Characterisation of a novel interaction between Influenza A Virus protein M2 and human PP4R3α/SMEK1 protein

A thesis submitted for the degree of Master of Philosophy

By

Yasmin Nicole Hansrod

Division of Biological Sciences, Department of Life Sciences, College of Health and Life Sciences, Brunel University London

2017

Registration Number - 1101325

Table of Contents

Declaration

I declare that I carried out this body of work under the supervision of Dr. Beatrice Nal-Rogier. Other contributions have been indicated, acknowledged and referenced clearly.

Acknowledgements

I would like to thank my supervisory team, especially Dr Beatrice Nal-Rogier and Dr Sabrina Tosi for their academic and emotional support in the production of this thesis. My family, notably my Father, Cassim Hansrod and my Grandparents, Beverley and Barrington Jervis for their love and care throughout my studies, and my Aunts and Uncles for keeping me sane. My colleagues at Brunel University London for their friendship and collegiality: Stephen Hare, Chrissi Tomiwa Oluwadare and Haroon Hussain, as well as the rest of the staff in the Biosciences Department at Brunel University London.

Abstract

Discovering new viral-host cell interactions is fundamental to the discovery of novel methods to target viral infections. Due to the parasitic nature in which virus's replicate, it is important to identify the exact mechanisms by which they interact and hijack host cell protein machinery. A novel interaction between a human cellular protein and a viral proton channel protein was identified through a yeast two-hybrid screen. This interaction is between viral protein M2 and a protein phosphatase regulatory subunit. Protein phosphatases remove phosphate groups from proteins; this usually results in deactivation of proteins. Activation and deactivation of proteins by phosphorylation is fundamental for signalling cascades. The regulatory subunit PP4R3α (SMEK1) regulates protein phosphatase 4 (PP4), which is a complex formed with other subunits (PP4R2 and PP4c). PP4 has many functions within human cells. PP4 has been identified to be involved in the type-1 interferon signalling cascade, which plays a crucial role in viral recognition and immune response stimulation. The hypothesis was that this interaction has important implications for the host and the virus. The interaction between SMEK1 and M2 was confirmed by co-immunoprecipitation. It was also found that influenza A virus (IAV) infection caused an increase in PP4c expression in A549 lung epithelial cells. By understanding the way in which IAV M2 interacts with host cell systems and hijacks these systems we may identify novel strategies to combat or prevent infection.

Chapter 1: Introduction

1.1 Influenza A Virus

1.1.1 Epidemiology

The Influenza A virus (IAV) presents a real threat to world health. Viruses of H1N1 and H3N2 subtypes cause seasonal epidemics in humans. Symptoms are usually mild but can be severe in populations at risk (elderly, young children and pregnant women). 200,000 to 500,000 people die every year from these respiratory infections (WHO, 2015). The threat is due to the emergence of highly pathogenic strains from the animal reservoir as well as the ability of influenza viruses to mutate and recombine.

Adaption of IAV to new hosts depends on the sialic acid (SA) receptors on the surface of host cells that haemagglutinin (HA) binds to (Zeng *et al.*, 2013; Couceiro *et al.*, 1993; Baum and Paulson, 1990). There are two types of SA receptors, human influenza strains prefer α2,6-linked SA, however avian influenza prefers α2,3-linked SA. Tropism and infectivity of influenza depends on these two receptors. α2,6-linked SA are found on epithelial cells including ciliated cells in the upper respiratory tract, whereas α2,3-linked SA are primarily found on non-ciliated bronchial cells and alveolar type 2 cells in the lower respiratory tract (Zeng *et al.*, 2013; Couceiro *et al.*, 1993; Baum and Paulson, 1990). This explains why seasonal human influenza primarily infects the upper respiratory tract, whereas pandemic highly pathogenic strains that have recombined from avian influenza are found to infect the lower respiratory tract.

1.1.2 Structure/Genome

IAV are enveloped, negative strand RNA Orthomyxiviridae viruses, which are composed of a host-derived membrane. The membrane contains 3 viral proteins: haemagglutinin (HA), neuraminidase (NA) and small amounts of matrix proton channel protein (M2). HA and NA are used to classify different strains of IAV (Muller *et al.*, 2012; Nayak *et al.*, 2009). The viral genome consists of eight segments of negative sense RNA, which are replicated in the host nucleus alongside host DNA (Figure 1) (Roberts *et al.*, 2013; Nayak *et al.*, 2009; Chen *et al.*, 2008a).

The genome encodes for up to 12 proteins: the surface proteins HA and NA, matrix protein 1 (M1), matrix protein 2 proton channel (M2), nucleoprotein (NP), non-structural protein 1 (NS1), nuclear export protein (NEP), polymerase acid protein (PA), polymerase basic proteins 1 and 2 (PB1 and PB2), NS1 and PB1-F2 which are only found during viral replicaition (Muller *et al.*, 2012). PB1, PB2 and PA form RNA-dependent RNA polymerase (RdRp), in combination with NP and vRNA form viral ribonucleoprotein (vRNP) (Muller *et al.*, 2012). The structure of IAV, including the 8 viral RNA segments is shown in Figure 1.
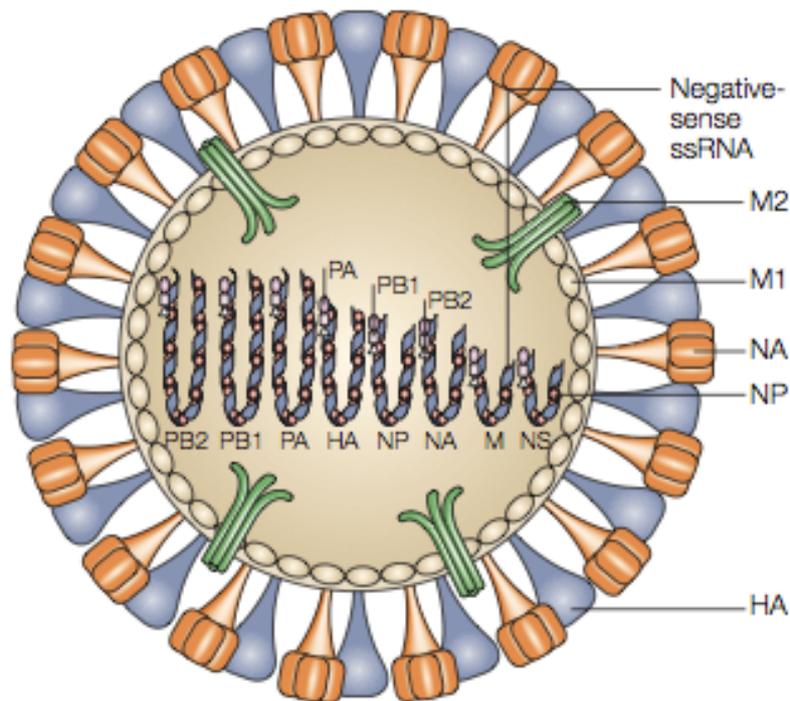
Figure 1. Schematic diagram of IAV structure. Shown are membrane proteins HA and NA, M2 spanning the viral membrane, M1 on the interior of the membrane. PA, PB1 and PB2 associated with the 8 labelled segments of viral RNA (Horimoto and Kawaoka, 2005).

1.1.3 Life cycle

The viral lifecycle depends on utilisation of host cell protein machinery to reproduce and generate progeny virus. The influenza lifecycle consists of 5 main events: entry into the host cell; vRNP entry into the host nucleus; viral genome transcription and replication; vRNP export from the nucleus; virion assembly; and budding at the host cell plasma membrane (Figure 2) (Samji, 2009).

1.1.3.1 Entry into host cell

IAV entry into cells is dependent on virion shape (circular or filamentous).This is via clathrin-mediated endocytosis or macropinocytosis; viral HA proteins bind to terminal sialic acid moieties on receptors on the host cell membrane (Muller *et al.*, 2012; de Vries *et al.*, 2011; Eierhoff *et al.*, 2010). Viral entry is determined by the host processing proteases of HA, which activates membrane fusion activity (Kido *et al.*, 2008). HA-sialic acid binding stimulates reorganisation of the plasma membrane, which leads to lipid raft clustering, that then acts as a signalling platform for viral internalisation (Muller *et al.*, 2012; Eierhoff *et al.*, 2010). The epidermal growth factor receptor (EGFR) cascade is also stimulated by HA-sialic acid binding; this promotes endocytosis of viral particles (Muller *et al.*, 2012). Eierhoff *et al., (2010)* presented data that suggests that when IAV binds to sialic acid, viral uptake is promoted by IAV clustering; this allows activation of EGFR and of other receptor tyrosine kinases. EGFR stimulation alone is not sufficient to induce viral uptake. Evidence suggests that this process also requires the activation of cell signals such as phosphatidylinositol-3 kinase (PI3K), protein kinase C (PKC) and extracellular signal-regulated kinase (ERK) (Eierhoff *et al.*, 2010). Activation of PKC occurs at the point at which HA binds to the cell membrane; PKC isoform βII then sorts the virus into late endosomes (Eierhoff *et al.*, 2010).

8

Spherical viral particles are sorted into clathrin-coated pits forming at attachment sites, for viral endocytosis (de Vries *et al.*, 2011). It has been shown that there are alternative pathways for viral entrance to the host cell through non-coated pits by micropinocytosis (Muller *et al.*, 2012; de Vries *et al.*, 2011). It was suggested that filamentous virions use macropinocytosis as, due to their larger size, they are unable to fit into the clathrin coated pits that spherical virions induce through the binding of tyrosine kinases receptors (Rossman and Lamb, 2011).

Seasonal IAV targets upper respiratory tract epithelial cells for infection; highly pathogenic strains target the lower respiratory tract due to differential tropism of viral strains (Zeng *et al.*, 2013; Couceiro *et al.*, 1993; Baum and Paulson, 1990). These cells are coated with a thick layer of mucous, which consists of small particles and glycoproteins (de Vries *et al.*, 2011). Environment can also affect the mode of entry into host cells used by IAV; it is most likely that virion shape is the most important deciding factor (de Vries *et al.*, 2011).

1.1.3.2 Acidification in endocytosis and entry into host nucleus,

Once inside the cell the virus is exposed to low pH (5.5-5.0), which changes the conformation of viral HA. This allows endosomal membrane binding for release of the capsid into the cytoplasm (Banerjee *et al.*, 2013). M2 possess a pH-activated channel that can selectively gate protons into the virion during the uncoating process (Roberts *et al.*, 2013; Sun *et al.*, 2010). The M2 protein promotes viral interior acidification that facilitates viral disassembly (Wharton *et al.*, 1994).

Acidification of the viral interior causes M1 and vRNP separate from each other. vRNP localises to and is imported into the nucleus whilst M1 is released into the cytosol (Banerjee *et al.*, 2013; Perez *et al.*, 2010). Host cellular factor RanBP5 has been found to interact with the vRNP protein PB1 or PA-PB1 heterodimer to aid PA-PB1 dimer enter the host nucleus to form PA-PB1-PB2 complex (Li *et al.*, 2011).

1.1.3.3 Transcription and replication

The viral ribonucleoprotein (vRNP) complex is responsible for the transcription and replication of the viral genome (Perez *et al.*, 2010). This complex consists of the 8 RNA segments of the viral genome as well as RNA-dependent RNA polymerase, which consist of 3 subunits (PA, PB1 and PB2). Each of the subunits has very important roles within the polymerase complex (Perez *et al.*, 2010). The promoter is one of the non-coding regions of the viral RNA that forms a corkscrew-like structure recognised by RdRp (Li *et al.*, 2011; Perez *et al.*, 2010). Viral polymerase uses host RNA oligonucleotides as primers for viral RNA transcription by seizing host mRNA, cleaving around 10-13 bases downstream of the 5' cap with PA for use by PB1 as a viral primer (Li *et al.*, 2011; Perez *et al.*, 2010). RdRp associates with the non-coding region of the gene, initiating transcription at the first 3' cytosine (Perez *et al.*, 2010). Transcription termination occurs at a sequence of 5-7 uridines 15-17 nucleotides upstream from the 5' end of the vRNA templates (Li *et al.*, 2011).

Replication requires the synthesis of a positive sense RNA copy of the genome as a template to produce progeny vRNA (Li *et al.*, 2011). Hsp90 interacts with PB2/ PB1-PB2 dimer to excite vRNA polymerase activity and promote entry into the nucleus to form the viral polymerase complex (Li *et al.*, 2011).

1.1.3.4 Export from the nucleus

IAV RNA associates with nucleotide proteins and RNA polymerase subunits to form the viral ribonucleocapsid (vRNP); this exits via nuclear pores and is transported to the plasma membrane on host microtubules (Roberts *et al.*, 2013). vRNA could piggyback onto Rab11 positive vesicles for transport to the plasma membrane, as monesin treatment during infection inhibits vRNA movement from the ER (Amorim *et al.*, 2011). It is hypothesised that Microtubules are used for long-distance transport of RNPs; a degree of functional redundancy would explain minor effect of cytoskeletal positions on viral titre (Amorim *et al.*, 2011). Cytoplasmic trafficking of IAV genome involves Rab11. It is hypothesised that there is a direct interaction between the RNA polymerase complex protein PB2 and Rab11, although there could be Rab11 related proteins involved that bridge the interaction (Amorim *et al.*, 2011; Rossman and Lamb, 2011). M1, Rab11 and Rab11-interacting proteins are also required for formation of filamentous virions (Roberts *et al.*, 2013).

1.1.3.5 Assembly and budding

At the plasma membrane, the vRNP associates with matrix proteins M1, along with M2, HA and NA where it forms the budding virion (Roberts *et al.*, 2013). M1 forms a matrix protein layer that links the outer membrane to the inner helical vRNP (Rossman and Lamb, 2011; Nayak *et al.*, 2009). M1 is recruited to the cytoplasmic tails of HA and NA underneath the plasma membrane. Here it associates with the lipid raft and recruits M2 to the plasma membrane to sites of viral budding (Gorai *et al.*, 2012; Rossman and Lamb, 2011). M2 modifies M1, which enables it to polymerise and create the inner structure of the budding virion (Gorai *et al.*, 2012; Rossman and Lamb, 2011). M1 is recruited to lipid raft domains by HA cytoplasmic domains and NA residue (Chen *et al.*, 2008a). M1 recruits M2 to the periphery of lipid rafts where it is mostly excluded from the budding virion (Roberts *et al.*, 2013). This stabilises the budding site in the initial stages to allow polymerisation of the matrix protein for filamentous virion formation by elongation of the matrix protein (Gorai *et al.*, 2012; Rossman and Lamb, 2011).

The M1 binding region on M2 overlaps with its supposed cholesterol-binding domain. This could explain why M2 is excluded from the cholesterol rich lipid-rafts that M1 is recuited to. M2 is therefore recruited to the budding zone via binding to M1, causing exclusion from budding zone. M2 also contains an effector domain for effective virion formation and proper morphology (McCown and Pekosz, 2006). M2 mediated packaging of budding virions is facilitated by binding to the M1 proteins in the vRNP complex, then M1 interacting with other M2 cytoplasmic tails for efficient genome packaging (Chen *et al.*, 2008a; McCown and Pekosz, 2006).

M2 is only found to be recruited around the edges of the HA and NA rich lipid raft domains due to M1 cholesterol binding domain overlap with M2 binding domain (Chen *et al.*, 2008a). It is possible that IAV may use its genome to drive forward the interactions required for virion formation at budding sites (McCown and Pekosz, 2006).
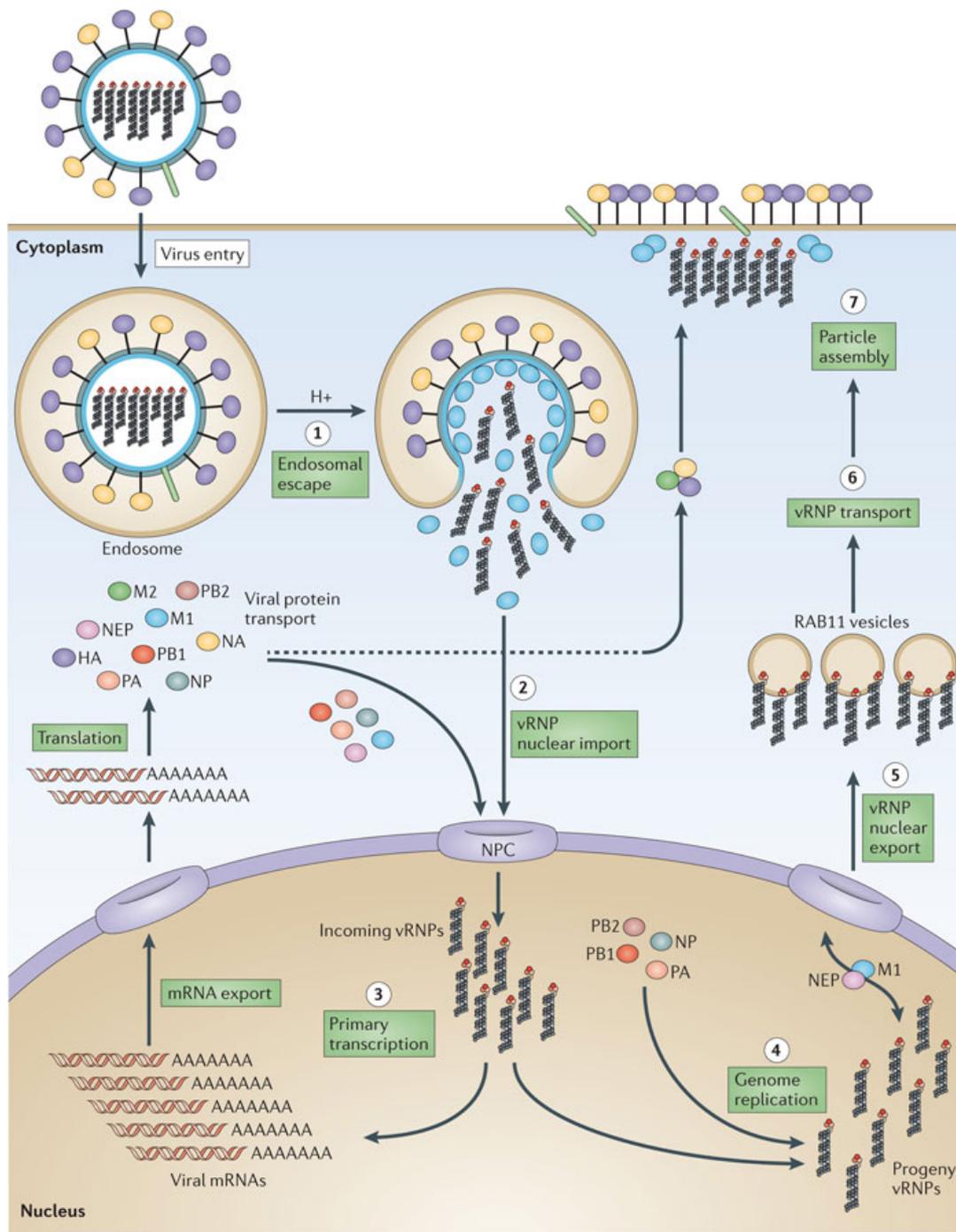
*Figure 2. IAV lifecycle. **Virus entry:** Viral protein HA binds to host receptor molecules at the plasma membrane stimulating IAV entry by endocytosis. **1:** Endosomal acidification leads to conformational changes in HA; this facilitates virion and endosomal membrane fusion, which allows the viral genome access to the cytoplasm. **2:** M2 promotes virion interior acidification, which causes M1 to dissociate from the viral genome; this releases vRNPs to travel to the nucleus via nuclear pore complex. **3:** Primary transcription of viral proteins occurs and viral mRNA is exported to the cytoplasm where cellular ribosomes translate viral proteins. **4:** Viral proteins are transported to the plasma membrane; viral genome replication occurs. **5:** vRNP are exported to the cytoplasm with the help of M1 and NEP. **6:** vRNPs are trafficked to the plasma membrane via Rab11. **7:** vRNPs are incorporated into the assembly virion progeny* (Eisfeld, Neumann and Kawaoka, 2015).

1.1.4 Proton channel viral protein M2

IAV M2 protein plays a crucial role in IAV life cycle (Figure 2). M2 is a protein of 97 amino-acids, with a 24-residue extracellular domain, 20-residue transmembrane domain and 53-amino-acid cytosolic tail (CT) (Rossman and Lamb, 2011; Sun *et al.*, 2010). M2 forms a proton channel, which is activated at acidic pH (Figure 3). M2 is important (i) for virus uncoating at early stages of cell infection through acidification of the interior of viral particles after endocytosis by target cells, (ii) for transporting of HA at the cell surface at later stages; when M2 proton channel activity enables de-acidification of the trans-Golgi lumen to prevent premature activation of HA, and finally (iii) for budding and releasing of nascent virions at the cell surface (Roberts *et al.*, 2013; Sun *et al.*, 2010).

The amphipathic helix residues 46-62 of the cytosolic tail are highly important in producing infectious virions. They aid the packaging of viral RNA, and are involved in the efficiency of budding (Rossman and Lamb, 2011; McCown and Pekosz, 2006). As a multifunctional matrix protein, M2 has a role in determining where on the plasma membrane virion assembly takes place and where it associates with the membrane as a homotetramer (Figure 3) (Roberts *et al.*, 2013; Sun *et al.*, 2010). M2, specifically its amphipathic helix, induces negative membrane curvature; this may stabilise the budding site which enables filament formation (Rossman and Lamb, 2011). The M2 amphipathic helix is involved in localisation to the host membrane, thus having a role in budding and scission (Roberts *et al.*, 2013). The hydrophobic face of the amphipathic helix contains residues that insert into the plasma membrane. This secures it and allows polar residues to induce membrane curvature (Roberts *et al.*, 2013). The amphipathic helix of M2 has only a weak interaction with the viral protein HA as seen in mutated M2 proteins. Therefore, other interactions are required for full localisation of M2 to the viral budzone at a lipid raft (Roberts *et al.*, 2013).

A

25 44 97

N — | Ecto | TM | Cytoplasmic Tail | — C

DRL**FF**KC**IY**R**FF**EHGLKRGPSTEGVPESMREEYRKEQQSAVDADDSHFVSIELE

Amphipathic Helix

---CRAC---

B
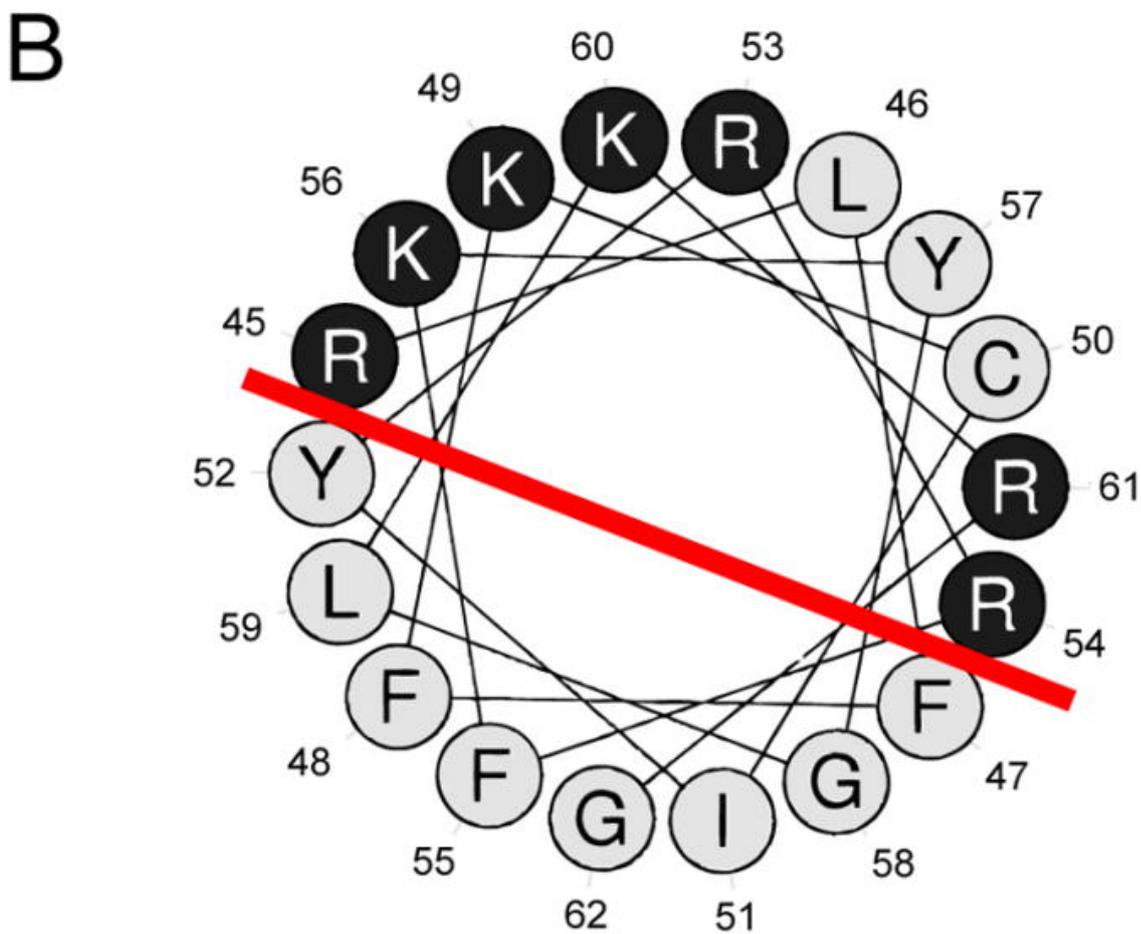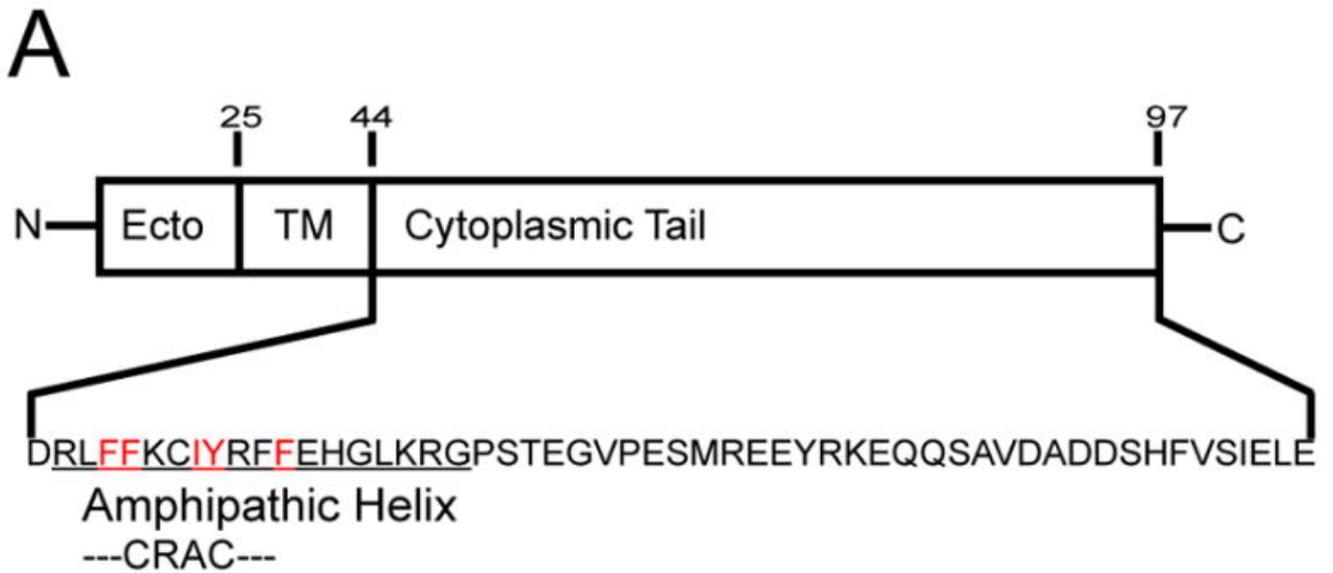
Figure 3. This figure shows the M2 domains and structural motifs. Panel A shows the M2 protein with CRAC motif, Ecto (ectodomain), TM (transmembrane) domain and amphipathic helix of cytoplasmic tail of M2 sequence identified. B shows the helical wheel plot of the amphipathic helix, generated by Rossman et al, (2010) at http://heliquest.ipmc.cnrs.fr. Hydrophobic residues are in grey and hydrophilic residues are shown in black, the red line differentiates the two sides of the helix (Rossman et al., 2010).

1.1.5 Immunological anti-viral response

1.1.5.1 Initiation of response

Recognition of pathogen-associated molecular patterns (PAMPs) by pattern recognition receptors (PRRs) initiates the host immune response (Yoo *et al.*, 2013; Holt *et al.*, 2008). This causes a cascade of events that leads to the stimulation of the inflammasome, which is an innate intracellular protein complex that has a key role in innate and adaptive immune response to viral infection (Yoo *et al.*, 2013; Takeuchi and Akira, 2009).

1.1.5.1.1 Type 1 Interferons (IFN-1)

Type 1 interferons are activated following PRR recognition of pathogen components (Yoo *et al.*, 2013). IFN-1 induce transcription of many essential infection resistance genes. These include components of the innate and adaptive immunity, such as antigen presentation and cytokine production which activate T cells, B cells and natural killer (NK) cells (Yoo *et al.*, 2013). IRF3 transcription factor is phosphorylated and activated by viral infection and by dsRNA (Killip *et al.*, 2014). IRF3 then travels to the nucleus to switch on IFN expression (amongst other genes), IFN is secreted and diffuses out of the cell (Killip *et al.*, 2014). This creates an anti-viral environment in neighbouring cells (Yoo *et al.*, 2013)**.** IFN-1 then binds to receptors on neighbouring cells to induce formation of ribonuclease (RNase), protein kinase, and 2-5A synthetase (the last two are also stimulated by dsRNA). 2-5A synthetase then activates ribonuclease to degrade RNA (Killip *et al.*, 2014). Protein kinases phosphorylate and inactivates protein chain initiation factors; this prevents the formation of proteins (Killip *et al.*, 2014). The combination of RNA degradation and inhibition of protein synthesis prevents viral replication following viral infection.

1.2. Protein Phosphatases

Protein phosphatases (PSPs) act in parallel to protein kinases. Protein kinases add phosphate groups to proteins from ATP through hydrolysis of ATP to ADP and a phosphate group; protein phosphatases act to remove these phosphate groups (McAvoy and Nairn, 2010). There are around 30 PSPs that are divided into 3 main groups, the phosphoprotein phosphatases (PPPs), metal dependent phosphatases (PPMs) and aspartate-based phosphatases (FCP/SCPs) (Shi, 2009). PSPs form holoenzymes composed of a catalytic and a regulatory subunit (Shi, 2009). There are limited numbers of phosphatases and a wide range of substrates; therefore, there are many regulatory subunits that confer substrate specificity to the holoenzyme (Shi, 2009). The addition of phosphate groups to proteins activates them. Serine, threonine and tyrosine are the amino acids most commonly phosphorylated. Within the PPP group of phosphatases there are two major Ser, Thr and Tyr phosphatase groups; these are structurally and mechanistically different from each other (McAvoy and Nairn, 2010).

