (Chapter 5.1 and 5.2). Here, STAT92E was still phosphorylated in response to Upd2-GFP when endocytosis was inhibited, therefore suggesting that STAT92E activation is independent of receptor endocytosis. This creates a model whereby the receptor can activate JAK/STAT signalling and induce transcriptional outputs when at the cell surface, yet signalling is not competent for the transcription of specific targets without internalisation.

Knockdown of further downstream endocytic components revealed striking differences between the expression of *10xSTATluciferase* and *SOCS36E* during trafficking. Hrs and TSG101 are members of the ESCRT-0 and ESCRT-1, respectively, and are involved in the formation of ILVs at MVBs prior to lysosomal degradation. Knockdown of Hrs and TSG101 does not alter the expression of the *10xSTATluciferase* reporter (Figure 3.2), suggesting that only internalisation but not further trafficking is required for its expression. In contrast, knockdown of Hrs but not TSG101 inhibits *SOCS36E* expression (Figure 3.4). This indicates that for *SOCS36E* expression the receptor needs to be internalised and trafficked to a Hrs positive endosome but not further association with the ESCRT-1 complex. Therefore, endocytosis appears to regulate the transcription of specific JAK/STAT targets through distinct mechanisms.

In relation to published data concerning the endocytic regulation of the *Drosophila* JAK/STAT pathway, my data may provide insight into the discrepancies between these studies. Vidal et al., 2010, suggested that endocytosis negatively regulates JAK/STAT signalling, whereas Devergne et al., 2007, demonstrated positive regulation of JAK/STAT signalling. The two studies used different biological systems to study these effects, Vidal et al., 2010, used the *Drosophila* Kc_{167} cell line, whilst Devergne et al., 2007, utilised the *Drosophila* egg chamber. Therefore, differences in results may be as a result of the varying cellular environments, suggesting that endocytosis is crucial in the regulation of signalling in a context-dependent manner. Discrepancies may also

be due to the complexity of the regulation occurring, as my study has demonstrated that specific subsets of targets are regulated by endocytosis. Vidal et al., 2010, demonstrated negative endocytic regulation utilising a distinct luciferase reporter (*6x2xDrafLuc*), which is based on the *Draf* gene and contains 12 STAT92E binding sites (Müller et al., 2005). As this reporter differs from the *10xSTATluciferase* used in my study, this may also demonstrate the complexity of endocytic regulation, i.e. specific targets are not altered by endocytosis, whereas some are negatively regulated and others are positively regulated.

Although the transcriptome analysis carried out in the project did not reveal other targets that undergo endocytic regulation (Chapter 3.2), it also did not identify SOCS36E or other JAK/STAT targets. Hence, we cannot rule out that other targets are endocytically regulated, as there appears to have been a fundamental issue with the microarray. Therefore, it would be beneficial to carry out further transcriptomic analysis, for example RNAseq which has a greater dynamic range (Chapter 3.4.1), in order to expand our understanding of the target specific regulation of endocytosis.

6.2.1.1 Distinct STAT92E DNA binding sites

Investigating the ability of STAT92E to bind specific DNA response elements when endocytosis is inhibited may also provide key information as to the genes STAT92E regulates during trafficking. STAT92E binds to the consensus sequence TTC(n)GAA, where n can be 3 or 4 amino acids. Binding to 3n sites occurs at higher affinity than to 4n sites (Rivas et al., 2008). Both *10xSTATluciferase* and *SOCS36E* have 3n binding sites, however the number of binding sites and the surrounding sequence may be important for distinct transcriptional activity. The *10xSTATluciferase* reporter contains five repeats of a 441-bp enhancer region upstream of *SOCS36E*, which has two 3n binding sites. As this sequences was taken from the *SOCS36E* gene, the fact that endocytosis differentially regulates the reporter and the gene may initially raise concerns. However, the reporter contains tandem repeats of 10 STAT92E binding sites (Baeg et al., 2005), whilst the *SOCS36E* gene region contains 19 potential 3n STAT92E binding sites which are irregularly spaced along the DNA (Karsten et al., 2002). Therefore, although *10xSTATluciferase* contains two of the same STAT binding sites as SOCS36E, the surrounding context is very different and may account for their distinct expression patterns during endocytosis.

Cooperative binding of mammalian STAT dimers, via N-terminal interactions, is important for the recognition of weaker STAT binding sites (Xu et al., 1996). Tetramerisation of STATs enables binding to adjacent imperfect binding motifs, with the variable spacing between these sites providing specificity for the expression of specific targets (Soldaini et al., 2000). In mice STAT5 dimers are essential for viability and normal development, whereas STAT5 tetramers are vital for normal immune function (Lin et al., 2012). Therefore, cooperative STAT binding increases the number of genes regulated by the JAK/STAT pathway, and may enable regulation of distinct subsets. Although there is no direct evidence that STAT92E forms tetramers, this phenomenon could allow for distinct targets to be activated through endocytosis, due to modulation of STAT92E dimer interactions.

6.2.2. STAT92E manipulation by endocytosis.

As endocytosis plays key roles in defining the outcomes of the JAK/STAT signalling pathway, I next wanted to determine the molecular mechanism by which this occurred. I first aimed to specify the component of the JAK/STAT pathway that was undergoing endocytic regulation. Ligand induced activation of the receptor is transduced through the associated JAK to phosphorylate STAT92E at the conserved tyrosine (704). Both western blot analysis (Figure 5.2) and MS (Figure 5.13) revealed that Upd2-GFP induced STAT92E phosphorylation remains unchanged following the knockdown of endocytic components. This result is also complemented by the fact that *lama* and *chinmo* are still expressed when endocytosis is inhibited (Figure 3.21). Therefore, the receptor and kinase are capable of activating STAT92E when Dome is retained at the cell surface. These data also suggests that although phosphorylation of Y^{704} is required for STAT92E activity, it is not sufficient for expression of all targets (i.e. *10xSTATluciferase* and *SOCS36E*).

STAT-GFP was also capable of entering the nucleus in response to ligand when endocytic components were knocked down (Figure 5.6), however only a single time point was examined and therefore differences in translocation rate cannot be ruled out. It is also important to note that due to a lack of phospho-STAT92E antibodies, I was unable to confirm that the STAT-GFP accumulating in the nucleus is phosphorylated. The timescale of localisation was, however, comparable to nuclear accumulation of phosphorylated mammalian STATs (McBride et al., 2002). Ligand stimulation can also result in an increase in non-canonical unphosphorylated STAT dimers (Chapter

1.1.2.2), however this requires upregulation of STAT expression and so the timescale of this assay (30 mins) is unlikely to allow for this.

Changes to the environment of proteins during trafficking can facilitate distinct protein-protein interactions, and has been observed for the epidermal growth factor receptor (EGFR) (Teis et al., 2002) and the transforming growth factor β receptor (TGFβ) (Di Guglielmo et al., 2003). In both pathways specific endosomal adaptor proteins have been identified, MP1 and SARA respectively, enabling specific protein interactions and pathway activation during endocytosis. Proteomic studies of the EGFR also revealed changes to PTMs during trafficking, which may result in differences in downstream signalling (Tong et al, 2014). In the case of cytokine signalling, CME of the IL-6R is required for STAT3 to be transcriptionally competent. However, for maximal transcriptional activity S727 phosphorylation is required, which occurs via cross-talk with ERK1/2 at the LE (German et al., 2011). This serine residue is also important in STAT1, STAT4 and STAT5 function, and so is a possible modification that could link endocytosis to signal regulation. Therefore, further PTMs or binding partner(s) of STAT92E that alter during endocytosis may be responsible for its ability to activate distinct targets. Intriguingly, when examining STAT92E phosphorylation by western blotting, multiple phosphorylated STAT92E species were revealed (Figure 5.3), implying that there are multiple STAT92E phosphorylation sites, which have not yet been identified. Therefore, I carried out MS experiments to investigate changes to the STAT92E PTMs and interacting partners during endocytosis.

Although MS revealed novel phosphorylation sites (discusses in Chapter 6.4) I was unable to identify if these were altered during endocytic trafficking (see Chapter 5.3). I also did not detect serine phosphorylation in the transactivation domain of STAT92E, suggesting this modification is not conserved. As expected Dome and endogenous STAT92E were enriched in samples treated with Upd2-GFP, and AP2 and Hrs were absent in samples where they had been treated with targeting dsRNA. A portion of known STAT92E interactors, such as Mor, His1 and Su(var)205, were identified by MS but were not enriched after treatment or knockdown of endocytic components.

6.2.2.1 STAT92E dimerisation during endocytosis

Intriguingly, when AP2 and Hrs are knocked down, there is an increase in the association of endogenous STAT92E with STAT-GFP (Figure 5.15C,D and Figure 5.18), suggesting that when endocytosis is perturbed there is greater STAT92E

dimerisation or oligomerisation. In the canonical pathway of STAT92E activation tyrosine phosphorylation results in dimerisation, however this remains unchanged when AP2 and Hrs are knocked down (Figure 5.13). Therefore, this increase in STAT92E dimerisation appears independent of phosphorylation. Dissociation from the DNA has been shown to result in a reorientation from a parallel STAT1 dimer to an antiparallel STAT1 dimer, which allows the nuclear phosphatase TC45 to dephosphorylate STAT1, therefore forming an unphosphorylated STAT1 dimer (Zhong et al., 2005). Unphosphorylated STATs have also been demonstrated to oligomerise in the cytoplasm prior to activation (Braunstein et al., 2003), and U-STAT dimers are important in the regulation of specific transcriptional targets (see Chapter 1.1.2.2) (Chatterjee-Kishore et al., 2000; Yang et al., 2007). Therefore, the increased STAT dimerisation suggested by MS may be due to the increased formation of unphosphorylated STAT92E dimers. In *Drosophila*, unphosphorylated STAT92E is important in genome stability (Shi et al., 2008), but may also be capable of regulating subsets of genes as in mammalian STATs (Chapter 1.1.2.2). The increased formation of STAT92E dimers could be due to prolonged PTMs or interactions with binding partners that have not yet been identified. Although there was no change in overall phosphorylation levels of STAT92E following inhibition of endocytosis, I did not examine changes to the rate of phosphorylation/dephosphorylation cycle. If there are changes to the residency time of STAT92E bound to DNA, due to other PTMs or STAT92E oligomerisation, this could reduce the rate of dephosphorylation and prevent further round of STAT92E activation. This could therefore result in decreased transcription even though the overall level of phosphorylated STAT92E at a given time remains the same.

6.2.3. *SOCS36E* expression requires Hrs

The localisation of Upd2/Dome during endocytic trafficking appears to be critical for *SOCS36E* expression. Upon Hrs knockdown, delivery of Upd2-GFP to Rab5 endosomes is delayed, yet at later timepoints Upd2-GFP accumulates in large Rab5 positive endosomes (Figure 4.13) and seems to be unable to traffic efficiently to Rab7 positive endosomes (Figure 4.14). This suggests a block in Upd2-GFP trafficking, in an environment which is not competent for *SOCS36E* expression. This may be because Upd2/Dome cannot traffic to a signalling competent endosome, or that and interaction with Hrs itself is crucial in *SOCS36E* expression. Upd2-GFP induced *SOCS36E* expression requires Hrs but not TSG101 (Figure 3.4).

In *Drosophila*, the precise localisation of the Notch receptor in the endocytic pathway is also crucial for distinct signalling. Although activation of Notch by ligands occurs at the cell surface, trafficking to a late endosomal compartment is required for ligand-independent, Deltex-mediated signalling. Expression of a dominant-negative form of Rab5 prevented Deltex-mediated signalling, whereas a constitutively active form of Rab7 increased signalling (Wilkin et al., 2004, 2008). Shrub, an ESCRT-III component, was identified as a key modulator of Notch activation at the late endosome (Hori et al., 2011; Baron, 2012). Therefore, the precise location within the endocytic trafficking pathway, and an association with ESCRT components, also appears to be key for Notch activity.

Hrs is a component of the ESCRT-0 complex, which binds to ubiquitinated cargo and initiates their incorporation into ILVs. Hrs acts as a heterodimer with STAM, which can also bind ubiquitinated cargo (Mizuno et al., 2003) and has been demonstrated to positively regulate JAK/STAT signalling in mammals (Lohi and Lehto, 2001). Therefore, it is plausible that an association between JAK/STAT signalling components and STAM is required for *SOCS36E* expression, and that upon Hrs knockdown this association is abolished. In mammals, STAMs are phosphorylated in response to a range of cytokines and growth factors (Pandey et al., 2000). STAM1 and STAM2 interact with JAKs via their immunoreceptor tyrosine-based activation motif (ITAM) domain and increase JAK/STAT signalling (Lohi and Lehto, 2001; Hou et al., 2002). However, *Drosophila* STAM does not contain a ITAM domain and therefore its role in JAK/STAT signalling remains unclear (Hou et al., 2002). The Hrs/STAM heterodimer remains an interesting link between signalling and endocytosis, as it has been demonstrated to have both positive and negative roles in the regulation of RTK signalling in *Drosophila*, and is dependent on the specific tissue and developmental stage (Chanut-Delalande et al., 2010).

The MS data produced during my PhD suggests that Hrs interacts with STAT92E. Both ESCRT-0 components (Hrs and STAM) are enriched in samples where STAT-GFP was transfected, in comparison to samples expressing eGFP. This is evident in Figure 5.16A, where STAM and Hrs are found on the left hand side of the volcano plot, demonstrating enrichment in the Cntrl- sample of the comparison. This enrichment in STAT-GFP samples was also independent of ligand treatment (Figure 5.16B). When Hrs is knocked down (Figure 5.15D, 5.16D, 5.17D), STAM levels are also reduced in the sample. This indicates a physical interaction between STAT92E and Hrs, which in turn interacts with STAM, or that STAT92E cannot reach a STAM positive endosome

when Hrs is absent. Therefore, Hrs may be required for Upd2/Dome/STAT92E to associate with STAM, and therefore regulate *SOCS36E* expression.

Although knockdown of Hrs appears to cause accumulation of Upd2-GFP within Rab5 positive endosomes and block trafficking to Rab7 endosomes, previous studies have demonstrated that depletion of these ESCRT components leads to increased recycling. For example, RNAi targeting Hrs and TSG101 in Hela cells increase the rate of EGFR endocytosis and recycling (Raiborg et al., 2008). Vidal et al., 2010, demonstrated that Upd2/Dome did not colocalise with Rab11, a marker of recycling endosomes, in KC_{167} cells. However, it would be interesting to understand if this is altered when Hrs is knocked down, as increased recycling may suggest the removal of the receptor from a signalling competent environment.

6.2.3.1 Membrane microdomains

Another interesting point regarding the specific localisation required for *SOCS36E* expression is that Hrs is restricted to microdomains at the early endosomal membrane. PI3P is enriched at the EE, yet is observed in concentrated puncta along the membrane. These areas are rich in Rab5 and PI3P interactors such as EEA1 and SARA, but are distinct from Hrs, another PI3P interactor (Gillooly et al., 2003). As well as binding to ubiquitinated cargo, Hrs contains a clathrin box in its C-terminal domain which facilitates interactions with clathrin heavy chain. This interaction results in a flat clathrin coat at the endosomal membrane, that localises Hrs into microdomains absent of EEA1. These Hrs/Clathrin microdomains are key in the degradation of EGFR (Raiborg et al., 2006) and in the association with the ESCRT-1 complex, allowing for efficient ILV formation (Wenzel et al., 2018). Further studies in *C. elegans* demonstrated that retromer associated proteins, important in facilitating recycling to the trans-golgi network, are also found in distinct microdomains (Norris et al., 2017). Therefore, these distinct patches on the limiting membrane provide a single endosome with areas of differing functionalities. For example, EEA1 is important in membrane fusion, whereas Hrs targets cargo for degradation (Raiborg et al., 2002).

In this study Upd2-GFP was observed in discrete, yet uncharacterised, microdomains within Rab5 endosomes (Figure 4.8). Therefore, this confinement may be key for signalling that is competent for *SOCS36E* transcription. Hence I propose a model whereby localisation to a Hrs-positive endosome causes Upd2/Dome to be restricted into a Hrs microdomain that facilitates interaction with STAM or other unidentified proteins, allowing STAT92E to become competent for *SOCS36E* expression. As

localisation of Hrs into a microdomain is important for ILV formation (Wenzel et al, 2018) this may also provide temporal regulation to SOCS36E-competent signalling. Upd2/Dome trafficking to a Hrs endosome is required for *SOCS36E* expression, however following ILV formation the signalling portion of the receptor will be intraluminal and no longer able to interact with STAT92E, hence providing a short window in which SOCS36E expression is regulated by the JAK/STAT pathway.

To understand the microdomain in which Upd2/Dome is restricted it would firstly be useful to produce a Dome antibody in order to verify that the receptor is also found within these Upd2-GFP microdomains. In Figure 4.17A Dome-flag colocalises with Upd2 in a microdomain, however Dome-flag is also identified in other uncharacterised structures (Figure 4.17B) that may represent inefficient trafficking. Therefore, antibody staining would also provide insight into whether this is a transfection specific issue and enable examination of endogenous protein location. To understand the composition of these sub-compartments the colocalisation of Upd2/Dome with a wider range of endosomal and lysosomal markers could be examined. In this study I aimed to minimise disruption to endocytosis and JAK/STAT signalling and therefore used antibodies against Rab5, Hrs and Rab7. However, there is a range available plasmids for the expression of fluorescently-tagged markers (Rabs, ESCRTs, Lamp1) that could be utilised to investigate endosomal subdomains. Further MS analysis would also provide information on changes in protein interactors. Defining this microdomain may reveal information about the environment required for STAT92E to become competent for *SOCS36E* gene activation.

6.2.3.2 Endosomal subpopulations

When investigating the localisation of Upd2-GFP, the ligand appeared to be not only restricted to endosomal microdomains, but also to subpopulations of Rab5 positive endosomes (Figure 4.8). As discussed in Chapter 1.3.2.1, Rab5 endosome subpopulations provide distinct signalling capabilities to EE's. In zebrafish, localisation to APPL1 endosomes was shown to enable differential Akt signalling (Schenck et al., 2008), suggesting a distinct signalling platform. Endosomal subpopulations have also been described for the activation of mammalian STAT3 by c-Met. As discussed in Chapter 1.3.4.1, OSM elicits strong and rapid phosphorylation of STAT3, which is independent of receptor endocytosis, whereas c-Met requires endocytosis to cause STAT3 nuclear accumulation. The delivery of c-Met to a population of perinuclear endosomes is required for STAT3 transcriptional activity. This suggests that localisation of c-Met into specific endosomes that are spatially close to the nucleus, is

key to overcome the weak STAT3 activity (Kermorgant and Parker, 2008). Therefore, localisation of Upd2-GFP to distinct subsets of endosomes may provide signal specificity due to association with a specific signalling platform and/or spatial restriction.

6.3 Internalisation and trafficking of the receptor

6.3.1. Dome internalisation motif

Although Dome has been shown to undergo CME (Devergne et al., 2007; Vidal et al., 2010), little is known about the sequence motif that permits internalisation. During this project I have demonstrated that a conserved dileucine box, [D/E]xxxL[L/I], that is crucial for internalisation of the mammalian gp130 receptor (Dittrich et al., 1996), is also essential for Dome internalisation (Chapter 4.3). This motif is well documented to bind to the α-σ2 hemicomplex of AP2 (Doray et al., 2007; Kelly et al., 2008), and mutation of the SESSKLLL (WT) motif to AAASKAAL (allA) in the cytoplasmic tail of Dome-flag prevented its internalisation (Figure 4.21A, D). This data supports previous studies (Devergne et al., 2007; Vidal et al., 2010), validating that Dome is indeed internalised through an AP2 and Clathrin-dependent mechanism. With the aim of identifying the essential residues within this dileucine box, individual amino acids were mutated. Mutation of the dileucines alone did not affect Dome-flag internalisation compared to wild-type. This was surprising as the binding of the dileucine to the σ2 subunit of AP2 is considered the key recognition event in the interaction of the adaptor protein and [D/E]xxxL[L/I] motif (Kelly et al., 2008). Further mutation of surrounding residues decreased, but did not completely abolish, internalisation when compared to the 'all A' mutant, suggesting multiple residues within this motif are important for Dome-flag internalisation (Figure 4.21D). To determine the role of these individual residues a more detailed examination, e.g. changes to rate of internalisation, may be required. During this study I have only investigated uptake after 15minutes, therefore any changes to early internalisation events or the rate of uptake would have not been identified.

6.3.1.1 Dome is constitutively endocytosed

Whilst investigating the Dome-flag internalisation mutants, it was evident that the receptor can be endocytosed in the absence of ligand (Figure 4.20). Ligand-

independent Dome internalisation has also been documented by Stec et al 2013, in KC_{167} cells, and by Devergne et al 2007, who observed membrane accumulation of Dome after mutation of clathrin across the *Drosophila* egg chamber that was not restricted to regions where Upd is expressed (Devergne et al., 2007). As KC_{176} and S2R+ cells secrete Upd ligands into their culture media, I replaced media an hour prior to biotinylation of cell-surface proteins with the aim of removing secreted Upd in the media and allowing ligand occupied Dome to be internalised. However, this made no difference to the unstimulated internalisation of Dome (Chapter 4.3). This data therefore suggests that Dome undergoes constitutive internalisation both *in vitro* and *in vivo*, which according to this project, occurs via CME as the AP2-dependent Dome-flag internalisation mutant is not internalised in absence of ligand. This indicates that Dome is persistently associated with AP2. Unfortunately, I have not been able to investigate Dome-AP2 interactions during this project, yet gp130 also undergoes ligand independent internalisation and interacts with AP2 continuously. This pre-associated complex may provide a mechanism to quickly internalise ligand-bound, signalling receptors as a way to regulate signalling through delivery to signalosomes and to the lysosome for degradation. Therefore, it would be interesting to understand if ligand stimulation increased the rate of Dome internalisation, and to determine differences, if any, between constitutive and stimulated endocytosis. Windpipe was recently identified as a target of JAK/STAT pathway activation, and was demonstrated to interact with Dome, promoting internalisation and degradation in a ligand-independent manner. This resulted in negative regulation of signalling due to increased receptor degradation (Ren et al., 2015) and provides a role for ligand-independent endocytosis is signal regulation by modulating cell-surface Dome levels.

6.4 Context specific roles of endocytic regulation

6.4.1. Is the mechanism of internalisation dependent on ligand concentration?

Dome appears to be primarily internalised via CME when stimulated with 3nM Upd2-GFP (Figure 4.15). However, when stimulated with higher concentrations of Upd2-GFP, clathrin knockdown only partially inhibits Upd2-GFP uptake (Figure 1.6). This suggests that Dome is internalised via multiple endocytic portals that are dependent on ligand concentration. This may allow cells to respond and limit signalling when concentration of extracellular ligand are high. This has been previously demonstrated

for the EGFR which is internalised via CME and trafficked for recycling at low doses of EGF, yet at high doses of EGF the receptor is internalised via CIE and trafficked for degradation (Sigismund et al., 2005).

Treating S2R+ cells with increased concentrations of Upd2-GFP (from 3nM to 40nM) abolished the effect of AP2 and Hrs knockdown on *SOCS36E* expression (Figure 3.4 and Figure 3.6). At 3nM Upd2-GFP, knockdown of Hrs inhibits *SOCS36E* expression, whereas at 40nM Upd2-GFP knockdown of Hrs does not alter *SOCS36E* expression (Figure 3.6B). This may suggest that the level of Hrs knockdown is not sufficient to overcome increased levels of JAK/STAT signalling, or that at high ligand concentrations Dome is trafficked through a CIE route that enables expression of *SOCS36E* expression from receptors at the cell surface. Physiologically it would be beneficial to upregulate the level of SOCS36E for a sustained periods when there are high levels of ligand. As SOCS36E is a negative regulator of the pathway, its increased expression could be important in limiting the level of JAK/STAT signalling. In contrast, at low levels of ligand, CME and trafficking may cause SOCS36E expression to be more transient. This would result in the production of less SOCS36E and maintain JAK/STAT signalling after weak signals. Therefore, internalisation via different mechanisms may be essential in fine-tuning signalling so that cells can interpret changes to their environment and respond accordingly. In line with the fact that Dome can internalise independent of ligand (discussed in Chapter 6.2.1.1), this may provide a model whereby endocytosis can regulate the level and specificity of JAK/STAT signalling in response to environmental cues (Figure 6.1). When there is no ligand present endocytosis occurs via CME, but does not activate STAT92E. At low levels of ligand Dome is internalised via CME and JAK/STAT signalling is differentially regulated, requiring trafficking to a Hrs-positive endosome for *SOCS36E* expression. At high levels of ligand Dome appears to be internalised also via CIE, whereby *SOCS36E* expression is not dependent on Hrs and its sustained activation may lead to downregulation of the pathway.

Evidently, more work is required to investigate this hypothesis. The major drawback for these studies is the production of Upd2-GFP at high concentrations. Therefore, it may be useful to produce a stable Upd2-GFP expressing S2R+ cell line, instead of transiently transfecting the plasmid, or expressing the protein in bacteria. One mechanism to study Dome internalisation directly is via cell-surface biotinylation, to investigate whether high levels of Upd2 can stimulate uptake of the Dome internalisation mutant, which would therefore confirm clathrin independent

internalisation. Analysis by the anti-GFP ELISA would also be useful to examine Upd2-GFP uptake over a wide range of concentrations. It would also be interesting to investigate how other JAK/STAT targets are regulated by ligand concentrations, for example some targets that are dependent on AP2 expression may only be expressed at low Upd2-GFP levels.

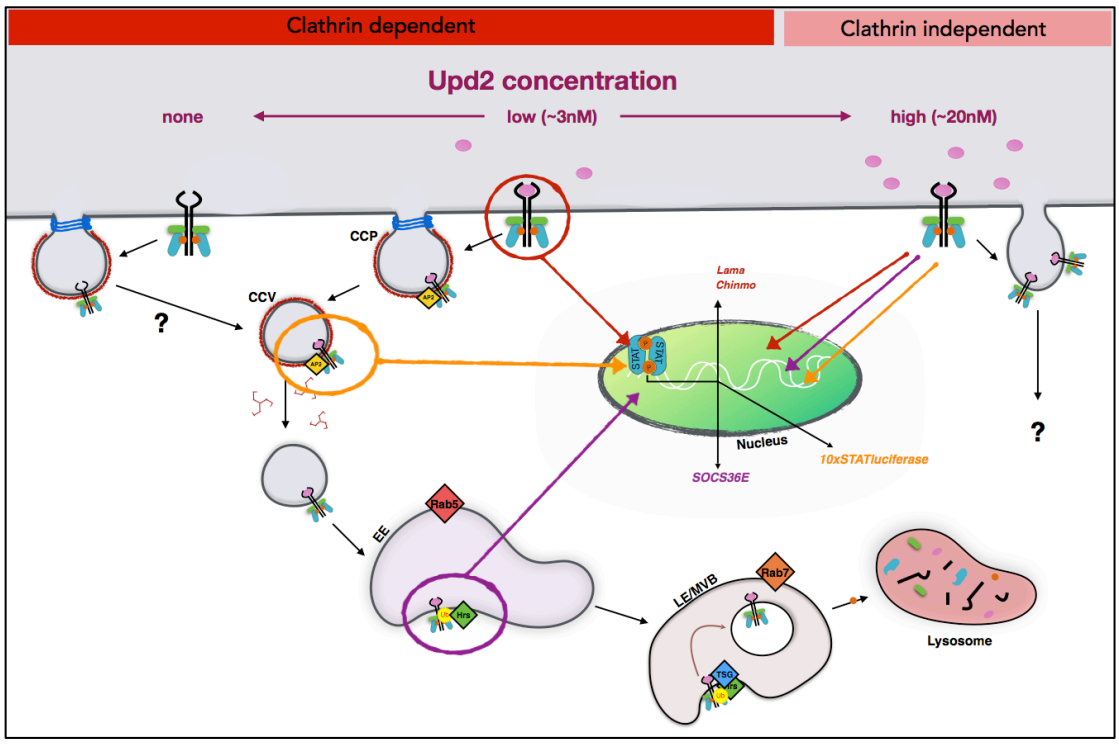


Figure 6.1: Model of Dome internalisation at varying ligand concentrations. Dome is constitutively endocytosed via CME when no ligand is present, yet does not activate transcriptional targets. At low levels of Upd2-GFP, Dome is internalised via CME and endocytic trafficking causes differential regulation of transcriptional targets. At high levels of Upd2-GFP, Dome appears to undergo CIE whereby transcriptional targets are expressed at the cell surface.

6.4.2. How is endocytosis of the JAK/STAT receptor altered if stimulated with different ligands?

The EGFR can be activated by multiple ligands that differentially regulate EGFR phosphorylation, the mechanism of EGFR uptake, intracellular trafficking and the receptor fate (Roepstorff et al., 2009; Henriksen et al., 2013). Firstly, EGFR activated by EGF and TGF-α undergo different endocytic trafficking fates. Due to the strong associated of EGF to the receptor, the ligand remains bound in the endosome, leading to ubiquitination and degradation. In contrast, TGF-α dissociates from the receptor at the low pH of the endosome, causing the receptor to be recycled. These different

endocytic routes are key for the duration of downstream signalling, and means that TGF-α is more potent than EGF for specific biological processes (Decker, 1990; Ebner and Derynck, 1991). Further studies demonstrated that heparin-binding EGF-like growth factor (HB-EGF) and betacellulin (BTC) treatment causes higher levels of EGFR degradation than stimulation with EGF (Roepstorff et al., 2009). All ligands appear to induce CME of EGFR, but HB-EGF and BTC also appear to stimulate internalisation via a CIE method (Henriksen et al., 2013), although concentration dependence of this finding was not tested. Therefore, different ligands can differentially regulate uptake and trafficking of the same receptor, leading to variability in downstream signalling and biological outcomes. This may be especially important in the tumour microenvironment, where specific ligands are dysregulated and therefore understanding their differential role on endocytosis may be key to understanding signalling defects.

The *Drosophila* genome contains three ligands capable of activating Dome, Upd, Upd2 and Upd3. Considering the differential roles of EGFR ligands, investigating trafficking of Dome after stimulation with varying Upd-ligands may provide information regarding how Upd2 produces longer lasting signals (Wright et al., 2011) (Chapter 1.1.4.1). Upd2 stimulated Dome endocytosis has been characterised in detail, using biochemical and microscopy techniques (Chapter 4.1) (Vidal et al., 2010). CME of Dome and colocalisation with Rab5 and Rab7 in follicle cells of the *Drosophila* egg chamber (Devergne et al., 2007) demonstrates that the endocytic route of Dome upon Upd2 stimulation is also physiologically relevant. However, all ligands are present *in vivo* and therefore ligand-specific roles cannot be defined. Therefore, it would be interesting to carry out biochemical and cell-based microscopy assays to understand if Upd and Upd3 stimulate differential internalisation and trafficking of the receptor. As there appears to be distinct targets activated by each ligand (Yang et al., 2015), it would also be interesting to understand how endocytosis regulates expression of these specific outputs.

6.4.3. Crosstalk of JAK/STAT signalling with other pathways

The JAK/STAT pathway is a relatively straightforward signal transduction pathway, with four main components. However, its interactions and crosstalk with various other signalling pathways adds complexity to understanding this pathway. Although I have not investigated the relationship between signal crosstalk and endocytic regulation of

JAK/STAT signalling, it is important to consider how this may alter the context of signalling. As mentioned briefly in Chapter 1.3.4.1, some RTKs activate JAK/STAT signalling. EGFR and platelet derived growth factor receptor (PDGFR) can tyrosine phosphorylate STATs through Src kinase activity, that is independent of JAK activity (David et al., 1996; Olayioye et al., 1999; Wang et al., 2000). Activation of MAPKs downstream of RTKs also results in the phosphorylation of STATs at serine residues (Chung et al., 1997) which alters their transcriptional activity in a STAT-specific manner (discussed in Chapter 1.1.3.2). We did not detect phosphorylation of serine residues in the C-terminus of STAT92E (Chapter 5.3) however, and there is no evidence to suggest that *Drosophila* EGFR is capable of activating STAT92E independently of Dome. However, phostag gels revealed the presence of phospho-STAT species without stimulation. Whether this is a 'background' level of phosphorylation, or low level of activation due to ligands in the media is unknown. However, it could also be due to a dome-independent activation mechanism, that has not been identified to-date.

The *Drosophila* JAK/STAT pathway interacts with the MAPK pathway via Hopscotch (the JAK). After activation Hop can directly interact with, and activate, *Drosophila* Raf to initiate a Raf/MEK/MAPK cascade (Luo et al., 2002). Although this does not appear to affect STAT92E activity, *Draf* has been identified as a transcriptional target of STAT92E (Kwon et al., 2000). A study by Ammeux et al., 2016, examined the crosstalk of signalling pathways in a variety of *Drosophila* cell lines. They found that SOCS36E and JAK/STAT ligands are upregulated in response to insulin, EGFR and JNK pathway stimulation. Interestingly, the expression of Upd2 required activation of both the JAK/STAT pathway and BMP pathway, whereas upd and upd3 were expressed when the JAK/STAT pathway was activated alone. There also appeared to be an inhibitory relationship between the JAK/STAT and insulin pathways, as fewer transcriptional targets were expressed than if the pathways were activated independently (Ammeux et al., 2016). Therefore, the crosstalk of pathways enables modulation of signalling outcomes, and may be influenced by endocytic regulation.

6.5 Phosphorylation of threonine 702 is essential for STAT92E activity

MS experiments carried out during this study revealed three novel STAT92E phosphorylation sites; threonine 47, serine 227 and threonine 702. As other PTMs have key roles in mammalian STAT activity (see Chapter 1.1.3.2) I decided to examine the role of these newly identified phosphorylation sites on STAT92E activity. To do this I produced a S2R+ cell line with low levels of endogenous STAT92E using CRISPR/Cas9 genome engineering (Chapter 5.4.1). S2R+ cells cannot grow in isolation and therefore I was unable to produce a clonal cell line. However, the population of transfected cells had low STAT92E levels (Figure 5.20), and signalling could be rescued by the addition of STAT-GFP (Figure 5.25). Mutants of STAT92E where threonine 47 and serine 227 could not be phosphorylated, $STAT^{T47V}$-GFP and $STAT^{S227A}$-GFP, did not alter expression of *10xSTATluciferase* compared to $STAT^{WT}$-GFP, when expressed in crSTAT cells (Figure 5.27). In contrast, mutation of threonine 702 to a unphosphorylatable valine residue (T702V) within STAT-GFP, inhibited JAK/STAT signalling to a similar extent as expression of $STAT^{T704F}$-GFP (Figure 5.27). This indicated a key role for the threonine in STAT92E activity. Phosphomimetic mutations at T^{702} rescued this signalling defect (Figure 5.34), suggesting a specific role for threonine phosphorylation. Unfortunately, this phosphorylation site was not identified in MS experiments where AP2 and Hrs where knocked down, and therefore I could not determine if it is endocytically regulated.

6.5.1. Dimerisation of unphosphorylated STAT92E

Further study revealed that $STAT^{T702V}$-GFP cannot enter the nucleus in response to ligand (Figure 5.30), hinting that it is unresponsive to receptor and JAK activation. The dimerisation of STAT proteins is key for their nuclear translocation, DNA binding and thus transcriptional activity. Therefore, to understand the requirement of T^{702} phosphorylation on STAT92E activity I carried out coimmunoprecipitation experiments to investigate the dimerisation of STAT-GFP with endogenous STAT92E. $STAT^{WT}$-GFP is capable of interacting with both unphosphorylated and phosphorylated, identified by a shift in SDS-PAGE mobility (Figure 5.31). $STAT^{Y704F}$-GFP appears to associate with only unphosphorylated STAT92E, which is expected due to the lack of phosphorylation required for parallel dimer formation. In contrast $STAT^{T702V}$-GFP does not appear to interact with either species of STAT92E. This may suggest that the phosphorylation of T^{702} is required for the formation of dimers in the absence of tyrosine phosphorylation prior to ligand stimulation. Unphosphorylated dimers have been speculated to be a prerequisite for STAT binding to their cognate receptor (Ota et al., 2004; Mao et al., 2005), and therefore may represent an essential stage in the priming

of STATs for activation in response to stimulation. As shown by MS, $STAT^{T702V}$-GFP is not efficiently phosphorylated at Y^{704} (Figure 5.33). Therefore, inhibition of STAT dimerisation, by the T702V mutation, may prevent tyrosine phosphorylation by inhibiting STAT from binding the receptor and hence becoming activated.