The two groups of serine/threonine phosphatases, (PPM), include PP2A and PPP. The PPP subfamily includes phosphatases PP1 and PP2A. The PP2A subfamily of phosphatases includes PP2A, PP4 and PP6, which are more closely related by sequence than other phosphatases (Wang *et al.*, 2008). The PP2A family, composed of PP4c (catalytic subunit) and R2/R3 (regulatory subunits), forms the holoenzyme shown in Figure 4 (Gingras *et al.*, 2005).

1.2.1 Human PP4

1.2.1.1 The PP4 complex

The PP4 (protein serine/threonine phosphatase 4) core complex contains a catalytic unit (C) and a scaffold subunit (A). The subunit interacts with a wide range of other regulatory subunits and proteins that target the AC core dimer to specific substrates and subcellular locations (Wang *et al.*, 2008). PP4 AC core dimer (PP4c) forms 3 mutually exclusive complexes, PP4c-PP4R1, PP4c-PP4R2-PP4R3α (SMEK1) and PP4c-PP4R2-PP4R3β (Wang *et al.*, 2008). The PP4c-PP4R2-PP4R3α complex is shown in Figure 4. PP4R3α (ortholog of SMEK1 in *Dictyostelium* and FALAFEL in *Drosophila*) is a subunit of PP4 as are PP4R3β and SMEK2. PP4 holoenzyme is implicated in the following cellular processes: microtubule growth and organization (Toyo-oka *et al.*, 2008); apoptosis (Wolff *et al.*, 2006; Mihindukulasuriya *et al.*, 2004; Mourtada-Maarabouni *et al.*, 2003); TNF signaling (Mihindukulasuriya *et al.*, 2004); and pre-T-cell receptor signaling (Shui, Hu and Tan, 2007). siRNA knockdown of PP4 results in embryonic lethality in mice (Wang *et al.*, 2008). This means the functions of PP4 are probably essential for mammalian development.

1.2.1.1.1 T cell response

It has been shown that PP4 plays an essential role in thymocyte development and pre-T-cell receptor signalling (Shui, Hu and Tan, 2007). By deleting PP4 in the T-cell lineage Shui *et al.* demonstrated that PP4 has an essential role in thymocyte development and efficacy of positive selection. It was also shown the PP4 has a role in calcium flux and phospholipase C-Y-1-extracellular signal-related kinase

activation (Shui, Hu and Tan, 2007). It has also been noted that PP4 negatively regulates AMP-activated protein kinase (AMPK) (Liao *et al.*, 2016). AMPK is an important serine/threonine kinase that regulates cellular metabolism through inhibiting mTOR and cell cycle progression to inhibit proliferation (Liao *et al.*, 2016). By inhibiting AMPK, PP4 allows T-cell proliferation and expansion through mTOR.

1.2.1.1.2 Negative regulation of type I interferon (IFN)

Type 1 IFN (IFN-1) has a central role in the clearance of viral infection. Its production is highly regulated as overexpression can lead to autoimmune disorders; thus, careful regulation is required to maintain homeostasis (Zhan *et al.*, 2015). Activated TANK-binding kinase1 (TBK1) phosphorylates and thus activates IRF3. PP4 specifically inhibits RNA virus induced IFN-β activity by dephosphorylation of TBK1 after infection at Ser 172, shown in Figure 5 (Zhan *et al.*, 2015). PP4c has been found to maintain homeostasis of TBK1 in macrophages and dendritic cells (Zhan *et al.*, 2015). It has been reported that PP4 expression is increased in pandemic H1N1 infected KY/136 lung bronchial epithelial cells (Zhan *et al.*, 2015). Further investigation with other viral strains and cell lines would be beneficial to clarify the role of PP4 in anti-viral response.

1.2.1.1.3 Inhibition of NFκB activation

There has been evidence of interactions between PP4c and signalling pathways involving NFκB (Brechmann *et al.*, 2012; Martin-Granados *et al.*, 2008). Dephosphorylation by PP2A at Ser (exact Ser residue unknown) inhibits NFκB p65 subunit, whereas dephosphorylation by PP4c of Thr 435 activates NFκB (Martin-Granados *et al.*, 2008; Yeh *et al.*, 2004). PP4 is activated by TNFα-induced Ser/Thr phosphorylation; TNFα is also a NFκB inducer (Yeh *et al.*, 2004; Zhou *et al.*, 2002). TNFα may be inducing through involvement with PP4c. Thr435 is the targeting residue of the interaction between NFκB and PP4c, residues 281:444 are the NFκB amino acids responsible for the PP4c- NFκB interaction (Yeh *et al.*, 2004). SMEK1/ PP4R3α has been found to downregulate NFκB when combined with germacitabine, increasing apoptosis (Dong *et al.*, 2012; Byun *et al.*, 2012).
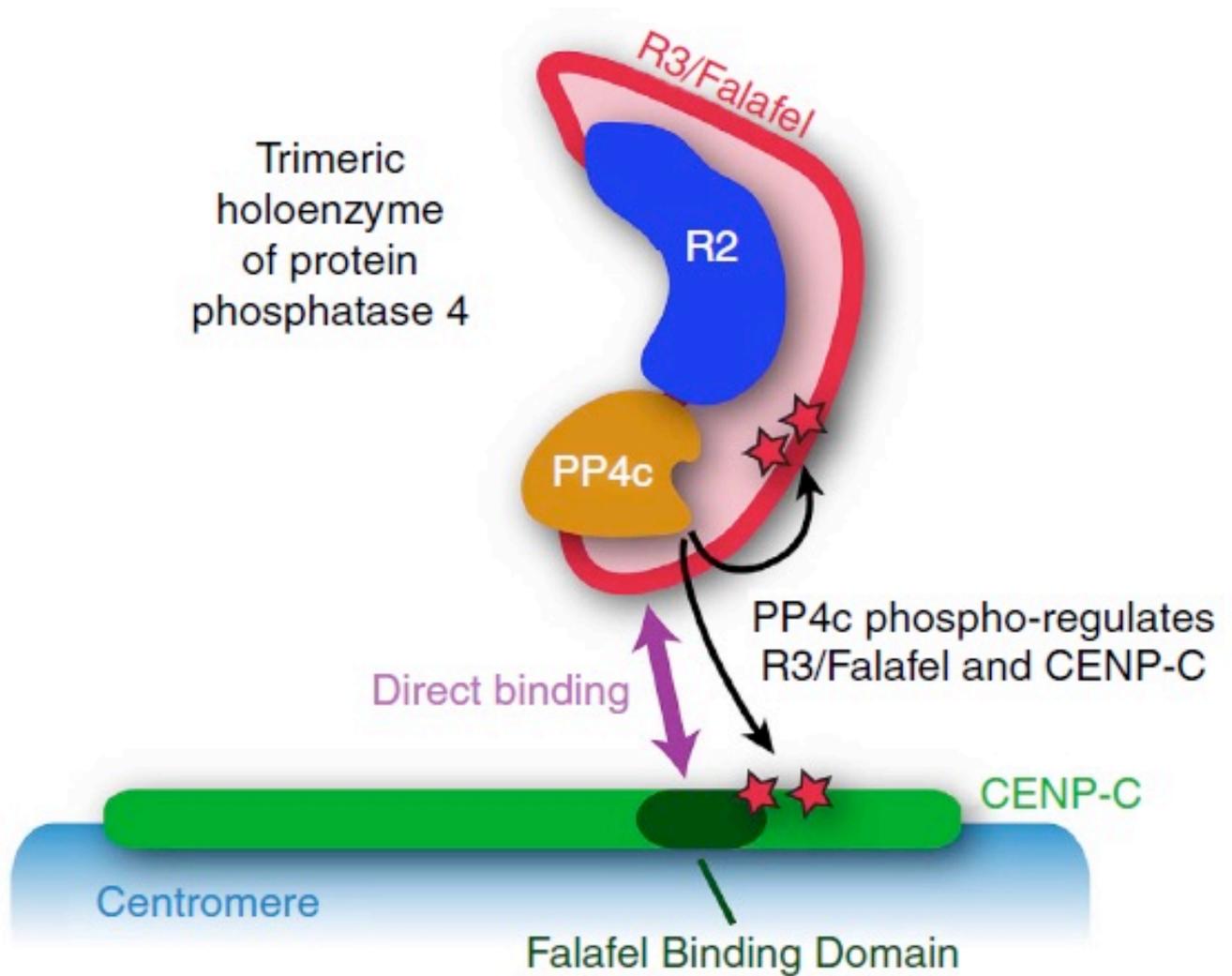
Figure 4. Hypothetical PP4 complex formation and localisation to centromere in Drosophila cells, SMEK1 interacts with Falafel Binding Domain of CENP-C (violet arrow). The trimeric holoenzyme of PP4; PP4c, R2 and R3 (Falafel in Drosophila) binds via PP4c dephosphorylation of R3 and CENP-C binding (red stars). PP4 activity at centromeres is important for maintaining mitotic centromere integrity (Eisfeld, Neumann and Kawaoka, 2015)

RIG-1/MDA5 +RNA Virus

Recruits IFN-β Promoter Stimulator 1 (IPS1)

IPS1 + TNFR-associated factor (TRAF) 3

Phosphorylation of TBK1 at Ser-172

PP4c Dephosphorylation

TBK1 Activation

TBK1 inhibition

Phosphorylation of IFN-regulatory factors IRF3 and IRF7

Induction of type 1 IFN and IFN-stimulated genes

Inhibition of viral replication and clearance of infected cells
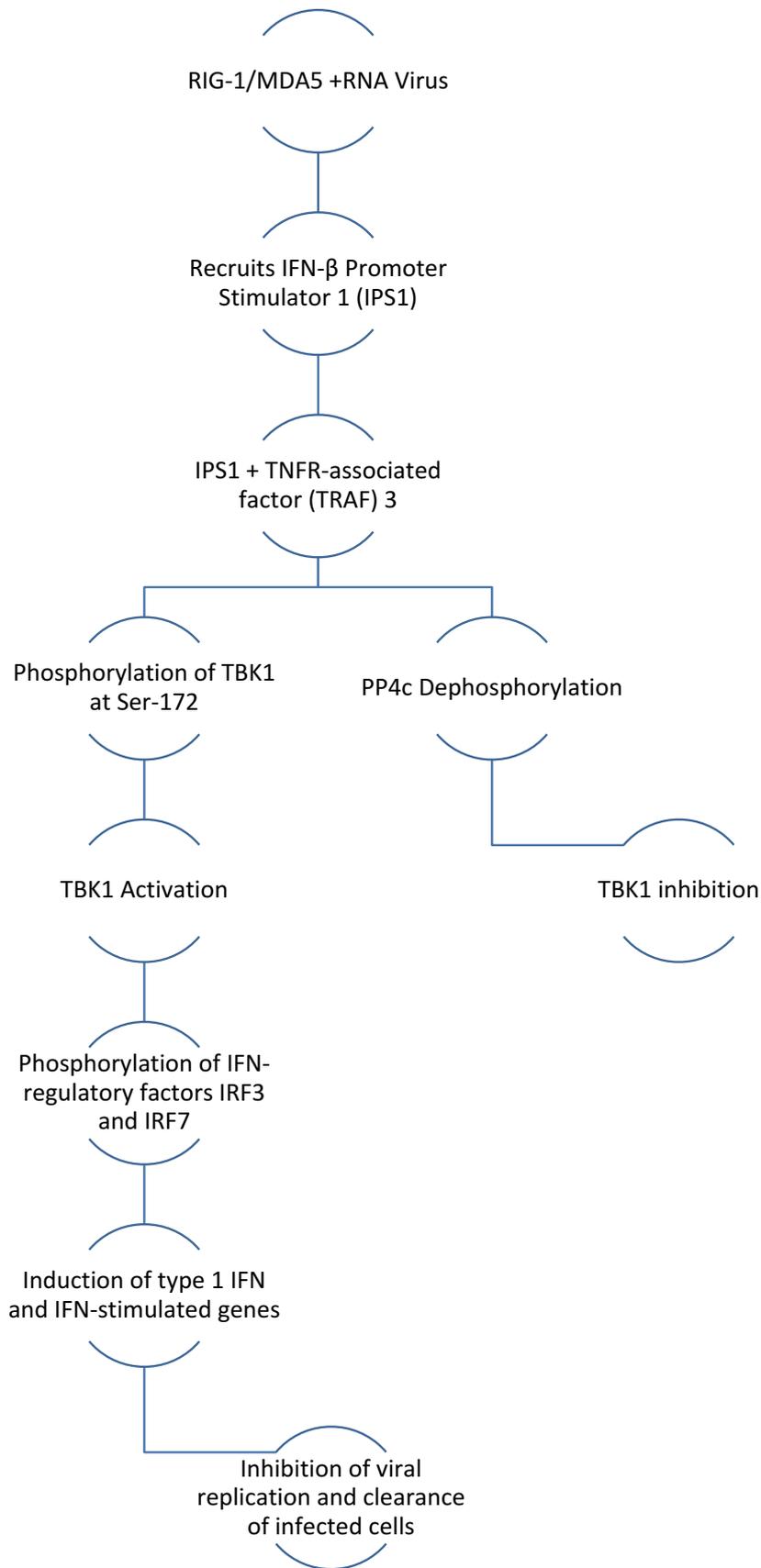
Figure 5. Schematic diagram of the IFN pathway and interaction of PP4c based on research by Zhan et al. (2015). Shown is where PP4c dephosphorylates TBK1 at Ser-172, leading to TBK1 inhibition and downstream signalling. Without dephosphorylation IFN-regulatory factors are phosphorylated leading to inhibition of viral replication and clearance of infected cells.

1.2.1.1.4 F-actin interactions

There are potential links to F-actin interactions in cell extensions in sub-confluent regions due to decreased F-actin immunostaining in 70% PP4c depleted cells (Martin-Granados et al., 2008).

1.2.1.2 PP4R1 function

There is evidence that PP4R1 targets the PP4c holoenzyme to the IκB Kinase (IKK) complex. IKK activation (via phosphorylation) results in subsequent phosphorylation of inhibitory κB proteins (IκB proteins); these binds to and deactivate NFκB within the cytoplasm (Brechmann *et al.*, 2012). Research has not demonstrated whether SMEK1 forms part of the holoenzyme with PP4 and PP4R1; that was one of the areas of investigation in the study by *Brechmann et al (2012)*.

1.2.1.4 PP4R3α and PP4R3β function

PP4R3α (ortholog of SMEK1 in *Dictyostelium* and FALAFEL in *Drosophila*) is a subunit of PP4. The R3 subunit (SMEK) is known to target PP4c to its substrate; SMEK1 Ran Binding Domain (RanBD) contains domains that bind PP4c to SMEK1 (Lyu *et al.*, 2013). This regulates PP4c structure and function through centromere targeting, shown in Figure 4. The way these subunits cooperate and mediate substrate specificity, sub-cellular localisation and holoenzyme stability is not well understood (Lipinszki *et al.*, 2015). R3 orthologues are conserved throughout evolution with similar domain architecture and have been found in a wide variety of organisms from yeasts to humans (Lipinszki *et al.*, 2015). The modular activity of its subunits regulates PP4 function. SMEK1 (PP4R3α) and SMEK2 (PP4R3β) are human isoforms; SMEK2 is involved in regulation of S phase. The SMEK2 complex requires R2 and R3β to target the PP4 complex to centrosomes (Lipinszki *et al.*, 2015)

SMEK1 is involved in a wide variety of activities (Table 1): suppression of MEK (Mendoza *et al.*, 2005), H2AX phosphorylation and histone H3 and H4 acetylation (Lyu, Jho and Lu, 2011). In *Dictyostelium* SMEK1 function is regulated by its localisation. Nuclear localisation is important for some of PP4c functions following the induction of stress signals; SMEK1 translocates PP4c to the cytoplasm through hypothesised interactions with the EVH1 domain(Mendoza *et al.*, 2007). Overexpression of SMEK1 leads to increased nuclear accumulation of PP4c in *Dictyostelium*; however, further clarification on this is required (Mendoza *et al.*, 2005). The localisation function of SMEK1 is mediated by a nuclear localisation signal (NLS) at the C-terminal (Figure 4) (Mendoza *et al.*, 2007).

Function	Targets	Reference	Cell Type
Cell cycle progression	Cyclin D1, cyclin dependent kinase 4, p27, centrosomes and mitotic spindles	(Dong *et al.*, 2012; Martin-Granados *et al.*, 2008)	SKOV-3 cells, HEK293 cells
Cytokinesis	Cell cortex	(Mendoza *et al.*, 2007)	*Dictyostelium* KAx-3 cells
Mitosis	Par3	(Lyu *et al.*, 2013)	E14 mouse embryonic stem cells
Microtubule growth and organisation	Centromeric protein C	(Lipinszki *et al.*, 2015)	D.Mel-2 cells, mouse embryonic fibroblasts
Centrosome maturation and cell migration	RhoGTPases	(Martin-Granados *et al.*, 2008)	HEK293 cells
Chemotaxis in development.	*MEK1* effectors	(Mendoza *et al.*, 2005)	*Dictyostelium* KAx-3 cells
DNA repair	Histones H3 and H4, H2AX, HDAC1	(Lyu, Jho and Lu, 2011; Martin-Granados *et al.*, 2008)	Embryonic stem cells, HEK293 cells
TNF signalling	Insulin receptor substrate-4	(Mihindukulasuriya *et al.*, 2004)	W7.2 mouse cell line (derived from WEHI-105.726)
Type 1 interferon	TANK-binding kinase 1	(Zhan *et al.*, 2015)	C57BL/6 mice
Pre-T-cell signalling	AMPK	(Shui, Hu and Tan, 2007)	Embryonic stem cells,
NFkB activation	p21, p53	(Byun *et al.*, 2012)	OVCAR-3 cells
Apoptosis	Bcl-2, Bax	(Byun *et al.*, 2012; Mourtada-Maarabouni and Williams, 2009; Mihindukulasuriya *et al.*, 2004)	OVCAR-3 cells, SKOV-3 cells HEK293/T cells, W7.2 mouse cell line (derived from WEHI-105.726)

Table 1. Shown are the various cellular functions that the PP4 complex (PP4c-PP4R3α-PP4R2) is involved in and the targets of the PP4 complex within these processes. For reference, papers that these targets were identified in are included, as well as the cell types.

1.2.2 SMEK1 Gene and protein structure

SMEK1 gene is found on chromosome 14, band q32.12. The gene is 53,074 bp long and composed of 17 exons (Weizmann Institute of Science, 2015). It encodes a protein of 833 amino acids with a weight of 95368 Da (Weizmann Institute of Science, 2015). SMEK1 is a ubiquitous regulatory subunit of PP4C with 3 isoforms in humans; this associates with the structural protein PP4R2 (Lipinszki *et al.*, 2015; Martin-Granados *et al.*, 2008).

1.2.3 SMEK1 functions

1.2.3.1 Cell cycle

SMEK1 affects cell growth inhibition through apoptosis via caspase activation, downregulation of Bcl-2 and upregulation of Bax expression (Byun et al., 2012; Dong et al., 2012). SMEK1 addition has shown suppression of the cell cycle, causing G1 growth arrest through reducing expression of Cyclin D1 and CDK4; this is due to inhibition by p21 and p27 (a Cyclin D1 inhibitor) (Dong et al., 2012). This pathway has yet to be described. 85% depletion of PP4c resulted in many cells progressing slowly through G2. This suggested that PP4c depletion slowed progression through G2 or prevented normal mitotic progression (Martin-Granados et al., 2008). This was caused by interactions with centrosomes and mitotic spindles (Martin-Granados *et al.*, 2008). PP4c–R2–R3 complex regulation of RhoGTPases may co-ordinate centrosome maturation and cell migration (Martin-Granados et al., 2008). PP4c is also involved in prometaphase. Depleted cells were halted at prometaphase, which further confirmed involvement of PP4C with mitotic spindles (Martin-Granados *et al.*, 2008). The R3β subunit (SMEK2) is involved in S-phase progression, behaving like a chromatin protein that has been phosphorylated by Cyclin-dependent kinases (Lipinszki et al., 2015). Mitotic centromere integrity is reliant on PP4, which requires SMEK1 interactions to deliver the catalytic subunit. Therefore, the PP4c-R2-R3 complex regulates maturation of the centrosome (Figure 4) (Martin-Granados et al., 2008).

SMEK1 was distributed in larger amounts in the nucleus in comparison to the cytoplasm during interphase; there was no intense staining at the centrosome. R3β was similarly distributed in larger amounts, but was localised to the centrosomes during mitosis (Martin-Granados *et al.*, 2008). R3 subunits use their influence in DNA repair pathways and the subsequent role of these repair pathways influence cell cycle progression (Lipinszki et al., 2015). PP4c is critical for microtubule growth and organisation; this is due to its interactions with regulatory proteins (Chen *et al.*, 2008b). SMEK1 is essential for cytokinesis during vegetative growth, when it is localised to the cell cortex (via an EHV1 domain). Following the initiation of development, SMEK1 translocates to the nucleus due to a NLS at the C terminus (Toyo-oka *et al.*, 2008; Mendoza *et al.*, 2005). PP4c negatively regulates cyclin dependent kinase 1 (CDK1) activity at the centrosome; loss of PP4c leads to unscheduled activations of CDK1 in interphase (Toyo-oka *et al.*, 2008). NDEL1 is a substrate of CDK1. Phosphorylation of NDEL1 recruits katanin p60 to control microtubule dynamics at centrosome and facilitates microtubule remodelling

(Toyo-oka *et al.*, 2008). PP4c is functionally dominant over CDK1; it may therefore be a checkpoint protein (Toyo-oka *et al.*, 2008).

1.2.3.2 Development

In *Dictyostelium* SMEK1 localises to the cell cortex via its EVH1 domain at the N terminus during vegetative growth, where it is required for proper cytokinesis. There is a 72% similarity between the EVH1 domain in *Dictyostelium* EVH1 and the 2 human homologs. At the start of development of embryos SMEK1 localises to the nucleus via the NLS at the C terminus (Mendoza *et al.*, 2007).

In *Dictyostelium* SMEK1 acts independently of *MEK1* but regulates *MEK1* effectors in an opposite manner to the *MEK1/ERK1* pathways (Mendoza *et al.*, 2007). SMEK1 functions in a complex with other factors regulated by the *MEK1/ERK1* pathway (Mendoza *et al.*, 2005).

SMEK1 is excluded from the nucleus in Neural Stem/Progenitor Cells undergoing mitosis. This allows it to interact with Par3 (a cell polarity protein) and mediate dephosphorylation by PP4c (Lyu *et al.*, 2013).

1.2.3.3 DNA repair

HDAC1 depletion increases histone H3 phosphorylation and deacetylation, which changes chromatin structure. The loss of PP4c complex also induces histone H3 phosphorylation/deacetylation (Lyu, Jho and Lu, 2011). SMEK1 recruits HDAC1, which forms the SMEK1/PP4c/HDAC1 complex; this performs histone acetylation in embryonic stem cells (Lyu, Jho and Lu, 2011; Martin-Granados *et al.*, 2008). HDAC1 complex formation requires PP4c and SMEK1. When all three are combined, they suppress brachyury promoter in undifferentiated embryonic stem cells by deacetylating histones H3 and H4 (Lyu, Jho and Lu, 2011). SMEK1 recruits PP4c and HDAC1 to the promoter to mediate promoter suppression (Lyu, Jho and Lu, 2011).

1.2.3.4 Apoptosis

SMEK1 is a pro-apoptotic protein that is primarily located in the nucleus; however, its pathways and mechanisms are not well characterised (Byun *et al.*, 2012; Chen *et al.*, 2008b). It binds to proteins involved in apoptosis, and histone deacetylation (Byun *et al.*, 2012). The balance of phosphate status of proteins regulates apoptosis (Mourtada-Maarabouni *et al.*, 2003). Therefore, the balance between kinases and phosphatases controls the regulation of apoptosis.

As SMEK1 is a regulatory protein of a phosphatase there are implications for PP4c in apoptotic pathways. SMEK1 activates transcription of p21 and p53 and targets the FAK/Akt/mTOR signalling pathway, which inhibits tumour growth and metastasis (Byun et al., 2012). Combining SMEK1 with germacitabine increases apoptosis by increasing Bax and reducing Bcl-xL and Bcl-2. It also up-regulates p53, and down-regulates NFκB (Dong *et al.*, 2012; Byun *et al.*, 2012). This works in contrast to PP4 activation of NFκB (Martin-Granados *et al.*, 2008; Yeh *et al.*, 2004).

1.2.4 Other PPPs

1.2.4.1 PP1

PP1α and PP1γ have been shown to co-localise with RIG-1 in HeLa cells upon infection with Sendai Virus (SeV) (Wies *et al.*, 2013). Interactions between RIG-1 and PP1α and PP1γ were also shown via co-immunoprecipitation in HEK239T cells post-SeV infection. This interaction was also confirmed in primary human lung fibroblasts infected with ΔNS1 influenza virus (Wies *et al.*, 2013). Through knock out and directed mutagenesis experiments it was found that RIG-1 is crucial for IFN-inducing ability in response to viral infection. Dephosphorylation of RIG-1 by PP1 is essential for the ubiquitination of the CARD domain of RIG-1, and the resulting signal-transducing activity(Wies *et al.*, 2013). However, this has not been confirmed *in vivo*, therefore there is a question over the relevance of this interaction.

1.2.4.2 PP2A

PP2A holds strong homology with PP4 (65% amino acid sequence). Both protein phosphatases are okadaic acid sensitive. Previous studies have used okadaic acid, which is a small molecule inhibitor produced by marine dionoflagelates (Swingle, Ni and Honkanen, 2007). Okadaic acid was widely used in studies to inhibit PP1 and PP2A, however it also inhibits PP4, PP5 and probably PP6 (Swingle, Ni and Honkanen, 2007). As PP4 has been more recently characterised, it has been said that some of the functions attributed to PP2A may belong to PP4 (Shui, Hu and Tan, 2007).

1.2.5 Background to this research

A yeast two-hybrid genome-wide screen was performed to identify human factors that interact with IAV M2 cytoplasmic domain that could have implications for our understanding of IAV lifecycle and influence on host cells and vice versa. Two other interactions with IAV M2 CT have been identified using data from the same screening. The human proteins already confirmed to interact with IAV M2 are Annexin A6 and Cyclin D3 (Fan *et al.*, 2017; Ma *et al.*, 2012). This research will investigate a third interaction between M2 CT domain and human SMEK1/PP4R3α. Understanding the interaction could give insight into IAV processes within infected cells. This is especially important as SMEK1 is ubiquitously expressed, especially in lung tissues (Weizmann Institute of Science, 2015).

1.2.6 Hypothesis and Aims

1.2.6.1 Hypothesis

No research has been published on SMEK1 regarding its roles during IAV infection. SMEK1 may be utilised by IAV to prevent early apoptosis during infection, though its interactions with Bax, Bcl-xl, Bcl-2, p53 and NFκB and the subsequent regulation of cell apoptosis. There is also the potential ability of SMEK1 to control cytokine response through its interactions with NFκB and IFN-1. Inhibition of SMEK1 could lead to the development of antivirals with the ability to prevent IAV control of the cell cycle.

If an interaction between M2 and Cyclin D3 arrests cells in G1/S phase, there are implications for PP4c function in S/G2 (Fan *et al.*, 2017). If SMEK1 stimulation causes Cyclin D1 and CDK4 suppression,

which arrests cells in G1, this could have further implications on the Cyclin D3 pathway. The halt of the cell cycle by IAV in G1/S phase could have consequences if PP4 is also implicated in this.

Negative regulation of type 1 interferon by PP4 could have serious implications due to the interaction between M2 and SMEK1. M2 could bind to SMEK1 and regulate the PP4 complex function through the regulatory subunit. IAV can thus negatively regulate type 1 interferon viral response through this interaction. This presents a novel pathway by which IAV interferes with host cell immune response.

The hypothesis is that the interaction between SMEK1 and IAV M2 protein influences viral replication. This is due to the wide range of processes that the PP4 complex is involved in during IAV infection. The interaction with the SMEK1 subunit may result in the alteration of PP4 complex functions.

If IAV can inhibit regulation of type 1 IFN process through its interactions with SMEK1, then it presents a pathway through which it evades the innate response to infection.

1.2.6.2 Aims

To test the hypothesis, the first step is to clarify the results of the yeast two-hybrid screen. For this, an *in-silico* analysis of the clones produced by the screen is required. The clones will then need to be aligned to identify homologous areas. The M2 binding domain will need to be identified as well as the other domains found on SMEK1. A literature search for known domains on SMEK1 and its homologs will need to be carried out to understand the differential functions of SMEK1 and its homologs.

SMEK1 and PP4c expression need to be confirmed in A549 lung alveolar epithelial cells to understand the implications of interactions in lung cells. Cell lysates will be run on an SDS-PAGE gel and western blotting will be used to describe the expression of SMEK1 and PP4c in A549 cells. It is important to confirm the interaction of SMEK1 and IAV M2 by co-immunoprecipitation. A549 human lung epithelial cells will be infected with A/WSN/33 IAV and tested for SMEK1 and PP4c expression. Using Immunofluoresence assays and western blotting it is hoped that an effect will be seen on SMEK1 and PP4c expression levels and cellular localisation following infection.

The overall aim of my project was to confirm an interaction between SMEK1 and M2, and to try to gain an understanding of the implications of this interaction. To achieve this, the set objectives were:

(i)   Analyse the cDNA clones identified through the Y2H screening to determine which SMEK1 isoform and which specific motif interacted with M2 CT.

(ii)   Confirm the presence of SMEK1 and PP4c in the A549 lung epithelial cells used in the lab.

(iii)   Confirm the interaction between SMEK1 and M2 by co-immunoprecipitation

(iv)   Create an expression profile of SMEK1 and PP4c throughout IAV infection.

Chapter 2: Materials and Method

2.1 Yeast two-hybrid screening

The screening procedure performed by Hybrigenics, S.A., Paris, France
(http://www.hybrigenics.com/homepage.html), was described by Ma *et al*. (2012), used bait cloning from influenza A virus M2 domains. IAV M2 C-terminal (CT) domain (bait) was sub-cloned into plasmid vector pB27 to enable fusion with DNA binding domain. The vector was then transformed into the yeast strain L40ΔGAL4. A random-primed human placenta cDNA library was subcloned into the pLexA plasmid vector in frame with a DNA activation domain. The vector was then transformed into the yeast strain L40ΔGAL4. A random-primed human placenta cDNA library was transformed into a vector containing a DNA activation domain. The human placenta cDNA library contained 10 million independent fragments. This vector was transformed into the Y187 yeast strain. Y187 and L40ΔGAL4 were mated; using a specific mating method increased efficiency. A small-scale screening was used to allow the bait to adapt to selective pressure (selective mediums) (Rain *et al.*, 11). Toxicity and auto activation were not seen. The full screening was then performed to allow a minimum of 50 million interactions to be tested. This covered five times the primary complexity of the yeast-transformed cDNA library (Rain *et al.*, 11).

273 fragments were picked following selection on a medium lacking leucine, tryptophan and histidine. These fragments were subjected to PCR and sequencing at the 5" and 3" ends. The sequences were filtered, divided into a contig and compared to the GenBank database using BLASTN. Each interaction was given a predicted biological score to define the reliability of the interaction (Ma *et al.*, 2012; Millet *et al.*, 2012; Teoh *et al.*, 2010). The predicted biological score (PBS) was defined by accounting for redundancy and independency of prey fragments; other factors were the distributions of reading frames and stop codons in overlapping fragments. False positives were given a score of "E". These scores have been shown to correlate to the significance of the interactions (Rain *et al.*, 11). 20 cDNA clones were identified to correspond to the KIAA2010 (SMEK1) protein.