6.5.2. Conservation in mammalian STATs

The T^{702} residue in STAT92E is conserved in mammalian STAT1 and is a phosphomimetic (aspartic acid) in STAT5 (Figure 6.2A), suggesting a conserved role for this residue. Using the ExPASy SWISS-MODEL (https://swissmodel.expasy.org/) I aligned the sequence of STAT92E (isoform C) with the crystal structure of STAT1 (PBD: 1bf5.1.A) (Figure 6.2B). From this alignment, T^{702} and Y^{704} are found in a flexible loop region (Figure 6.2C). Changes to the charge of residues, such as upon phosphorylation, may drastically alter the structure of this region and potentially change STAT92E interactions. A large conformation change could explain why $STAT^{T702V}$-GFP runs slower on an SDS-PAGE gel (Figure 5.28), compared to $STAT^{WT}$-GFP.

Two residues downstream of the conserved tyrosine, is a lysine that is also present in STAT92E, STAT1 and STAT5 (Figure 6.2A). Lysine residues have a basic charge, and therefore an interaction with the negative charge of the phosphorylated threonine or tyrosine within the same flexible loop is conceivable. This conserved lysine residue has been reported to be both acetylated and SUMOylated in mammalian STATs. STAT5a SUMOylation of lysine 696 reduces tyrosine phosphorylation, whereas acetylation is required for phosphorylation (Van Nguyen et al., 2012). Therefore, other residues in close proximity to the threonine residue and within the same flexible loop play significant roles in STAT activity. T^{702} is adjacent to the SH2 domain of STAT92E, which binds phosphorylated tyrosine's on the receptor. Therefore, a drastic change in the conformation near the SH2 domain may prevent STAT92E from binding receptor, and hence inhibit subsequent tyrosine phosphorylation and STAT92E activation. Thus, the location of the threonine falls within a region which has been previously identified to have roles in STAT activity.

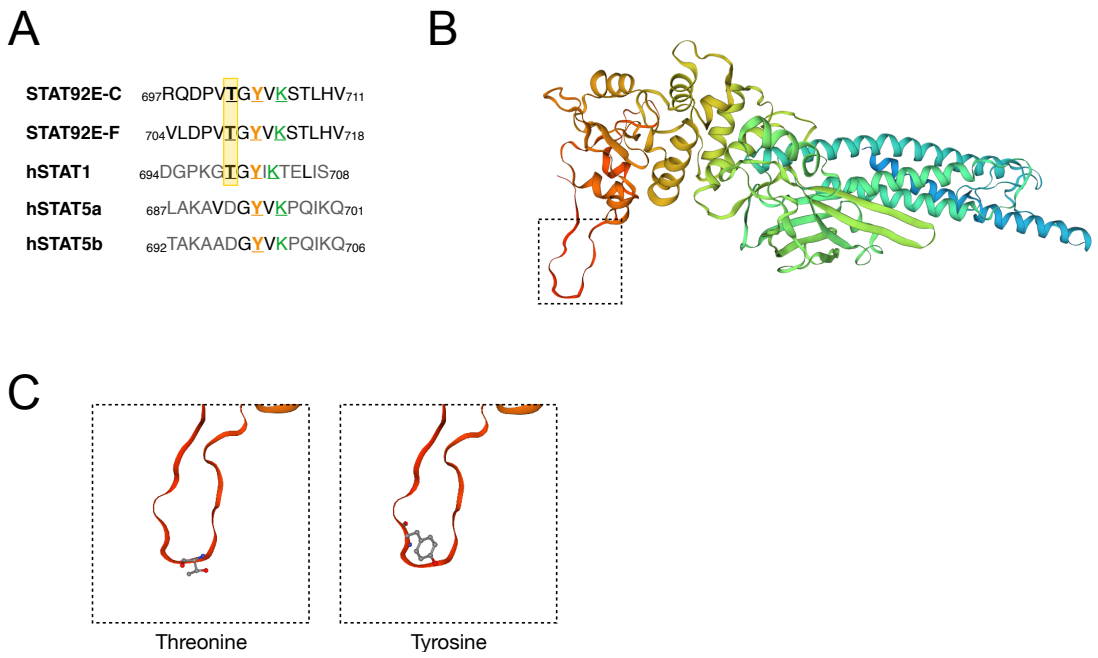


Figure 6.2: Threonine 702 conservation and location within STAT1 crystal strucutre. A) Alignment of sequences surrounding the conserved tyrosine in STAT92E-C (C isoform), STAT92E-F (long isoform), human STAT1, STAT5a and STAT5b. The conserved tyrosine is highlighted orange, and a conserved lysine highlighted in green. The threonine residue is in a yellow box. B) Crystal structure of STAT1 (1bf5.1.A). C) Location of the threonine and tyrosine residues within the STAT1 crystal structure.

Due to time limitations I was unable to investigate the role of the conserved threonine residue in the activity of mammalian STAT1 and STAT5. However, initial experiments would be to examine if mutation of the threonine, to non-phosphorylatable residue, also prevents tyrosine phosphorylation in STAT1 and STAT5. As there are tyrosine phosphorylation specific antibodies available for mammalian STATs, analysis of tyrosine phosphorylation in mutant STAT1 constructs can be carried out via western blotting. Following this, further experiments could then investigate phenotypic outputs of STAT1 and STAT5 activity. Both regulate proliferation, however STAT1 is antiproliferative (Bromberg et al., 1996) whereas STAT5 stimulates proliferation (Baskiewicz-Masiuk and Machalinski, 2004). Therefore, flow cytometry could be used to study the cell cycle progression of cells expressing either WT or the threonine mutant of STAT1 or STAT5.

6.5.3. Is the threonine phosphorylation constitutive or regulated?

If the role of threonine phosphorylation is found to be conserved in mammalian STATs, it would be interesting to understand whether the residue is constitutively phosphorylated or if it is regulated in response to, for example, MAPK activity. As mentioned previously (1.1.3.2), phosphorylation of a serine residue in the transactivation domain of STATs can occur via activation of multiple kinases. Although this serine was not identified in STAT92E (Chapter 5.3), it demonstrates that modulation of STAT activity can occur through the activity of distinct kinases. Therefore, I ran the sequence surrounding threonine 702 in STAT92E and STAT1 through kinase predictors, Group-based Prediction System (GPS) and NetPhos 3.1. These predictors only take into account sequence, and not protein structure, to produce a score that relates to the potential of the kinase to phosphorylate a specific residue. The top 6 scores when using GPS are the same for both STAT1 and STAT92E, indicating conservation of the surround residues. Only one kinase was identified using NetPhos3.1, and this was PKC. This is interesting as PKC-δ has been shown to be important in mediating type 1 IFN responses by phosphorylation serine 727 of STAT1 (Uddin et al, 2002). To determine which kinases are responsible for threonine phosphorylation, various kinase inhibitors could be used. If phosphorylation of the threonine is key for STAT1 tyrosine phosphorylation, as it appears to be in STAT92E, inhibiting the threonine kinase may subsequently prevent tyrosine phosphorylation and expression of transcriptional targets.

STAT	Kinase	GPS Score	Netphos score
STAT92E	AGC/GRK	28.5	
SKRQDPVTGYVKSTL	AGC/DMPK/ROCK	20.952	
	CK1/VRK/VRK2	15.25	
	Atypical/PIKK/FRAP	13.583	
	STE/STE7/MEK3/MAP2K3	13.111	
	AGC/PKC/PKCa/PRKCA	10.808	
	PKC		0.746
STAT1	AGC/DMPK/ROCK	36.565	
MLDGPKGTGYIKTEL	AGC/GRK/BARK/BARK1	20.381	
	STE/STE7/MEK3/MAP2K3	13.667	
	AGC/PKC/PKCa/PRKCA	13.212	
	CK1/VRK/VRK2	12.75	
	Atypical/PIKK/FRAP	8.306	
	PKC		0.600

Table 6.1: Output of kinase predictors for phosphorylation of threonine residue. A 15bp sequence surrounding T^{702} in STAT92E and T^{699} in STAT1 were loading into GPS and NetPhos3.1. The top 6 scoring kinases from the GPS predictor are shown; the higher the score the greater the potential of phosphorylation. NetPhos3.1 scores from 0-1, with 0.5 being classed as a positive prediction.

6.6 Conclusion and future directions

During this PhD project I aimed to increase understanding of how endocytosis regulates the outcome of a signalling pathways to produce a defined and appropriate response. Utilising the *Drosophila* system to examine JAK/STAT signalling, I have demonstrated that CME of Dome and subsequent endocytic trafficking enables the differential regulation of distinct JAK/STAT target genes. This appears to occur in a context-dependent manner, and suggests that endocytosis is a mechanism to modulate signalling output in response to the extracellular environment. Intriguingly, STAT92E is still phosphorylated at the conserved tyrosine residue when Dome is at cell-surface and throughout trafficking. This suggests that STAT92E is transcriptionally active, but is not competent for a subset of target genes. Further research should focus on investigating other JAK/STAT targets that are regulated by endocytosis, and defining the endosomal environments required for STAT92E to be competent for expression of specific target genes. This study has also revealed a novel STAT92E phosphorylation site that appears to be crucial for transcription factor activation. This threonine residue is conserved in STAT1 and is a phosphomimetic in STAT5.

Therefore, future work should investigate the role of threonine phosphorylation in mammalian STAT function.

In conclusion, this thesis has presented work that builds upon our current understanding of endocytic regulation. Although the *Drosophila* JAK/STAT is activate when the receptor is at the cell-surface, internalisation and trafficking to different endosomal compartments is required for the expression of distinct targets, demonstrating an essential role for endocytosis in qualitatively regulating signalling pathways.

Appendix

1. Pearson's Correlation Coefficient script

```
macro "test [f5]" {

name = getTitle();
upperROI = roiManager("count");

selectImage(name);
run("Split Channels");
redname = "C1-"+name;
greenname = "C2-"+name;

selectImage(redname);
run("Add...", "value=1");
selectImage(greenname);
run("Add...", "value=1");

	for (index = 0; index < upperROI; index++) {

		selectImage(redname);
		roiManager("Select", index);
		reddupl = "red"+index;
		run("Duplicate...", "title="+reddupl);		//red ROI to
correlate

		selectImage(greenname);
		roiManager("Select", index);
		greendupl = "green"+index;
		run("Duplicate...", "title="+greendupl);		//green ROI
to correlate

		roiManager("Deselect");
		selectImage(greendupl);
		getRawStatistics(ycount, ymean, ymin, ymax, ystd);
		run("32-bit");
		imageCalculator("Multiply", greendupl, reddupl);
		selectImage(greendupl);
		getRawStatistics(mcount, mmean, mmin, mmax, mstd);
		close(greendupl);
		selectImage(reddupl);
		getRawStatistics(xcount, xmean, xmin, xmax, xstd);
		close(reddupl);
		Rr = ((xcount*mmean)-(xcount*xmean*ymean))/((xcount-1)*xstd*ystd);
		setResult("Pearson", index, Rr);
	}

close(redname);
close(greenname);
```

2. Bioconductor scripts for use in R

```
# create shortcuts to working directories
home<-"~/Dropbox/Microarray_data/"
CEL<-"~/Dropbox/Microarray_data/CEL_files/"
results<-"~/Dropbox/Microarray_data/Results/"
raw_d<-"~/Dropbox/Microarray_data/Results/Raw/"

# CEL contains file_annotation.csv with following table
 Name          Condition   Date
 Control_1.CEL Control     17/12/2015
 Treated_1.CEL Treated     17/12/2015
 AP2_1.CEL     AP2         17/12/2015
 Hrs_1.CEL     Hrs         17/12/2015
 Tsg_1.CEL     TSG101      17/12/2015
 Control_2.CEL Control     18/01/2016
 Treated_2.CEL Treated     18/01/2016
 AP2_2.CEL     AP2         18/01/2016
 Hrs_2.CEL     Hrs         18/01/2016
 Tsg_2.CEL     TSG101      18/01/2016
 Control_3.CEL Control     16/03/2016
 Treated_3.CEL Treated     16/03/2016
 AP2_3.CEL     AP2         16/03/2016
 Hrs_3.CEL     Hrs         16/03/2016
 Tsg_3.CEL     TSG101      16/03/2016

# Install packages required for analysis from Bioconductor
source("www.bioconductor.org/biocLite.R")
biocLite("name of the package")
```

PUMA mmgmos script

```
# Load packages
library(affy)
library(gplots)
library(drosophila2cdf)

# Load data into Affybatch
setwd(CEL)
File_Info<-read.csv("File_annotation.csv", header=TRUE)
names<-t(data.frame(File_Info$Name))
affybatch_d<-ReadAffy(filenames=names)

# attach the pheno data
pData(affybatch_d) <-
 data.frame(condition=File_Info$Condition,row.names=rownames(pData(affybatch_d)))

# boxplot raw data
boxplot(affybatch_d)

# Median global scaling normalisation
library(puma)
raw_data_puma<-mmgmos(affybatch_d)

# write results
setwd(raw_d)
write.reslts(raw_data_puma, file="raw_drosphila_data.csv")

# boxplot mmgmos normalised data

boxplot(exprs(raw_data_puma))

# histogram of mmgmos normalised data
hist(exprs(raw_data_puma))

# PCA of mmgmos normalised data
puma_pca<-pumaPCA(raw_data_puma)
plot(puma_pca)

# Combination of repeats
data_puma_comb <- pumaComb(raw_data_puma)
save(data_puma_comb, file="raw_data_puma_comb.rda")
write.reslts(data_puma_comb, file="data_puma_comb.csv")

# Calculate differential expression
de_drosophila_comb <- pumaDE(data_puma_comb)
save(de_drosophila_comb, file="de_drosophila_comb.rda")
write.reslts(de_drosophila_comb, file="de_drosophila_comb.csv")
```

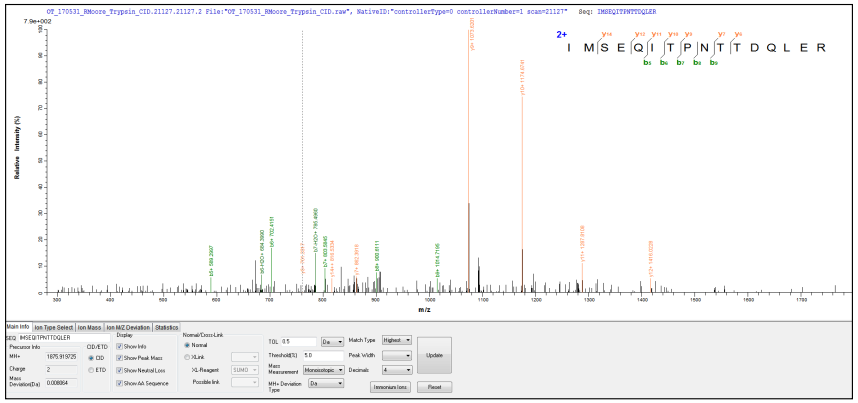
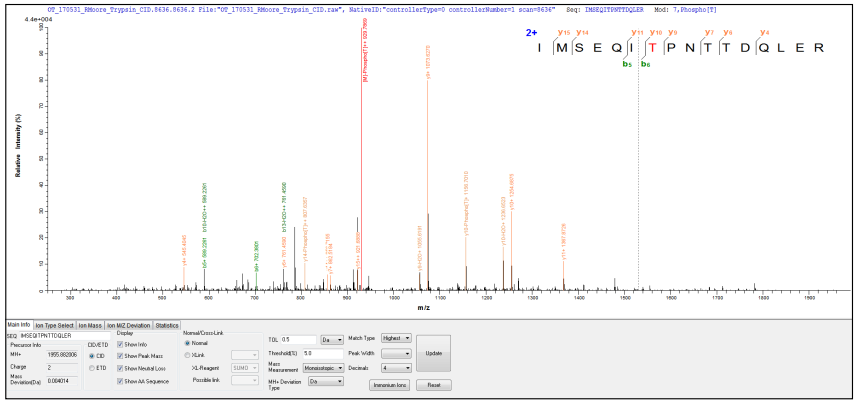
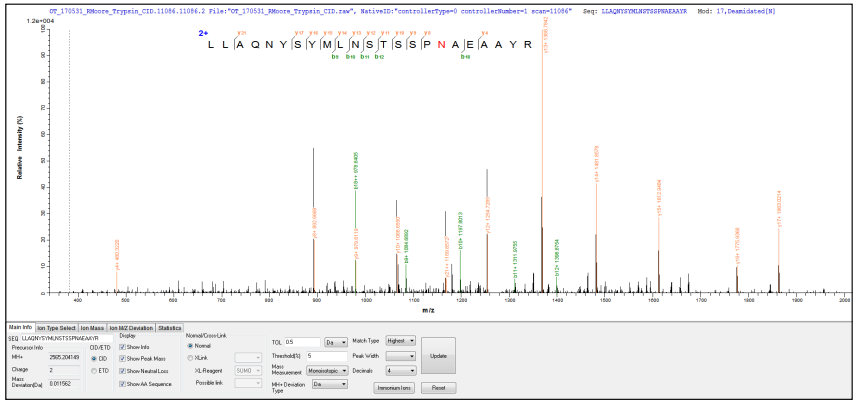
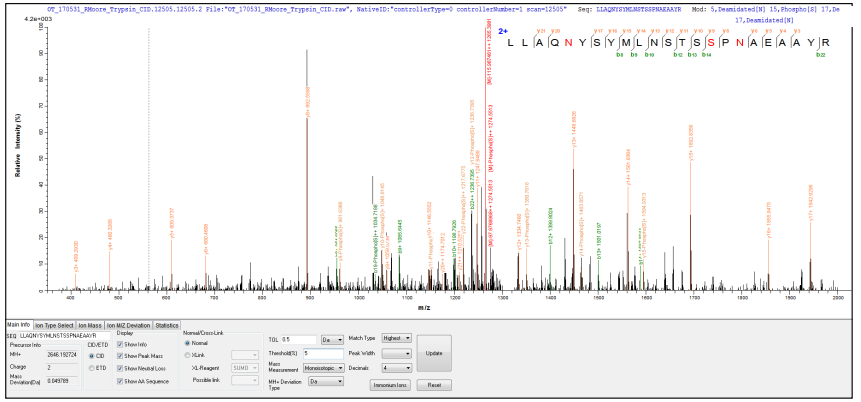
RMA script

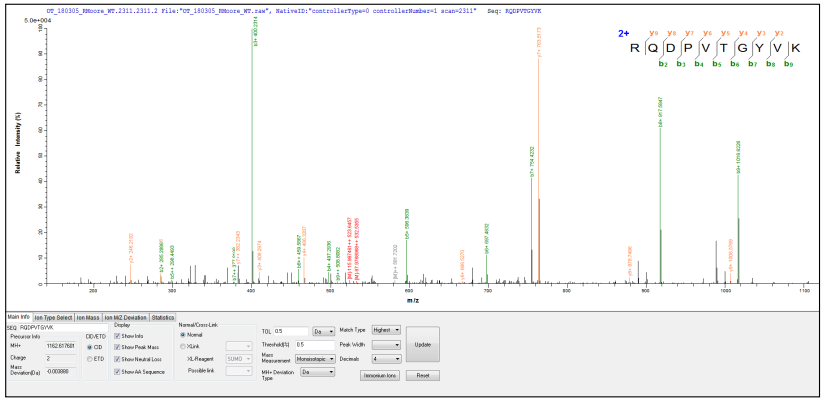
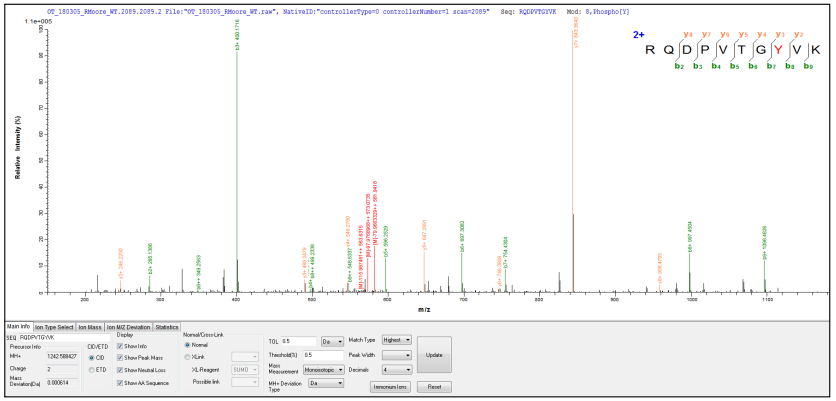
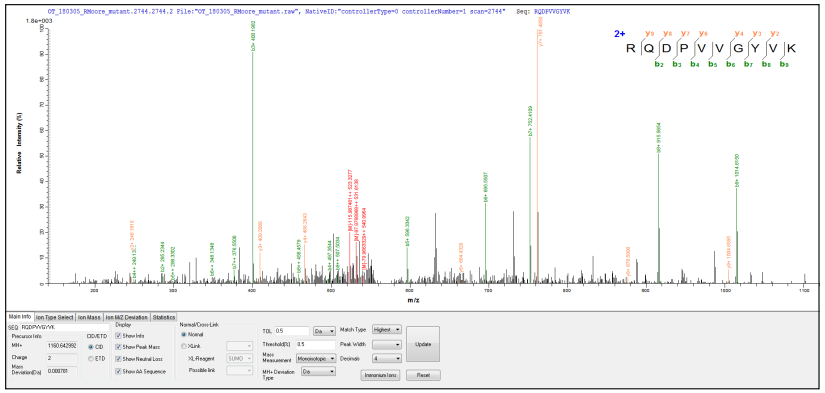
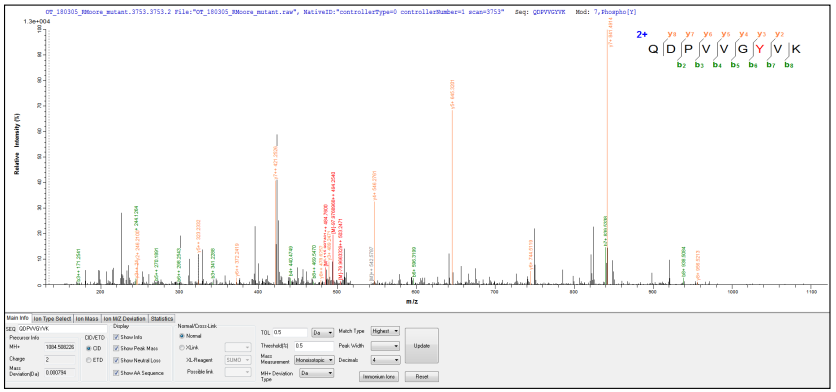
```
# Load packages
library(affy)
library(limma)
library(gplots)
library(drosophila2cdf)

# Load data into Affybatch
setwd(CEL)
File_Info<-read.csv("File_annotation.csv", header=TRUE)
names<-t(data.frame(File_Info$Name))
affybatch_d<-ReadAffy(filenames=names)

# attach the pheno data
pData(affybatch_d) <-
 data.frame(condition=File_Info$Condition,row.names=rownames(pData(affybatch_d)))

# normalise, background correct and calculate expression
raw_data<-rma(affybatch_d)
write.reslts(raw_data, file="raw_data.csv")

# boxplot RMA normalised data
boxplot(exprs(raw_data))

# histogram of RMA normalised data
hist(exprs(raw_data))

# construct matrix to compare differentially expressed genes
pd<-pData(raw_data)
design<-model.matrix(~0+condition, pd)
colnames(design)<-c("AP2","Control","Hrs","Treated","TSG101")
# should give table with correct 1 in each column

fit<-lmFit(raw_data,design)
constrast.matrix<-makeContrasts(Treated-Control,Treated-AP2,Treated-Hrs,Treated-
 TSG101,levels=design)
fit2<-contrasts.fit(fit,constrast.matrix)
ebFit2<-eBayes(fit2)

# Create top table of DE genes between treated and control
topTable(ebFit2,coef="Treated - Control", adjest.method="none", n=20,p.value,genelist=genes)
```

PMmmgmos script (use oligo package to read cel files)

```
# Load packages
library(affy)
library(oligo)
library(gplots)
library(drosophila2cdf)

setwd(CEL)
File_Info<-read.csv("File_annotation.csv",  header=TRUE)
names<-t(data.frame(File_Info$Name))
affybatch_d<-read.celfiles(filenames=names)

library(puma)
raw_data_PM<-PMmmgmos(affybatch_d)

# continue as with mmgmos analysis for combining repeats and for DE
```

3. MS/MS Spectra

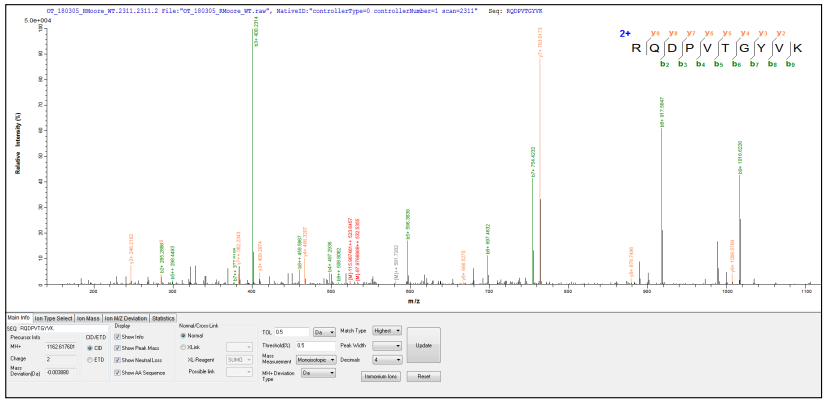
	unphospho	phosphoSTY
T47		
S227		

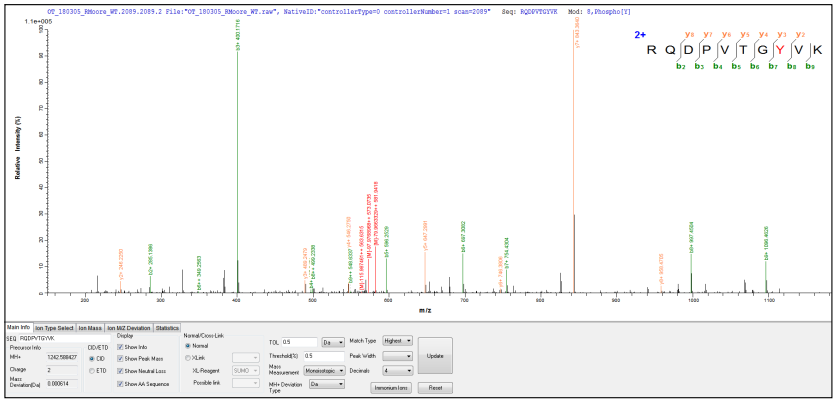
	unphospho	phosphoSTY
WT		
T702V		

unphospho

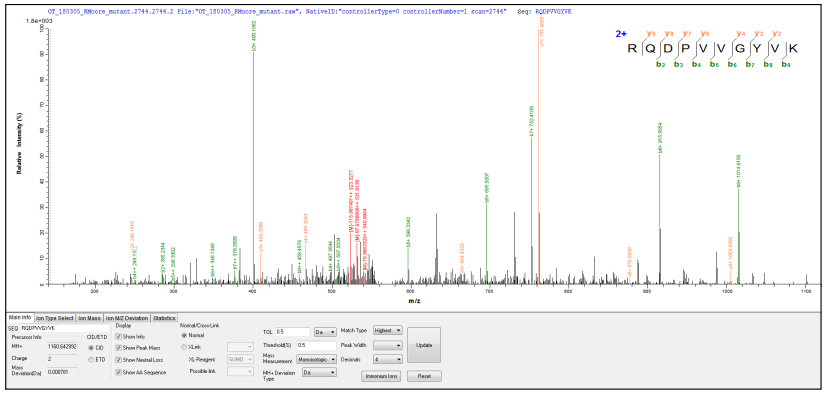
phosphoSTY

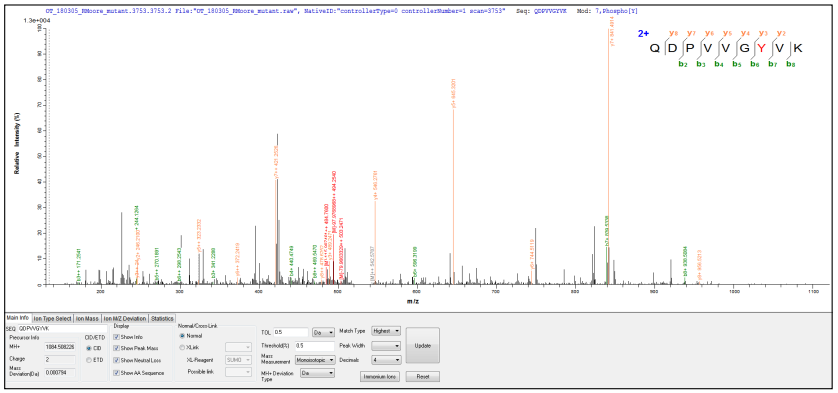
WT





T702V





References

Agaisse, H., Petersen, U. M., Boutros, M., Mathey-Prevot, B. and Perrimon, N. (2003) 'Signaling role of hemocytes in Drosophila JAK/STAT-dependent response to septic injury.' *Developmental Cell*, 5(3) pp. 441–450.

Allanach, K., Mengel, M., Einecke, G., Sis, B., Hidalgo, L. G., Mueller, T. and Halloran, P. F. (2008) 'Comparing microarray versus RT-PCR assessment of renal allograft biopsies: Similar performance despite different dynamic ranges.' *American Journal of Transplantation*, 8(5) pp. 1006–1015.

Altirriba, J., Barbera, A., Del Zotto, H., Nadal, B., Piquer, S., Sánchez-Pla, A., Gagliardino, J. J. and Gomis, R. (2009) 'Molecular mechanisms of tungstate-induced pancreatic plasticity: A transcriptomics approach.' *BMC Genomics*, 10 p. 406.

Ammeux, N., Housden, B. E., Georgiadis, A., Hu, Y. and Perrimon, N. (2016) 'Mapping signaling pathway cross-talk in Drosophila cells.' *PNAS*, 113(35) pp. 9940–9945.

Amoyel, M., Anderson, A. M. and Bach, E. A. (2014) 'JAK/STAT pathway dysregulation in tumors: A Drosophila perspective.' *Seminars in Cell and Developmental Biology*. Elsevier Ltd, 28 pp. 96–103.

Antonny, B., Burd, C., De Camilli, P., Chen, E., Daumke, O., Faelber, K., Ford, M., Frolov, V. A., Frost, A., Hinshaw, J. E., Kirchhausen, T., Kozlov, M. M., Lenz, M., Low, H. H., McMahon, H., Merrifield, C., Pollard, T. D., Robinson, P. J., Roux, A. and Schmid, S. (2016) 'Membrane fission by dynamin: what we know and what we need to know.' *The EMBO Journal*, 35(21) pp. 2270–2284.

Apel, A. R., Hoban, K., Chuartzman, S., Tonikian, R., Sidhu, S., Schuldiner, M., Wendland, B. and Prosser, D. (2017) 'Syp1 regulates the clathrin-mediated and clathrin-independent endocytosis of multiple cargo proteins through a novel sorting motif.' *Molecular Biology of the Cell*, 28(18) pp. 2434–2448.

Aranda, P., LaJoie, D. and Jorcyk, C. (2012) 'Bleach Gel: A Simple Agarose Gel for Analyzing RNA Quality.' *Electrophoresis*, 33(2) pp. 366–369.

Arbouzova, N. I., Bach, E. A. and Zeidler, M. P. (2006) 'Ken & Barbie selectively regulates the expression of a subset of JAK/STAT pathway target genes.' *Current Biology*, 16(1) pp. 80–88.

Bach, E. A., Vincent, S., Zeidler, M. P. and Perrimon, N. (2003) 'A Sensitized Genetic Screen to Identify Novel Regulators and Components of the Drosophila Janus Kinase / Signal Transducer and Activator of Transcription Pathway.' *Genetics Society of America*, 1166(November) pp. 1149–1166.

Baeg, G., Zhou, R. and Perrimon, N. (2005) 'Genome-wide RNAi analysis of JAK / STAT signaling components in Drosophila.' *Genes & Development*, 19 pp. 1861–1870.

Balogh, P., Katz, S. and Kiss, A. L. (2013) 'The role of endocytic pathways in TGF-β signaling.' *Pathology and Oncology Research*, 19(2) pp. 141–148.

Banninger, G. and Reich, N. C. (2004) 'STAT2 nuclear trafficking.' *Journal of Biological*

Chemistry, 279(38) pp. 39199–39206.

Baron, M. (2012) 'Seminars in Cell & Developmental Biology Endocytic routes to Notch activation.' *Seminars in Cell and Developmental Biology*. Elsevier Ltd, 23(4) pp. 437–442.

Baskiewicz-Masiuk, M. and Machalinski, B. (2004) 'The role of the STAT5 proteins in the proliferation and apoptosis of the CML and AML cells.' *European Journal of Haematology*, 72 pp. 420–429.

Bassett, A. R., Tibbit, C., Ponting, C. P. and Liu, J.-L. (2014) 'Mutagenesis and homologous recombination in Drosophila cell lines using CRISPR/Cas9.' *Biology Open*, 3(1) pp. 42–49.

Bazan, J. F. (1990) 'Structural design and molecular evolution of a cytokine receptor superfamily.' *PNAS*, 87(September) pp. 6934–6938.

Becker, S., Groner, B. and Muller, C. W. (1998) 'Three-dimensional structure of the Stat3beta homodimer bound to DNA.' *Nature*, 394(6689) pp. 145–151.

Betz, A., Lampen, N., Martinek, S., Young, M. W. and Darnell, J. E. (2001) 'A Drosophila PIAS homologue negatively regulates stat92E.' *PNAS*, 98(17).

Bhuin, T. and Roy, J. K. (2014) 'Rab proteins: The key regulators of intracellular vesicle transport.' *Experimental Cell Research*, 8 pp. 1–19.

Bild, A. H., Turkson, J. and Jove, R. (2002) 'Cytoplasmic transport of Stat3 by receptor-mediated endocytosis.' *The EMBO journal*, 21(13) pp. 3255–63.

Bina, S. (2009) *Identification and analysis of JAK/STAT pathway target genes in Drosophila melanogaster (Unpublished Doctoral Thesis)*. Georg August University Göttingen.

Bina, S., Wright, V. M., Fisher, K. H., Milo, M. and Zeidler, M. P. (2010) 'Transcriptional targets of Drosophila JAK/STAT pathway signalling as effectors of haematopoietic tumour formation.' *EMBO reports*, 11(3) pp. 201–7.

Binari, R. and Perrimon, N. (1994) 'Stripe-specific regulation of pair-rule genes by hopscotch, a putative Jak family tyrosine kinase in Drosophila.' *Genes and Development*, 8(3) pp. 300–312.

Blaszczyk, K., Nowicka, H., Kostyrko, K., Antonczyk, A., Wesoly, J. and Bluyssen, H. A. R. (2016) 'The unique role of STAT2 in constitutive and IFN-induced transcription and antiviral responses.' *Cytokine and Growth Factor Reviews*. Elsevier Ltd, 29 pp. 71–81.

Blouin, C. M., Hamon, Y., Gonnord, P., Boularan, C., Kagan, J., Viaris de Lesegno, C., Ruez, R., Mailfert, S., Bertaux, N., Loew, D., Wunder, C., Johannes, L., Vogt, G., Contreras, F. X., Marguet, D., Casanova, J. L., Galès, C., He, H. T. and Lamaze, C. (2016) 'Glycosylation-Dependent IFN-γR Partitioning in Lipid and Actin Nanodomains Is Critical for JAK Activation.' *Cell*, 166(4) pp. 920–934.

Boccaccio, C., Ando, M., Tamagnone, L., Bardelli, A., Michieli, P., Battistini, C. and Comoglio, P. M. (1998) 'Induction of epithelial tubules by growth factor HGF depends on the STAT pathway.' *Nature Letters*, 391 pp. 285–288.

Böhmer, F.-D. and Friedrich, K. (2014) 'Protein tyrosine phosphatases as wardens of STAT signaling.' *Jak-Stat*, 3(1) p. e28087.

Bose, P. and Verstovsek, S. (2017) 'JAK2 inhibitors for myeloproliferative neoplasms: What is next?' *Blood*, 130(2) pp. 115–125.

Braunstein, J., Brutsaert, S., Olson, R. and Schindler, C. (2003) 'STATs Dimerize in the Absence of Phosphorylation.' *Journal of Biological Chemistry*, 278(36) pp. 34133–34140.

Bromberg, J. F., Horvath, C. M., Wen, Z., Schreibert, R. D. and Darnell, J. E. (1996) 'Transcriptionally active Statl is required for the antiproliferative effects of both interferon a and interferon y.' *PNAS*, 93 pp. 7673–7678.