2.2 *In silico* analysis

2.2.1 Alignment of cDNA sequences

The nucleotide sequences corresponding to the 20 cDNA clones identified in the yeast two-hybrid screen were saved in FASTA format and aligned using CLUSTAL OMEGA (http://www.ebi.ac.uk/Tools/msa/clustalo/). cDNA sequences were entered; parameters were set for the output format as CLUSTAL without numbers. This produced a contig which determined the overlapping regions in these cDNAs and identified the longest and shortest clones. Subsequent protein sequence alignments were also performed using CLUSTAL OMEGA.

2.2.2 Translation of cDNA into protein sequence

Following alignment of the 20 clones, the M2-binding domain was identified as the common region between contig. The common region was translated using ExPASY translation tool (http://web.expasy.org/translate/).

2.2.3 Protein sequence alignments

The translated protein sequence was searched for non-redundant protein sequence matches using BLASTp database (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). This is where conserved proteins and isoforms were found. The isoforms of SMEK1 in *Homo sapiens* were searched in NCBI database (https://www.ncbi.nlm.nih.gov/); isoforms' protein sequences were identified and then aligned using CLUSTAL OMEGA database to identify deletions. The Gene Cards database (http://www.genecards.org/) was also searched for any information on isoforms.

2.3 Cells and viruses

Human alveolar basal epithelial cells (A549), Madin-Darby canine kidney cells (MDCK) and cervical cancer epithelial HeLa cells (HeLa), were grown in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 100µ/ml penicillin streptomycin at 37^oC in humid atmosphere supplemented with $5\%CO_2$.

2.3.1 Viral Amplification

Influenza A virus strains used were A/WSN/33 (H1N1) and A/Udorn/72 (H3N2).

The virus was propagated in MDCK cells serum-free DMEM medium supplemented with: 1µg/ml tusylsulfonyl phenylalnyl chloromethyl ketrone (TPCK)-treated trypsin, 100µ/ml penicillin streptomycin and 0.5% BSA.

2.3.2 Viral Titration

Viral titre was determined by tissue culture infectious dose assay in 50% of inoculated cells ($TCID_{50}$) assay on MDCK cells. A dilution of 10^{-2} virus/ml was prepared, which was serially diluted across the 96 well plate. This created a serial dilution of $1/2\log_{10}$ (10^{-2}, $10^{-2.5}$, 10^{-3}, $10^{-3.5}$, 10^{-4}, $10^{-4.5}$, to $10^{-7.5}$); column 12 was left as a negative infection control. Virus stock was added and left for 1 hour to attach to the 96 well plate, before $1x10^5$ cells were added to each well. 100µl of infection medium was added to each well and left for 3 days. After three days, the numbers of infection-positive wells at each dilution were counted using the Reed-Muench algorithm to calculate the concentration of the virus. The principle behind the algorithm calculates the $TCID_{50}$ of the starting viral dilution by measuring the concentration at which 50% of wells are infected (limiting dilution). This algorithm can calculate the number of plaque-forming units (pfu) in a suspension of virus. The wells were counted to find a dilution where 50% of wells were infected (another algorithm was used to calculate proportional distance if this did not occur); this dilution was the limiting dilution. The theory was that if only 50% of wells were infected at that dilution, infected wells were infected by one or two pfu.

$TCID_{50}$/ml = 20 X Limiting Dilution

To calculate proportional distance =
$$\frac{\% \text{ positive wells} >50\% - 50}{\% \text{ positive wells} >50\% - \% \text{ positive value} <50\%}$$

Proportional distance was added to dilution showing >50% position wells.

To get pfu instead of $TCID_{50}$ = $TCID_{50}$ / 0.7

2.4 Infection of A549 cells for western blot and IFA

Sterile glass coverslips were placed in $10cm^3$ petri dishes. They were coated with 0.01% collagen in ddH$_2$0 and left for 2 hours at 37^oC or overnight at 4^oC. Cells were plated with 5ml complete medium. They grew to 80% confluence. Once they had reached this confluency they were washed twice with PBS before an appropriate amount of viral suspension was added to the petri dish. The amount of viral suspension added depended on the viral titre and the multiplicity of infection (MOI) required for the infection. The virus could bind to the surface of the cells for 1 hour at 37^oC (tilted every 15 minutes to ensure complete coverage). After binding the viral suspension was removed and 3 ml of infectious medium were added to the petri dishes to allow the virus to enter the cells. The dishes were left in an incubator at 37^oC for the allotted time for each time point. Fixation of cells was dependent on the method of analysis.

2.5 Antibodies

The following antibodies were used during analysis. Dilutions are indicated for relevant techniques:

−   Rabbit anti-SMEK1 ab70635, IgG polyclonal antibody corresponding to 770-820 (C-terminus) on protein ID NP_115949.1. Immunoprecipitation concentration 2μg per mg of lysate, Western blot concentration 1:2000 (Abcam, Cambridge).

−   Rabbit anti-PP4c catalogue number A300-835A-M, IgG polyclonal antibody corresponding to residue 1-50 on protein ID NP_002711.1, gene ID 5531. Immunoprecipitation concentration, Immunofluorescence concentration 1:50, Western blot concentration 1:1000 (Bethyl).

−   Rabbit anti-SMEK1 catalogue number A300-840A-M, IgG polyclonal antibody corresponding to residue 770-820 (C-terminus) on protein ID NP_002711.1, gene ID 5531. Immunoprecipitation concentration, Immunofluorescence concentration 1:50, Western blot concentration 1:2000 (Bethyl).

−   Mouse anti-M2, ab5416, IgG1, monoclonal clone 14C2. Western blot concentration 1:1000 (Abcam, Cambridge).

−   Mouse anti-GAPDH, mabcam 9484, monoclonal ab9484. Western blot concentration 1:1000 (Abcam, Cambridge).

−   Rabbit anti-Annexin-A6 gene ID 309 polyclonal antibody, catalogue number PAB18085 Immunoprecipitation concentration (Abnova).

−   Rabbit anti-NFκB (c-20) catalogue number Sc372, corresponding to c-terminus of p65 subunit, concentration for immunoprecipitation (Santa Cruz).

−   Mouse anti-M1, MCA401, IgG, Immunofluorescence concentration 1:100 (AbD serotec).

−   Anti-rabbit HRP from donkey, NA 934-1ML, western blotting concentration 1:1000 (Amersham, GE Healthcare).

−   Anti-mouse HRP from sheep, NA931-1ML, western blotting concentration 1:1000 (Amersham, GE Healthcare).

–   Alexa 488 goat anti-rabbit, Invitrogen UK, polyclonal IgG (H+L), Immunofluorescence concentration 1:100.

–   Texas Red rabbit anti-mouse, Jackson ImmunoResearch, polyclonal IgM μ Chain Specific, Immunofluorescence concentration 1:100.

2.6 BCA Assay

Samples were analysed for protein concentration with a nanodrop spectrophotometer (Nanodrop2000). A working reagent was prepared at a 50:1 ratio of kit reagents A and B (BCA kit, ThermoFisher, 23225). For each sample, 200μl of working reagent was added to 10μl of sample into clean Eppendorf tubes. In parallel, standards were prepared to make a standard curve. For this, 200μl of working reagent was added to 7 tubes (A-G). 10μl of pre-diluted BSA (2mg/ml) was added to tube A. 10μl was serially diluted from tube A-G (1/20 dilutions), ensuring that tubes were well mixed. BSA was diluted serially to final BSA concentrations of 2000, 1500, 1000, 750, 500, 250, 125, 25 and 0 μg/ml. The standards and samples were incubated at 37^oC for 30 minutes. 2μl of standard/sample was vortexed, added to the lens and analysed by Nanodrop2000 at 280nm wavelength.

2.7 Western blot

For western blot analysis for fractionated lysates, cells were washed twice with PBS, then lysed with 400μl of lysis buffer (Table 2) on ice for 15 minutes. The cells were then scraped and the lysis buffer, containing cell lysates, was collected. Cell lysates were centrifuged at 13000 rpm for 1 hour at 4^oC. The supernatant was removed and the pellet was re-suspended in 200μl of lysis buffer. The supernatant contained the soluble fraction of the cell lysates; this contained cytosol and solubilised membranes of lysed cells. The re-suspended pellet contained the insoluble fraction of the cell lysates; this contained the nuclear debris, insoluble membranes and cytoskeleton components.

Using the western blot lysis technique for whole cell lysates, cells were washed twice with PBS, trypsinised and centrifuged twice at 12000rpm for 5 minutes. Cells were then counted using a haemocytometer and 100μl of lysis buffer (Table 2) was added for every $1x10^6$ cells. The lysates were centrifuged and heated for 5 minutes at 100^oC. Anti-protease cocktail (Fisher Bioreagents BPE970-1) was added immediately after heating at a 1:1000 dilution. Lysates were then frozen at -20^oC until required.

For western blotting, fractionated protein samples were solubilised in loading buffer (Table 2). 10mM dithiothreitol (DTT) was added and samples were boiled at 95^oC for 5 minutes. They were then loaded into 8% Bis-Tris mini gel (Table 2 for resolving and stacking buffer) for separation by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The samples were run at room temperature at 250V until the loading buffer ran to the end of the gel. Hybond-P polyvinylidene difluoride (PVDF) membrane was dehydrated in 100% ethanol and rehydrated in ddH_2O. Separated proteins were transferred to a PVDF membrane (Amersham Biosciences) overnight at 4^oC in transfer buffer (Table 2) at 30V. The membrane was blocked for 1 hour in 5% semi-skimmed milk in PBS-T (PBS containing 0.01% Tween 20). The membrane was hybridised with primary antibody in PBS-T 5% semi-skimmed

milk (SMEK1, M1, M2, Annexin A6 and NFκB) or 5% BSA (GAPDH, PP4c) for 1 hour. It was washed in PBS-T 5% semi-skimmed milk and hybridised with secondary antibody in PBS-T 5% semi-skimmed milk for 1 hour. Following washing, the membrane was covered in Amersham ECL Prime western blotting detection reagent for 2 minutes, exposed to Amersham Hyperfilm ECL X-ray film and developed.

Lysis Buffer (for soluble/insoluble fractions)	1% Triton X100
	20mM Tris-HCl ph 7.5
	150mM NaCl
	1mM EDTA
	(prepared as a stock)
	Anti-protease cocktail added immediately before use at 1:1000
Lysis Buffer (for whole cell lysates)	1ml 1M tris-HCl pH 6.8
	4ml 10% SDS
	2ml Glycerol
	0.5ml β-mercaptoethanol
	Dab of Bromophenol blue powder
	Topped up with ddH_2O to 10ml and aliquoted
Resolving Buffer (for 2 gels)	3ml 40% Acrylamide/Bis,
	3ml Tris Base HCl 2M ph 8.8
	150μl SDS
	7.8ml ddH_2O
	100μl APS
	25μl TEMED
Stacking Buffer (for 2 gels)	0.5ml 40% Acrylamide,
	0.5ml Tris Base HCl 1M ph 6.8
	25μl SDS
	4.2ml ddH_2O
	25μl APS
	10μl TEMED
Running Buffer	Tris Base 4.5g,
	Glycine 34g,
	SDS 1.5g
Loading Buffer 5X (not required for whole cell lysates)	Tris (base) 6.8 pH 30mM
	SDS 15%
	Glycerol 50%
	Bromophenol blue
	ddH_2O
Transfer Buffer	2.4g Tris base 20mM
	11.6g Glycine 150mM,
	10ml SDS 0.2%
	200ml Ethanol 20%
	1.29L ddH_2O

Table 2. Various buffers used for SDS PAGE electrophoresis analysis.

2.8 Dot Blot

PVDF membrane was dehydrated in 100% ethanol and rehydrated in ddH$_2$O. The membrane could dry for approximately 1 hour before adding varying amounts of crude cell lysates to the membrane. This was left to dry again for approximately 1 hour before blocking the membrane for 1 hour in 5% BSA in PBS-T (PBS containing 0.01% Tween 20). The membrane was hybridised with primary antibody in PBS-T 5% semi-skimmed milk (SMEK1) or 5% BSA (GAPDH, PP4c) for 1 hour. It was washed in PBS-T 5% semi-skimmed milk and hybridised with secondary antibody in PBS-T 5% semi-skimmed milk for 1 hour. Following washing, the membrane was covered in ECL western blotting reagent for 2 minutes, exposed to Amersham Hyperfilm ECL X-ray film and developed.

2.9 Immunofluorescence assay

For the immunofluorescence assay (IFA), cells were plated on autoclaved coverslips within the petri dish (preparation described above). These were washed twice with PBS and placed in a 6-well plate with 1ml of 4% PFA (parafolmadehyde) for 2 hours at room temperature or overnight at 4^{0}C. The coverslips were then washed twice with PBS before IFA labelling could occur.

Coverslips were quenched with 50mM NH$_4$Cl for 15 minutes and permeabilised with TritonX100 0.1% for 5 minutes on ice. The next steps took place on parafilm attached to a flat surface; cells were placed facing down on drops of solution. Coverslips were washed again in PBS before blocking with 40μl of 10% Normal Donkey Serum (NDS) for 1 hour. They were washed in PBS before being added to 5% NDS solution containing various primary antibodies for 1 hour (rabbit anti-SMEK1 from Bethyl, rabbit anti-PP4c, mouse anti-M2). Additionally, some coverslips were incubated in 5% NDS without antibodies as a secondary antibody control. Coverslips were washed again in PBS and added to 5% NDS with various secondary antibodies for 45 minutes (Alexa 488 anti-rabbit Invitrogen 1:100, Texas Red anti-mouse Jackson ImmunoResearch 1:100). Following that, coverslips were washed in PBS and ddH$_2$0 before being placed on pre-washed microscope slides with 9μl drops of DABCO or Vectashield mounting medium and left to dry. Once dry they were sealed with clear nail varnish to prevent movement and viewed using HF14 Leica DM4000 microscope. Images were acquired using secondary only labelled coverslips as a control for measuring background signal. The same acquisition conditions (exposure time etc.) were used for each immunofluorescence channel for all images in one experiment.

Image analysis was performed on Image J. On the 'analyse' menu, 'set measurements' was selected, ensuring area, integrated density and mean grey value were chosen. The cell of interest was selected using the freeform drawing tool. The cell was measured by selecting 'measure' from the 'analyse' menu. Background was measured by using the circle-drawing tool to measure three areas surrounding the cell of interest. This was repeated for all cells in the image. For non-infected cells, or where no cells could be seen, background readings were taken in place of cell area. Values for each cell were copied into Excel and the following formula was used to calculate corrected total cell fluorescence (CTCF).

CTCF = Integrated Density – (Area of selected cell X Mean fluorescence of background readings)

The average CTCF was taken and the values used in a graph.

2.10 Co-Immunoprecipitation Assay

Co-Immunoprecipitation assay was performed using Thermo Fisher Pierce Classic Immunoprecipitation kit.

A549 cells were infected and lysed in the same way as for western blot analysis, except that the lysis buffer provided by the manufacturer (Thermo Fisher) was used. Lysate was transferred to microcentrifuge tubes and centifuged at 13000g for 10 minutes to pellet cellular debris. Some supernatant was transferred to a new tube for protein concentration determination. 20μl total lysate was saved as a control. Per 1mg of cell lysate 80μl of control agarose resin slurry (40μl of settled slurry) was added into a spin column. The column was centrifuged to remove the storage buffer. 100μl of 0.1M sodium phosphate and 0.15M sodium chloride pH7.2 were added to the column, centrifuged and the flow-through was discarded. 1mg of lysate was added to the column containing resin and incubated at 4°C for 30 minutes to 1 hour with gentle end-over-end mixing. The column was centrifuged for 1 minute at 1000g. The column containing resin was discarded and flow-through saved (20μl of precleared lysate was saved as a control). The amount of sample added was dependent on the antibody-antigen system. 2-10μg of affinity purified antibody was combined with pre-cleared lysate in a microcentrifuge tube (total protein per IP reaction should be 500μg). Antibody/lysate solution was diluted with IP lysis/wash buffer and incubated overnight at 4°C to form the immune complex. The solution was combined with loading buffer and run on an SDS-PAGE gel. The gel was transferred and analysed using the described western blotting technique.

Chapter 3: Results and Discussion

3.1 Results

3.1.1 In Silico analysis of human SMEK1 gene, protein and homologs

An *in silico* analysis of SMEK1 was conducted. This was because there were no review articles on SMEK1 or PP4 that could provide comprehensive analysis or information on SMEK1. The purpose of the *in silico* analysis was to clarify the domains, size, functions and isoforms of SMEK1. This was important, as it allowed an understanding the structure of the SMEK1 isoform identified, how it compares with other isoforms, how domains are organised in the protein product and what is the domain involved in M2 interaction. We could then develop a full understanding of the implications of any alterations to normal expression that might have been seen following IAV infection. Without understanding normal SMEK1 expression, nothing could be accurately deduced from IAV altered expression.

3.1.1.1 Analysis of SMEK1 cDNA clones and production of cDNA contig

Twenty clones corresponding to isoform 1 were produced by the yeast two-hybrid screen, performed by Hybrigenics (Paris) (Table 3). The longest and shortest clones were identified by CLUSTALW alignment of all clones; they were then translated using ExPASY translation. The longest translated clone sequence was searched in the BLASTp database. The shortest cDNA was clone number 79, which encoded a polypeptide that is 369 amino acids in length and is found at position 372-741 on the full-length protein. Its protein sequence can be found in Table 4 Appendix 9 (i). Three clones were identified to be the longest, but clone 294 had the highest percentage identity (Table 3) as per the results of the screen and was thus selected as the longest clone. Clone 294 encodes a protein of 479 amino acids in length and can be found at 340-819 on the full-length protein. Its protein sequence can be found in Table 4 Appendix 9 (i). These clones were then aligned again to confirm the M2 binding domain; the alignment is shown in Figure 6. The area found to be common between all clones is the area that has successful bound to M2 during the yeast two-hybrid screen. Therefore, the area common to all clones can be identified as the M2 binding domain.

The first 3 results of the protein database search with 100% identity were an unnamed protein accession number BAB14877.1 (cDNA sequence project), KIAA2010 accession number EAW81454.1 (human genome sequence) and PP4R3α accession number NPO01271209. The shortest clone was also searched through BLASTp; this clone is very highly conserved and produced a high volume of matches. The longest and shortest clones were then aligned with PP4R3α sequence found through the BLASTp search, to identify the location of these clones on the full-length protein. Through aligning the longest and shortest clones (294 and 129), the common region was translated with ExPASY translation tool. A contig of the translated short and long clones and full protein can be found in Figure 6. This protein sequence was searched in the BLAST database to find conserved proteins and isoforms. The isoforms of SMEK1 were searched in NCBI database to allow for a run of alignment of protein sequences on CLUSTAL OMEGA and to identify deletions.

Clone Name	Type of sequence	Contig(s) Name	Global PBS score	% Id 5p/3p	Start	Stop	Reading Frame
PLA_RP_hgx1492v1_pB27_A-67-3p	3p	18635238	A	88.3	No Data	2336	??
PLA_RP_hgx1492v1_pB27_A-180-5p	5p	18635238	A	98.2	1020	No Data	IF
PLA_RP_hgx1492v1_pB27_A-68	5p 3p	18635238 / 18635398	A	88.9 / 73.1	1020	2418	IF
PLA_RP_hgx1492v1_pB27_A-294	5p 3p	18635238	A	99.9 / 99.7	1020	2418	IF
PLA_RP_hgx1492v1_pB27_A-137	5p 3p	18635238	A	99.7 / 99.9	1020	2418	IF
PLA_RP_hgx1492v1_pB27_A-339	5p 3p	18635238	A	97.6 / 99.7	1020	2418	IF
PLA_RP_hgx1492v1_pB27_A-142	5p 3p	18635238	A	99.5 / 99.4	1062	2337	IF
PLA_RP_hgx1492v1_pB27_A-264	5p 3p	18635238	A	99.3 / 99.9	1062	2337	IF
PLA_RP_hgx1492v1_pB27_A-240	5p 3p	18635238	A	98.6 / 100.0	1062	2337	IF
PLA_RP_hgx1492v1_pB27_A-207	5p 3p	18635238	A	98.6 / 99.2	1062	2337	IF
PLA_RP_hgx1492v1_pB27_A-39	5p 3p	18635238	A	91.0 / 93.0	1071	2337	IF
PLA_RP_hgx1492v1_pB27_A-280	5p 3p	18635238	A	98.4 / 99.9	1071	2337	IF
PLA_RP_hgx1492v1_pB27_A-373	5p 3p	18635238	A	98.5 / 100.0	1071	2337	IF
PLA_RP_hgx1492v1_pB27_A-276	5p 3p	18635238	A	99.7 / 99.6	1071	2337	IF
PLA_RP_hgx1492v1_pB27_A-102	5p 3p	18635238	A	99.3 / 99.9	1071	2337	IF
PLA_RP_hgx1492v1_pB27_A-88	5p 3p	18635238	A	90.5 / 91.7	1071	2337	IF
PLA_RP_hgx1492v1_pB27_A-79	5p 3p	18635238	A	97.2 / 92.0	1116	2186	IF
PLA_RP_hgx1492v1_pB27_A-123	5p 3p	18635238	A	98.8 / 99.6	1116	2186	IF
PLA_RP_hgx1492v1_pB27_A-128	5p 3p	18635238	A	99.6 / 99.1	1116	2186	IF
PLA_RP_hgx1492v1_pB27_A-135	5p 3p	18635238	A	98.7 / 99.7	1116	2186	IF
PLA_RP_hgx1492v1_pB27_A-129	5p 3p	18635238 / 18635526	A	98.8 / 84.2	1119	2117	IF

Table 3. The information regarding clones identified through the yeast two-hybrid screen. The clone name comprises of the programme used (PLA_RP), screen name (hgx1492v1), vector (pB27) and individual clone number beginning with A. Type of sequence describes whether the sequence is 3" or 5". Global PBS score describes the confidence in the interaction, A shows very high confidence in the interaction. %id shows percentage identity between both 5" and 3" clones and best matched gene (Human KIAA2010).

```
PP4R3a_iso1    MTDTRRRVKVYTLNEDRQWDDRGTGHVSSGYVERLKGMSLLVRAESDGSLLLESKINPNT    60
pB27_A-294     ------------------------------------------------------------     0
pB27_A-129     ------------------------------------------------------------     0


PP4R3a_iso1    AYQKQQDTLIVWSEAENYDLALSFQEKAGCDEIWEKICQVQGKDPSVDITQDLVDESEEE   120
pB27_A-294     ------------------------------------------------------------     0
pB27_A-129     ------------------------------------------------------------     0


PP4R3a_iso1    RFDDMSSPGLELPSCELSRLEEIAELVASSLPSPLRREKLALALENEGYIKKLLELFHVC   180
pB27_A-294     ------------------------------------------------------------     0
pB27_A-129     ------------------------------------------------------------     0


PP4R3a_iso1    EDLENIEGLHHLYEIIKGIFLLNRTALFEVMFSEECIMDVIGCLEYDPALSQPRKHREFL   240
pB27_A-294     ------------------------------------------------------------     0
pB27_A-129     ------------------------------------------------------------     0


PP4R3a_iso1    TKTAKFKEVIPISDPELKQKIHQTYRVQYIQDMVLPTPSVFEENMLSTLHSFIFFNKVEI   300
pB27_A-294     ------------------------------------------------------------     0
pB27_A-129     ------------------------------------------------------------     0


PP4R3a_iso1    VGMLQEDEKFLTDLFAQLTDEATDEEKRQELVNFLKEFCAFSQTLQPQNRDAFFKTLSNM   360
pB27_A-294     --------------------------------------XFSQTLQPQNRDAFFKTLSNM    21
pB27_A-129     ------------------------------------------------------------     0


PP4R3a_iso1    GILPALEVILGMDDTQVRSAATDIFSYLVEYNPSMVREFVMQEAQQNDDDILLINLIIEH   420
pB27_A-294     GILPALEVILGMDDTQVRSAATDIFSYLVEYNPSMVREFVMQEAQQNDDDILLINLIIEH    81
pB27_A-129     ------------XDTQVRSAATDIFSYLVEYNPSMVREFVMQEAQQNDDDILLINLIIEH    48
                           ************************************************


PP4R3a_iso1    MICDTDPELGGAVQLMGLLRTLVDPENMLATANKTEKTEFLGFFYKHCMHVLTAPLLANT   480
pB27_A-294     MICDTDPELGGAVQLMGLLRTLVDPENMLATANKTEKTEFLGFFYKHCMHVLTAPLLANT   141
pB27_A-129     MICDTDPELGGAVQLMGLLRTLVDPENMLATANKTEKTEFLGFFYKHCMHVLTAPLLANT   108
               ************************************************************


PP4R3a_iso1    TEDKPSKDDFQTAQLLALVLELLTFCVEHHTYHIKNYIINKDILRRVLVLMASKHAFLAL   540
pB27_A-294     TEDKPSKDDFQTAQLLALVLELLTFCVEHHTYHIKNYIINKDILRRVLVLMASKHAFLAL   201
pB27_A-129     TEDKPSKDDFQTAQLLALVLELLTFCVEHHTYHIKNYIINKDILRRVLVLMASKHAFLAL   168
               ************************************************************


PP4R3a_iso1    CALRFKRKIIGLKDEFYNRYIMKSFLFEPVVKAFLNNGSRYNLMNSAIIEMFEFIRVEDI   600
pB27_A-294     CALRFKRKIIGLKDEFYNRYIMKSFLFEPVVKAFLNNGSRYNLMNSAIIEMFEFIRVEDI   261
pB27_A-129     CALRFKRKIIGLKDEFYNRYIMKSFLFEPVVKAFLNNGSRYNLMNSAIIEMFEFIRVEDI   228
               ************************************************************


PP4R3a_iso1    KSLTAHVIENYWKALEDVDYVQTFKGLKLRFEQQRERQDNPKLDSMRSILRNHRYRRDAR   660
pB27_A-294     KSLTAHVIENYWKALEDVDYVQTFKGLKLRFEQQRERQDNPKLDSMRSILRNHRYRRDAR   321
pB27_A-129     KSLTAHVIENYWKALEDVDYVQTFKGLKLRFEQQRERQDNPKLDSMRSILRNHRYRRDAR   288
               ************************************************************


PP4R3a_iso1    TLEDEEEMWFNTDEDDMEDGEAVVSPSDKTKNDDDIMDPISKFMERKKLKESEEKEVLLK   720
pB27_A-294     TLEDEEEMWFNTDEDDMEDGEAVVSPSDKTKNDDDIMDPISKFMERKKLKESEEKEVLLK   381
pB27_A-129     TLEDEEEMWFNTDEDDMEDGEAVVSPSDKTKNDDDIMDPISKFME--------------   333
               ********************************************


PP4R3a_iso1    TNLSGRQSPSFKLSLSSGTKTNLTSQSSTTNLPGSPGSPGSPGSPGSPGSVPKNTSQTAA   780
pB27_A-294     TNLSGRQSPSFKLSLSSGTKTNLTSQSSTTNLPGSPGSPGSPGSPGSPGSVPKNTSQTAA   441
pB27_A-129     ------------------------------------------------------------   333


PP4R3a_iso1    ITTKGGLVGLVDYPDDDEDDDEDEDKEDTLPLSKKAKFDS 820
pB27_A-294     ITTKGGLVGLVDYPDDDEDDDEDEDK--------------  467
pB27_A-129     ----------------------------------------  333
```

Figure 6. Alignment of PP4R3α isoform 1, clone 294 (longest clone) and clone 129 (shortest clone). This was performed using BLASTN alignment tool. Overlapping areas of the sequences identify the proposed M2 binding domain.

The isoform found through this yeast two-hybrid screen was identified as isoform 1. This has a deletion of 12 amino acids from 410-422 of VSKKLTEQKITSK, which is a poly-lysine insert. Another isoform, isoform 3, has a larger deletion, 67 – 305; this resulted in a small 594 amino acid chain. The result of the loss of this sequence would cause deletion of much of the DUF625 domain; this could negate the

function of SMEK1 to recognise PP4c for binding. CLUSTAL OMEGA alignment of isoforms allowed identification of deleted regions found in Figure 7.

3.1.1.2 SMEK1 isoform 1 was isolated through the screen

The yeast two-hybrid screen identified isoform 1, an 820-amino acid sequence with a deletion of 409:422 (as found on longest isoform sequence). This is different to the sequence found on the Gene Cards database, which is an 833-amino acid sequence. The variety of identified isoforms led to a search for other isoforms of SMEK1; 3 isoforms were found on NCBI protein database. Isoform 1 is missing 409:422, isoform 3 is missing 67:305; isoform 4 is the full-length protein. SMEK1 *sapiens* was then searched in NCBI nucleotide database to identify the transcript variants that led to these isoforms. Transcript variant 2 is missing 410:420, transcript variant 5 is also missing 410:420, but is an incomplete non-translatable variant. NCBI database, Gene Cards database and published papers were then searched for the domains of SMEK1, shown in Figure 7.

3.1.1.2 Protein Domains

There were five domains on SMEK1: Pleckstrin-homology domain (PH domain), Domain of unknown function (DUF625), Armadillo-type fold (ARM), Ran binding domain (RanBD) and M2 binding domain (M2 BD). These domains were found through alignment of yeast two-hybrid screening clones (Weizmann Institute of Science, 2015).

The RanBD domain was found through literature analysis (Lyu *et al.*, 2013; Mendoza *et al.*, 2007). This was the most conserved portion of SMEK1, which contained a 115-amino acid fragment at the N-terminal.

DUF625 is crucial to SMEK1 function and SMEK1 participation in neurogenesis. DUF625 binds to the Par3 c-terminus. Par3 is a negative regulator of neurogenesis, which is in turn negatively regulated by PP4. This is important for neuronal differentiation as Par3 critically regulates proliferation as opposed to differentiation during cortical development(Lyu *et al.*, 2013). The central DUF625 mediates SMEK1 function and participates in neurogenesis alongside the RanBD domain (N-terminal Ran-binding domain) (Lyu *et al.*, 2013; Mendoza *et al.*, 2007). The DUF625 is also important for Par3/ SMEK1 interaction with the coiled-coil region of Par3, enabling PP4c dephosphorylation of Par3 (Lyu *et al.*, 2013). Following the DUF625 is a variable amount of armadillo (ARM) repeats; these regulate neuronal differentiation through PP4c (Lipinszki *et al.*, 2015; Lyu *et al.*, 2013).This explains why PP4 knock out mice are embryonically lethal (Wang *et al.*, 2008). Without proliferation, embryos are not able to develop. RanBD (Ran binding domain) mediates SMEK1 function in cell cycle, through binding to Ran. SMEK1 cytoplasmic/cortical localisation during mitosis is essential for PP4 to target Par3 during development (Lyu *et al.*, 2013)

Poly-lysine inserts can be found at 410:422, as found in the amino acid sequence: VSKKLTEQKITSK. This insert was found in all isoforms except isoform 1.

The EVH1 domain in SMEK1 is required for SMEK1 localisation to the cell cortex during vegetative growth where is it is required for proper cytokinesis (Mendoza *et al.*, 2007).

3.1.1.2.3 Definition of the SMEK1 M2-binding domain

Shown in Figure 7 are the longest and shortest clones found through the yeast two-hybrid screen. They are arranged to show their position relative to PP4R3α/SMEK1 isoforms. Clone 294 is the longest clone and includes the serine stretch, which contains 6 serine's at positions 741, 768, 771, 774, 777 and 780. Clone 129 is the shortest clone and was identified as the M2-binding domain. It does not contain the serine rich area found in the longest clone and all isoforms. Both clones are homologous to isoform 1 in the respect that they also lack the poly-lysine insert, which is found in isoform 4 and 3. It was by this logic that it was deduced that the SMEK1 isoform identified in the yeast two-hybrid screen was isoform 1. This could mean that M2 can only bind this isoform 1. It could also mean that the placenta library was enriched in isoform 1 clones or the placenta library did not contain clones for other isoforms, thus no other isoform clones were identified.

In silico identification of the M2 binding domain through clone alignment identified an overlap with the ARM domain. It was found that the M2 binding domain was adjacent to the DUF 625 domain shown in

Figure 7. The ARM domain is required for PP4-SMEK1 complex formation as a recognition site for PP4(Lyu *et al.*, 2013). SMEK1 is important for PP4 localisation, therefore the ARM domain acts as a binding site for PP4; this is essential for nuclear localisation of the PP4 complex (Lyu *et al.*, 2013). The M2 binding domain overlaps with the ARM repeats which are important for PP4 recognition of SMEK1. This could have implications for PP4 and SMEK1 binding following IAV infection, although this is an area yet to be researched.
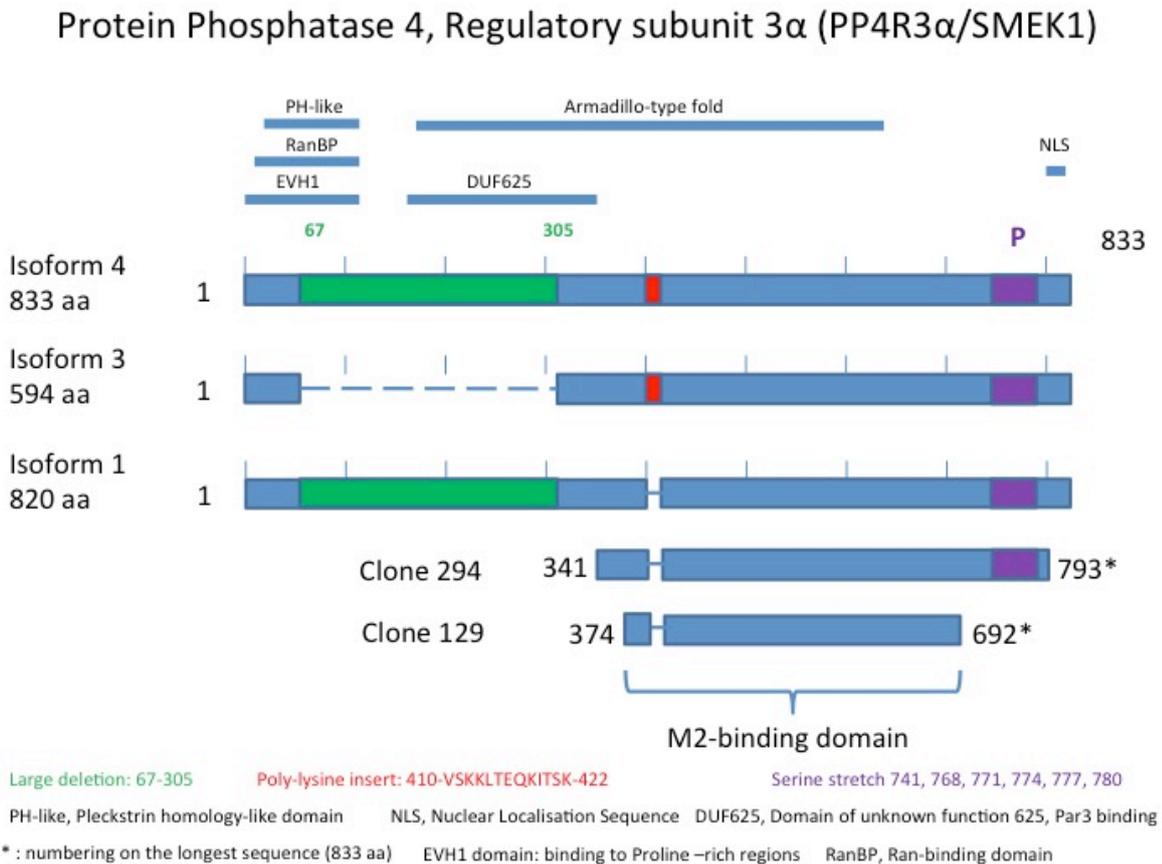
Figure 7. Full length Protein SMEK1 and clones 294 and 129. M2 binding domain is shown, as well as the amino acid position of each domain on the full-length protein and the longest and shortest clone fragments. Important domains found the literature search has identified RanBD, the EVH1 domain, the PH (Pleckstrin-homology) domain, which also overlaps with the RanBD and EVH1 domains at the N-terminus and a poly-lysine insert with potential phosphorylation sites. DUF625 (domain of unknown function 625, also referred to as the SMEK1 domain), overlaps with the ARM (armadillo-type fold) repeat. At the C-terminus there is a nuclear localisation sequence (NLS). Below this are the longest (294) and shortest (129) clones of 369 and 479 amino acids in length respectively.

3.1.1.4 Similarity of human SMEK1 with orthologs

An alignment was performed using: SMEK1 (Human) accession number AAH72409.1, clone 294, FALAFEL isoform E (*Drosophila*) accession number NP_731850.2 and PP4R3α isoform 1 (Human) accession number NP_001271209.1, shown in Figure 8. The phylogenic tree showed that FALAFEL and PP4R3α were more homologous than clone 294 and SMEK1.

```
NP_731850.2       MTTDTRRRVKLYALNAERQWDDRGTGHVSSTYVERLKGISLLVRAESDGSLLLESKIQPD
AAH72409.1        -MTDTRRRVKVYTLNEDRQWDDRGTGHVSSGYVERLKGMSLLVRAESDGSLLLESKINPN
pB27_A-294        ------------------------------------------------------------
NP_001271209.1    -MTDTRRRVKVYTLNEDRQWDDRGTGHVSSGYVERLKGMSLLVRAESDGSLLLESKINPN

NP_731850.2       TAYQKQQDTLIVWSEGDNFDLALSFQEKAGCDEIWEKICQVQGKDPSVEITQDIVEESED
AAH72409.1        TAYQKQQDTLIVWSEAENYDLALSFQEKAGCDEIWEKICQVQGKDPSVDITQDLVDESEE
pB27_A-294        ------------------------------------------------------------
NP_001271209.1    TAYQKQQDTLIVWSEAENYDLALSFQEKAGCDEIWEKICQVQGKDPSVDITQDLVDESEE

NP_731850.2       ERFEDMSDTAPPIELPPCELSRLEDISETIQSCLSTPLRKEKLSMALESESYIKKLLNLF
AAH72409.1        ERFDDMS--SPGLELPSCELSRLEEIAELVASSLPSPLRREKLALALENEGYIKKLLELF
pB27_A-294        ------------------------------------------------------------
NP_001271209.1    ERFDDMS--SPGLELPSCELSRLEEIAELVASSLPSPLRREKLALALENEGYIKKLLELF

NP_731850.2       HVCEDLDNTEGLHHLFEIFKNIFLLNKNALFEIMFADDTIFDVVGCLEYDPSVSQPKKHR
AAH72409.1        HVCEDLENIEGLHHLYEIIKGIFLLNRTALFEVMFSEECIMDVIGCLEYDPALSQPRKHR
pB27_A-294        ------------------------------------------------------------
NP_001271209.1    HVCEDLENIEGLHHLYEIIKGIFLLNRTALFEVMFSEECIMDVIGCLEYDPALSQPRKHR

NP_731850.2       QYLKQLAKFREAVPIKNLDLLAKIHQTFRVQYIQDIILPTPSVFVEDNMLNTLSSFIFFN
AAH72409.1        EFLTKTAKFKEVIPISDPELKQKIHQTYRVQYIQDMVLPTPSVFE-ENMLSTLHSFIFFN
pB27_A-294        ------------------------------------------------------------
NP_001271209.1    EFLTKTAKFKEVIPISDPELKQKIHQTYRVQYIQDMVLPTPSVFE-ENMLSTLHSFIFFN

NP_731850.2       KVEIVTMIQDDERYLLDVFAVLTDPTTGDAKRRDTVLFLKEFCNYAQNLQPQGKDSFYKT
AAH72409.1        KVEIVGMLQEDEKFLTDLFAQLTDEATDEEKRQELVNFLKEFCAFSQTLQPQNRDAFFKT
pB27_A-294        --------------------------------------------XFSQTLQPQNRDAFFKT
NP_001271209.1    KVEIVGMLQEDEKFLTDLFAQLTDEATDEEKRQELVNFLKEFCAFSQTLQPQNRDAFFKT
                                                              ::*.****.:*:*:**

NP_731850.2       LTCLGILQALELTLVMNDKKTKSASIDILTAIVEFSPLVVRNYTLNQANRP---------
AAH72409.1        LSNMGILPALEVILGMDDTQVRSAATDIFSYLVEYNPSMVREFVMQEAQQNDDVSKKLTE
pB27_A-294        LSNMGILPALEVILGMDDTQVRSAATDIFSYLVEYNPSMVREFVMQEAQQNDD-------
NP_001271209.1    LSNMGILPALEVILGMDDTQVRSAATDIFSYLVEYNPSMVREFVMQEAQQNDD-------
                  *: :*** ***: * *:*.:.:**: **:: :**:.* :**::.:::*::

NP_731850.2       ----EVERMLLNIAIEQMLNDSEPELGIAVQLMGIVKILLEPENMLT-----EKGDFLNF
AAH72409.1        QKITSKDILLINLIIEHMICDTDPELGGAVQLMGLLRTLVDPENMLATANKTEKTEFLGF
pB27_A-294        ------DILLINLIIEHMICDTDPELGGAVQLMGLLRTLVDPENMLATANKTEKTEFLGF
NP_001271209.1    ------DILLINLIIEHMICDTDPELGGAVQLMGLLRTLVDPENMLATANKTEKTEFLGF
                       : :*:*: **:*: *:::**** ******::: *::*****:    ** :**.*

NP_731850.2       FYKYSVQTLVAPVILNTIGDRPQNEDYQTAQLLGIVLDILSFCVEHHSYHIKNFLLQKDL
AAH72409.1        FYKHCMHVLTAPLLANTTEDKPSKDDFQTAQLLALVLELLTFCVEHHTYHIKNYIINKDI
pB27_A-294        FYKHCMHVLTAPLLANTTEDKPSKDDFQTAQLLALVLELLTFCVEHHTYHIKNYIINKDI
NP_001271209.1    FYKHCMHVLTAPLLANTTEDKPSKDDFQTAQLLALVLELLTFCVEHHTYHIKNYIINKDI
                  ***:.::.*.**::*: **  *:*.::*:******.:**::*:*******:*****::::**:

NP_731850.2       LKRILVLMKSTHTFLVLGALRLLRKIIALKDEFYNRHIVKCNLFAPVVDAFIRNNGRYNL
AAH72409.1        LRRVLVLMASKHAFLALCALRFKRKIIGLKDEFYNRYIMKSFLFEPVVKAFLNNGSRYNL
pB27_A-294        LRRVLVLMASKHAFLALCALRFKRKIIGLKDEFYNRYIMKSFLFEPVVKAFLNNGSRYNL
NP_001271209.1    LRRVLVLMASKHAFLALCALRFKRKIIGLKDEFYNRYIMKSFLFEPVVKAFLNNGSRYNL
                  *:*:**** *.*:*.* ***: ****.********:*:*. ** ***.**:.*..****

NP_731850.2       LESAILELFEFIKLEDIRTLCVYFVENFSKIFDEIEYVQTFKYLKNRYDQYQDRLKDRDK
AAH72409.1        MNSAIIEMFEFIRVEDIKSLTAHVIENYWKALEDVDYVQTFKGLKLRFEQQRERQDN---
pB27_A-294        MNSAIIEMFEFIRVEDIKSLTAHVIENYWKALEDVDYVQTFKGLKLRFEQQRERQDN---
NP_001271209.1    MNSAIIEMFEFIRVEDIKSLTAHVIENYWKALEDVDYVQTFKGLKLRFEQQRERQDN---
                  ::***:*:****::***::* .:.:**: * ::::::****** ** *::* ::* .:

NP_731850.2       MENRTDGGLPIIRSGGRFRRDQRQMEEEEEMWFNEEDDFTEEIDTYNNVMKSVSEKNG--
AAH72409.1        ---PKLDSMRSILRNHRYRRDARTLEDEEEMWFNTDEDDMEDGEAVVSPSDKTKNDDDIM
pB27_A-294        ---PKLDSMRSILRNHRYRRDARTLEDEEEMWFNTDEDDMEDGEAVVSPSDKTKNDDDIM
NP_001271209.1    ---PKLDSMRSILRNHRYRRDARTLEDEEEMWFNTDEDDMEDGEAVVSPSDKTKNDDDIM
                   . ..:  *  . *:*** * :*:******* ::*  *: ::  .  ....:.:.

NP_731850.2       ------------------------------------------------------------
AAH72409.1        DPISKFMERKKLKESEEKEVLLKTNLSGRQSPSFKLSLSSGTKTNLTSQSSTTNLPGSPG
pB27_A-294        DPISKFMERKKLKESEEKEVLLKTNLSGRQSPSFKLSLSSGTKTNLTSQSSTTNLPGSPG
NP_001271209.1    DPISKFMERKKLKESEEKEVLLKTNLSGRQSPSFKLSLSSGTKTNLTSQSSTTNLPGSPG

NP_731850.2       ------------------------------------------------------------
AAH72409.1        SPGSPGSPGSPGSVPKNTSQTAAITTKGGLVGLVDYPDDDEDDDEDEDKEDTLPLSKKAK
pB27_A-294        SPGSPGSPGSPGSVPKNTSQTAAITTKGGLVGLVDYPDDDEDDDEDEDK-----------
NP_001271209.1    SPGSPGSPGSPGSVPKNTSQTAAITTKGGLVGLVDYPDDDEDDDEDEDKEDTLPLSKKAK

NP_731850.2       ---
AAH72409.1        FDS
pB27_A-294        ---
NP_001271209.1    FDS
```

```
                                            ┌──────────── NP_731850.2 0.41957
                          ┌─────────────────┤
┌─────────────────────────┤                 └──────────── NP_001271209.1 -0.03003
│                         └──────────────────────────────── AAH72409.1 -0.02915
└─────────────────────────────────────────────────────────── pB27_A-294 0.03129
```

Figure 8. Alignment of: Drosophila FALAFEL isoform E (>NP_731850.2), Human SMEK1 isoform 1 (>AAH72409.1), Clone 294 from the yeast two-hybrid screen (> pB27_A-294) and Human PP4R3α (> NP_001271209.1). A phylogenic tree below shows similarity between sequences.

3.1.2 Expression of SMEK1 and PP4c in human A549 pneumocytes

The expression of SMEK1 in non-infected A549 lung pneumocytic epithelial cells needed to be confirmed. The cell line used for the work on IAV was an immortalised A549 lung epithelial cell. It was noted on the Gene Cards website that A549 expresses SMEK1 (http://www.genecards.org/), however confirmation of expression was needed before continuing with the project. Human cervix epithelial cells (HeLa) cells were included as a positive control. Soluble and insoluble fractions of cell lysates were collected from non-infected A549 and HeLa cells (as a positive control) and analysed via dot blot with specific antibodies. A dot blot was performed as shown in Figure 9. By adding crude cell lysates directly to the PVDF membrane, PP4c and SMEK1 expression was detected in A549 and HeLa cells.
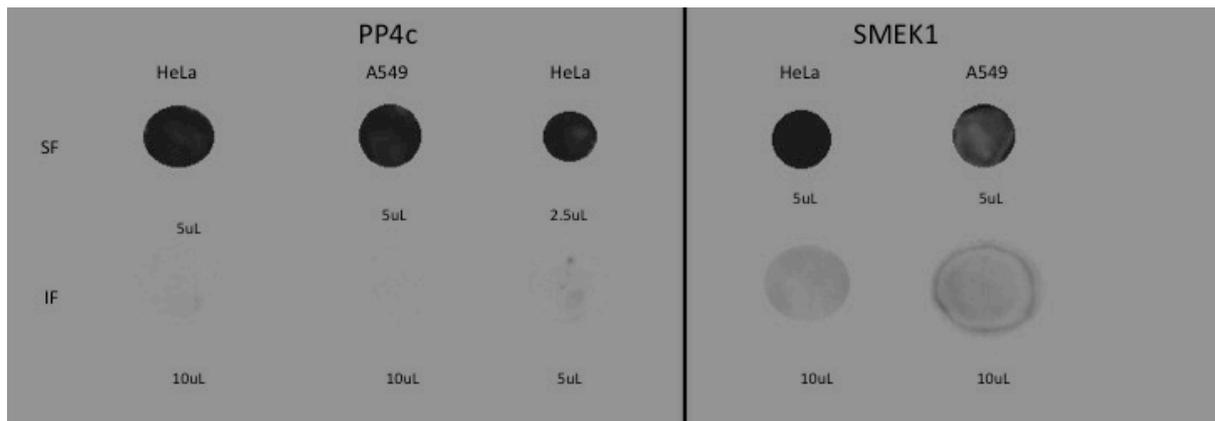
Figure 9. Analysis of detection of SMEK1 and PP4c in A549 and HeLa cell lines by dot blot for optimisation of SMEK1 and PP4c detection after 10 minutes' exposure. Antibodies used were rabbit anti-SMEK1 (abcam) and anti-PP4c antibody (Bethyl Labs) in HeLa insoluble fractions (IF) and soluble fractions (SF) and A549 SF and IF). There is clear expression of SMEK1 and PP4c protein in SF of both A549 and HeLa cells. This confirms SMEK1 expression detection with rabbit anti-SMEK1 antibody (abcam). (Exposure time and full membranes can be found in Appendix 9 i).

The results from the dot blot show that in non-infected cells PP4c is expressed in both A549 and HeLa cells. There is a higher presence of PP4c in the cytosol and solubilised membranes (the soluble fraction) than in nuclear debris, insoluble membranes and cytoskeleton components (insoluble fraction). There is PP4c expression at higher levels in the soluble fraction than in the insoluble fraction. There is also a higher level of SMEK1 expression in the soluble fraction rather than the insoluble fraction. There is more SMEK1 expression in the insoluble fraction in comparison to PP4c. SMEK1 is primarily found in the nucleus, as seen in Figures 13, 14 and 15. PP4c, however, is expressed more ubiquitously.
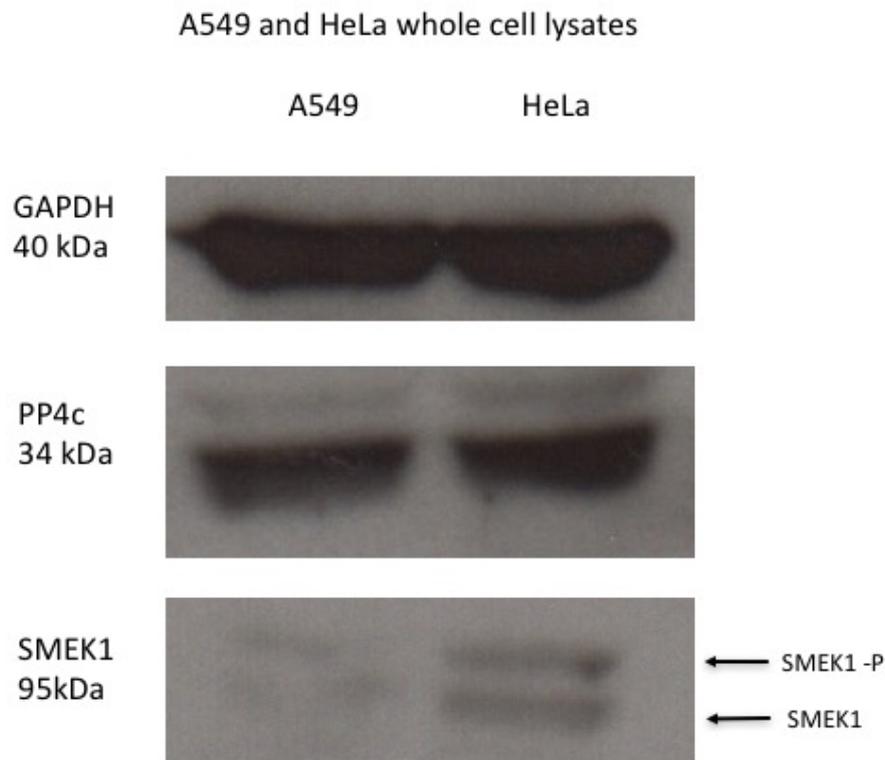
Figure 10. Expression of SMEK1 and PP4c using whole cell lysis technique (Chapter 2.7 Table 2). 10μl of cell lysates was loaded for this SDS-PAGE gel run. There is a higher level of expression of SMEK1 in HeLa cells compared to A549 lung epithelial cells. There are bands displaying PP4c in similar amounts in both HeLa and A549. (Exposure time and full membranes can be found in Appendix 10 i, ii and iii).

Non-infected A549 and HeLa whole cell lysates were run on an SDS-PAGE gel and analysed by western blot to confirm the detection of SMEK1 and PP4c. Figure 10 shows PP4c expression in both A549 and HeLa cells. GAPDH, a house-keeping protein, was used as a loading control (upper panel). There is a faint band for SMEK1 in A549, but bands for HeLa. This is could be due to lower SMEK1 expression in A549 cells.

There was also clear presence of double bands for SMEK1. These showed the presence of both phosphorylated forms of SMEK1, as described by Lipinzki *et al*. (2015), as the higher molecular weight band on western blot analysis of the SDS-PAGE gel. The presence of two bands indicated the expression of both phosphorylated and dephosphorylated forms of SMEK1. This denoted the presence of PP4c in cells, as it is known to dephosphorylate SMEK1 in a 1:1 ratio. This suggests dephosphorylation of SMEK1 by PP4c (Lipinszki *et al.*, 2015).

3.1.2.3 Effect of denaturation and reduction of disulphide bonds on detection of SMEK1 and PP4c

It was necessary to study the forms of SMEK1 and PP4c detected in different conditions through SDS-PAGE and western blot analysis. Using whole cell lysis technique (described in Chapter 2.7), the conditions of the lysates were altered to enable optimal detection. It was then possible to determine the effect of heat-denaturation and reduction of disulphide bonds on detection of monomers and PP4c/SMEK1 complexes. Lysates were prepared without reducing agents in the lysis buffer. Samples were then subjected to different conditions.
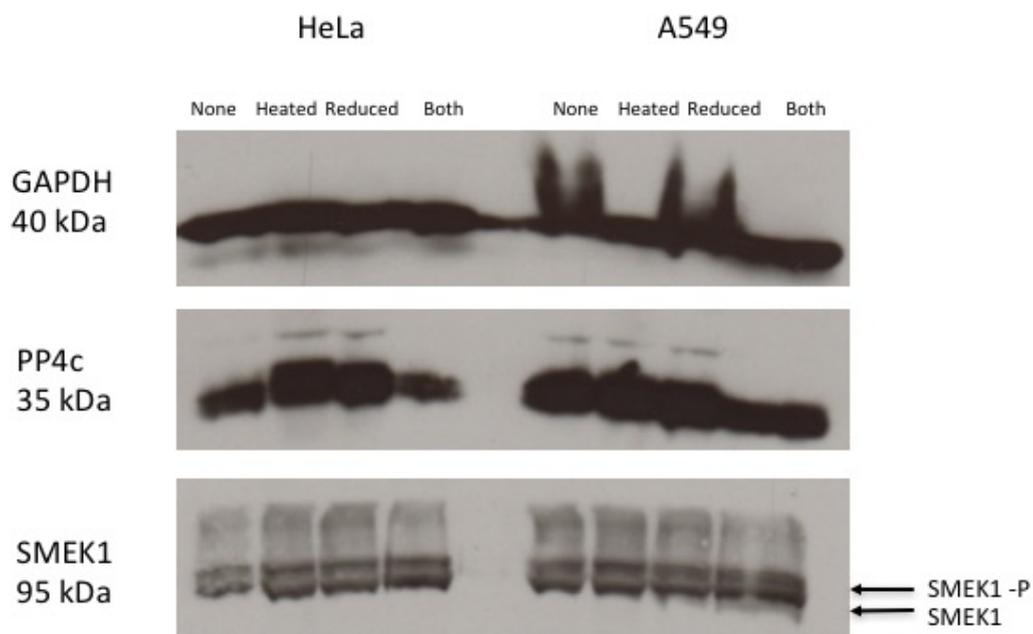
Figure 11. Western blot analysis of SDS-PAGE gel for differing treatments of A549 and HeLa lysate treatments. 'None' refers to no treatment except lysis buffer); 'Heated' indicates that samples were heated to 100°C for 5 minutes before loading; 'Reduced' samples were treated with B-mercaptoethanol to reduce disulphide bonds; 'Both' indicates samples were subjected to both heating and reducing treatments. (Exposure time and full membranes can be found in Appendix 11 i, ii and iii).

Figure 11 shows further analysis of SMEK1 and PP4c in A549 cells. There was no noticeable difference in SMEK1 levels across the differing treatments of A549 cell lysates. Figure 11 also confirmed the double bands shown in Figure 10; here this was also accompanied by a smear at a higher molecular weight. The larger of the double bands was due to the presence of both phosphorylated and non-phosphorylated forms of SMEK1 (Lipinszki *et al.*, 2015).

There appeared to be some reduction in signal in HeLa cell lysates when no treatment was performed on lysates, however this was not the case for A549 cells without treatment. This was most likely due to the difference in protein structure. Without heating or reducing treatments the tertiary structure of these proteins may have prevented antibody binding site exposure.

The presence of higher molecular weight bands in the SMEK1 labelled lysates could mean identification of high weight molecular complexes containing SMEK1 and/or PP4c in absence of denaturation or disulphide bonds reduction.

3.1.3 Effect of influenza A virus infection on SMEK1 and PP4c expression

PP4c and SMEK1 expression patterns were analysed following IAV A/WSN/33 infection of A549 cells over varying time points. The time points were chosen to enable visualisation of protein expression after viral binding to cellular membrane, post viral entry and after cycles of replication. It was hoped that IAV infection would influence the pattern of expression of SMEK1 and PP4c. This was described by both western blot and immunofluorescence.

3.1.3.1 Infection down-regulates SMEK1 expression in A549 cells

Lysates were taken using the fractionation lysis technique; the soluble fraction ran on SDS-PAGE gel.
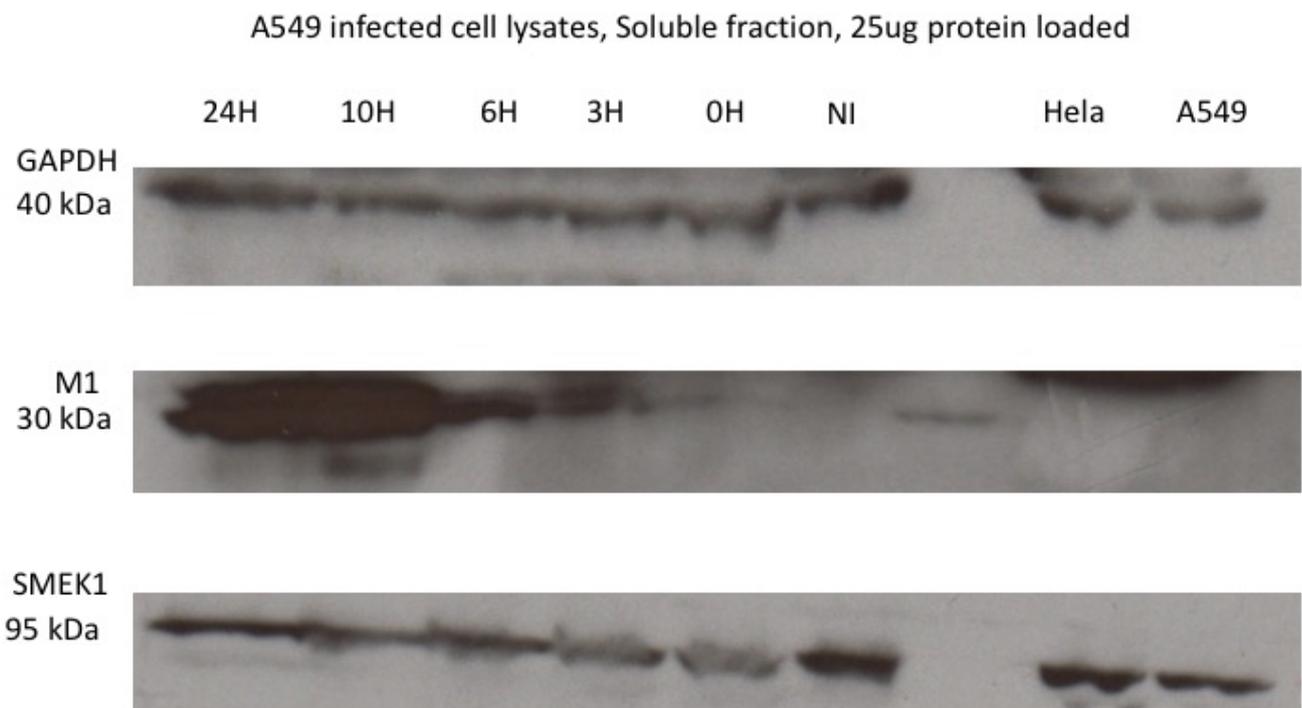
Figure 12. Western blot analysis of A549 cells infected at an MOI of 3 for 24 hours. Shown here are western blots for GAPDH (loading control), M1 (viral protein), SMEK1 (protein of interest). Antibodies used were: Mouse anti-GAPDH 1:2000 (Abcam), Mouse anti-M1 (1;1000), Rabbit anti-SMEK1 (1:2000) (Abcam). (Exposure time and full membranes can be found in Appendix 12 i, ii and iii).

Cell lysates were collected using fractionation lysis buffer after fixing the coverslips in PFA. Shown in Figure 12 is an increase in M1 as infection continued (from 3H to 24H). There was a drop in SMEK1 expression at 0H (where IAV binds to the cellular membrane, but does not enter the cell), and an expression recovered as infection continued to 24H. This also coincided with an increase in M1 from 0H

to 24H. This suggested that as infection continued and viral protein load increased with viral replication, there was also an increase in SMEK1 expression.

No double bonds for SMEK1 indicated that there was either a large amount of PP4c present, or a lack of PP4c. This could also be due to the different lysis buffer used for this technique. Double bands for SMEK1 indicated that there were both phosphorylated and dephosphorylated forms of SMEK1. There was a faint double band for SMEK1 for HeLa cell sample; this is most likely due to mid-transfer movement of the gel/membrane sandwich. However, it could indicate both phosphorylated forms of SMEK1 were present in only this sample. Further analysis is required to clarify this.