Brown, S. (2003) 'Novel level of signalling control in the JAK/STAT pathway revealed by in situ visualisation of protein-protein interaction during Drosophila development.' *Development*, 130(14) pp. 3077–3084.

Brown, S., Hu, N. and Hombría, J. C. G. (2001) 'Identification of the first invertebrate interleukin JAK/STAT receptor, the Drosophila gene domeless.' *Current Biology*, 11(21) pp. 1700–1705.

Buckley, C. M. and King, J. S. (2017) 'Drinking problems: mechanisms of macropinosome formation and maturation.' *FEBS Journal*, 284(22) pp. 3778–3790.

Bulut, G. B., Sulahian, R., Ma, Y., Chi, N. W. and Huang, L. J. S. (2011) 'Ubiquitination regulates the internalization, endolysosomal sorting, and signaling of the erythropoietin receptor.' *Journal of Biological Chemistry*, 286(8) pp. 6449–6457.

Bulut, G. B., Sulahian, R., Yao, H. and Huang, L. J. S. (2013) 'Cbl ubiquitination of p85 is essential for Epo-induced EpoR endocytosis.' *Blood*, 122(24) pp. 3964–3972.

Chang, H. C., Zhang, S., Oldham, I., Naeger, L., Hoey, T. and Kaplan, M. H. (2003) 'STAT4 requires the N-terminal domain for efficient phosphorylation.' *Journal of Biological Chemistry*, 278(34) pp. 32471–32477.

Chanut-Delalande, H., Jung, A. C., Baer, M. M., Lin, L., Payre, F. and Affolter, M. (2010) 'The Hrs/Stam Complex Acts as a Positive and Negative Regulator of RTK Signaling during Drosophila Development.' *PLoS ONE*, 5(4).

Chatterjee-Kishore, M., Wright, K. L., Ting, J. P. Y. and Stark, G. R. (2000) 'How Stat1 mediates constitutive gene expression: A complex of unphosphorylated Stat1 and IRF1 supports transcription of the LMP2 gene.' *EMBO Journal*, 19(15) pp. 4111–4122.

Chen, H.-W., Chen, X., Oh, S.-W., Marinissen, M. J., Gutkind, J. S. and Hou, S. X. (2002) 'mom identifies a receptor for the Drosophila JAK/STAT signal transduction pathway and encodes a protein distantly related to the mammalian cytokine receptor family.' *Genes & development*, 16(3) pp. 388–98.

Chen, H., Chen, X., Oh, S., Marinissen, M. J., Gutkind, J. S. and Hou, S. X. (2002) 'mom identifies a receptor for the Drosophila JAK / STAT signal transduction pathway and encodes a protein distantly related to the mammalian cytokine receptor family.' *Genes & Development*, 16(3) pp. 388–398.

Chen, X., Vinkemeier, U., Zhao, Y., Jeruzalmi, D., Darnell, J. E. and Kuriyan, J. (1998) 'Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA.' *Cell*, 93(5) pp. 827–839.

Cherbas, L., Willingham, A., Zhang, D., Yang, L., Zou, Y., Eads, B. D., Carlson, J. W., Landolin, J. M., Kapranov, P., Dumais, J., Samsonova, A., Choi, J. H., Roberts, J., Davis, C. A., Tang, H., Van Baren, M. J., Ghosh, S., Dobin, A., Bell, K., Lin, W., Langton, L., Duff, M. O., Tenney, A. E., Zaleski, C., Brent, M. R., Hoskins, R. A., Kaufman, T. C., Andrews, J., Graveley, B. R., Perrimon, N., Celniker, S. E., Gingeras, T. R. and Cherbas, P. (2011) 'The transcriptional diversity of 25 Drosophila cell lines.' *Genome Research*, 21(2) pp. 301–314.

Chiba, T., Amanuma, H. and Todokoro, K. (1992) 'Tryptophan residue of Trp-Ser-X-Trp-Ser motif in extracellular domains of erythropoietin receptor is essential for signal transduction.' *Biochemical and Biophysical Research Communications*, 184(1) pp. 485–490.

Chmiest, D., Sharma, N., Zanin, N., Viaris de Lesegno, C., Shafaq-Zadah, M., Sibut, V., Dingli, F., Hupé, P., Wilmes, S., Piehler, J., Loew, D., Johannes, L., Schreiber, G. and Lamaze, C. (2016) 'Spatiotemporal control of interferon-induced JAK/STAT signalling and gene transcription by the retromer complex.' *Nature Communications*, 7 p. 13476.

Chung, J., Uchida, E., Grammer, T. C. and Blenis, J. (1997) 'STAT3 Serine Phosphorylation by ERK-Dependent and -Independent Pathways Negatively Modulates Its Tyrosine Phosphorylation,' 17(11) pp. 6508–6516.

Conner, S. D. and Schmid, S. L. (2003) 'Regulated portals of entry into the cell.' *Nature*, 422(6927) pp. 37–44.

Costa-Pereira, A. (2011) 'Dysregulation of janus kinases and signal transducers and activators of transcription in cancer.' *American Journal of Cancer Research*, 1(6) pp. 806–816.

Cui, X., Zhang, L., Luo, J., Rajasekaran, A., Hazra, S., Cacalano, N. and Dubinett, S. M. (2007) 'Unphosphorylated STAT6 contributes to constitutive cyclooxygenase-2 expression in human non-small cell lung cancer.' *Oncogene*, 26(29) pp. 4253–4260.

Daaka, Y., Luttrell, L. M., Ahn, S., Rocca, G. J. Della, Ferguson, S. S. G., Caron, M. G. and Lefkowitz, R. J. (1998) 'Essential Role for G Protein- coupled Receptor Endocytosis in the Activation of Mitogen- activated Protein Kinase *' pp. 685–689.

Dagil, R., Knudsen, M. J., Olsen, J. G., O'Shea, C., Franzmann, M., Goffin, V., Teilum, K., Breinholt, J. and Kragelund, B. B. (2012) 'The WSXWS motif in cytokine receptors is a molecular switch involved in receptor activation: Insight from structures of the prolactin receptor.' *Structure*, 20(2) pp. 270–282.

Dalma-Weiszhausz, D. D., Warrington, J., Tanimoto, E. Y. and Miyada, C. G. (2006) '[1] The Affymetrix GeneChip® Platform: An Overview.' *Methods in Enzymology*, 410(06) pp. 3–28.

Damke, H. (1996) 'Dynamin and receptor-mediated endocytosis.' *FEBS Letters*, 389(1) pp. 48–51.

Danson, C., Brown, E., Hemmings, O. J., Mcgough, I. J., Yarwood, S., Heesom, K. J., Carlton, J. G., Martin-serrano, J., May, M. T., Verkade, P. and Cullen, P. J. (2013) 'SNX15 links clathrin endocytosis to the PtdIns3 P early endosome independently of the APPL1

endosome.'

Darnell, J. E. (1997) 'STATs and gene regulation.' *Science*, 277(5332) pp. 1630–1635.

David, M., Wong, L., Flavell, R., Thompson, S. A., Wells, A., Larner, A. C. and Johnson, G. (1996) 'STAT Activation by Epidermal Growth Factor (EGF) and Amphiregulin.' *Journal of Biological Chemistry*, 271(16) pp. 9185–9188.

Decker, J. (1990) 'Epiderma Growth Factor and Transforming Growth Factor-alpha Induce Differential Processing of the Epidermal Growth Factor Receptor.' *Biochemical and biophysical research communications*, (30) pp. 615–621.

Devergne, O., Ghiglione, C. and Noselli, S. (2007) 'The endocytic control of JAK/STAT signalling in Drosophila.' *Journal of cell science*, 120(Pt 19) pp. 3457–3464.

Dickensheet, H. L., Venkataraman, C., Schindler, U. and Donnelly, R. P. (1999) 'Interferons inhibit activation of STAT6 by interleukin 4 in human monocytes by inducing SOCS-1 gene expression.' *PNAS*, 96 pp. 10800–10805.

Dittrich, E., Haft, C. R., Muys, L., Heinrich, P. C. and Graeve, L. (1996) 'A di-leucine motif and an upstream serine in the interleukin-6 (IL-6) signal transducer gp130 mediate ligand-induced endocytosis and down-regulation of the IL-6 receptor.' *Journal of Biological Chemistry*, 271(10) pp. 5487–5494.

Dittrich, E., Rose-John, S., Gerhartz, C., Müllberg, J., Stoyan, T., Yasukawa, K., Heinrich, P. C. and Graeve, L. (1994) 'Identification of a region within the cytoplasmic domain of the interleukin-6 (IL-6) signal transducer gp130 important for ligand-induced endocytosis of the IL-6 receptor.' *Journal of Biological Chemistry*, 269(29) pp. 19014–19020.

Donaldson, J. G., Johnson, D. L. and Dutta, D. (2016) 'Rab and Arf G proteins in endosomal trafficking and cell surface homeostasis.' *Small GTPases*. Taylor & Francis, 7(4) pp. 247–251.

Doray, B., Lee, I., Knisely, J., Bu, G., Kornfeld, S. and Louis, S. (2007) 'The γ/σ1 and α/σ2 Hemicomplexes of Clathrin Adaptors AP-1 and AP-2 Harbor the Dileucine Recognition Site.' *Molecular Biology of the Cell*, 18(May) pp. 1887–1896.

Dunn, J., Reid, G. and Bruening, M. (2010) 'Techniques for Phosphopeptide Enrichement Prior to Analsus by Mass Spectrometry.' *Mass Spectromety Reviews*, 29 pp. 29–54.

Dunn, K. W., Kamocka, M. M. and Mcdonald, J. H. (2011) 'A practical guide to evaluating colocalization in biological microscopy.' *American Journal of Physiology Cell Physiology* pp. 723–742.

Ebner, R. and Derynck, R. (1991) 'Epidermal growth factor and transforming growth factor-alpha: differential intracellular routing and processing of ligand-receptor complexes.' *Cell regulation*, 2(8) pp. 599–612.

Ekas, L. A., Cardozo, T. J., Flaherty, M. S., McMillan, E. A., Gonsalves, F. C. and Bach, E. A. (2010) 'Characterization of a dominant-active STAT that promotes tumorigenesis in Drosophila.' *Developmental Biology*, 344(2) pp. 621–636.

Feng, J., Witthuhn, B. a, Matsuda, T., Kohlhuber, F., Kerr, I. M. and Ihle, J. N. (1997) 'Activation of Jak2 catalytic activity requires phosphorylation of Y1007 in the kinase activation loop.'

Molecular and cellular biology, 17(5) pp. 2497–501.

Ferrao, R. and Lupardus, P. J. (2017) 'The Janus Kinase (JAK) FERM and SH2 domains: Bringing specificity to JAK-receptor interactions.' *Frontiers in Endocrinology*, 8 pp. 1–11.

Ferreira, A. P. A. and Boucrot, E. (2018) 'Mechanisms of Carrier Formation during Clathrin-Independent Endocytosis.' *Trends in Cell Biology*. Elsevier Ltd, 28(3) pp. 188–200.

Fisher, K. H., Fragiadaki, M., Pugazhendhi, D., Bausek, N., Brown, S. and Zeidler, M. (2018) 'A genome-wide RNAi screen identifies MASK as a positive regulator of cytokine receptor stability.' *Journal of Cell Science*, 131(13).

Fisher, K. H., Stec, W., Brown, S. and Zeidler, M. P. (2016) 'Mechanisms of JAK/STAT pathway negative regulation by the short coreceptor Eye Transformer/Latran.' *Molecular Biology of the Cell*, 27(3) pp. 434–441.

Flaherty, M. S., Salis, P., Evans, C. J., Ekas, L. A., Marouf, A., Zavadil, J., Banerjee, U. and Bach, E. A. (2010) 'Chinmo Is a Functional Effector of the JAK/STAT Pathway that Regulates Eye Development, Tumor Formation, and Stem Cell Self-Renewal in Drosophila.' *Developmental Cell*. Elsevier Ltd, 18(4) pp. 556–568.

Flaherty, M., Zavadil, J., Ekas, L. A. and Bach, E. A. (2009) 'Genome-wide expression profiling in the Drosophila eye reveals unexpected repression of Notch signaling by the JAK/STAT pathway.' *Developmental Dynamics*, 238(9) pp. 2235–2253.

Fossey, S. L., Bear, M. D., Kisseberth, W. C., Pennell, M. and London, C. A. (2011) 'Oncostatin M promotes STAT3 activation, VEGF production, and invasion in osteosarcoma cell lines.' *BMC Cancer*. BioMed Central Ltd, 11(1) p. 125.

Friedbichler, K. and Kerenyi, M. (2010) 'Stat5a serine 725 and 779 phosphorylation is a prerequisite for hematopoietic transformation.' *Blood*, 116(9) pp. 1548–1559.

Gäbler, K., Behrmann, I. and Haan, C. (2013) 'JAK2 mutants (e.g., JAK2V617F) and their importance as drug targets in myeloproliferative neoplasms.' *Jak-Stat*, 2(3) p. e25025.

Garuti, R., Jones, C., Li, W., Michaely, P., Herz, J., Gerard, R. D., Cohen, J. C. and Hobbs, H. H. (2005) 'The Modular Adaptor Protein Autosomal Recessive Hypercholesterolemia (ARH) Promotes Low Density Lipoprotein Receptor Clustering into Clathrin-coated Pits.' *The Journal of Biological Chemistry*, 280(49) pp. 40996–41004.

German, C. L., Sauer, B. M. and Howe, C. L. (2011) 'The STAT3 beacon: IL-6 recurrently activates STAT 3 from endosomal structures.' *Experimental Cell Research*. Elsevier Inc., 317(14) pp. 1955–1969.

Ghiglione, C. (2002) 'The Drosophila cytokine receptor Domeless controls border cell migration and epithelial polarization during oogenesis.' *Development*, 129(23) pp. 5437–5447.

Gilbert, M. M., Beam, C. K., Robinson, B. S. and Moberg, K. H. (2009) 'Genetic interactions between the Drosophila tumor suppressor gene ept and the stat92E transcription factor.' *PLoS ONE*, 4(9).

Gilbert, M. M., Weaver, B. K., Gergen, J. P. and Reich, N. C. (2005) 'A novel functional activator of the Drosophila JAK/STAT pathway, unpaired2, is revealed by an in vivo reporter of pathway activation.' *Mechanisms of Development*, 122(7–8) pp. 939–948.

Gillooly, D. J., Raiborg, C. and Stenmark, H. (2003) 'Phosphatidylinositol 3-phosphate is found in microdomains of early endosomes.' *Histochemistry and Cell Biology*, 120(6) pp. 445–453.

Gough, D. J., Koetz, L. and Levy, D. E. (2013) 'The MEK-ERK pathway is necessary for serine phosphorylation of mitochondrial STAT3 and ras-mediated transformation.' *PLoS ONE*, 8(11) pp. 1–9.

Grivennikov, S. and Karin, M. (2011) 'Dangerous liaisons: STAT3 and NF-κB collaboration and crosstalk in cancer.' *Cytokine*, 21(1) pp. 11–19.

Gronholm, J., Ungureanu, D., Vanhatupa, S., Ramet, M. and Silvennoinen, O. (2012) 'Sumoylation of Drosophila Transcription Factor STAT92E.' *Journal of innate immunity*, 2(6) pp. 618–624.

Grönholm, J., Vanhatupa, S., Ungureanu, D., Väliaho, J., Laitinen, T., Valjakka, J. and Silvennoinen, O. (2012) 'Structure-function analysis indicates that sumoylation modulates DNA-binding activity of STAT1.' *BMC Biochemistry*, 13(1) pp. 1–12.

Di Guglielmo, G. M., Baass, P. C., Ou, W. J., Posner, B. I. and Bergeron, J. J. (1994) 'Compartmentalization of SHC, GRB2 and mSOS, and hyperphosphorylation of Raf-1 by EGF but not insulin in liver parenchyma.' *The EMBO journal*, 13(18) pp. 4269–77.

Di Guglielmo, G. M., Le Roy, C., Goodfellow, A. F. and Wrana, J. L. (2003) 'Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover.' *Nature cell biology*, 5(5) pp. 410–21.

Hamdan, F. F., Rochdi, M. D., Breton, B., Fessart, D., Michaud, D. E., Charest, P. G., Laporte, S. A. and Bouvier, M. (2007) 'Unraveling G protein-coupled receptor endocytosis pathways using real-time monitoring of agonist-promoted interaction between β-arrestins and AP-2.' *Journal of Biological Chemistry*, 282(40) pp. 29089–29100.

Harrison, D. a, Mccoon, P. E., Binari, R., Gilman, M. and Perrimon, N. (1998) 'Drosophila unpaired encodes a secreted protein that activates the Drosophila unpaired encodes a secreted protein that activates the JAK signaling pathway.' *Genes and Development*, 12 pp. 3252–3263.

He, K., Yan, X., Li, N., Dang, S., Xu, L., Zhao, B., Li, Z., Lv, Z., Fang, X., Zhang, Y. and Chen, Y. G. (2015) 'Internalization of the TGF-β type i receptor into caveolin-1 and EEA1 double-positive early endosomes.' *Cell Research*, 25(6) pp. 738–752.

Hemalatha, A., Prabhakara, C. and Mayor, S. (2016) 'Endocytosis of Wingless via a dynamin-independent pathway is necessary for signaling in *Drosophila* wing discs.' *Proceedings of the National Academy of Sciences* p. 201610565.

Henriksen, L., Grandal, M. V., Knudsen, S. L. J., van Deurs, B. and Grøvdal, L. M. (2013) 'Internalization Mechanisms of the Epidermal Growth Factor Receptor after Activation with Different Ligands.' *PLoS ONE*, 8(3).

Henriksen, M. A., Betz, A., Fuccillo, M. V. and Darnell, J. E. (2002) 'Negative regulation of STAT92E by an N-terminally truncated STAT protein derived from an alternative promoter site.' *Genes and Development*, 16(18) pp. 2379–2389.

Heppler, L. N. and Frank, D. A. (2017) ‘Targeting Oncogenic Transcription Factors : Therapeutic Implications of Endogenous STAT Inhibitors.’ *Trends in Cancer*. Elsevier Inc., 3(12) pp. 816–827.

Hilton, D. J., Watowich, S. S., Katz, L. and Lodish, H. F. (1996) ‘Saturation Mutagenesis of the WS X WS Motif of the Erythropoietin Receptor.’ *The Journal of Biological Chemistry*, 271(9) pp. 4699–4708.