3.1.3.2 Immunofluorence Assay of Infected A549 cells

A549 lung epithelial cells were infected at varying multiplicities of infection (MOI) for varying time points to ascertain the effect of ongoing infection. Using immunofluorescence, it was possible to observe the localisation of expression.
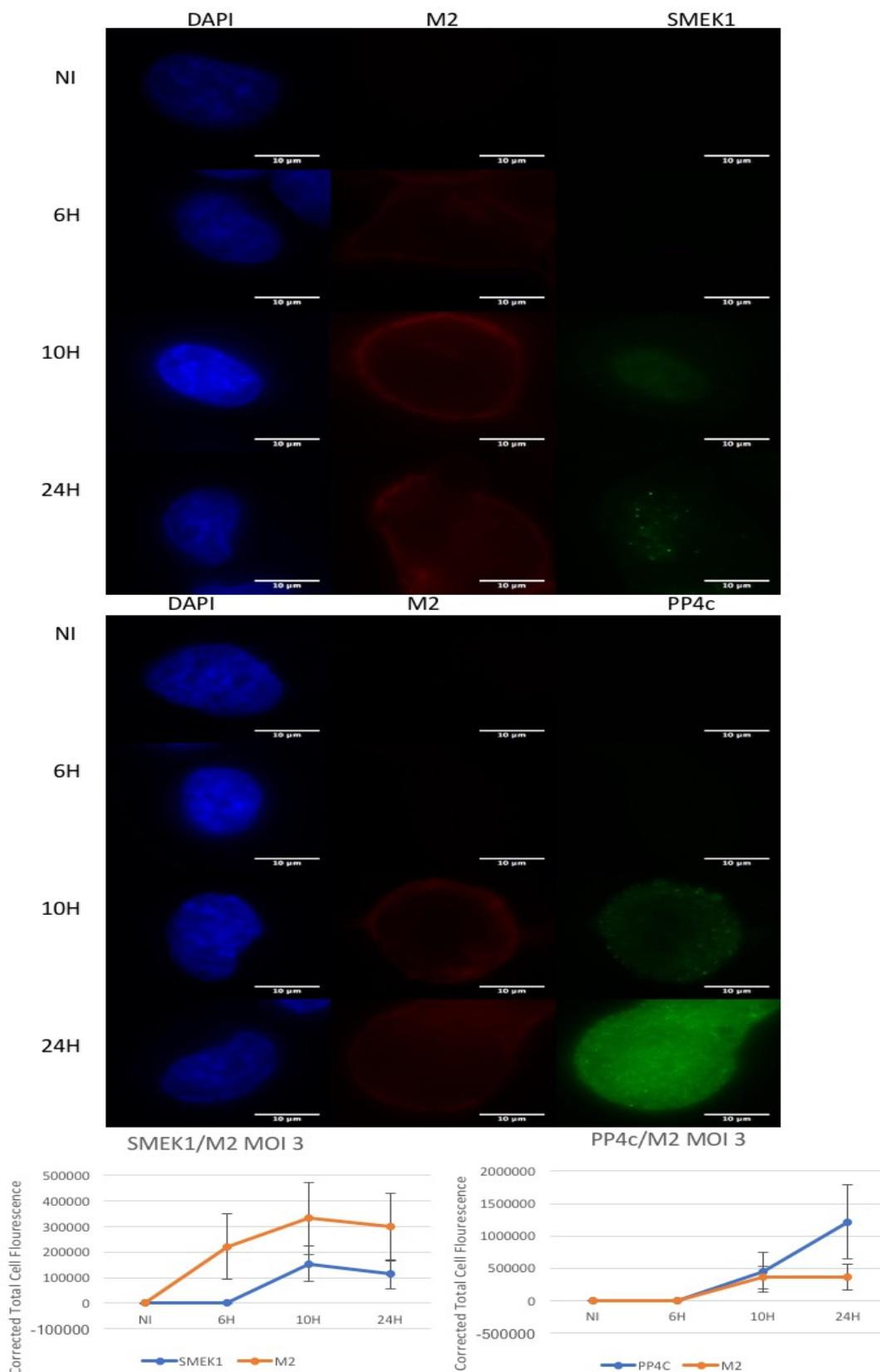
Figure 13. IFA analysis of MOI 3 infection of A549 lung epithelial cells at varying time points post-infection. Shown here is SMEK1 expression and PP4c expression at varying infection time points at MOI 3. DAPI staining of the nucleus is shown on the left, M2 stained with mouse anti-M2 primary antibody and Texas Red secondary antibody, SMEK1 stained with rabbit anti-pp4r3α (SMEK1) primary anti-body and anti-rabbit Alexa 488 secondary antibody, PP4c stained with Rabbit anti-PP4c and anti-rabbit Alexa 488 secondary antibody. Corrected total cell fluorescence was measured and plotted in the two bottom panels. Error bars show standard deviation of average corrected total cell fluorescence (CTCF).

The samples analysed by western blot (Figure 12) were studied by fluorescence microscopy (Figure 13). Little to no signal for PP4c and SMEK1 was found in non-infected (NI) and after 6 hours (6H) of infection. As this was seen in both SMEK1 and PP4c samples it is possible that this is due to very low relative expression levels of PP4c and SMEK1 in non-infected samples.

Interestingly, a sharp increase in PP4c expression was observed at 10 hours' post-infection and a further increase at 24 hours' post-infection (Figure 13). The A/WSN/33 lifecycle is approximately 6-8 hours. These results suggest that after one cycle of infection (at the 10-hour time point) infection caused an increase in expression of PP4c across the cell. There was a larger increase in expression at 24 hours, following approximately 3 cycles of infection. The expression pattern of PP4c at 10H is mainly cytosolic, however at 24H it appears that there is concentration of PP4c in the nucleus. This result requires confirmation, as a later IFA (Figure 14 and 15) did not show such a clear increase in expression. This may be due to differing MOI, or a different mounting medium being used (Vectashield vs Dabco), which allowed for better immunofluorescence and detection of PP4c expression in NI cells.

As shown in Figure 13, SMEK1 expression started to increase at 6 hours' post-infection. There was M2 expression, which corresponded with increase in SMEK1 expression. At 10 hours, post-infection there was a large increase in SMEK1 expression. Most of the SMEK1 was found in the nucleus (corresponding with DAPI staining), with some in the cytosol. At 24 hours, the level of expression of SMEK1 in the nucleus decreased, whilst there appeared to be an increase in SMEK1 in the cytosol. This suggested SMEK1 had translocated out of the nucleus. This would explain the decrease in corrected total cell fluorescence, as this measured overall cell fluorescence. Therefore, lower concentrated expression would appear to have reduced average expression.

The error bars on the bottom right panel in Figure 13 show that as PP4c expression levels increase so does the error. This coincides with an increase in error as M2 levels increase. This shows the variety of increase, most likely dependent on infection efficacy. Expression levels are dependent on efficacy of infection of all cells. If all cells are not infected, then this will skew results. This is also the case for M2 expression in the bottom left panel. This shows the variety in levels of infection.

This experiment was repeated, using a higher MOI of 10 (Figure 14). Cells were stained for M1 rather than M2 structural protein due to its higher prevalence at shorter time points; it therefore served as a better positive control of infection. Staining with M1 rather than M2 meant it would not be possible to see any co-localisation between M2 and SMEK1.
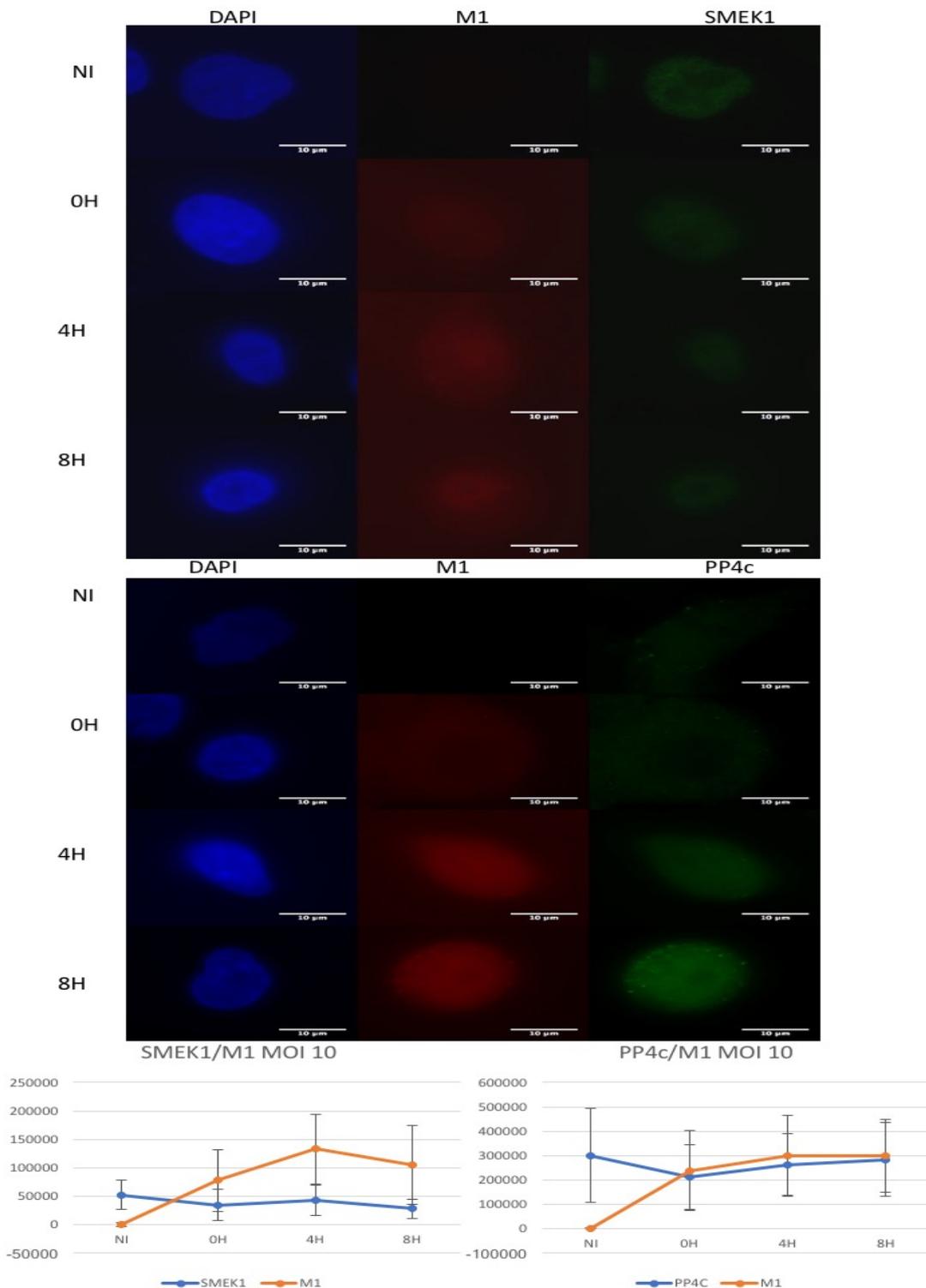
Figure 14. IFA analysis of MOI 10 infection of A549 lung epithelial cells at varying time points post-infection. Shown is PP4c and SMEK-1 expression NI (non-infected), 0H (viral binding only), 4H (4 hours post-infection), 8H (8 hours post-infection,) at MOI 10. DAPI staining of the nucleus is on the left. M1 was stained as an infection positive control with mouse anti-M1 primary antibody and Texas Red secondary antibody. PP4c was stained in the right-hand panels with rabbit anti-PP4c primary antibody and anti-rabbit Alexa 488 secondary antibody. SMEK1 was stained with rabbit anti-PP4R3α (SMEK1) primary antibody and anti-rabbit Alexa 488 secondary antibody. Corrected total cell fluorescence was measured and plotted in the two bottom panels. Error bars show standard deviation of average corrected total cell flourescence (CTCF).

In Figure 14 translocation of PP4c from the nucleus to the cytosol can be seen. The staining of PP4c from NI to 0H decreases, then increases after 8H but with similar distribution to M1. It was interesting to see detection of M1 from viral particles added to cells at 0H. This showed that an MOI of 10 had ensured that all cells were infected and showed the high ratio of viral particles on cells; this confirmed binding had been effective. At 0H time point the virus could bind to the cells, but not enter them. Binding of IAV appeared to influence PP4c expression within cells. This could be due to IAV triggering a cascade of events by binding to PRRs at the cell membrane (Yoo *et al.*, 2013).

At 4H and 8H post-infection, the pattern of expression of PP4c had changed. The expression pattern was like that of M1; this could be seen at higher levels in the nucleus at 4H. The expression of PP4c in NI cells was much more distinctive; there was defined expression in the cytosol and less expression in the nucleus. Following infection this changed. PP4c expression became more nucleic; it was possible to identify where the nucleus was by looking at PP4c expression. Here we can clearly see nuclear and cytosolic localisation of PP4c.

SMEK1 expression followed a similar pattern of increase as PP4c during IAV infection. The levels to which they increased were not as dramatic as PP4c. There was little translocation of SMEK1 from the nucleus to the cytosol. To quantify the degree of translocation would require further analysis. There was, however, less SMEK1 expression in the cytosol in NI cells when compared to 8H of infection. This did not seem to be as significant as that seen in PP4c cells.

It appeared there was some similar distribution of M1 and PP4c post-infection. However, this would need to be analysed by confocal microscopy to confirm and give better resolution. Similar distribution patterns are likely to have been due to leakage between channels due to the lack of similar expression patterns of M1 in the SMEK1 samples. There were also similar artefacts found in both samples. This experiment will need to be repeated to confirm these results. If there is indeed co-localisation between M1 and PP4c, this presents a novel interaction that may have further implications for IAV infection.

In this experiment, influence of infection on SMEK1 expression was only detected after one cycle of infection was completed. However, for the effect on PP4c to be noticed, only one cycle of IAV infection was required to see a difference in expression patterns.

The error bars (bottom left panel) for M1 are larger than SMEK1; this represents the variety in levels of infection. Cells allowing better viral replication will show higher expression levels of M1. The error bars for PP4c (bottom right panel) are much larger, again representing much larger variety in expression, especially in NI cells.
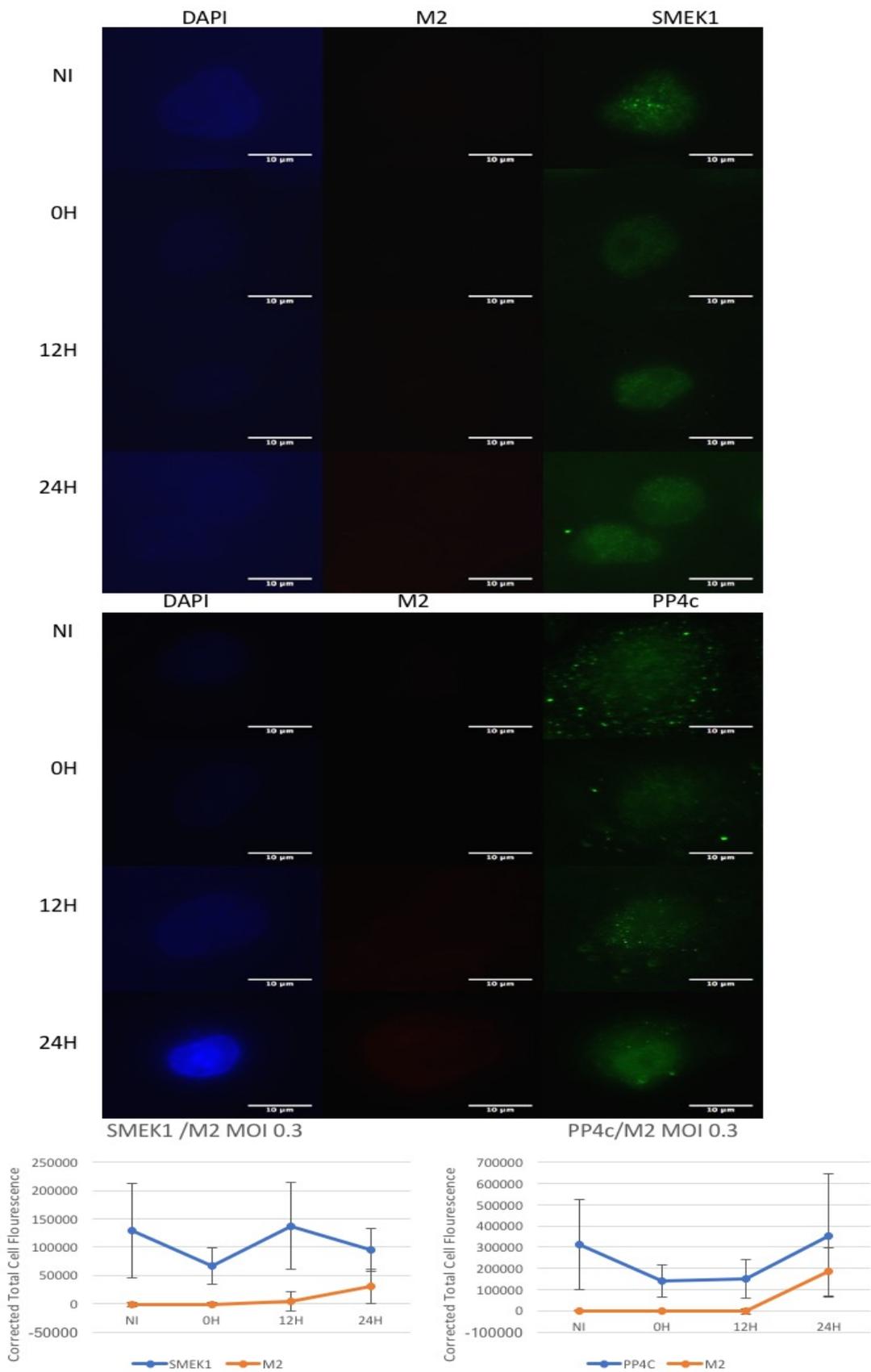
Figure 15. IFA analysis of MOI 0.3 infection of A549 lung epithelial cells at varying time points post-infection. Shown are viral protein M2 and human protein PP4c and SMEK1 expression across 4 time points: non-infected cells (NI), viral binding (0H), 12 hours' post-infection (12H) and 24 hours' post-infection (24H). Corrected total cell fluorescence was measured and plotted in the two bottom panels Error bars show standard deviation of average corrected total cell fluorescence (CTCF).

To investigate the effect of infection using a low MOI, a similar IFA was repeated using an MOI of 0.3. As expected, M2 signal could only be efficiently detected after several rounds of infection at 24H (Figure 15). This experiment cells confirmed an increase in PP4c expression at 24 hours' post-infection, after a decrease following viral binding to cell surface. At 24H of infection PP4c expression levels were slightly higher than NI and much higher than at 0H. This showed an elevated level of PP4c in cells expressing more of the viral protein M2, even at a lower MOI than shown before (MOI 0.3 as opposed to MOI 3 shown in Figure 10). The advantage of this lower MOI 0.3 for this experiment was to show an increase in PP4c correlated to the level of viral proteins within infected cells.

The levels of SMEK1 expression were very interesting. Following IAV binding at 0H SMEK1 and PP4c levels decreased. After 12H of infection SMEK1 levels increased, but PP4c levels did not. This followed an increase in M2 expression levels from 12H to 24H. At 24H post-infection, SMEK1 expression levels decreased. The increase in M2 expression levels was not as dramatic as that seen in the PP4c sample and could be partially responsible for the decrease in SMEK1. This could be because the coverslip with cells for this sample did not experience a similar level of infection of IAV.

Figure 15 confirmed western blot analysis (Figure 12) of SMEK1 expression patterns; following viral binging to cell surface, SMEK1 levels are down-regulated. As infection continues SMEK1 levels increase. However, in Figure 15 SMEK1 levels decreased again at 24H. PP4c expression patterns were also confirmed by IFA in Figures 13 and 15; non-infected expression is mainly cytosolic, following 24H of infection, PP4c is found in both the cytosol and the nucleus. The expression increase of PP4c found at 24H of infection was found to be relative to the MOI of infection. A higher MOI caused a more dramatic increase in PP4c expression. This suggests that the increase is due to viral infection.

3.1.4 Co-Immunoprecipitation of SMEK1 and M2 in infected A549 cells

An *in vivo* interaction needed to be proved. IAV infected A549 cell lysates were probed for SMEK1 and M2 binding. Cell lysates were incubated with IgG antibodies anti-M2, anti-PP4R3α (SMEK1), anti-AnnexinA6, anti-PP4c and anti-NFkB. Annexin A6 and NFkB were included as controls. PP4c is known to bind to SMEK1 as it forms a complex. NFkB and Annexin A6 were used as control rabbit antibodies. Complexes were recovered by adsorption on protein A/G-sepharose beads and analysis by SDS-page gel electrophoresis and western blot. This allowed confirmation of the results of the yeast two-hybrid screen *in vivo*.
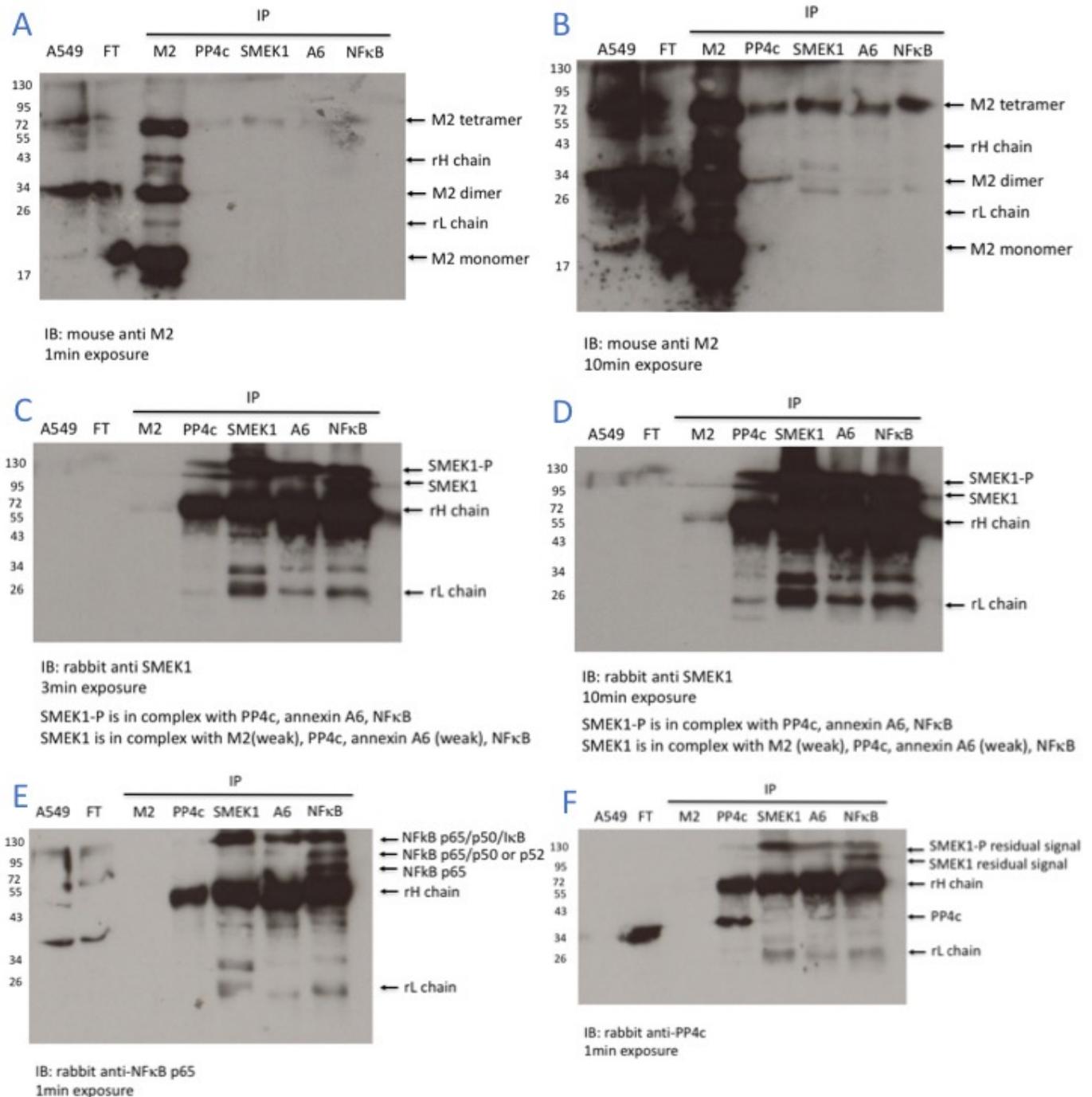
Figure 16. Western Blot analysis following SDS-PAGE gel co-immunoprecipitation of 24 hours infected A549 lung epithelial cell lysates. The lanes indicated are (left – right): A549 cell lysates (A549); flow through from pre-clearing of infected A549 cell lysates (FT); then samples that were co-immunoprecipitated (IP) by incubation with mouse anti-M2 (M2); rabbit anti-PP4c (PP4c); rabbit anti-SMEK1 (SMEK1); rabbit anti-Annexin A6 (A6); and rabbit anti-NFkB (NFkB) antibodies. Membranes were then immunoblotted with anti-M2 (A and B), anti-SMEK1 (C and D), anti-NFkB (E), and anti-PP4c (F). (Full membranes can be found in Appendix 16 i, ii, iii and iiii).

M2 tetramer co-immunoprecipitates with PP4c, SMEK1, Annexin A6 and NFkB

Immunoblot with anti-M2 antibody shows the proteins by which M2 is co-immunoprecipitated with, there are bands in all samples for M2 tetramer and dimer (Figure 16A and B). There are stronger bands for

the M2 tetramer, this shows poor denaturation of the M2 complex whilst processing samples and a more efficient co-immunoprecipitation of the tetramer. Interestingly all proteins tested could immunoprecipitate M2 tetramers. This suggests formation of multimolecular complexes between IAV M2 and PP4c, SMEK1, Annexin A6 and NFkB.

Differential interaction of SMEK1 and phosphorylated SMEK1 with M2, PP4c, Annexin A6 and NFkB

Immunoblot with anti-SMEK1 antibody shows the proteins SMEK1 is co-immunoprecipitated with (Figure 16C). There was a band at 95 kDa for SMEK1 in the sample incubated (16D) with mouse anti-M2 antibody, which indicates interaction. After longer exposure (Figure 16D), this band is clearer and shows dephosphorylated SMEK1 co-immunoprecipitates with M2.

This result indicates that M2 only interacts with de-phosphorylated SMEK1. This confirms an interaction between SMEK1 and M2 in A549 cells. Previous work by *Ma et al. (2012)* required up-regulation of host protein expression to get clearly defined bands for co-immunoprecipitation. Therefore, it may be necessary to repeat this following upregulation of SMEK1 for clearer results.

It was interesting to note that Annexin A6 only interacts with the phosphorylated form of SMEK1, while PP4c and NFkB bind to both phosphorylated and non-phosphorylated SMEK1. This result further confirms the formation of multimolecular complexes and indicates that SMEK1 and phosphorylated SMEK1 are differentially involved in these complexes.

NFkB interacts with SMEK1 and Annexin A6; PP4c interacts with the Annexin A6 and NFkB

Interestingly, NFkB was immunoprecipitated with SMEK1 and Annexin A6 (Figure 16E). A high molecular weight band could be detected with the anti-NFkB p65 antibody. It was reasoned that these bands correspond to the complexes involving the p65 subunit of NFkB, most likely p65/p50 or p52 and p65/p50 or p52 and IkB.

Weak bands could be observed for PP4c in lanes corresponding to immunoprecipitation of SMEK1, Annexin A6 and NFkB. These data not only confirm the interaction between M2 and SMEK1 (non-phosphorylated form) in infected cells, but also present novel interactions between the PP4 complex, Annexin A6 and NFkB.

To ensure that the detection of SMEK1 in M2 co-immunoprecipitated samples was not due to leakage from neighbouring lanes, the samples were run again on a SDS-PAGE gel with lanes left empty between each sample (Figure 17).
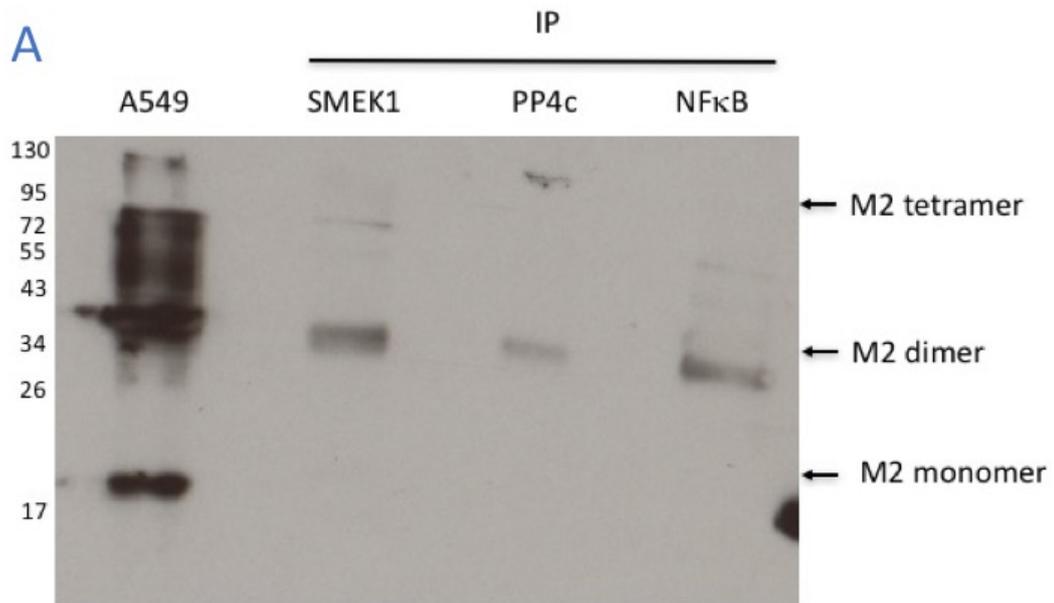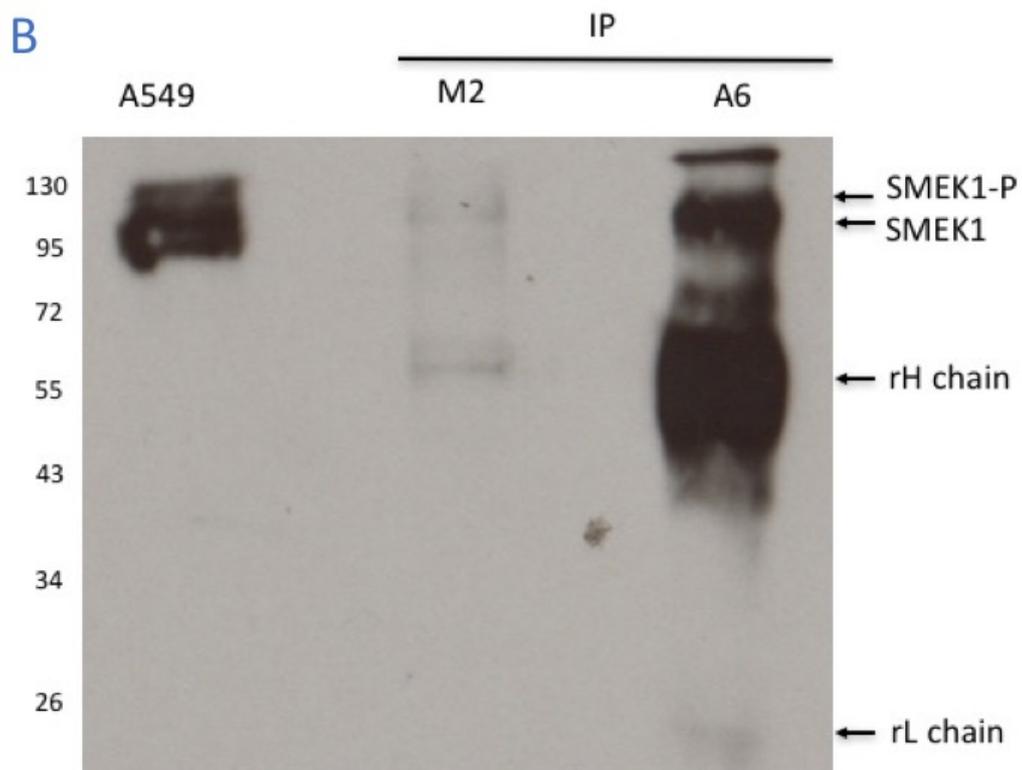
Figure 17. Western blot analysis following SDS-PAGE gel of co-immunoprecipitation. The lanes indicated are (left – right): Panel A: infected A549 cell lysates (A549); rabbit anti-SMEK1 (SMEK1); rabbit anti-PP4c (PP4c); rabbit anti-NFkB (NFkB) antibodies. Immunoblotted with mouse anti-M2 monoclonal antibody. Panel B: infected A549 cell lysates (A549); incubated with mouse anti-M2 (M2); and mouse anti-Annexin A6 (A6). Immunoblotted with rabbit anti-SMEK1 monoclonal antibody. (Exposure time and full membranes can be found in Appendix 17 i and ii).