Hoeve, J., Ibarra-Sanchez, M., Wei Zhu, Y. F., Tremblay, M., David, M. and Shuai, K. (2002) ‘Identification of a nuclear Stat1 protein tyrosine phosphatase.’ *Molecular and Cellular Biology*, 22(16) pp. 5662–5668.

Hombría, J. C. and Brown, S. (2002) ‘The Fertile Field of Drosophila JAK / STAT Signalling.’ *Current Biology*, 12(02) pp. 569–575.

Hombría, J. C. G., Brown, S., Häder, S. and Zeidler, M. P. (2005) ‘Characterisation of Upd2, a Drosophila JAK/STAT pathway ligand.’ *Developmental Biology*, 288(2) pp. 420–433.

Hori, K., Sen, A., Kirchhausen, T. and Artavanis-Tsakonas, S. (2011) ‘Synergy between the ESCRT-III complex and Delex defines a ligand-independent Notch signal.’ *The Journal of Cell Biology*, 195(6) pp. 1005–1015.

Horn, T. and Boutros, M. (2010) ‘E-RNAi: A web application for the multi-species design of RNAi reagents-2010 update.’ *Nucleic Acids Research*, 38 pp. 332–339.

Horn, T., Sandmann, T. and Boutros, M. (2010) ‘Design and evaluation of genome-wide libraries for RNAi screens.’ *Genome biology*, 11(6) p. R61.

Hornbeck, P. V, Zhang, B., Murray, B., Kornhauser, J. M., Latham, V. and Psp, P. R. (2015) ‘PhosphoSitePlus, 2014 : mutations , PTMs and recalibrations.’ *Nucleic Acids Research*, 43 pp. 512–520.

Hou, S. X., Zheng, Z., Chen, X. and Perrimon, N. (2002) ‘The Jak/STAT pathway in model organisms: emerging roles in cell movement.’ *Developmental Cell*, 3 pp. 765–778.

Hou, X. S., Melnick, M. B. and Perrimon, N. (1996) ‘marelle acts downstream of the Drosophila HOP/JAK kinase and encodes a protein similar to the mammalian STATs.’ *Cell*, 84(3) pp. 411–419.

Houseley, J. and Tollervey, D. (2009) ‘The Many Pathways of RNA Degradation.’ *Cell*. Elsevier Inc., 136(4) pp. 763–776.

Hu, X., Dutta, P., Tsurumi, A., Li, J., Wang, J., Land, H. and Li, W. X. (2013) ‘Unphosphorylated STAT5A stabilizes heterochromatin and suppresses tumor growth.’ *PNAS*, 110(25) pp. 3–8.

Huang, G., Yan, H., Ye, S., Tong, C. and Ying, Q. L. (2014) ‘STAT3 phosphorylation at tyrosine 705 and serine 727 differentially regulates mouse esc fates.’ *Stem Cells*, 32 pp. 1149–1160.

Huang, H., Tsoi, S., Sun, Y. and Li, S. (1998) ‘Identification and characterisation of the smt3 cDNA and gene encoding ubiquitin-like protein from Drosophila Melanogaster.’ *Biochemistry and Molecular Biology International*, 46(4) pp. 775–785.

Huber, W., Heydebreck, A. Von and Vingron, M. (2003) ‘Analysis of microarray gene

expression data.' *In Handbook of Statistical Genetics*, pp. 1–37.

Huotari, J. and Helenius, A. (2011) 'Endosome maturation.' *The EMBO journal*. Nature Publishing Group, 30(17) pp. 3481–500.

Hurley, J. and Emr, S. (2006) 'NIH Public Access.' *Annual Review of Biophysics and Biomolecular Structure*, 35 pp. 277–298.

Hutagalung, A. H. and Novick, P. J. (2011) 'Role of Rab GTPases in Membrane Traffic and Cell Physiology.' *Physiological Reviews*, 91(1) pp. 119–149.

Ihle, J. N. (2001) 'The Stat family in cytokine signaling.' *Current Opinion in Cell Biology*, 13(2) pp. 211–217.

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U. and Speed, T. P. (2003) 'Exploration, normalization, and summaries of high density oligonucleotide array probe level data.' *In Biostatistics*, pp. 249–264.

Isaacs, A. and Lindenmann, J. (1957) 'Virus interference . I . The interferon.' *Proceedings of the Royal Society B Biological Sciences*, (147) pp. 258–267.

Jaksik, R., Iwanaszko, M., Rzeszowska-Wolny, J. and Kimmel, M. (2015) 'Microarray experiments and factors which affect their reliability.' *Biology Direct*. Biology Direct, 10(1) pp. 1–14.

Jékely, G., Sung, H. H., Luque, C. M. and Rørth, P. (2005) 'Regulators of endocytosis maintain localized receptor tyrosine kinase signaling in guided migration.' *Developmental Cell*, 9(2) pp. 197–207.

Jiang, X. and Sorkin, A. (2002) 'Coordinated Traffic of Grb2 and Ras during Epidermal Growth Factor Receptor Endocytosis Visualized in Living Cells.' *Molecular biology of the cell*, 13(6) pp. 2170–2179.

Jovic, M., Sharma, M., Rahajenng, J. and Caplan, S. (2010) 'The early endosome: a busy sorting station for proteins at the crossroads.' *Histology and Histopathology*, 25(1) pp. 99–112.

Kalaidzidis, I., Miaczynska, M., Brewi, M., Hupalowska, A., Ferguson, C., Parton, R. G., Kalaidzidis, Y. and Zerial, M. (2015) 'APPL endosomes are not obligatory endocytic intermediates but act as stable cargo-sorting compartments,' 211(1).

Kallio, J., Myllymaki, H., Gronholm, J., Armstrong, M., Vanha-aho, L.-M., Makinen, L., Silvennoinen, O., Valanne, S. and Ramet, M. (2010) 'Eye transformer is a negative regulator of Drosophila JAK/STAT signaling.' *The FASEB Journal*, 24(11) pp. 4467–4479.

Kamakura, S., Oishi, K., Yoshimatsu, T., Nakafuku, M., Masuyama, N. and Gotoh, Y. (2004) 'Hes binding to STAT3 mediates crosstalk between Notch and JAK-STAT signalling.' *Nature Cell Biology*, 6(6) pp. 547–554.

Kanaseki, T. and Kadota, K. (1969) 'The "Vesicle in a Basket" - A Morphological Study of the Coated Vesicle Isolated from the Never Endings of the Guinea Pig Brain, with Specail Reference to the Mechanism of Membrane Movements.' *The Journal of Cell Biology*, 42 pp. 202–220.

Karsten, P. (2007) *Mutational analysis of Drosophila STAT92E (Unpublished Doctoral Thesis)*.

Brunswick Institute of Technology.

Karsten, P., Häder, S. and Zeidler, M. P. (2002) 'Cloning and expression of Drosophila SOCS36E and its potential regulation by the JAK/STAT pathway.' *Mechanisms of Development*, 117(1–2) pp. 343–346.

Karsten, P., Plischke, I., Perrimon, N. and Zeidler, M. P. (2006) 'Mutational analysis reveals separable DNA binding and trans-activation of Drosophila STAT92E.' *Cellular Signalling*, 18(6) pp. 819–829.

Kelly, B. T., Graham, S. C., Liska, N., Dannhauser, P. N., Höning, S., Ungewickell, E. J. and Owen, D. J. (2014) 'AP2 controls clathrin polymerization with a membrane-activated switch.' *Science*, 345(6195) pp. 459–463.

Kelly, B. T., Mccoy, A. J., Spate, K., Miller, S. E., Evans, P. R. and Owen, D. J. (2008) 'A structural explanation for the binding of endocytic dileucine motifs by the AP2 complex.' *Nature Letters*, 456(18).

Kermorgant, S. and Parker, P. J. (2008) 'Receptor trafficking controls weak signal delivery: A strategy used by c-Met for STAT3 nuclear accumulation.' *Journal of Cell Biology*, 182(5) pp. 855–863.

Kile, B. T., Schulman, B. A., Alexander, W. S., Nicola, N. A., Martin, H. M. E., Hilton, D. J., Kile, B. T., Alexander, W. S., Nicola, N. A., Martin, H. M. E. and Hilton, D. J. (2002) 'The SOCS box : a tale of destruction and degradation.' *Trends in Biochemical Sciences*, 27(5) pp. 235–241.

Krämer, O. H. and Heinzel, T. (2010) 'Phosphorylation-acetylation switch in the regulation of STAT1 signaling.' *Molecular and Cellular Endocrinology*, 315(1–2) pp. 40–48.

Krebs, D. L. and Hilton, D. J. (2001) 'SOCS Proteins: Negative Regulators of Cytokine Signaling.' *Stem Cells*, 19(5) pp. 378–387.

Kumar, A., Commane, M., Flickinger, T. W., Horvath, C. M. and Stark, G. R. (1997) 'Defective TNF-α– Induced Apoptosis in STAT1-Null Cells Due to Low Constitutive Levels of Caspases.' *Science*, 278(November) pp. 1630–1633.

Kurgonaite, K., Gandhi, H., Kurth, T., Pautot, S., Schwille, P., Weidemann, T. and Bokel, C. (2015) 'Essential role of endocytosis for interleukin-4-receptor-mediated JAK/STAT signalling.' *Journal of Cell Science*, 128(20) pp. 3781–3795.

Kwon, E. J., Park, H. S., Kim, Y. S., Oh, E. J., Nishida, Y., Matsukage, A., Yoo, M. A. and Yamaguchi, M. (2000) 'Transcriptional regulation of the Drosophila raf proto-oncogene by drosophila STAT during development and in immune response.' *Journal of Biological Chemistry*, 275(26) pp. 19824–19830.

Lauvrak, S. U., Torgersen, M. L. and Sandvig, K. (2004) 'Efficient endosome-to-Golgi transport of Shiga toxin is dependent on dynamin and clathrin.' *Journal of Cell Science*, 117(11) pp. 2321–2331.

Leung, K. F., Dacks, J. B. and Field, M. C. (2008) 'Evolution of the multivesicular body ESCRT machinery; retention across the eukaryotic lineage.' *Traffic*, 9(10) pp. 1698–1716.

Lim, C. P. and Cao, X. (2006) 'Structure, function, and regulation of STAT proteins.' *Molecuar*

Biosystems, (2) pp. 536–550.

Lin, J., Li, P., Liu, D., Jin, H. T., He, J., Ata, M., Rasheed, U., Rochman, Y., Wang, L., Cui, K., Liu, C., Kelsall, B. L., Ahmed, R. and Leonard, W. J. (2012) 'Article Critical Role of STAT5 Transcription Factor Tetramerization for Cytokine Responses and Normal Immune Function.' *Immunity*, 36 pp. 585–599.

Liu, L., McBride, K. M. and Reich, N. C. (2005) 'STAT3 nuclear import is independent of tyrosine phosphorylation and mediated by importin-α3.' *PNAS*, 102(23) pp. 8150–8155.

Liu, X., Gao, Z., Zhang, L. and Rattray, M. (2013) 'puma 3 . 0 : improved uncertainty propagation methods for gene and transcript expression analysis.' *BMC Bioinformatics*, 14(29).

Liu, X., Milo, M., Lawrence, N. D. and Rattray, M. (2005) 'A tractable probabilistic model for Affymetrix probe-level analysis across multiple chips.' *Bioinformatics*, 21(18) pp. 3637–3644.

Lloyd, T. E., Atkinson, R., Wu, M. N., Zhou, Y., Pennetta, G. and Bellen, H. J. (2002) 'Hrs Regulates Endosome Membrane Invagination and Tyrosine Kinase Receptor Signaling in Drosophila.' *Cell*, 108(2) pp. 261–269.

Lohi, O. and Lehto, V. P. (2001) 'STAM/EAST/Hbp adapter proteins - Integrators of signalling pathways.' *FEBS Letters*, 508(3) pp. 287–290.

Luo, H., Rose, P., Roberts, T. and Dearolf, C. (2002) 'The Hopscotch Jak kinase requires the Raf pathway to promote blood cell activation and differentiation in Drosophila.' *Molecular Genetics and Genomics*, 267(1) pp. 57–63.

Maib, H., Smythe, E. and Ayscough, K. (2017) 'Forty years on: clathrin-coated pits continue to fascinate.' *Molecular Biology of the Cell*, 28(7) pp. 843–847.

Makki, R., Meister, M., Pennetier, D., Ubeda, J. M., Braun, A., Daburon, V., Krzemień, J., Bourbon, H. M., Zhou, R., Vincent, A. and Crozatier, M. (2010) 'A short receptor downregulates JAK/STAT signalling to control the Drosophila cellular immune response.' *PLoS Biology*, 8(8) pp. 33–34.

Mao, X., Ren, Z., Parker, G. N., Sondermann, H., Pastorello, M. A., Wang, W., McMurray, J. S., Demeler, B., Darnell, J. E. and Chen, X. (2005) 'Structural bases of unphosphorylated STAT1 association and receptor binding.' *Molecular Cell*, 17(6) pp. 761–771.

Marchetti, M., Monier, M.-N., Fradagrada, A., Mitchell, K., Baychelier, F., Eid, P., Johannes, L. and Lamaze, C. (2006) 'Stat-mediated signaling induced by type I and type II interferons (IFNs) is differentially controlled through lipid microdomain association and clathrin-dependent endocytosis of IFN receptors.' *Molecular biology of the cell*, 17(7) pp. 2896–909.

Maritzen, T., Koo, S. J. and Haucke, V. (2012) 'Turning CALM into excitement: AP180 and CALM in endocytosis and disease.' *Biology of the Cell*, 104(10) pp. 588–602.

Martínez, A., Varadé, J., Márquez, A., Cénit, M. C., Espino, L., Perdigones, N., Santiago, J. L., Frenández-Arquero, M., Calle, H., Arroyo, R., Mendoza, J. L., Frenández-Gutiérrez, B., Concha, E. G. and Urcelay, E. (2008) 'Association of the STAT4 Gene With Increased Susceptibility for Some Immune-Mediated Diseases.' *Arthritis & Rheumatism*, 58(9) pp.

2598–2602.

Maurer, M. E. and Cooper, J. A. (2006) 'The adaptor protein Dab2 sorts LDL receptors into coated pits independently of AP-2 and ARH.' *Journal of Cell Science*, 119(20) pp. 4235–4246.

Mayor, S., Parton, R. G., Donaldson, J. G., Cossart, P., Helenius, A., Settembre, C., Ballabio, A., Maxfield, F. R., Wideman, J. G., Leung, K. F., Field, M. C., Bökel, C., Brand, M., Paolo, P., Fiore, D., Zastrow, M. Von, Kirchhausen, T., Owen, D., Stephen, C. and Weigert, R. (2014) 'Clathrin-Independent Pathways of Endocytosis Clathrin-Independent Pathways of Endocytosis.' *Cold Spring Harbor Perspectives in Biology* pp. 1–20.

McBride, K. M., Banninger, G., McDonald, C. and Reich, N. C. (2002) 'Regulated nuclear import of the STAT1 transcription factor by direct binding of importin-alpha.' *EMBO Journal*, 21(7) pp. 1754–1763.

Mclornan, D., Percy, M. and McMullin, M. F. (2006) 'JAK2 V617F: A single mutation in the myeloproliferative group of disorders.' *Ulster Medical Journal*, 75(2) pp. 112–119.

McMahon, H. T. and Boucrot, E. (2011) 'Molecular mechanism and physiological functions of clathrin-mediated endocytosis.' *Nature reviews. Molecular cell biology*. Nature Publishing Group, 12(8) pp. 517–33.

Meissner, T., Krause, E., Lödige, I. and Vinkemeier, U. (2004) 'Arginine Methylation of STAT1.' *Cell*, 119(5) pp. 587–589.

Mettlen, M., Chen, P., Srinivasan, S., Danuser, G. and Schmid, S. L. (2018) 'Regulation of Clathrin-Mediated Endocytosis.' *Annual Review of Biochemistry*, 87 pp. 871–896.

Meyer, T. and Vinkemeier, U. (2004) 'Nucleocytoplasmic shuttling of STAT transcription factors.' *European Journal of Biochemistry*, 271(23–24) pp. 4606–4612.

Miaczynska, M., Christoforidis, S., Giner, A., Shevchenko, A., Uttenweiler-Joseph, S., Habermann, B., Wilm, M., Parton, R. G. and Zerial, M. (2004) 'APPL proteins link Rab5 to nuclear signal transduction via an endosomal compartment.' *Cell*, 116(3) pp. 445–456.

Mishra, S. K., Watkins, S. C. and Traub, L. M. (2002) 'The autosomal recessive hypercholesterolemia (ARH) protein interfaces directly with the clathrin-coat machinery.' *Proceedings of the National Academy of Sciences*, 99(25) pp. 16099–16104.

Mitra, A., Ross, J. A., Rodriguez, G., Nagy, Z. S., Wilson, H. L. and Kirken, R. A. (2012) 'Signal transducer and activator of transcription 5b (Stat5b) serine 193 is a novel cytokine-induced phospho-regulatory site that is constitutively activated in primary hematopoietic malignancies.' *Journal of Biological Chemistry*, 287(20) pp. 16596–16608.

Miura, S. and Mishina, Y. (2011) 'Hepatocyte growth factor-regulated tyrosine kinase substrate (Hgs) is involved in BMP signaling through phosphorylation of SMADs and TAK1 in early mouse embryo Shigeto.' *Developmental Dynamics*, 240(11) pp. 389–399.

Mizuno, E., Kawahata, K., Kato, M., Kitamura, N. and Komada, M. (2003) 'STAM Proteins Bind Ubiquitinated Proteins on the Early Endosome via the VHS Domain and Ubiquitin-interacting Motif.' *Molecular Biology of the Cell*, 14 pp. 3675–3689.

Mohr, A., Chatain, N., Domoszlai, T., Rinis, N., Sommerauer, M., Vogt, M. and Müller-Newen,

G. (2012) ‘Dynamics and non-canonical aspects of JAK/STAT signalling.’ *European Journal of Cell Biology*. Elsevier, 91(6–7) pp. 524–532.

Monahan, A. J. and Starz-gaiano, M. (2015) ‘Mechanisms of Development Socs36E limits STAT signaling via Cullin2 and a SOCS-box independent mechanism in the Drosophila egg chamber.’ *Mechanisms of Development*. Elsevier B.V., 138 pp. 313–327.