Figure 17 shows a rerun of the samples in Figure 16. Gaps were left between samples to ensure that there was no leakage of samples between lanes. There is co-immunoprecipitation of M2 dimer in SMEK1, PP4c and NFkB precipitated samples at 34 kDa (Figure 17A). This confirmed there was an interaction in infected A549 cells between SMEK1 and M2. This confirms the results shown in Figure 16B. As there is a band found at 34kDa in the A549 lysate sample (which contains no antibodies), this precludes to poor denaturisation of the samples. It is expected that the monomer of M2 would be separated even if denaturation had worked efficiently. The monomer is expected to be found at 17 kDa, therefore a dimer would be expected at around 34 kDa. Therefore, it is possible that instead of separating M2 as a monomer, a dimer has been separated in the gel. This would confirm the interaction between M2 and SMEK1, whilst also identifying an interaction between M2 and PP4c and NFkB. To confirm this, thorough denaturation of samples would be required.

Figure 17B also confirms the co-immunoprecipitation between SMEK1 and Annexin A6. This presents a novel interaction between Annexin A6 and SMEK1. Further research is required to confirm and describe the implications of this. Interaction with M2 was also detected, although it was weaker.

3.2 Discussion

3.2.1 Conclusions

This work is the first to describe an interaction between influenza A virus M2 and human SMEK1 in infected lung epithelial cells.

The overlap between the M2 binding domain and ARM repeat domain (shown in Figure 7), may influence the ability of SMEK1 to bind to PP4 in infected cells. The ARM repeat was shown to be important for PP4 recognition for SMEK1 binding. This requires further investigation.

There is some evidence of SMEK1 translocation from the nucleus to the cytosol (Figures 13, 14, 15). This requires further quantification. If this is proven it could have implications for the effect on IAV infection on expression patterns of SMEK1 after infections. Measurements of concentration of SMEK1 in the nucleus and in the cytosol would need to be taken before and after infection. Western blot analysis of nuclear fractionation of infected A549 cell lysates could confirm the translocation observed.

Western blot results showed that there was an initial decrease in SMEK1 expression, but this returned to normal levels as infection continued (Figure 12). Repeats are required to confirm these results. Western blot and immunofluorescence analysis of shorter infection time points could help explain and describe the expression profile of SMEK1 post infection.

The bands for M2 dimer and tetramers in the anti-SMEK1 antibody immunoprecipitated samples showed an interaction between SMEK1 and M2. The probability of this reaction being transient increased with the observation that SMEK1 was not co-immunoprecipitated as well with anti-M2 antibody incubated samples. The co-immunoprecipitation of SMEK1 by M2 gave signs of an interaction with the non-phosphorylated form of SMEK1. To confirm the presence of an interaction between M2 and SMEK1, it may be necessary to upregulate expression of SMEK1 in A549 cells. Subsequent co-immunoprecipitation would give better and clearer results; in theory, stronger bands would be produced after an increase in SMEK1 expression. From the same Yeast two-hybrid screen, Ma *et al (2012)* used this method of upregulation of human cellular factors to confirm co-immunoprecipitation with viral proteins.

Through co-immunoprecipitation a potential complex was identified that has not previously been described (Figure 18). Here we have hypothesised that the interactions identified may form part of a complex. It is known that the PP4 complex can interact with p65 subunit of the NFkB complex, and that complex can interact with Annexin A6. The interaction between SMEK1 and p65 subunit has identified how the PP4 and NFkB complexes interact. As it was shown that only the phosphorylated form of SMEK1 interacts with M2, NFkB and Annexin A6, it may be that SMEK1 interacts with these factors in the absence of PP4c. The dephosphorylated form of SMEK1 was identified to interact with NFkB, Annexin A6 and PP4c, therefore it may be that when PP4c is in complex with SMEK1, PP4c can dephosphorylate SMEK1 and cause dissociation from M2.
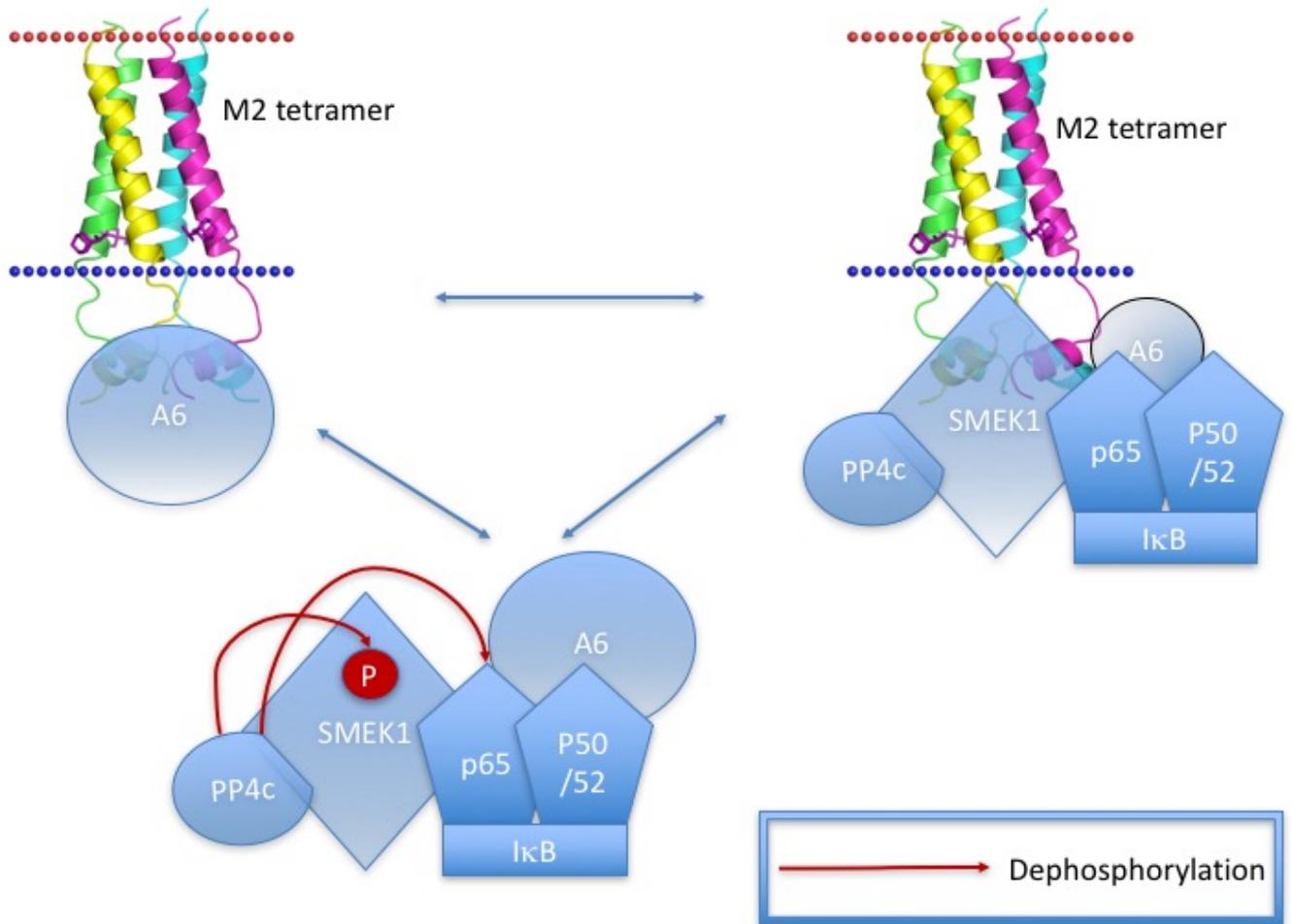
Figure 18. Schematic diagram of the proposed interactions as found by this study. Shown are the potential forms in which the cellular factors investigated could interact. Annexin A6 (A6) is shown to interact with the CT tail of transmembrane M2 protein. The proposed interaction between dephosphorylated SMEK1 and the CT tail of transmembrane M2 in complex with PP4c, the NFkB complex (IkB, p65 and p50/p52), and potentially Annexin A6. The phosphorylated form of SMEK1 is shown to interact with PP4c, Annexin A6, and the NFkB complex. Red arrows signify potential dephosphorylation by PP4c.

There was an impact of infection on the catalytic unit (PP4c) of the PP4 complex that SMEK1 forms. An increase in PP4c expression has been confirmed in infected A549 cells. The accumulation of PP4c may have an advantageous effect for the host. Increase in PP4c causes an increase in dephosphorylation of TBK1 at Ser-172, leading to TBK1 inhibition (Zhan *et al.*, 2015). A decrease in activated TBK1 would cause a decrease of phosphorylation of IRF3 and IRF7, this could decrease type-1 IFN and IFN-stimulated induction. This would then lead to a decreased inhibition of viral replication and inhibited clearance of infected cells (Figure 5). Therefore, this increase in PP4c is a negative host cell response to viral infection.

The increase in PP4 production post-IAV infection (Figure 13) needs to be confirmed; the same expression pattern was not seen at a lower MOI (Figure 15). If this increase were to be confirmed, it would present implications for the understanding of cellular reactions to IAV infection. PP4 is known to be involved in IFN signalling (Zhan *et al.*, 2015).

The implications of an increase in PP4 post-infection could have one of two effects. This could be due an accumulation of PP4, whereby the function is blocked by viral blockage of the cell cycle. It was described by *Fan et al. (2017)* that IAV halts the cell cycle. Due to PP4c role in the cell cycle, it is possible that IAV-induced cell cycle blockage has led to an accumulation of PP4c during infection.

There is also the possibility that the interaction between M2 and SMEK1 could be a cellular response to viral infection. Post-infection host cells increase PP4 production as a part of immune response to increase IFN expression and clear IAV infected cells by inducing an anti-viral state using IFN stimulated genes.

3.2.2 Future Directions

Following confirmation of the interaction, the reasons for M2 and SMEK1 binding would be investigated to discover the physiological relevance of this interaction. Interactions between SMEK1 and NFkB as well as SMEK1 and Annexin A6 were identified through this co-immunoprecipitation assay. SMEK1 has been found to down-regulate NFkB (Dong *et al.*, 2012; Byun *et al.*, 2012). However, PP4 was found to activate NFkB (Martin-Granados *et al.*, 2008; Yeh *et al.*, 2004). Further clarification of the roles of the subunits in NFkB regulation is required. It is possible SMEK1's role is independent of the PP4 complex, the role of SMEK1 and PP4 require further investigation. This co-immunoprecipitation has presented new interactions, not previously described in literature; and confirmed those already known. These interactions should be confirmed, presenting new projects to understand SMEK1 interactions in human cells.

The cause for the PP4c increase seen at 24 hours' post-infection should be investigated. This could be done by testing the production of cytokines related to IFN response, to viral infection, at varying time points by qPCR. This would clarify the implications of the increase in PP4 on viral immune response. A change in immunocytokine production could produce a very interesting and novel piece of work to clarify the host response to viral infection.

The effect of this bond on viral particle production could be ascertained by down-regulating PP4 expression in A549 cells. It would be possible to determine whether PP4 is a restricting factor for IAV infection. Then test the effects on viral pathogenesis using siRNA down-regulation of PP4 expression in A549 lung epithelial cell line. This would be done following testing by western blot to confirm down-regulation of PP4 expression. Following this a test for the effectiveness of IAV infection in both A549 and A549 PP4c-downregulated cells should be performed using several methods.

− $TCID_{50}$ would allow an assessment of the viral pathogenesis on these different cell lines. This would allow the calculation of the effect of PP4c-downregulation on IAV.

–   Western blot and flow cytometry can confirm the levels of viral proteins in both infected cell lines.

–   IFA could be used to visualize changes following PP4c-downregulation to clarify whether PP4c depletion influences viral progeny. IFA allows visualisation of the location of proteins within the cell. Therefore, any differences in viral protein localisation could be found easily.

It would then be necessary to elucidate the effects on other systems within the cell following IAV infection and subsequent binding to SMEK1. Protein phosphatase PP4 has many functions in many different cellular processes. There is a strongly defined role of PP4c in cell cycle and microtubule organization, as well as a role in apoptosis. This could be affected by the up-regulation in PP4 found at 24 hours' post-infection. Therefore, it would be prudent to see what effect increased PP4 has on these functions post-infection.

It may be possible to discover more about the function of SMEK1 and its role in the cell through its relationship with M2. By testing for the effects of an increase of PP4 on the functions of PP4, it may be possible to find new functions previously undiscovered. It is conceivable that a novel role for PP4 may be found in the host viral immune response.

3.2.2.1 Substrates

By analyzing the effect of this interaction, it is important to investigate the effect of infection on the substrates of PP4c. TBK1 is dephosphorylated by PP4c (Zhan *et al.*, 2015); an accumulation of PP4c post infection could has a negative effect on the signalling pathway for IFN-1 production. This would be beneficial for IAV as IFN-1 is essential for viral clearance.

The role of PP4 in the cell cycle is much better characterised than its role in apoptosis. p27 is a substrate of PP4c, p27 is a CD1 inhibitor. PP4 allows progression through the cell cycle by activating p27 and inhibiting CD1 and CDK4 (Dong *et al.*, 2012). An increase in PP4c could influence the progression of cells through G2. However, it is known that infection blocks progression of the cell cycle (Fan *et al.*, 2017). There is an interaction between PP4 and CDK1, involved in microtubule organisation. PP4 regulates CDK1 negatively; this prevents unscheduled activation of CDK1 throughout the cell cycle. It is thought that due to its dominance over CDK1, PP4 could be a checkpoint protein. NDEL1 (a substrate of CDK1) recruits katanin p60, which is heavily involved in microtubule dynamics and remodelling (Toyo-oka *et al.*, 2008). An increase of PP4c following infection could have a negative effect on CDK1. This could lead to cell cycle arrest

It is known that PP4 interacts with RhoGTPases. The regulation of RhoGTPases by PP4-R2-R3 (R3 being SMEK1) complex may co-ordinate centrosome maturation and cell migration (Martin-Granados et al., 2008). It may be hard to quantify an effect of the up-regulation of PP4c on these processes. However, the effect on RhoGTPases can be observed.

.

References

Amorim, M.J., Bruce, E.A., Read, E.K., Foeglein, A., Mahen, R., Stuart, A.D. and Digard, P. (2011) 'A Rab11- and microtubule-dependent mechanism for cytoplasmic transport of influenza A virus viral RNA', *Journal of virology,* 85(9), pp. 4143-4156.

Banerjee, I., Yamauchi, Y., Helenius, A. and Horvath, P. (2013) 'High-content analysis of sequential events during the early phase of influenza A virus infection', *PloS one,* 8(7), pp. e68450.

Baum, L. and Paulson, J. (1990) 'Sialyloligosaccharides of the respiratory epithelium in the selection of human influenza-virus receptor specificity', *Acta Histochemica,* , pp. 35-38.

Brechmann, M., Mock, T., Nickles, D., Kiessling, M., Weit, N., Breuer, R., Muller, W., Wabnitz, G., Frey, F., Nicolay, J.P., Booken, N., Samstag, Y., Klemke, C.D., Herling, M., Boutros, M., Krammer, P.H. and Arnold, R. (2012) 'A PP4 holoenzyme balances physiological and oncogenic nuclear factor-kappa B signaling in T lymphocytes', *Immunity,* 37(4), pp. 697-708.

Byun, H., Kim, B., Yoo, R., Park, S. and Rho, S.B. (2012) 'sMEK1 enhances gemcitabine anti-cancer activity through inhibition of phosphorylation of Akt/mTOR', *Apoptosis : An International Journal on Programmed Cell Death,* 17(10), pp. 1095-1103.

Chen, B.J., Leser, G.P., Jackson, D. and Lamb, R.A. (2008a) 'The influenza virus M2 protein cytoplasmic tail interacts with the M1 protein and influences virus assembly at the site of virus budding', *Journal of virology,* 82(20), pp. 10059-10070.

Chen, G.I., Tisayakorn, S., Jorgensen, C., D'Ambrosio, L.M., Goudreault, M. and Gingras, A.C. (2008b) 'PP4R4/KIAA1622 forms a novel stable cytosolic complex with phosphoprotein phosphatase 4', *The Journal of biological chemistry,* 283(43), pp. 29273-29284.

Couceiro, J.N.S.S., Couceiro, J.N.S.S., Paulson, J.C., Paulson, J.C., Baum, L.G. and Baum, L.G. (1993) 'Influenza virus strains selectively recognize sialyloligosaccharides on human respiratory epithelium; the role of the host cell in selection of hemagglutinin receptor specificity', *Virus research,* 29(2), pp. 155-165.

de Vries, E., Tscherne, D.M., Wienholts, M.J., Cobos-Jimenez, V., Scholte, F., Garcia-Sastre, A., Rottier, P.J. and de Haan, C.A. (2011) 'Dissection of the influenza A virus endocytic routes reveals macropinocytosis as an alternative entry pathway', *PLoS pathogens,* 7(3), pp. e1001329.

Dong, S.M., Byun, H.J., Kim, B.R., Lee, S.H., Trink, B. and Rho, S.B. (2012) 'Tumor suppressor BLU enhances pro-apoptotic activity of sMEK1 through physical interaction', *Cellular signalling,* 24(6), pp. 1208-1214.

Eierhoff, T., Hrincius, E.R., Rescher, U., Ludwig, S. and Ehrhardt, C. (2010) 'The epidermal growth factor receptor (EGFR) promotes uptake of influenza A viruses (IAV) into host cells', *PLoS pathogens,* 6(9), pp. e1001099.

Eisfeld, A., Neumann, G. and Kawaoka, Y. (2015) 'At the centre: influenza A virus ribonucleoproteins', *NATURE REVIEWS MICROBIOLOGY,* 13(1), pp. 28-41.

Fan, Y., Mok, C.K., Chan, M.C.W., Zhang, Y., Nal-Rogier, B., Kien, F., Bruzzone, R. and Sanyal, S. (2017) 'Cell Cycle Independent Role of Cyclin D3 in Host Restriction of Influenza Virus Infection', *Journal of Biological Chemistry,* , pp. jbc.M117.776112.

Gingras, A.C., Caballero, M., Zarske, M., Sanchez, A., Hazbun, T.R., Fields, S., Sonenberg, N., Hafen, E., Raught, B. and Aebersold, R. (2005) 'A novel, evolutionarily conserved protein phosphatase complex involved in cisplatin sensitivity', *Molecular & cellular proteomics : MCP,* 4(11), pp. 1725-1740.

Gorai, T., Goto, H., Noda, T., Watanabe, T., Kozuka-Hata, H., Oyama, M., Takano, R., Neumann, G., Watanabe, S. and Kawaoka, Y. (2012) 'F1Fo-ATPase, F-type proton-translocating ATPase, at the plasma membrane is critical for efficient influenza virus budding', *Proceedings of the National Academy of Sciences of the United States of America,* 109(12), pp. 4615-4620.

Holt, P.G., Strickland, D.H., Wikstrom, M.E. and Jahnsen, F.L. (2008) 'Regulation of immunological homeostasis in the respiratory tract', *Nature reviews.Immunology,* 8(2), pp. 142-152.

Horimoto, T. and Kawaoka, Y. (2005) 'Influenza: lessons from past pandemics, warnings from current incidents', *Nature reviews.Microbiology,* 3(8), pp. 591-600.

Kido, H., Okumura, Y., Takahashi, E., Pan, H.Y., Wang, S., Chida, J., Le, T.Q. and Yano, M. (2008) 'Host envelope glycoprotein processing proteases are indispensable for entry into human cells by seasonal and highly pathogenic avian influenza viruses', *Journal of molecular and genetic medicine : an international journal of biomedical research,* 3(1), pp. 167-175.

Killip, M., Smith, M., Jackson, D. and Randall, R. (2014) 'Activation of the Interferon Induction Cascade by Influenza A Viruses Requires Viral RNA Synthesis and Nuclear Export', *Journal of virology,* 88(8), pp. 3942-3952.

Li, G., Zhang, J., Tong, X., Liu, W. and Ye, X. (2011) 'Heat shock protein 70 inhibits the activity of Influenza A virus ribonucleoprotein and blocks the replication of virus in vitro and in vivo', *PloS one,* 6(2), pp. e16546.

Liao, F., Hsiao, W., Lin, Y., Chan, Y. and Huang, C. (2016) 'T cell proliferation and adaptive immune responses are critically regulated by protein phosphatase 4', *Cell Cycle,* 15(8), pp. 1073.

Lipinszki, Z., Lefevre, S., Savoian, M.S., Singleton, M.R., Glover, D.M. and Przewloka, M.R. (2015) 'Centromeric binding and activity of Protein Phosphatase 4', *Nature communications,* 6, pp. 5894.

Lyu, J., Jho, E.H. and Lu, W. (2011) 'Smek promotes histone deacetylation to suppress transcription of Wnt target gene brachyury in pluripotent embryonic stem cells', *Cell research,* 21(6), pp. 911-921.

Lyu, J., Kim, H., Yamamoto, V., Choi, S.H., Wei, Z., Joo, C. and Lu, W. (2013) 'Protein phosphatase 4 and Smek complex negatively regulate Par3 and promote neuronal differentiation of neural stem/progenitor cells', *Cell reports,* 5(3), pp. 593-600.

Ma, H., Kien, F., Maniere, M., Zhang, Y., Lagarde, N., Tse, K.S., Poon, L.L. and Nal, B. (2012) 'Human annexin A6 interacts with influenza a virus protein M2 and negatively modulates infection', *Journal of virology,* 86(3), pp. 1789-1801.

Martin-Granados, C., Philp, A., Oxenham, S.K., Prescott, A.R. and Cohen, P.T. (2008) 'Depletion of protein phosphatase 4 in human cells reveals essential roles in centrosome maturation, cell migration and the regulation of Rho GTPases', *The international journal of biochemistry & cell biology,* 40(10), pp. 2315-2332.

McAvoy, T. and Nairn, A.C. (2010) 'Serine/threonine protein phosphatase assays', *Current protocols in molecular biology / edited by Frederick M.Ausubel ...[et al.],* Chapter 18, pp. Unit18.18.

McCown, M.F. and Pekosz, A. (2006) 'Distinct domains of the influenza a virus M2 protein cytoplasmic tail mediate binding to the M1 protein and facilitate infectious virus production', *Journal of virology,* 80(16), pp. 8178-8189.

Mendoza, M.C., Du, F., Iranfar, N., Tang, N., Ma, H., Loomis, W.F. and Firtel, R.A. (2005) 'Loss of SMEK, a novel, conserved protein, suppresses MEK1 null cell polarity, chemotaxis, and gene expression defects', *Molecular and cellular biology,* 25(17), pp. 7839-7853.

Mendoza, M.C., Booth, E.O., Shaulsky, G. and Firtel, R.A. (2007) 'MEK1 and Protein Phosphatase 4 Coordinate Dictyostelium Development and Chemotaxis', *Molecular and cellular biology,* 27(10), pp. 3817-3827.

Mihindukulasuriya, K.A., Zhou, G., Qin, J. and Tan, T. (2004) 'Protein phosphatase 4 interacts with and down-regulates insulin receptor substrate 4 following tumor necrosis factor-α stimulation', *Journal of Biological Chemistry,* 279(45), pp. 46588-46594.

Millet, J.K., Peiris, J.S.M., Altmeyer, R.M., Nal, B., Kien, F., Cheung, C., Siu, Y., Chan, W., Li, H., Leung, H., Jaume, M. and Bruzzone, R. (2012) 'Ezrin interacts with the SARS coronavirus Spike protein and restrains infection at the entry stage', *PloS one,* 7(11), pp. e49566.

Mourtada-Maarabouni, M., Kirkham, L., Jenkins, B., Rayner, J., Gonda, T.J., Starr, R., Trayner, I., Farzaneh, F. and Williams, G.T. (2003) 'Functional expression cloning reveals proapoptotic role for protein phosphatase 4', *Cell death and differentiation,* 10(9), pp. 1016-1024.

Mourtada-Maarabouni, M. and Williams, G.T. (2009) 'Protein phosphatase 4 regulates apoptosis in leukemic and primary human T-cells', *Leukemia research,* 33(11), pp. 1539-1551.

Muller, K.H., Kakkola, L., Nagaraj, A.S., Cheltsov, A.V., Anastasina, M. and Kainov, D.E. (2012) 'Emerging cellular targets for influenza antiviral agents', *Trends in pharmacological sciences,* 33(2), pp. 89-99.

Nayak, D.P., Balogun, R.A., Yamada, H., Zhou, Z.H. and Barman, S. (2009) 'Influenza virus morphogenesis and budding', *Virus research,* 143(2), pp. 147-161.

Perez, J.T., Varble, A., Sachidanandam, R., Zlatev, I., Manoharan, M., Garcia-Sastre, A. and tenOever, B.R. (2010) 'Influenza A virus-generated small RNAs regulate the switch from transcription to replication', *Proceedings of the National Academy of Sciences of the United States of America,* 107(25), pp. 11525-11530.

Rain, J., Jean-Christophe Rain, Luc Selig, Hilde De Reuse and Veronique Battaglia (11) 'The protein-protein interaction map of Helicobacter pylori', *Nature (London),* 409(6817), pp. 211; 211.

Roberts, K.L., Leser, G.P., Ma, C. and Lamb, R.A. (2013) 'The amphipathic helix of influenza A virus M2 protein is required for filamentous bud formation and scission of filamentous and spherical particles', *Journal of virology,* 87(18), pp. 9973-9982.

Rossman, J.S., Jing, X., Leser, G.P. and Lamb, R.A. (2010) 'Influenza virus M2 protein mediates ESCRT-independent membrane scission', *Cell,* 142(6), pp. 902-913.

Rossman, J.S. and Lamb, R.A. (2011) 'Influenza virus assembly and budding', *Virology,* 411(2), pp. 229-236.

Samji, T. (2009) 'Influenza A: understanding the viral life cycle', *The Yale journal of biology and medicine,* 82(4), pp. 153-159.

Shi, Y. (2009) 'Serine/threonine phosphatases: mechanism through structure', *Cell,* 139(3), pp. 468-484.

Shui, J., Hu, M.C.-. and Tan, T. (2007) 'Conditional Knockout Mice Reveal an Essential Role of Protein Phosphatase 4 in Thymocyte Development and Pre-T-Cell Receptor Signaling', *Molecular and cellular biology,* 27(1), pp. 79-91.

Sun, L., Hemgard, G.V., Susanto, S.A. and Wirth, M. (2010) 'Caveolin-1 influences human influenza A virus (H1N1) multiplication in cell culture', *Virology journal,* 7, pp. 108-422X-7-108.

Swingle, M., Ni, L. and Honkanen, R.E. (2007) 'Small-molecule inhibitors of ser/thr protein phosphatases: Specificity, use and common forms of abuse', United States: , 23-38.

Takeuchi, O. and Akira, S. (2009) 'Innate immunity to virus infection', *Immunological reviews,* 227(1), pp. 75-86.

Teoh, K., Siu, Y., Chan, W., Schlüter, M.A., Liu, C., Peiris, J.S.M., Bruzzone, R., Margolis, B. and Nal, B. (2010) 'The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis', *Molecular biology of the cell,* 21(22), pp. 3838-3852.

Toyo-oka, K., Mori, D., Yano, Y., Shiota, M., Iwao, H., Goto, H., Inagaki, M., Hiraiwa, N., Muramatsu, M., Wynshaw-Boris, A., Yoshiki, A. and Hirotsune, S. (2008) 'Protein Phosphatase 4 Catalytic Subunit Regulates Cdk1 Activity and Microtubule Organization via NDEL1 Dephosphorylation', *The Journal of cell biology,* 180(6), pp. 1133-1147.

Wang, B., Ailian Zhao, A., Sun, L., Zhong, X., Zhong, J., Wang, H., Cai, M., Li, J., Xu, Y., Liao, J., Sang, J., Chowdhury, D.a., Pfeifer, G.P., Yen, Y. and Xu, X. (2008) 'Protein phosphatase PP4 is overexpressed in human breast and lung tumors', *细胞研究：英文版,* 18(9), pp. 974-977.

Weizmann Institute of Science (2015) *
SMEK1 Gene(Protein Coding)
SMEK Homolog 1, Suppressor Of Mek1 (Dictyostelium)
 .* Available at: http://www.genecards.org/cgi-bin/carddisp.pl?gene=smek1#drugs_compounds (Accessed: 10/27 2015).

Wharton, S.A., Belshe, R.B., Skehel, J.J. and Hay, A.J. (1994) 'Role of virion M2 protein in influenza virus uncoating: Specific reduction in the rate of membrane fusion between virus and liposomes by amantadine', *Journal of General Virology,* 75(4), pp. 945-948.

WHO (2015) *Warning signals from the volatile world of influenza viruses*. Available at: http://www.who.int/influenza/publications/warningsignals201502/en/ (Accessed: 02/27 2015).

Wies, E., Wang, M.K., Maharaj, N.P., Chen, K., Zhou, S., Finberg, R.W. and Gack, M.U. (2013) 'Dephosphorylation of the RNA Sensors RIG-I and MDA5 by the Phosphatase PP1 Is Essential for Innate Immune Signaling', *Immunity,* 38(3), pp. 437-449.

Wolff, S., Ma, H., Burch, D., Maciel, G.A., Hunter, T. and Dillin, A. (2006) 'SMK-1, an essential regulator of DAF-16-mediated longevity', *Cell,* 124(5), pp. 1039-1053.