Moore, C. A. C., Milano, S. K. and Benovic, J. L. (2007) ‘Regulation of Receptor Trafficking by GRKs and Arrestins.’ *Annual Review of Physiology*, 69 pp. 451–482.

Moore, R., Giralt-Pujol, M., Zhu, Z. and Smythe, E. (2018) ‘Interplay of endocytosis and growth factor receptor signalling.’ *In Endocytosis and Signaling*, pp. 181–202.

Mowen, K. A., Tang, J., Zhu, W., Schurter, B. T., Shuai, K., Herschman, H. R. and David, M. (2001) ‘Arginine methylation of STAT1 modulates IFNα/β-induced transcription.’ *Cell*, 104(5) pp. 731–741.

Müller, P., Kuttenkeuler, D., Gesellchen, V., Zeidler, M. P. and Boutros, M. (2005) ‘Identification of JAK/STAT signalling components by genome-wide RNA interference.’ *Nature*, 436(7052) pp. 871–875.

Murakami, M., Hibi, M., Nakagawa, N., Toshimasa, N., Kiyoshi, Y., Koichi, Y., Tetsuya, T. and Tadamitsu, K. (1993) ‘IL-6-Induced Homodimerization of gp 130 and Associated Activation of a Tyrosine Kinase.’ *Science*, 260(5115) pp. 1808–1810.

Murphy, J. E., Vohra, R. S., Dunn, S., Holloway, Z. G., Monaco, A. P., Homer-Vanniasinkam, S., Walker, J. H. and Ponnambalam, S. (2008) ‘Oxidised LDL internalisation by the LOX-1 scavenger receptor is dependent on a novel cytoplasmic motif and is regulated by dynamin-2.’ *Journal of Cell Science*, 121(13) pp. 2136–2147.

Naef, F., Lim, D. A., Patil, N. and Magnasco, M. (2002) ‘DNA hybridization to mismatched templates: A chip study.’ *The American Physical Society*, 65.

Nassoy, P. and Lamaze, C. (2012) ‘Stressing caveolae new role in cell mechanics.’ *Trends in Cell Biology*. Elsevier Ltd, 22(7) pp. 381–389.

Nazarewicz, R. R., Salazar, G., Patrushev, N., San Martin, A., Hilenski, L., Xiong, S. and Alexander, R. W. (2011) ‘Early Endosomal Antigen 1 (EEA1) is an obligate Scaffold for angiotensin II-induced, PKC-??-dependent Akt activation in endosomes.’ *Journal of Biological Chemistry*, 286(4) pp. 2886–2895.

Neculai, D., Neculai, A. M., Verrier, S., Straub, K., Klumpp, K., Pfitzner, E. and Becker, S. (2005) ‘Structure of the unphosphorylated STAT5a dimer.’ *Journal of Biological Chemistry*, 280(49) pp. 40782–40787.

Van Nguyen, T., Angkasekwinai, P., Dou, H., Lin, F. M., Lu, L. S., Cheng, J., Chin, Y. E., Dong, C. and Yeh, E. T. H. (2012) ‘SUMO-Specific Protease 1 Is Critical for Early Lymphoid Development through Regulation of STAT5 Activation.’ *Molecular Cell*. Elsevier, 45(2) pp. 210–221.

Nie, Y., Erion, D. M., Yuan, Z., Dietrich, M., Gerald, I., Horvath, T. L. and Gao, Q. (2009) ‘STAT3 inhibition of gluconeogenesis is downregulated by SirT1.’ *Nature Cell Biology*, 11(4) pp. 492–500.

Nielsen, E., Severin, F., Backer, J. M., Hyman, A. A. and Zerial, M. (1999) 'Rab5 regulates motility of early endosomes on microtubules.' *Nature Cell Biology*, 1 pp. 376–382.

Norris, A., Tammineni, P., Wang, S., Gerdes, J., Murr, A., Kwan, K. Y., Cai, Q. and Grant, B. D. (2017) 'SNX-1 and RME-8 oppose the assembly of HGRS-1/ESCRT-0 degradative microdomains on endosomes.' *Proceedings of the National Academy of Sciences*, 114(3) pp. E307–E316.

O'Shea, J. J., Schwartz, D. M., Villarino, A. V., Gadina, M., McInnes, I. B. and Laurence, A. (2015) 'The JAK-STAT Pathway: Impact on Human Disease and Therapeutic Intervention.' *Annual Review of Medicine*, 66(1) pp. 311–328.

Ohno, H., Stewart, J., Fournier, M., Bosshart, H., Miyatake, S., Saito, T., Gallusser, A., Kirchhausen, T., Bonifacino, S., Ohno, H., Stewart, J., Fournier, M., Bosshart, H., Rhee, I., Miyatake, S., Saito, T., Gallusser, A., Kirchhausen, T. and Bonifacino, J. S. (1995) 'Interaction of Tyrosine-Based Sorting Signals with Clathrin-Associated Proteins.' *Science*, 269(5232) pp. 1872–1875.

Olayioye, M. A., Beuvink, I., Horsch, K., Daly, J. M. and Hynes, N. E. (1999) 'ErbB receptor-induced activation of Stat transcription factors is mediated by Src tyrosine kinases.' *Journal of Biological Chemistry*, 274(24) pp. 17209–17218.

Ota, N., Brett, T. J., Murphy, T. L., Fremont, D. H. and Murphy, K. M. (2004) 'N-domain-dependent nonphosphorylated STAT4 dimers required for cytokine-driven activation.' *Nature Immunology*, 5(2) pp. 208–215.

Pálfy, M., Reményi, A. and Korcsmáros, T. (2012) 'Endosomal crosstalk: Meeting points for signaling pathways.' *Trends in Cell Biology*, 22(9) pp. 447–456.

Pandey, A., Fernandez, M. M., Steen, H., Blagoev, B., Nielsen, M. M., Roche, S., Mann, M. and Lodish, H. F. (2000) 'Identification of a novel immunoreceptor tyrosine-based activation motif-containing molecule, STAM2, by mass spectrometry and its involvement in growth factor and cytokine receptor signaling pathways.' *Journal of Biological Chemistry*, 275(49) pp. 38633–38639.

Pandey, K. N. (2009) 'Functional roles of short sequence motifs in the endocytosis of membrane receptors.' *Frontiers in Bioscience*, 14(1) p. 5339.

Pandey, K. N. (2010) 'Small Peptide Recognition Sequence for Intracellular Sorting.' *Current Opinion in Biotechnology*, 21(5) pp. 611–620.

Pearse, B. M. (1975) 'Coated vesicles from pig brain: purification and biochemical characterization.' *Journal of molecular biology*, 97(1) pp. 93–8.

Pearson, K. (1895) 'Mathematical Contributions to the Theory of Evolution - III. Regression, Heredity, and Panmixia.' *Philosophical Transactions of the Royal Society of London*.

Pearson, R., Liu, M., Rattray, M., Milo, M., Lawrence, N., Sanguinetti, G. and Zhang, L. (2009) 'puma: a Bioconductor package for Propogating Uncertainty in Microarray Analysis.' *BMC Bioinformatics*, 10.

Pitcher, C., Höning, S., Fingerhut, A., Bowers, K. and Marsh, M. (1999) 'Cluster of differentiation antigen 4 (CD4) endocytosis and adaptor complex binding require activation of the CD4

endocytosis signal by serine phosphorylation.' *Molecular biology of the cell*, 10(3) pp. 677–91.

Polo, S., Sigismund, S., Faretta, M., Guidi, M., Capua, M. R., Bossi, G., Chen, H., De Camilli, P. and Di Fiore, P. P. (2002) 'A single motif responsible for ubiquitin recognition and monoubiquitination in endocytic proteins.' *Nature Letters*, 416 pp. 451–455.

Putz, E. M., Gotthardt, D., Hoermann, G., Csiszar, A., Wirth, S., Berger, A., Straka, E., Rigler, D., Wallner, B., Jamieson, A. M., Pickl, W. F., Zebedin-Brandl, E. M., Müller, M., Decker, T. and Sexl, V. (2013) 'CDK8-mediated STAT1-S727 phosphorylation restrains NK cell cytotoxicity and tumor surveillance.' *Cell Reports*, 4(3) pp. 437–444.

Raiborg, C., Bache, K. G., Gillooly, D. J., Madshus, I. H., Stang, E. and Stenmark, H. (2002) 'Hrs sorts ubiquitinated proteins into clathrin-coated microdomains of early endosomes.' *Nature Cell Biology*, 4(5) pp. 394–398.

Raiborg, C., Bache, K. G., Mehlum, A., Stang, E. and Stenmark, H. (2001) 'Hrs recruits clathrin to early endosomes,' 20(17).

Raiborg, C., Malerød, L., Pedersen, N. M. and Stenmark, H. (2008) 'Differential functions of Hrs and ESCRT proteins in endocytic membrane trafficking.' *Experimental Cell Research*, 314(4) pp. 801–813.

Raiborg, C., Weshe, J., Malerød, L. and Stenmark, H. (2006) 'Flat clathrin coats on endosomes mediate degradative protein sorting by scaffolding Hrs in dynamic microdomains.' *Journal of Cell Science*, 119(12) pp. 2414–2424.

Rajan, A. and Perrimon, N. (2012) 'Drosophila cytokine unpaired 2 regulates physiological homeostasis by remotely controlling insulin secretion.' *Cell*. Elsevier Inc., 151(1) pp. 123–137.

Rawlings, J. S., Rennebeck, G., Harrison, S. M. W., Xi, R. and Harrison, D. A. (2004) 'Two Drosophila suppressors of cytokine signaling (SOCS) differentially regulate JAK and EGFR pathway activities.' *BMC Cell Biology*, 5 pp. 1–15.

Rawlings, J. S., Rosler, K. M. and Harrison, D. a (2004) 'The JAK/STAT signaling pathway.' *Journal of cell science*, 117 pp. 1281–1283.

Reich, N. C. (2013) 'STATs get their move on.' *Jak-Stat*, 2(4) p. e27080.

Reiner, C. and Nathanson, N. M. (2009) 'The internalization of the M2 and M4 muscarinic acetylcholine receptors involves distinct subsets of small G-proteins,' 82 pp. 718–727.

Ren, W., Zhang, Y., Li, M., Wu, L., Wang, G., Baeg, G.-H., You, J., Li, Z. and Lin, X. (2015) 'Windpipe Controls Drosophila Intestinal Homeostasis by Regulating JAK/STAT Pathway via Promoting Receptor Endocytosis and Lysosomal Degradation.' *PLOS Genetics*, 11 p. e1005180.

Riedel, F., Gillingham, A. K., Rosa-Ferreira, C., Galindo, A. and Munro, S. (2016) 'An antibody toolkit for the study of membrane traffic in *Drosophila melanogaster*.' *Biology Open*, 5(7) pp. 987–992.

Rink, J., Ghigo, E., Kalaidzidis, Y. and Zerial, M. (2005) 'Rab Conversion as a Mechanism of Progression from Early to Late Endosomes.' *Cell*, 122 pp. 735–749.

Rivas, M. L., Cobreros, L., Zeidler, M. P. and Hombría, J. C. G. (2008) 'Plasticity of Drosophila Stat DNA binding shows an evolutionary basis for Stat transcription factor preferences.' *EMBO Reports*, 9(11) pp. 1114–1120.

Rizk, A., Paul, G., Incardona, P., Bugarski, M., Mansouri, M., Niemann, A., Ziegler, U., Berger, P. and Sbalzarini, I. F. (2014) 'Segmentation and quantification of subcellular structures in fluorescence microscopy images using Squassh.' *Nature Protocols*, 9(3) pp. 586–596.

Robinson, M. S. (2015) 'Forty Years of Clathrin-coated Vesicles.' *Traffic*, 16(12) pp. 1210–1238.

Roepstorff, K., Grandal, M. V., Henriksen, L., Jeppe, S. L., Lerdrup, M., Grøvdal, L., Willumsen, B. M. and Deurs, B. Van (2009) 'Differential Effects of EGFR Ligands on Endocytic Sorting of the Receptor.' *Traffic*, 10 pp. 1115–1127.

Rogers, R. S., Horvath, C. M. and Matunis, M. J. (2003) 'SUMO modification of STAT1 and its role in PIAS-mediated inhibition of gene activation.' *Journal of Biological Chemistry*, 278(32) pp. 30091–30097.

Royle, S. J. (2012) 'The role of clathrin in mitotic spindle organisation.' *Journal of Cell Science*, 125(1) pp. 19–28.

Ruan, H., Zhang, Z., Wang, S., Nickels, L. M. and Tian, L. (2017) 'Tumor Necrosis Factor Receptor- Associated Factor 6 (TRAF6) Mediates Ubiquitination-Dependent STAT3 Activation upon Salmonella enterica Serovar Typhimurium Infection.' *Infection and Immunity*, 85(8) pp. 1–13.

Runyan, C. E., Schnaper, H. W. and Poncelet, A. C. (2005) 'The role of internalization in transforming growth factor β1-induced Smad2 association with Smad anchor for receptor activation (SARA) and Smad2-dependent signaling in human mesangial cells.' *Journal of Biological Chemistry*, 280(9) pp. 8300–8308.

Scheerlinck, E., Dhaenens, M., Van Soom, A., Peelman, L., De Sutter, P., Van Steendam, K. and Deforce, D. (2015) 'Minimizing technical variation during sample preparation prior to label-free quantitative mass spectrometry.' *Analytical Biochemistry*. Elsevier Inc, 490 pp. 14–19.

Schenck, A., Goto-Silva, L., Collinet, C., Rhinn, M., Giner, A., Habermann, B., Brand, M. and Zerial, M. (2008) 'The Endosomal Protein Appl1 Mediates Akt Substrate Specificity and Cell Survival in Vertebrate Development.' *Cell*, 133(3) pp. 486–497.

Schmid, S. L., Sorkin, A. and Zerial, M. (2014) 'Endocytosis : Past, Present, and Future.' *Cold Spring Harb Perspect Biol*, 3 pp. 1–9.

Schmidt, O. and Teis, D. (2012) 'The ESCRT machinery.' *Current Biology*. Elsevier, 22(4) pp. R116–R120.

Schneider-Brachert, W., Tchikov, V., Neumeyer, J., Jakob, M., Winoto-Morbach, S., Held-Feindt, J., Heinrich, M., Merkel, O., Ehrenschwender, M., Adam, D., Mentlein, R., Kabelitz, D. and Schütze, S. (2004) 'Compartmentalization of TNF receptor 1 signaling: Internalized TNF receptosomes as death signaling vesicles.' *Immunity*, 21(3) pp. 415–428.

Schöneberg, J., Lee, I., Iwasa, J. H. and Hurley, J. H. (2016) 'Reverse-topology membrane

scission by the ESCRT proteins.' *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 18(1) pp. 5–17.

Seibel, N. M., Eljouni, J., Nalaskowski, M. M. and Hampe, W. (2007) 'Nuclear localization of enhanced green fluorescent protein homomultimers.' *Analytical Biochemistry*, 368(1) pp. 95–99.

Semerdjieva, S., Shortt, B., Maxwell, E., Singh, S., Fonarev, P., Hansen, J., Schiavo, G., Grant, B. D. and Smythe, E. (2008) 'Coordinated regulation of AP2 uncoating from clathrin-coated vesicles by rab5 and hRME-6.' *Journal of Cell Biology*, 183(3) pp. 499–511.

Shah, M., Patel, K., Mukhopadhyay, S., Xu, F., Guo, G. and Sehgal, P. B. (2006) 'Membrane-associated STAT3 and PY-STAT3 in the cytoplasm.' *Journal of Biological Chemistry*, 281(11) pp. 7302–7308.

Shen, C. H. and Stavnezer, J. (1998) 'Interaction of stat6 and NF-kappaB: direct association and synergistic activation of interleukin-4-induced transcription.' *Molecular and cellular biology*, 18(6) pp. 3395–404.

Shi, S., Larson, K., Guo, D., Lim, S. J., Dutta, P., Yan, S. J. and Li, W. X. (2008) 'Drosophila STAT is required for directly maintaining HP1 localization and heterochromatin stability.' *Nature Cell Biology*, 10(4) pp. 489–496.

Shimizu, H., Woodcock, S. a., Wilkin, M. B., Trubenová, B., Monk, N. a M. and Baron, M. (2014) 'Compensatory flux changes within an endocytic trafficking network maintain thermal robustness of notch signaling.' *Cell*, 157(5) pp. 1160–1174.

Shuai, K. (2006) 'Regulation of cytokine signaling pathways by PIAS proteins.' *Cell Research*, 16(2) pp. 196–202.

Shuai, K. and Liu, B. (2003) 'Regulation of JAK-STAT signalling in the immune system.' *Nature Reviews Immunology*, 3(11) pp. 900–911.

Sigismund, S., Algisi, V., Nappo, G., Conte, A., Pascolutti, R., Cuomo, A., Bonaldi, T., Argenzio, E., Verhoef, L. G. G. C., Maspero, E., Bianchi, F., Capuani, F., Ciliberto, A., Polo, S. and Di Fiore, P. P. (2013) 'Threshold-controlled ubiquitination of the EGFR directs receptor fate.' *EMBO Journal*. Nature Publishing Group, 32(15) pp. 2140–2157.

Sigismund, S., Woelk, T., Puri, C., Maspero, E., Tacchetti, C., Transidico, P., Di Fiore, P. P. and Polo, S. (2005) 'Clathrin-independent endocytosis of ubiquitinated cargos.' *PNAS*, 102(8) pp. 2760–2765.

Silver, D. L., Geisbrecht, E. R. and Montell, D. J. (2005) 'Requirement for JAK / STAT signaling throughout border cell migration in Drosophila.' *Development*, (132) pp. 3483–3492.

Siupka, P., Hamming, O. T., Kang, L., Gad, H. H. and Hartmann, R. (2015) 'A conserved sugar bridge connected to the WSXWS motif has an important role for transport of IL-21R to the plasma membrane.' *Genes and Immunity*. Nature Publishing Group, 16(6) pp. 405–413.

Skvortsov, D., Abdueva, D., Curtis, C., Schaub, B. and Tavaré, S. (2007) 'Explaining differences in saturation levels for Affymetrix GeneChip® arrays.' *Nucleic Acids Research*, 35(12) pp. 4154–4163.