Yeh, P.Y., Yeh, K.H., Chuang, S.E., Song, Y.C. and Cheng, A.L. (2004) 'Suppression of MEK/ERK signaling pathway enhances cisplatin-induced NF-kappaB activation by protein phosphatase 4-mediated NF-kappaB p65 Thr dephosphorylation', *The Journal of biological chemistry,* 279(25), pp. 26143-26148.

Yoo, J.K., Kim, T.S., Hufford, M.M. and Braciale, T.J. (2013) 'Viral infection of the lung: host response and sequelae', *The Journal of allergy and clinical immunology,* 132(6), pp. 1263-76; quiz 1277.

Zeng, H., Goldsmith, C.S., Maines, T.R., Belser, J.A., Gustin, K.M., Pekosz, A., Zaki, S.R., Katz, J.M. and Tumpey, T.M. (2013) 'Tropism and Infectivity of Influenza Virus, Including Highly Pathogenic Avian H5N1 Virus, in Ferret Tracheal Differentiated Primary Epithelial Cell Cultures', *Journal of virology,* 87(5), pp. 2597-2607.

Zhan, Z., Cao, H., Xie, X., Yang, L., Zhang, P., Chen, Y., Fan, H., Liu, Z. and Liu, X. (2015) 'Phosphatase PP4 Negatively Regulates Type I IFN Production and Antiviral Innate Immunity by Dephosphorylating and Deactivating TBK1', *Journal of immunology (Baltimore, Md.: 1950),* 195(8), pp. 3849-3857.

Zhou, G., Mihindukulasuriya, K.A., MacCorkle-Chosnek, R.A., Van Hooser, A., Hu, M.C., Brinkley, B.R. and Tan, T.H. (2002) 'Protein phosphatase 4 is involved in tumor necrosis factor-alpha-induced activation of c-Jun N-terminal kinase', *The Journal of biological chemistry,* 277(8), pp. 6391-6398.

Appendices

A. In silico

(i) Clones from yeast two-hybrid screen

Clone Name	Frag theoretical Sequence
PLA_RP_hgx1492v1_pB27_A-67	GGAGGTTCATCGATTAAAAGATGACTTTCCAACCGTACTACTGACAGTTTCTGTTTAGACTGTAGTAGCATGTTCAGCATGATCCGGCTTCAATCAGATGCACTCATGCAATAATAGAGATGTCGACATATACTGAGAGTGAGGATATCAAAATCATTACTGCTCATGTAATTGAAAATACTGCAAGCACTGAAGGATGTAGCATTATGTACATACATTAAAGGATTAAAAGTGAGATTTATCAACAGAGAAAATGGCAGTATAATCCCAAACTCCACAGTGCGCGTTCCATTCCGTGGAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTTAAAATGACGATGATATTATGGATTCAATAAGTAAATTCACGGAAAGGAAGAAATTCAAAGAAAGTGAGGAAAAGGAAGTGCTTCCGAAAACAAACCTTTGTGGACGGCAGAGCTCAAGTTTCCAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAATCTGCCTGGTTCTCCGGGATCACCCGGATCCCCAGGATCTCCAGGCTCTCCCGGGTCCGTACCTAAAAATACATGTCAGAGGGCC
PLA_RP_hgx1492v1_pB27_A-180	GTTTTCCCAAACGCTACAGCCTCAAAACAGAGATGCTTTTTTCAAGACTTTGTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACAAGATTTTGTCATGCAGGAGGCACAACAGAATGATGATGATTTTTTGCTCATCAACCTCATTATAGAACATATTATTTGTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAAAGAACATGCTAGCCACTGCCAATAAAACAAAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAACAAAAAACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTTGTCATTATGTGCCCTTCGTTTTAAAAGAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGGAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCACCAATG
PLA_RP_hgx1492v1_pB27_A-68	GTTTTCCCAAACGCTACAGCCTCAAAACAGAGATGCTTTTTTCAAGACTTTGTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGCTACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCTCATGTAATTGAAAATTACTGGA

	AAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGAGATTTGAACAACAAAGAGAAAG GCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAGATATCGAAGAGATGCCAGAACA CTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCT CCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAATTCATGGAAAGGAAGAAATTAA AAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGCAGAGCCCAAGTTTCAAGCTTTC CCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAATCTGCCTGGTTCTCCGGGATCAC CTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCTCAGACGGCAGCTATTACTACA AAGGGAGGCCTCGTGGGTCTGGTAGATTATCCTGATGATGATGAAGATGATGATGAGGATGAAGATAAG
PLA_RP_ hgx1492 v1_pB27 _A-294	GTTTTCCCAAACGCTACAGCCTCAAAACAGAGATGCTTTTTTCAAGACTTTGTCAAACATGGGCATATTACCAGC TTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTG AATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAGAATGATGATGATATTTTGCTCATC AACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCT TCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGAAAAGACTGAATTTCTGGGTTTCT TCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAACAGAAGACAAACCTAGTAAAGAT GATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCAC ATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTG GCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGAA AAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGCTACAATCTGATGAACTCTGCCAT AATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCTCATGTAATTGAAAATTACTGGA AAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGAGATTTGAACAACAAAGAGAAAG GCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAGATATCGAAGAGATGCCAGAACA CTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCT CCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAATTCATGGAAAGGAAGAAATTAA AAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGCAGAGCCCAAGTTTCAAGCTTTC CCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAATCTGCCTGGTTCTCCGGGATCAC CTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCTCAGACGGCAGCTATTACTACA AAGGGAGGCCTCGTGGGTCTGGTAGATTATCCTGATGATGATGAAGATGATGATGAGGATGAAGATAAG
PLA_RP_ hgx1492 v1_pB27 _A-137	GTTTTCCCAAACGCTACAGCCTCAAAACAGAGATGCTTTTTTCAAGACTTTGTCAAACATGGGCATATTACCAGC TTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTG AATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAGAATGATGATGATATTTTGCTCATC AACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCT TCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGAAAAGACTGAATTTCTGGGTTTCT TCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAACAGAAGACAAACCTAGTAAAGAT GATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCAC ATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTG GCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGAA AAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGCTACAATCTGATGAACTCTGCCAT

	AATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCTCATGTAATTGAAAATTACTGGA AAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGAGATTTGAACAACAAAGAGAAAG GCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAGATATCGAAGAGATGCCAGAACA CTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCT CCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAATTCATGGAAAGGAAGAAATTAA AAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGCAGAGCCCAAGTTTCAAGCTTTC CCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAATCTGCCTGGTTCTCCGGGATCAC CTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCTCAGACGGCAGCTATTACTACA AAGGGAGGCCTCGTGGGTCTGGTAGATTATCCTGATGATGATGAAGATGATGATGAGGATGAAGATAAG
PLA_RP_ hgx1492 v1_pB27 _A-339	GTTTTCCCAAACGCTACAGCCTCAAAACAGAGATGCTTTTTTCAAGACTTTGTCAAACATGGGCATATTACCAGC TTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTG AATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAGAATGATGATGATATTTTGCTCATC AACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCT TCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGAAAAGACTGAATTTCTGGGTTTCT TCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAACAGAAGACAAACCTAGTAAAGAT GATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCAC ATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTG GCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGAA AAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGCTACAATCTGATGAACTCTGCCAT AATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCTCATGTAATTGAAAATTACTGGA AAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGAGATTTGAACAACAAAGAGAAAG GCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAGATATCGAAGAGATGCCAGAACA CTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCT CCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAATTCATGGAAAGGAAGAAATTAA AAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGCAGAGCCCAAGTTTCAAGCTTTC CCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAATCTGCCTGGTTCTCCGGGATCAC CTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCTCAGACGGCAGCTATTACTACA AAGGGAGGCCTCGTGGGTCTGGTAGATTATCCTGATGATGATGAAGATGATGATGAGGATGAAGATAAG
PLA_RP_ hgx1492 v1_pB27 _A-142	CAAGACTTTGTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAA GTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAG GCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAA CTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAA TAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGC AAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGT TAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTG CTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTA AAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAAT

	GGATCCCGCTACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATC ATTAACTGCTCATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGAT TAAAACTGAGATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAG GAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGA TGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCA ATAAGTAAATTCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTT TCTGGACGGCAGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATC TACAACAAATCTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAA AAATACATCTCAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-264	CAAGACTTTGTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAA GTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAG GCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAA CTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAA TAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGC AAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGT TAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTG CTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTA AAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAAT GGATCCCGCTACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATC ATTAACTGCTCATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGAT TAAAACTGAGATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAG GAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGA TGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCA ATAAGTAAATTCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTT TCTGGACGGCAGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATC TACAACAAATCTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAA AAATACATCTCAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-240	CAAGACTTTGTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAA GTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAG GCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAA CTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAA TAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGC AAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGT TAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTG CTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTA AAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAAT GGATCCCGCTACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATC

	ATTAACTGCTCATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGAT TAAAACTGAGATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAG GAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGA TGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCA ATAAGTAAATTCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTT TCTGGACGGCAGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATC TACAACAAATCTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAA AAATACATCTCAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-207	CAAGACTTTGTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAA GTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAG GCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAA CTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAA TAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGC AAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGT TAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGGAGAGTG CTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTA AAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAAT GGATCCCGCTACAATCTGATGAACTCTGCCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATC ATTAACTGCTCATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGAT TAAAACTGAGATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAG GAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGA TGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCA ATAAGTAAATTCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTT TCTGGACGGCAGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATC TACAACAAATCTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAA AAATACATCTCAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-39	GTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTA CTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAG AATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGA GCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGA AAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAAC AGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTG TGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGGAGAGTGCTAGTTCTTA TGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAG TTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGC TACAATCTGATGAACTCTGCCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCT CATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGA

	GATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAG ATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGA AGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAAT TCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGC AGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAAT CTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCT CAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-280	GTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTA CTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAG AATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGA GCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGA AAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAAC AGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTG TGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTA TGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAG TTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGC TACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCT CATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGA GATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAG ATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGA AGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAAT TCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGC AGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAAT CTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCT CAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-373	GTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTA CTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAG AATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGA GCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGA AAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAAC AGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTG TGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTA TGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAG TTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGC TACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCT CATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGA GATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAG

	ATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGA AGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAAT TCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGC AGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAAT CTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCT CAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-276	GTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTA CTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAG AATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGA GCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGA AAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAAC AGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTG TGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTA TGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAG TTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGC TACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCT CATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGA GATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAG ATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGA AGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAAT TCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGC AGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAAT CTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCT CAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-102	GTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTA CTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAG AATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGA GCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGA AAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAAC AGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTG TGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTA TGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAG TTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGC TACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCT CATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGA GATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAG ATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGA

	AGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAAT TCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGC AGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAAT CTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCT CAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-88	GTCAAACATGGGCATATTACCAGCTTTAGAAGTCATCCTTGGCATGGATGATACACAGGTGCGAAGTGCTGCTA CTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAGAGTTTGTCATGCAGGAGGCACAACAG AATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTTGTGATACAGATCCTGAACTTGGAGGA GCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAACATGCTAGCCACTGCCAATAAAACAGA AAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCTCACTGCTCCTTTACTAGCAAATACAAC AGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGGCACTTGTATTGGAATTGTTAACATTTTG TGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGGATATCCTCCGGAGAGTGCTAGTTCTTA TGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAAGAAAGATTATTGGATTAAAAGATGAG TTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTGAAAGCATTTCTCAACAATGGATCCCGC TACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGAGTGGAAGATATAAAATCATTAACTGCT CATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGTACAGACATTTAAAGGATTAAAACTGA GATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGTATGCGTTCCATTTTGAGGAATCACAG ATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGTTTAACACAGATGAAGATGACATGGA AGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGATGATATTATGGATCCAATAAGTAAAT TCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAGTGCTTCTGAAAACAAACCTTTCTGGACGGC AGAGCCCAAGTTTCAAGCTTTCCCTGTCCAGTGGAACGAAGACTAACCTCACCAGCCAGTCATCTACAACAAAT CTGCCTGGTTCTCCGGGATCACCTGGATCCCCAGGATCTCCAGGCTCTCCTGGATCCGTACCTAAAAATACATCT CAGACGGCA
PLA_RP_ hgx1492 v1_pB27 _A-79	GGATGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTAC GAGAGTTTGTCATGCAGGAGGCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATG ATTTGTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGA GAACATGCTAGCCACTGCCAATAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATG TTCTCACTGCTCCTTTACTAGCAAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTAT TGGCACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATA AGGATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTA AAAGAAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTA GTGAAAGCATTTCTCAACAATGGATCCCGCTACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATT AGAGTGGAAGATATAAAATCATTAACTGCTCATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATT ATGTACAGACATTTAAAGGATTAAAACTGAGATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGA CAGTATGCGTTCCATTTTGAGGAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATG TGGTTTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATG ATGATGATATTATGGATCCAATAAGTAAATTCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAG

	TGCTTCTGAAAACAAACCTTTCTGGACGGCAGAGCCC
PLA_RP_ hgx1492 v1_pB27 _A-123	GGATGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTAC GAGAGTTTGTCATGCAGGAGGCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATG ATTTGTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGA GAACATGCTAGCCACTGCCAATAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATG TTCTCACTGCTCCTTTACTAGCAAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTAT TGGCACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATA AGGATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTA AAAGAAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTA GTGAAAGCATTTCTCAACAATGGATCCCGCTACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATT AGAGTGGAAGATATAAAATCATTAACTGCTCATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATT ATGTACAGACATTTAAAGGATTAAAACTGAGATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGA CAGTATGCGTTCCATTTTGAGGAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATG TGGTTTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATG ATGATGATATTATGGATCCAATAAGTAAATTCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAG TGCTTCTGAAAACAAACCTTTCTGGACGGCAGAGCCC
PLA_RP_ hgx1492 v1_pB27 _A-128	GGATGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTAC GAGAGTTTGTCATGCAGGAGGCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATG ATTTGTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGA GAACATGCTAGCCACTGCCAATAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATG TTCTCACTGCTCCTTTACTAGCAAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTAT TGGCACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATA AGGATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTA AAAGAAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTA GTGAAAGCATTTCTCAACAATGGATCCCGCTACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATT AGAGTGGAAGATATAAAATCATTAACTGCTCATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATT ATGTACAGACATTTAAAGGATTAAAACTGAGATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGA CAGTATGCGTTCCATTTTGAGGAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATG TGGTTTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATG ATGATGATATTATGGATCCAATAAGTAAATTCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAG TGCTTCTGAAAACAAACCTTTCTGGACGGCAGAGCCC
PLA_RP_ hgx1492 v1_pB27 _A-135	GGATGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTAC GAGAGTTTGTCATGCAGGAGGCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATG ATTTGTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGA GAACATGCTAGCCACTGCCAATAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATG TTCTCACTGCTCCTTTACTAGCAAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTAT TGGCACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATA

	AGGATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTA AAAGAAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTA GTGAAAGCATTTCTCAACAATGGATCCCGCTACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATT AGAGTGGAAGATATAAAATCATTAACTGCTCATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATT ATGTACAGACATTTAAAGGATTAAAACTGAGATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGA CAGTATGCGTTCCATTTTGAGGAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATG TGGTTTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATG ATGATGATATTATGGATCCAATAAGTAAATTCATGGAAAGGAAGAAATTAAAAGAAAGTGAGGAAAAGGAAG TGCTTCTGAAAACAAACCTTTCTGGACGGCAGAGCCC
PLA_RP_ hgx1492 v1_pB27 _A-129	TGATACACAGGTGCGAAGTGCTGCTACTGATATATTCTCATACTTGGTTGAATATAATCCATCCATGGTACGAG AGTTTGTCATGCAGGAGGCACAACAGAATGATGATGATATTTTGCTCATCAACCTCATTATAGAACATATGATTT GTGATACAGATCCTGAACTTGGAGGAGCAGTCCAGCTTATGGGCCTGCTTCGAACTTTAGTTGACCCAGAGAA CATGCTAGCCACTGCCAATAAAACAGAAAAGACTGAATTTCTGGGTTTCTTCTACAAGCACTGTATGCATGTTCT CACTGCTCCTTTACTAGCAAATACAACAGAAGACAAACCTAGTAAAGATGATTTTCAGACTGCCCAACTATTGG CACTTGTATTGGAATTGTTAACATTTTGTGTGGAGCACCATACCTACCACATAAAGAACTACATTATTAATAAGG ATATCCTCCGGAGAGTGCTAGTTCTTATGGCCTCGAAGCATGCTTTCTTGGCATTATGTGCCCTTCGTTTTAAAA GAAAGATTATTGGATTAAAAGATGAGTTTTACAACCGCTACATAATGAAAAGTTTTTTGTTTGAACCAGTAGTG AAAGCATTTCTCAACAATGGATCCCGCTACAATCTGATGAACTCTGCCATAATAGAGATGTTTGAATTTATTAGA GTGGAAGATATAAAATCATTAACTGCTCATGTAATTGAAAATTACTGGAAAGCACTGGAAGATGTAGATTATGT ACAGACATTTAAAGGATTAAAACTGAGATTTGAACAACAAAGAGAAAGGCAAGATAATCCCAAACTTGACAGT ATGCGTTCCATTTTGAGGAATCACAGATATCGAAGAGATGCCAGAACACTAGAAGATGAAGAAGAGATGTGGT TTAACACAGATGAAGATGACATGGAAGATGGAGAAGCTGTAGTGTCTCCATCTGACAAAACTAAAAATGATGA TGATATTATGGATCCAATAAGTAAATTCATGGAAAG

Table 4 of clones identified through the yeast two-hybrid screen. Shown are the theoretical fragment sequences based on the 3" and 5" sequences identified through the screen. The first two clones PLA_RP_hgx1492v1_pB27_A-180 and PLA_RP_hgx1492v1_pB27_A-67 had only 3" and 5" sequences respectively. Therefore shown in this table is the 3" sequence for clone PLA_RP_hgx1492v1_pB27_A-180 shows the 3" sequence and clone PLA_RP_hgx1492v1_pB27_A-67 shows 5" sequence. During the yeast two-hybrid screen analysis these sequences were used for alignment as fragment theoretical sequences were not provided.

B Full western blot membrane images

Figure 9 (i)
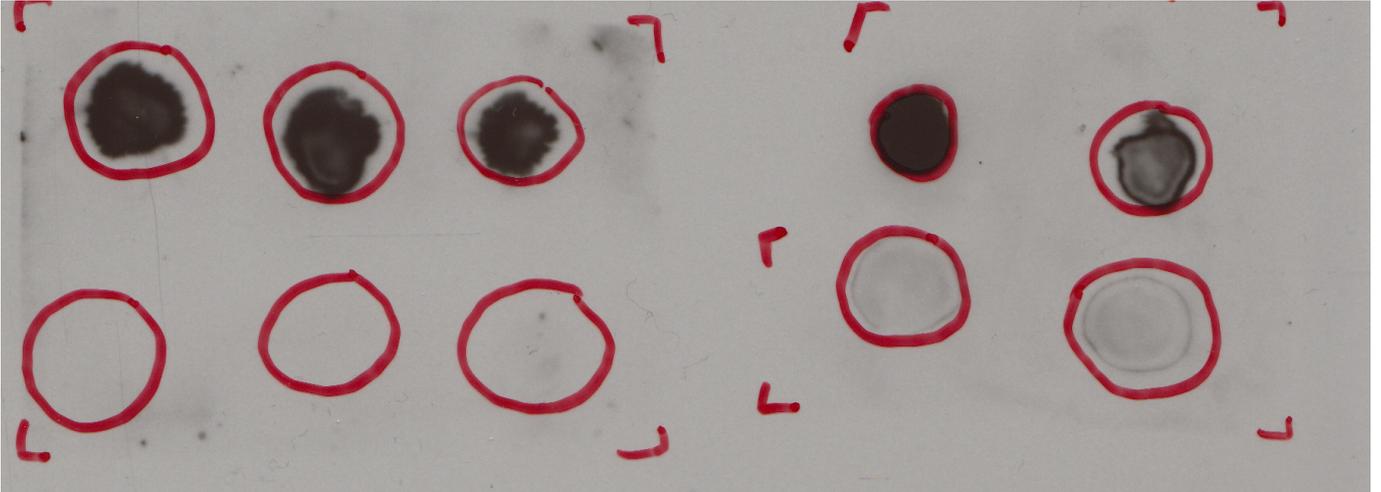
Full scanned X-ray film for dot blot analysis of SMEK1 and PP4c expression in A549 and HeLa cells.

Figure 10 (i)
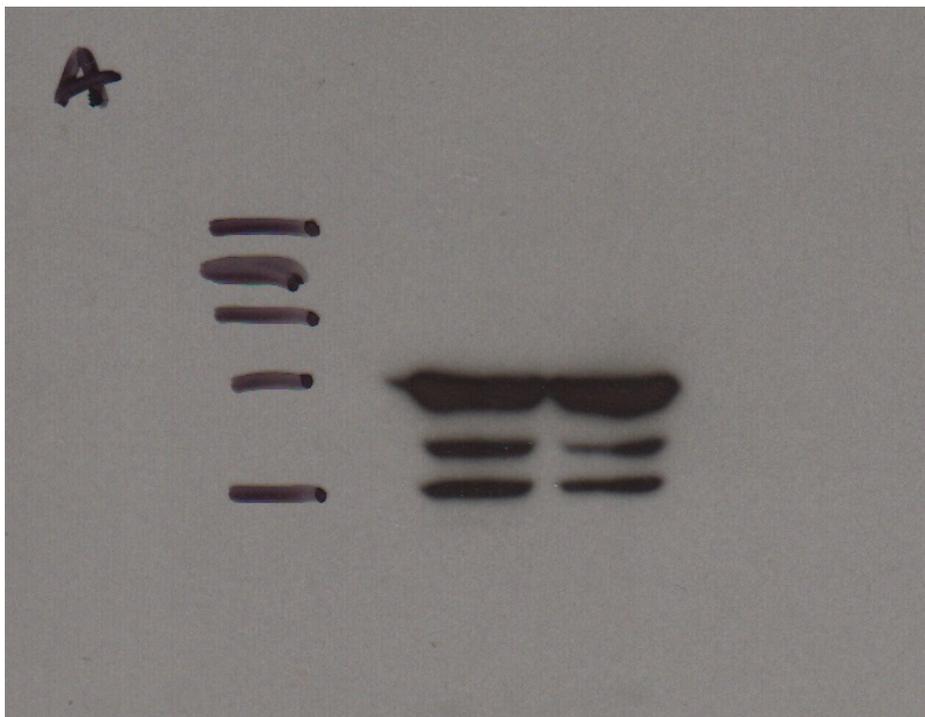
Full scanned X-ray film for western blot analysis of GAPDH labelling of A549 and HeLa samples, following 10 seconds of exposure.
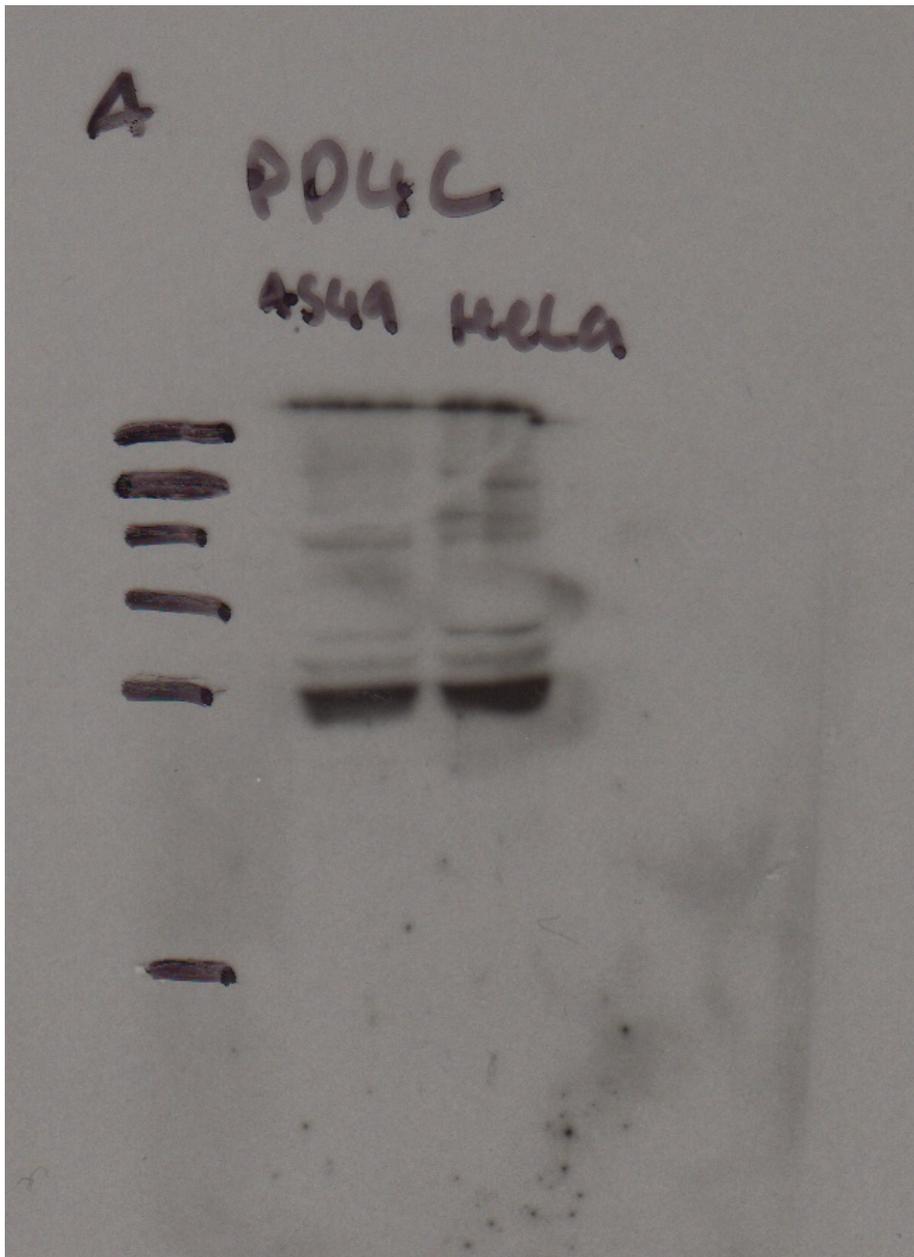
Figure 10 (ii)

Full scanned X-ray film for western blot analysis of PP4c labelling of A549 and HeLa samples, after 10 minutes of exposure.

Figure 10 (iii)
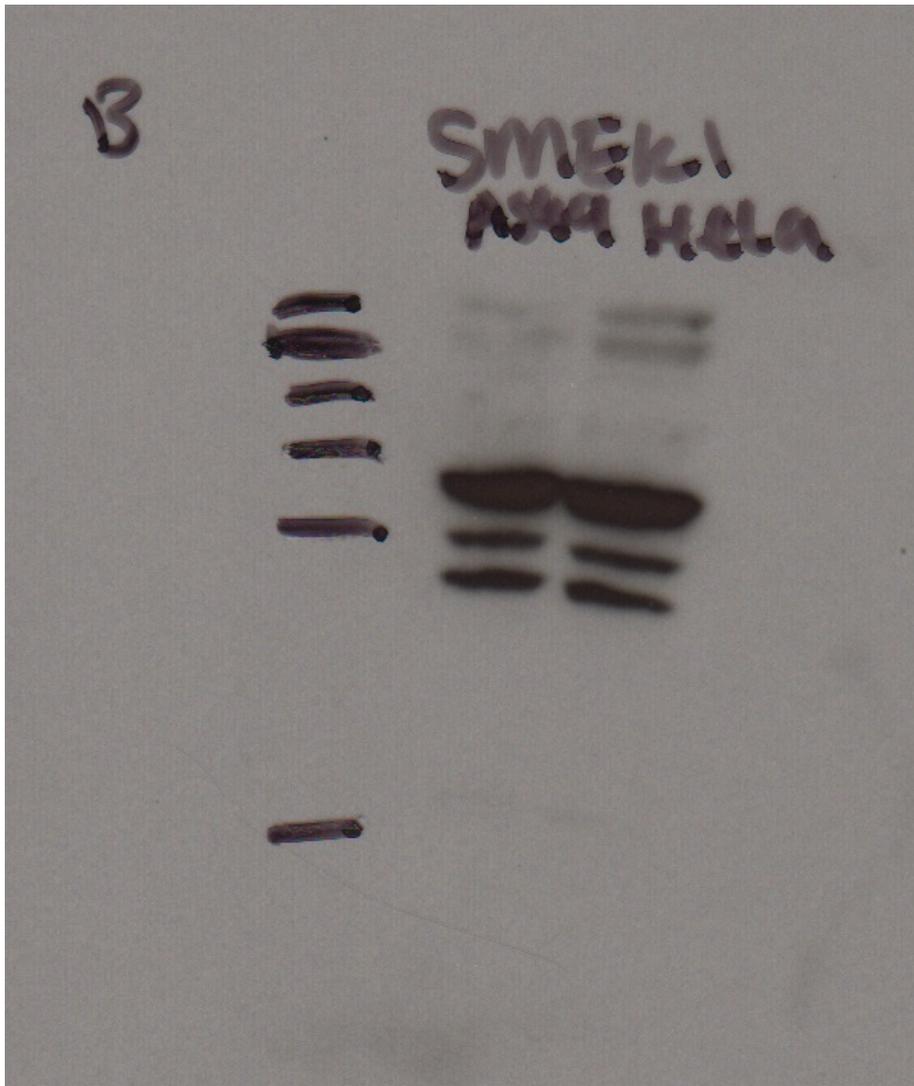
Full scanned X-ray film for western blot analysis of SMEK1 labelling of A549 and HeLa samples, following 10 minutes of exposure.
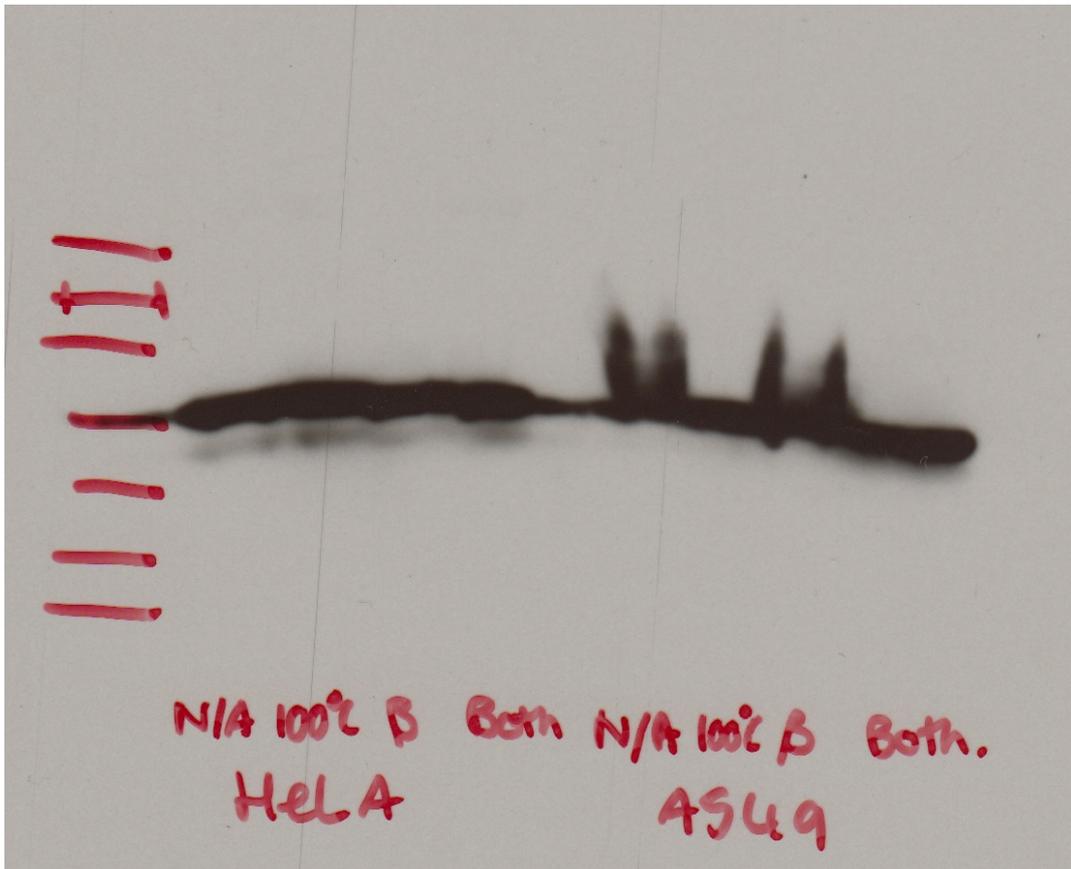
Figure 11(i)

Full scanned X-ray film for western blot analysis of GAPDH labelling of A549 and HeLa samples after 10 seconds of exposure.
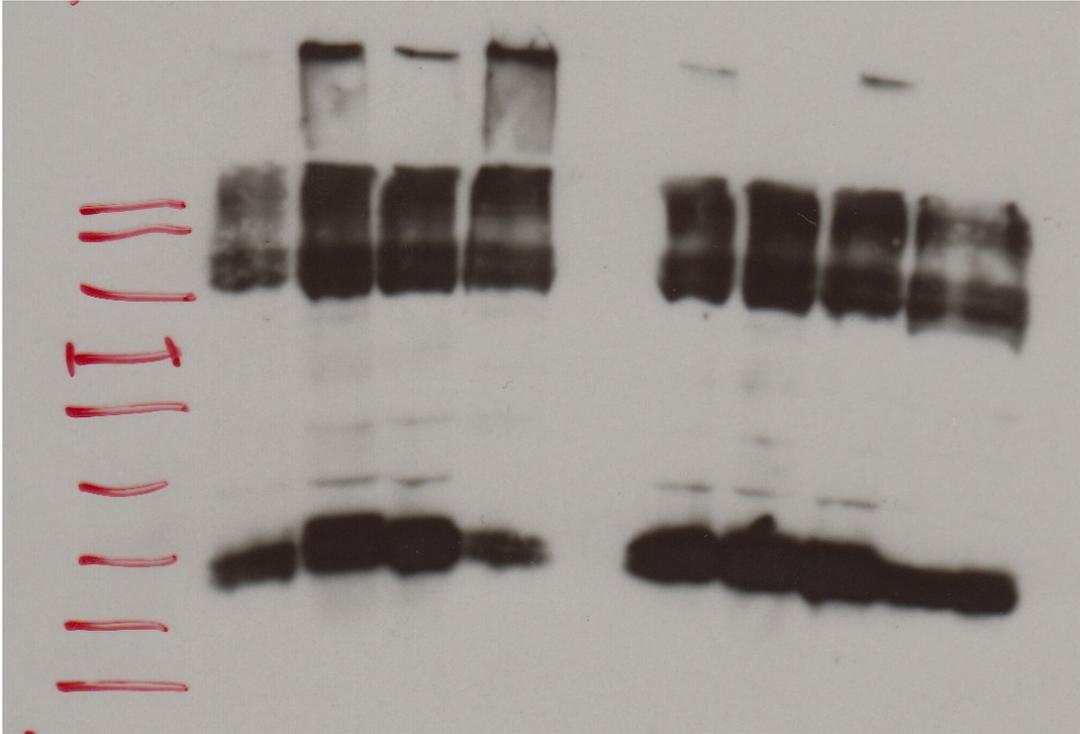
Figure 11(ii)

Full scanned X-ray film for western blot analysis of PP4c labelling of A549 and HeLa samples, after 5 minutes of exposure.