Slessareva, J. E., Routt, S. M., Temple, B., Bankaitis, V. A. and Dohlman, H. G. (2006)

‘Activation of the Phosphatidylinositol 3-Kinase Vps34 by a G Protein α Subunit at the Endosome.’ *Cell*, 126(1) pp. 191–203.

Smith, C. L., Debouck, C., Rosenberg, M. and Culp, J. S. (1989) ‘Phosphorylation of serine residue 89 of human adenovirus E1A proteins is responsible for their characteristic electrophoretic mobility shifts, and its mutation affects biological function.’ *Journal of Virology*, 63(4) pp. 1569–1577.

Soldaini, E., John, S., Moro, S., Bollenbacher, J., Schindler, U. and Leonard, W. J. (2000) ‘DNA Binding Site Selection of Dimeric and Tetrameric Stat5 Proteins Reveals a Large Repertoire of Divergent Tetrameric Stat5a Binding Sites.’ *Molecular and Cellular Biology*, 20(1) pp. 389–401.

Sorkin, A. and Zastrow, M. Von (2009) ‘Endocytosis and signalling: intertwining molecular networks.’ *Nature Reviews Molecular Cell Biology*, 10(9) pp. 609–622.

Stahl, N., Farruggella, T., Boulton, T., Zhong, Z., Darnell, J. and Yancopoulos, G. (1994) ‘Choice of STATs and Other Substrates Specified by Modular Tyrosine-based Motifs in Cytokine Receptors.’ *Science*, 264(5155) pp. 95–98.

Stark, G. R. and Darnell, J. E. (2012) ‘The JAK-STAT Pathway at Twenty.’ *Immunity*. Elsevier Inc., 36(4) pp. 503–514.

Stark, G. R., Kerr, I. M., Williams, B. R., Silverman, R. H. and Schreiber, R. D. (1998) ‘How cells respond to interferons.’ *Annual review of biochemistry*, 67 pp. 227–64.

Stec, W. J. and Zeidler, M. P. (2011) ‘Drosophila SOCS Proteins.’ *Journal of signal transduction*, 2011, January, p. 894510.

Stec, W., Vidal, O. and Zeidler, M. P. (2013) ‘Drosophila SOCS36E negatively regulates JAK/STAT pathway signaling via two separable mechanisms.’ *Molecular Biology of the Cell*, 24(18) pp. 3000–3009.

Stenmark, H., Parton, R. G., Steele-mortimer, O. and Lutcke, A. (1994) ‘Inhibition of rab5 GTPase activity stimulates membrane fusion.’ *EMBO Journal*, 13(6) pp. 1287–1296.

Strehlow, I. and Schindler, C. (1998) ‘Amino-terminal Signal Transducer and Activator of Transcription (STAT) Domains Regulate Nuclear Translocation and STAT Deactivation.’ *The Journal of Biological Chemistry*, 273(43) pp. 28049–28056.

Sun, W., Yan, Q., Vida, T. A. and Bean, A. J. (2003) ‘Hrs regulates early endosome fusion by inhibiting formation of an endosomal SNARE complex.’ *The Journal of Cell Biology* pp. 125–137.

Taga, T., Hibi, M., Hirata, Y., Yamasaki, K., Yasukawa, K., Matsuda, T., Hirano, T. and Kishimoto, T. (1989) ‘Interleukin-6 triggers the association of its receptor with a possible signal transducer, gp130.’ *Cell*, 58(3) pp. 573–581.

Takenouchi, O., Yosh, H. and Ozawa, T. (2018) ‘Unique Roles of β -Arrestin in GPCR Trafficking Revealed by Photoinducible Dimerizers.’ *Scientific Reports*, 8.

Tamiya, T., Kashiwagi, I., Takahashi, R., Yasukawa, H. and Yoshimura, A. (2011) ‘Suppressors of cytokine signaling (SOCS) proteins and JAK/STAT pathways: Regulation of T-cell inflammation by SOCS1 and SOCS3.’ *Arteriosclerosis, Thrombosis, and Vascular*

Biology, 31(5) pp. 980–985.

Tanaka, T., Soriano, M. A. and Grusby, M. J. (2005) 'SLIM is a nuclear ubiquitin E3 ligase that negatively regulates STAT signaling.' *Immunity*, 22(6) pp. 729–736.

Teis, D., Wunderlich, W. and Huber, L. A. (2002) 'Localization of the MP1-MAPK scaffold complex to endosomes is mediated by p14 and required for signal transduction.' *Developmental Cell*, 3(6) pp. 803–814.

Thiel, S., Dahmen, H., Martens, A., Müller-Newen, G., Schaper, F., Heinrich, P. C. and Graeve, L. (1998) 'Constitutive internalization and association with adaptor protein-2 of the interleukin-6 signal transducer gp130.' *FEBS Letters*, 441(2) pp. 231–234.

Thomas, S. J., Snowden, J. A., Zeidler, M. P. and Danson, S. J. (2015) 'The role of JAK/STAT signalling in the pathogenesis, prognosis and treatment of solid tumours.' *British Journal of Cancer*. Nature Publishing Group, 113(3) pp. 365–371.

Tian, X., Kang, D. S. and Benovic, J. L. (2014) 'β-arrestins and G Protein-Coupled Receptor Trafficking.' *Handbook of Experimental Phamacology*, 219 pp. 173–186.

Timofeeva, O. A., Chasovskikh, S., Lonskaya, I., Tarasova, N. I., Khavrutskii, L., Tarasov, S. G., Zhang, X., Korostyshevskiy, V. R., Cheema, A., Zhang, L., Dakshanamurthy, S., Brown, M. L. and Dritschilo, A. (2012) 'Mechanisms of Unphosphorylated STAT3 Transcription Factor Binding to DNA.' *The Journal of Biological Chemistry*, 287(17) pp. 14192–14200.

Tognon, E., Wollscheid, N., Cortese, K., Tacchetti, C. and Vaccari, T. (2014) 'ESCRT-0 is not required for ectopic notch activation and tumor suppression in Drosophila.' *PLoS ONE*, 9(4) pp. 1–11.

Townsend, P. A., Scarabelli, T. M., Davidson, S. M., Knight, R. A., Latchman, D. S. and Stephanou, A. (2004) 'STAT-1 Interacts with p53 to Enhance DNA Damage-induced Apoptosis.' *The Journal of Biological Chemistry*, 279(7) pp. 5811–5820.

Traub, L. M. and Bonifacino, J. S. (2013) 'Cargo recognition in clathrin-mediated endocytosis.' *Cold Spring Harbor Perspectives in Biology*, 5(11) pp. 1–24.

Trengove, M. C. and Ward, A. C. (2013) 'SOCS proteins in development and disease.' *American journal of clinical and experimental immunology*, 2(1) pp. 1–29.

Turriziani, B., Garcia-Munoz, A., Pilkington, R., Raso, C., Kolch, W. and von Kriegsheim, A. (2014) 'On-Beads Digestion in Conjunction with Data-Dependent Mass Spectrometry: A Shortcut to Quantitative and Dynamic Interaction Proteomics.' *Biology*, 3(2) pp. 320–332.

Uddin, S., Sassano, A., Deb, D. K., Verma, A., Majchrzak, B., Rahman, A., Malik, A. B., Fish, E. N. and Platanias, L. C. (2002) 'Protein kinase C-δ (PKC-δ) is activated by type I interferons and mediates phosphorylation of Stat1 on serine 727.' *Journal of Biological Chemistry*, 277(17) pp. 14408–14416.

Ungureanu, D., Vanhatupa, S., Grönholm, J., Palvimo, J. J. and Silvennoinen, O. (2005) 'SUMO-1 conjugation selectively modulates STAT1-mediated gene responses.' *Blood*, 106(1) pp. 224–226.

Usacheva, A., Sandoval, R., Domanski, P., Kotenko, S. V., Nelms, K., Goldsmith, M. A. and

Colamonici, O. R. (2002) 'Contribution of the Box 1 and Box 2 motifs of cytokine receptors to Jak1 association and activation.' *Journal of Biological Chemistry*, 277(50) pp. 48220–48226.

Vidal, O. M., Stec, W., Bausek, N., Smythe, E. and Zeidler, M. P. (2010) 'Negative regulation of Drosophila JAK-STAT signalling by endocytic trafficking.' *Journal of cell science*, 123(Pt 20) pp. 3457–66.

Vieira, a V, Lamaze, C. and Schmid, S. L. (1996) 'Control of EGF receptor signaling by clathrin-mediated endocytosis.' *Science (New York, N.Y.)*, 274(5295) pp. 2086–9.

Vinkemeier, U., Moarefi, I., Jr, J. E. D. and Kuriyan, J. (1998) 'Structure of the Amino-Terminal Protein Interaction Domain of STAT-4.' *American Association for the Advancement of Science*, 279(5353) pp. 1048–1052.

Vogt, M., Domoszlai, T., Kleshchanok, D., Lehmann, S., Schmitt, A., Poli, V., Richtering, W. and Muller-Newen, G. (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.

Wan, M., Zhang, W., Tian, Y., Xu, C., Xu, T., Liu, J. and Zhang, R. (2015) 'Unraveling a molecular determinant for clathrin-independent internalization of the M2 muscarinic acetylcholine receptor.' *Scientific Reports*. Nature Publishing Group, 5(June) pp. 1–16.

Wandinger-Ness, A. and Zerial, M. (2014) 'Rab Proteins and the Compatmentalization of the Endosomal System.' *Cold Spring Harbor Perspectives in Biology*, 6(11) p. a022616.

Wang, D., Li, Z., Messing, E. M. and Wu, G. (2005) 'The SPRY Domain-containing SOCS Box Protein 1 (SSB-1) Interacts with MET and Enhances the Hepatocyte Growth Factor-induced Erk-Elk-1-Serum Response Element Pathway.' *The Journal of Biological Chemistry*, 280(16) pp. 16393–16401.

Wang, R., Cherukuri, P. and Luo, J. (2005) 'Activation of Stat3 sequence-specific DNA binding and transcription by p300/CREB-binding protein-mediated acetylation.' *Journal of Biological Chemistry*, 280(12) pp. 11528–11534.

Wang, X., Lupardus, P., LaPorte, S. L. and Garcia, K. C. (2009) 'Structural Biology of Shared Cytokine Receptors.' *Annual review of Immunology*, 27 pp. 29–60.

Wang, Y., Miao, Z., Pommier, Y., Kawasaki, E. S. and Player, A. (2007) 'Gene expression Characterization of mismatch and high-signal intensity probes associated with Affymetrix genechips.' *Bioinformatics*, 23(16) pp. 2088–2095.

Wang, Y., Pennock, S., Chen, X., Wang, Z., Al, W. E. T. and Iol, M. O. L. C. E. L. L. B. (2002) 'Endosomal Signaling of Epidermal Growth Factor Receptor Stimulates Signal Transduction Pathways Leading to Cell Survival.' *Molecular and Cellular Biology*, 22(20) pp. 7279–7290.

Wang, Y. Z., Wharton, W., Garcia, R., Kraker, a, Jove, R. and Pledger, W. J. (2000) 'Activation of Stat3 preassembled with platelet-derived growth factor beta receptors requires Src kinase activity.' *Oncogene*, 19(17) pp. 2075–85.

Wei, J., Yuan, Y., Jin, C., Chen, H., Leng, L., He, F. and Wang, J. (2012) 'The Ubiquitin Ligase

TRAF6 Negatively Regulates the JAK-STAT Signaling Pathway by Binding to STAT3 and Mediating Its Ubiquitination.' *PLoS ONE*, 7(11).

Welsch, S., Habermann, A., Jäger, S., Müller, B., Krijnse-Locker, J. and Kräusslich, H. G. (2006) 'Ultrastructural analysis of ESCRT proteins suggests a role for endosome-associated tubular-vesicular membranes in ESCRT function.' *Traffic*, 7(11) pp. 1551–1566.

Wen, Z., Zhong, Z. and Darnell, J. E. (1995) 'Maximal activation of transcription by statl and stat3 requires both tyrosine and serine phosphorylation.' *Cell*, 82(2) pp. 241–250.

Wenta, N., Strauss, H., Meyer, S. and Vinkemeier, U. (2008) 'Tyrosine phosphorylation regulates the partitioning of STAT1 between different dimer conformations.' *Proceedings of the National Academy of Sciences*, 105(27) pp. 9238–9243.

Wenzel, E. M., Schultz, S. W., Schink, K. O., Pedersen, N. M., Nähse, V., Carlson, A., Brech, A., Stenmark, H. and Raiborg, C. (2018) 'Concerted ESCRT and clathrin recruitment waves define the timing and morphology of intraluminal vesicle formation.' *Nature Communications*. Springer US, 9(1) p. 2932.

Wilkin, M. B., Carbery, A., Fostier, M., Aslam, H., Mazaleyrat, S. L., Higgs, J., Myat, A., Evans, D. A. P., Cornell, M. and Baron, M. (2004) 'Regulation of Notch Endosomal Sorting and Signaling by Drosophila Nedd4 Family Proteins.' *Current Biology*, 14 pp. 2237–2244.

Wilkin, M., Tongngok, P., Gensch, N., Clemence, S., Motoki, M., Yamada, K., Hori, K., Taniguchi-kanai, M., Franklin, E., Matsuno, K. and Baron, M. (2008) 'Article Drosophila HOPS and AP-3 Complex Genes Are Required for a Deltex-Regulated Activation of Notch in the Endosomal Trafficking Pathway.' *Developmental Cell*. Elsevier Ltd, 15(5) pp. 762–772.

Winnebeck, E. C., Millar, C. D. and Warman, G. R. (2010) 'Why Does Insect RNA Look Degraded?' *Journal of Insect Science*, 10(159) pp. 1–7.

Wright, V. M., Vogt, K. L., Smythe, E. and Zeidler, M. P. (2011) 'Differential activities of the Drosophila JAK/STAT pathway ligands Upd, Upd2 and Upd3.' *Cellular Signalling*. Elsevier Inc., 23(5) pp. 920–927.

Xu, X., Sun, Y. L. and Hoey, T. (1996) 'Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain.' *Science (New York, N.Y.)*, 273(5276) pp. 794–797.

Yamaoka, K., Saharinen, P., Pesu, M., Holt III, V. E., Silvennoinen, O. and O'Shea, J. J. (2004) 'The Janus kinases (Jaks).' *Genome Biology*, 5(12) p. 253.

Yamashita, H., Nevalainen, M. T., Xu, J., LeBaron, M. J., Wagner, K. U., Erwin, R. a, Harmon, J. M., Hennighausen, L., Kirken, R. a and Rui, H. (2001) 'Role of serine phosphorylation of Stat5a in prolactin-stimulated beta-casein gene expression.' *Molecular and cellular endocrinology*, 183(1–2) pp. 151–63.

Yan, R., Small, S., Desplan, C., Dearolf, C. R. and Darnell, J. E. (1996) 'Identification of a Stat gene that functions in Drosophila development.' *Cell*, 84(3) pp. 421–430.

Yan, S.-J., Lim, S. J., Shi, S., Dutta, P. and Li, W. X. (2011) 'Unphosphorylated STAT and

heterochromatin protect genome stability.' *The FASEB Journal*, 25(1) pp. 232–241.

Yang, H., Kronhamn, J., Ekstro m, J.-O., Korkut, G. G. and Hultmark, D. (2015) 'JAK/STAT signaling in Drosophila muscles controls the cellular immune response against parasitoid infection.' *EMBO reports*, 16(12) pp. 1664–1672.

Yang, J., Huang, J., Dasgupta, M., Sears, N., Miyagi, M., Wang, B., Chance, M. R., Chen, X., Du, Y., Wang, Y., An, L., Wang, Q., Lu, T., Zhang, X., Wang, Z. and Stark, G. R. (2010) 'Reversible methylation of promoter-bound STAT3 by histone-modifying enzymes.' *PNAS*, 107(50) pp. 21499–21504.

Yang, J., Liao, X., Agarwal, M. K., Barnes, L., Auron, P. E. and Stark, G. R. (2007) 'Unphosphorylated STAT3 accumulates in response to IL-6 and activates transcription by binding to NFκB.' *Genes and Development*, 21(11) pp. 1396–1408.

Yang, J. and Stark, G. R. (2008) 'Roles of unphosphorylated STATs in signaling.' *Cell Research*, 18(4) pp. 443–451.

Yoshida, Y., Kumar, A., Koyama, Y., Peng, H., Arman, A., Boch, J. A. and Auron, P. E. (2004) 'Interleukin 1 Activates STAT3/Nuclear Factor-κB Cross-talk via a Unique TRAF6- and p65-dependent Mechanism.' *Journal of Biological Chemistry*, 279(3) pp. 1768–1776.

Zhang, J. J., Vinkemeier, U., Gu, W., Chakravarti, D., Horvath, C. M. and Darnell, J. E. (1996) 'Two contact regions between Stat1 and CBP/p300 in interferon gamma signaling.' *PNAS*, 93(26) pp. 15092–15096.

Zhang, T., Kee, W. H., Seow, K. T., Fung, W. and Cao, X. (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.

Zhang, Y., Cho, Y. Y., Petersen, B. L., Zhu, F. and Dong, Z. (2004) 'Evidence of STAT1 phosphorylation modulated by MAPKs, MEK1 and MSK1.' *Carcinogenesis*, 25(7) pp. 1165–1175.

Zhao, S., Fung-Leung, W. P., Bittner, A., Ngo, K. and Liu, X. (2014) 'Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells.' *PLoS ONE*, 9(1).

Zhong, M., Henriksen, M. A., Takeuchi, K., Schaefer, O., Liu, B., Hoeve, J. t., Ren, Z., Mao, X., Chen, X., Shuai, K. and Darnell, J. E. (2005) 'Implications of an antiparallel dimeric structure of nonphosphorylated STAT1 for the activation-inactivation cycle.' *Proceedings of the National Academy of Sciences*, 102(11) pp. 3966–3971.

Zhou, J., Liu, H., Zhou, S., He, P. and Liu, X. (2016) 'Adaptor protein APPL1 interacts with EGFR to orchestrate EGF-stimulated signaling.' *Science Bulletin*. Elsevier B.V. and Science China Press, 61(19) pp. 1504–1512.

Zhu, W., Mustelin, T. and David, M. (2002) 'Arginine methylation of STAT1 regulates its dephosphorylation by T cell protein tyrosine phosphatase.' *Journal of Biological Chemistry*, 277(39) pp. 35787–35790.

Zhuang, S. (2013) 'Regulation of STAT Signaling by Acetylation.' *Cellular Signalling*, 25(9) pp. 1924–1931.