Figure 11(iii)
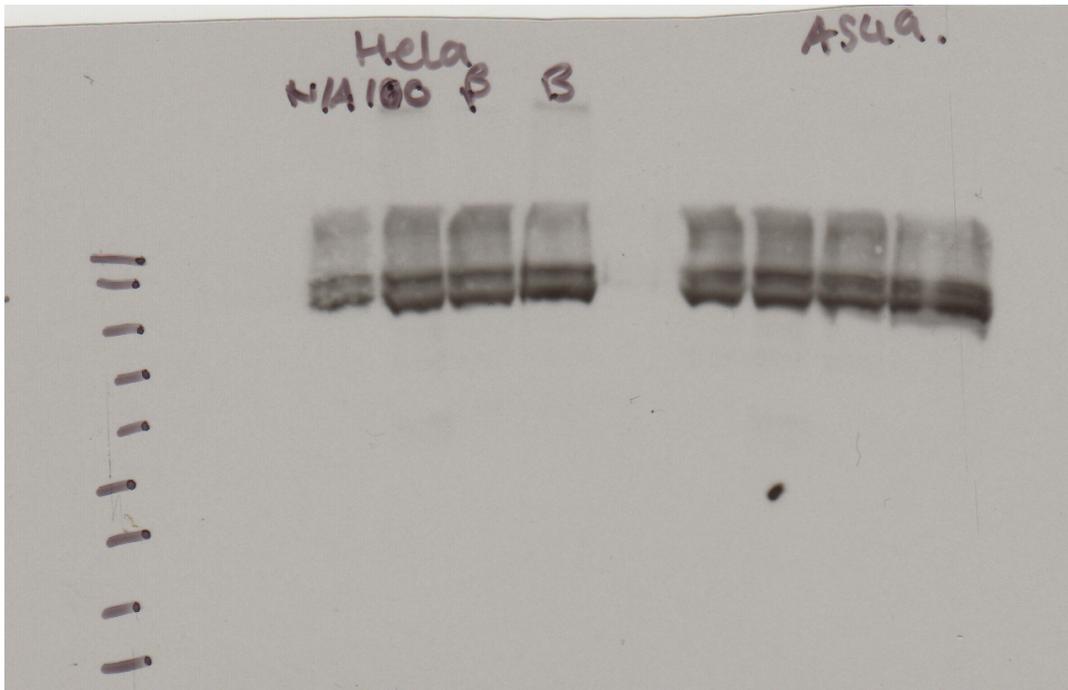
Full scanned X-ray film for western blot analysis of SMEK1 labelling of A549 and HeLa samples, following 1 minute of exposure.
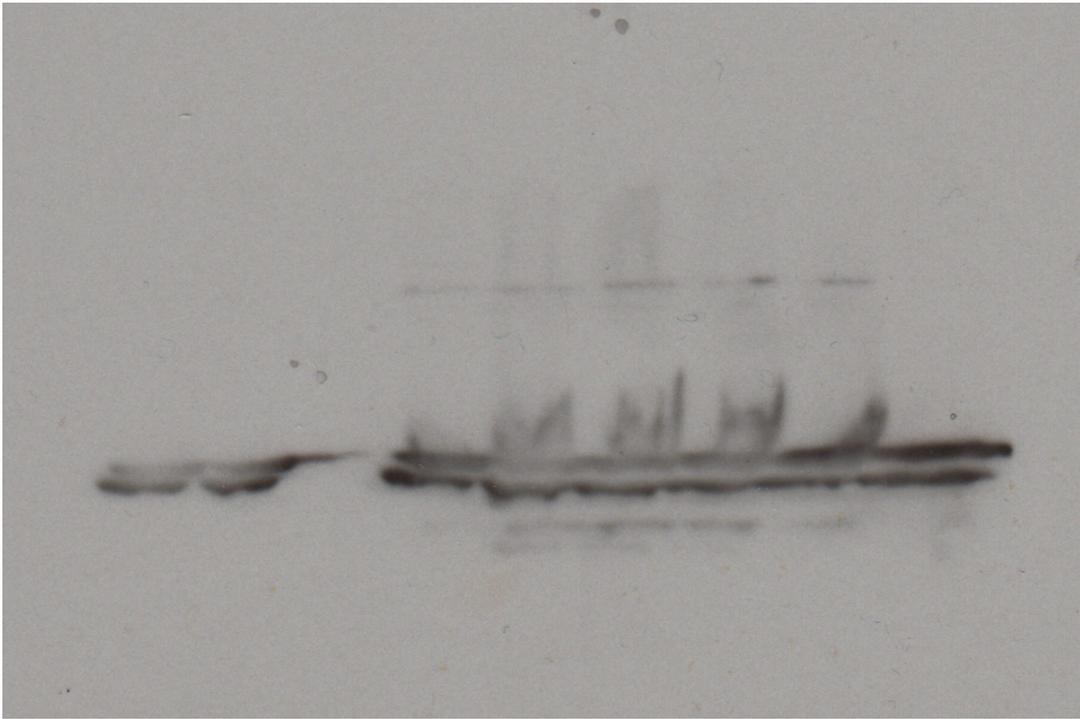
Figure 12 (i)

Full scanned X-ray film for western blot analysis of GAPDH labelling of A549, HeLa, and infected A549 samples (NI, 0H, 3H, 6H, 12H, 24H) after 10 seconds of exposure.
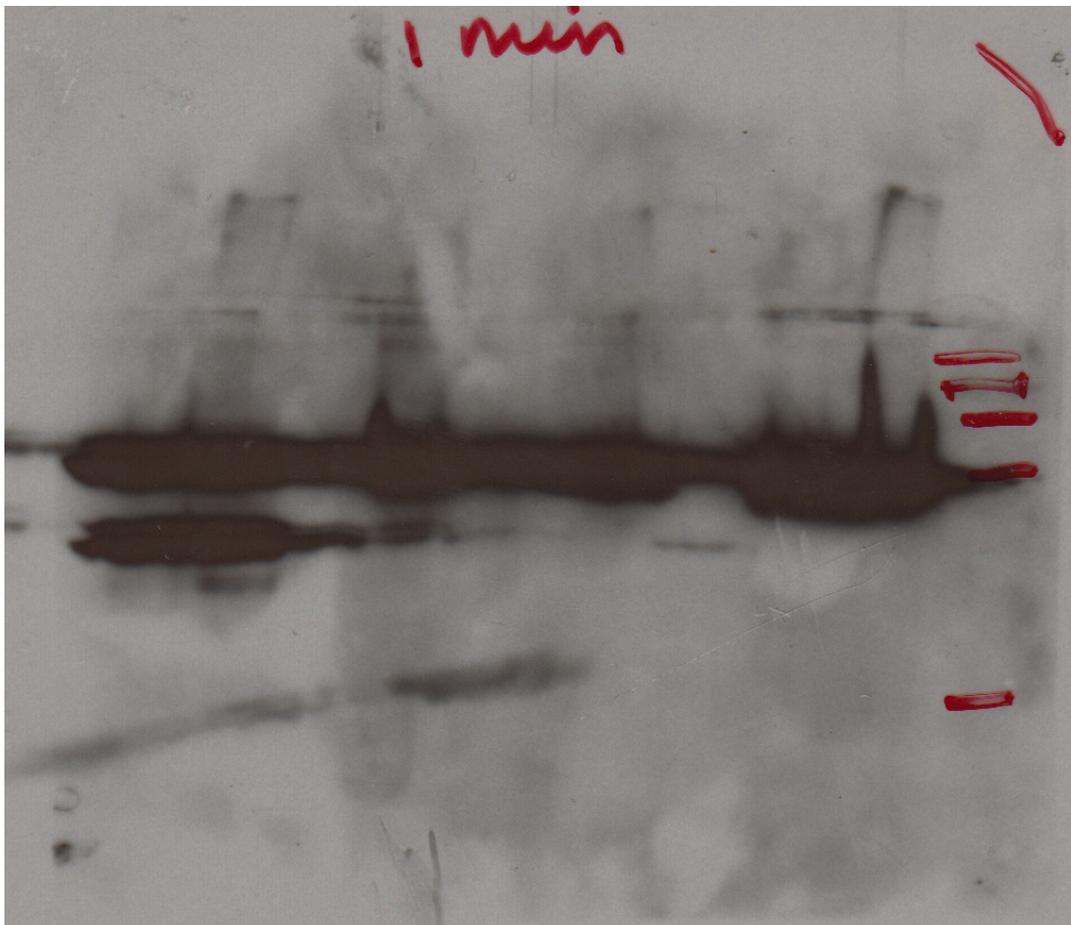
Figure 12 (ii)

Full scanned X-ray film for western blot analysis of M1 labelling of infected A549 samples (24H, 12H, 6H, 3H, 0H, NI), HeLa and A549, after 1 minute of exposure.

Figure 12 (iii)
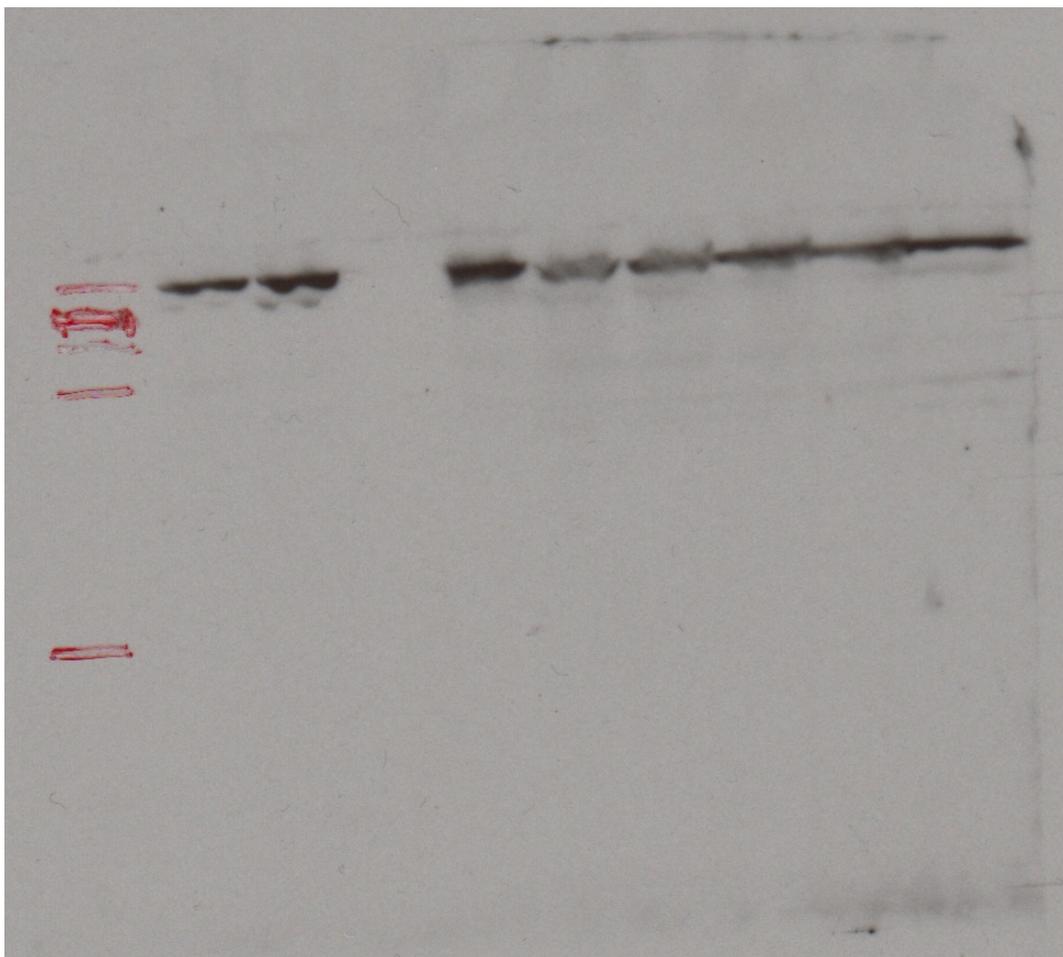
Full scanned X-ray film for western blot analysis of SMEK1 labelling of A549, HeLa and infected A549 samples (NI, 0H, 3H, 6H, 12H, 24H), following 10 minutes of exposure.
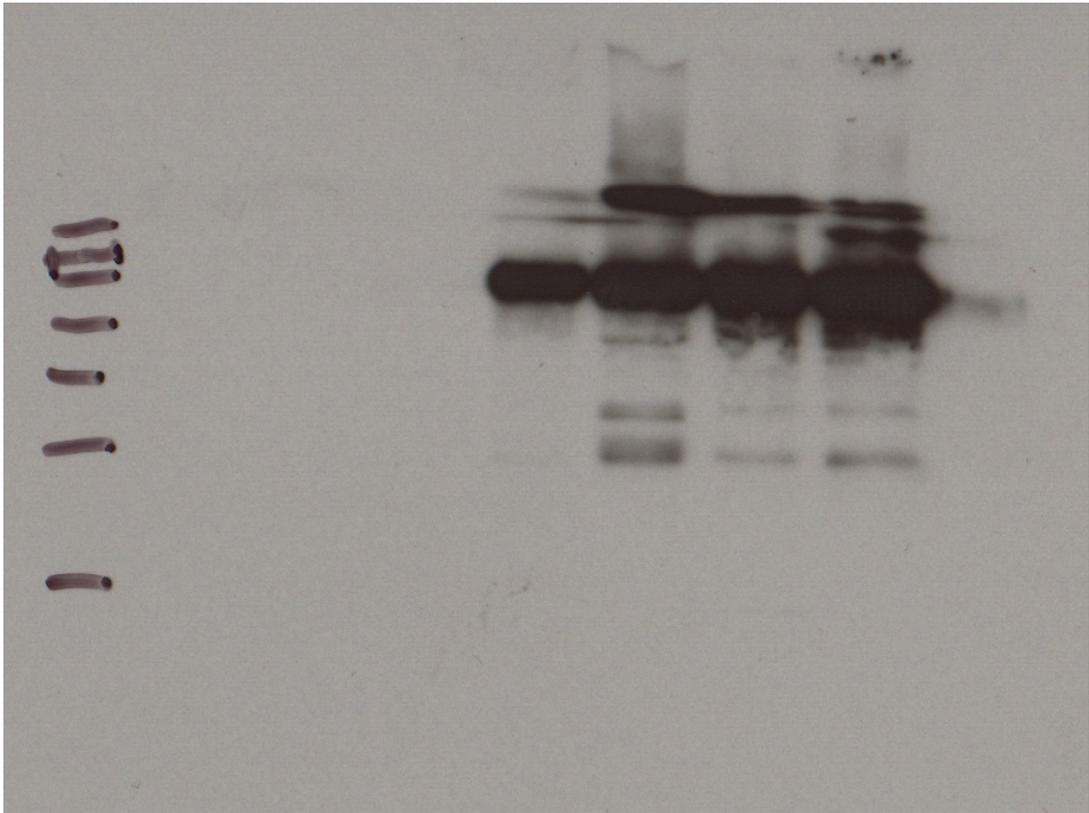
Figure 16 (i)

Full scanned X-ray film for western blot analysis of SMEK1 staining of A549 lysate, Flow through, M2, PP4c, SMEK1, Annexin A6 and NFκB incubated co-immunoprecipitation samples after 1 minute of exposure.
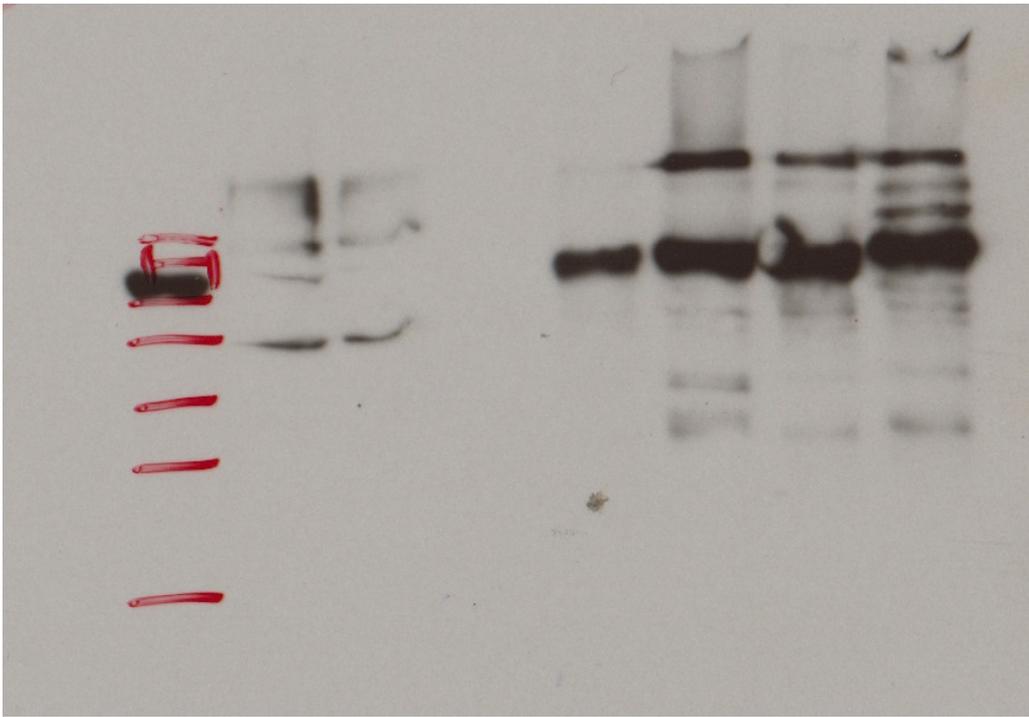
Figure 16 (ii)

Full scanned X-ray film for western blot analysis of NFκB staining of A549 lysate, Flow through, M2, PP4c, SMEK1, Annexin A6 and NFκB incubated co-immunoprecipitation samples after 30 seconds of exposure.
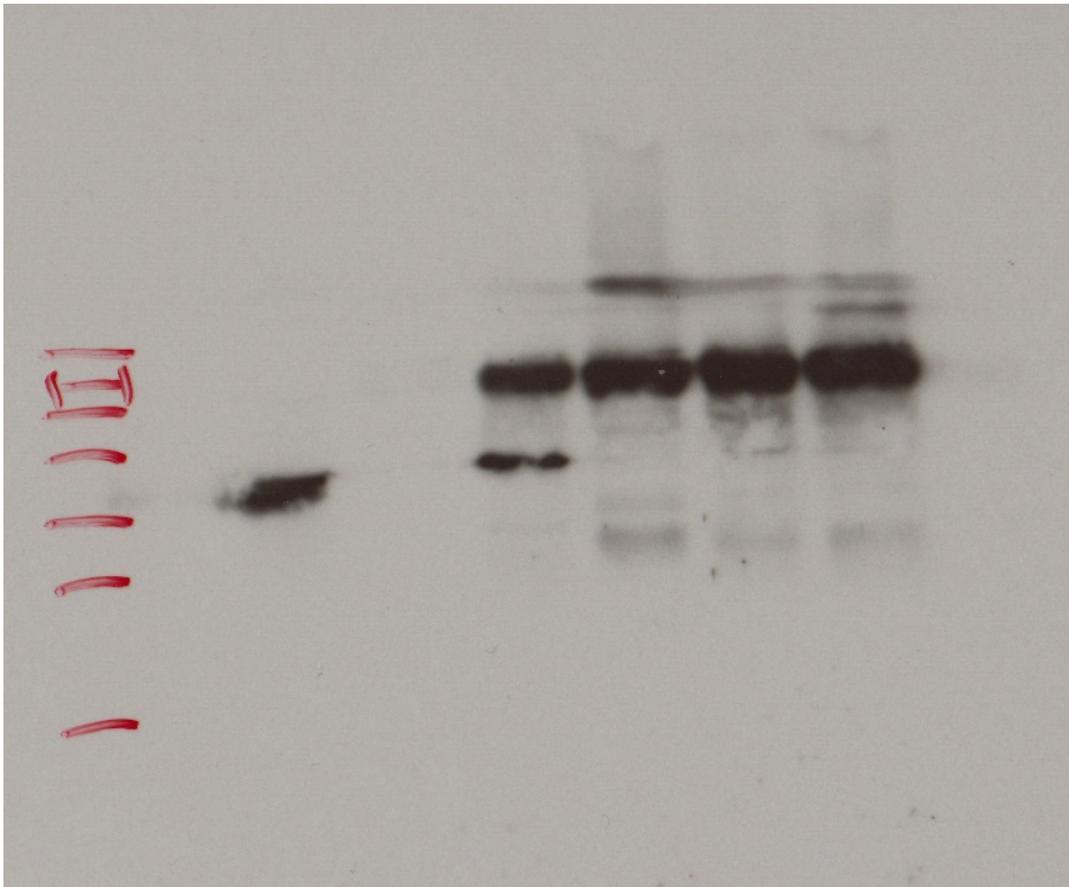
Figure 16 (iii)

Full scanned X-ray film for western blot analysis of PP4c staining of A549 lysate, Flow through, M2, PP4c, SMEK1, Annexin A6 and NFκB incubated co-immunoprecipitation samples after 30 seconds of exposure.

Figure 16 (iiii)
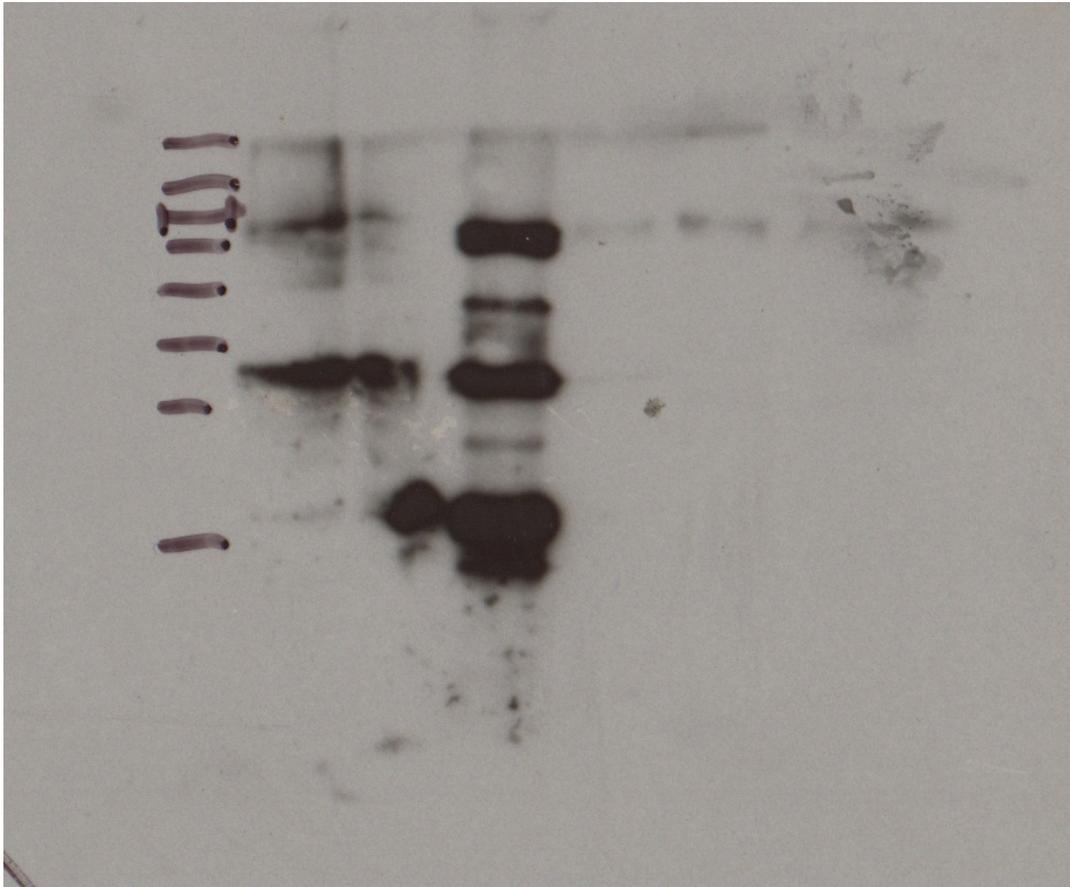
Full scanned X-ray film for western blot analysis of M2 staining of A549 lysate, Flow through, M2, PP4c, SMEK1, Annexin A6 and NFκB incubated co-immunoprecipitation samples after 1 minute of exposure.
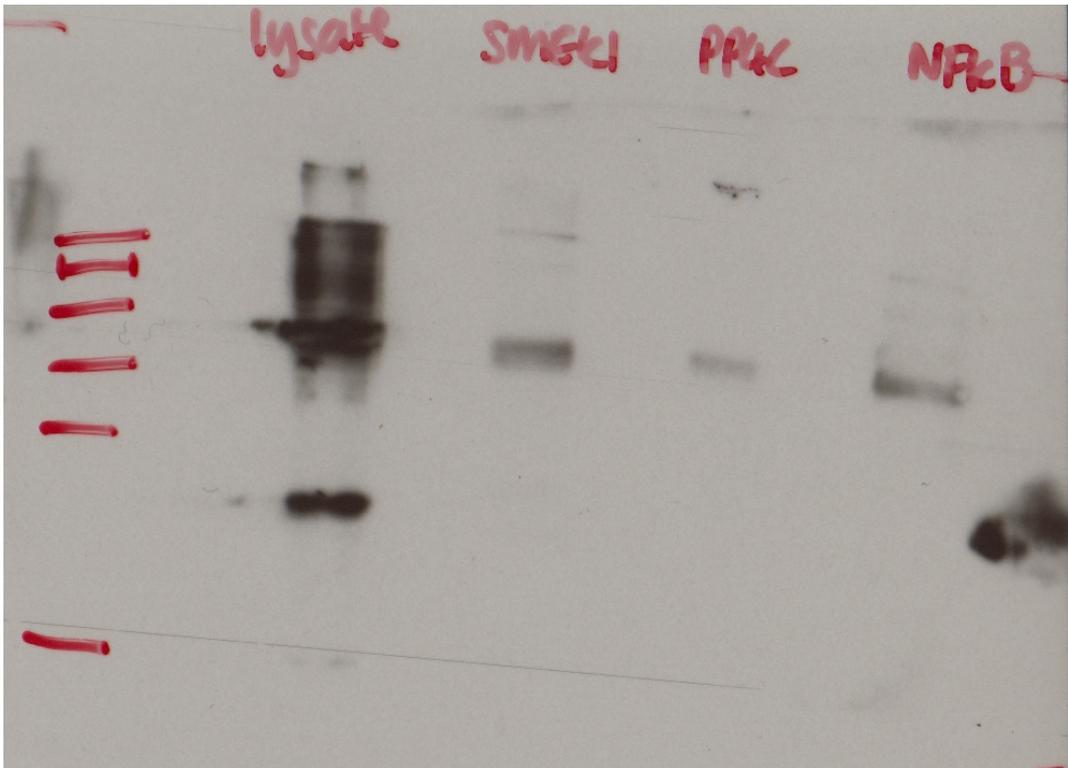
Figure 17(i)

Full scanned X-ray film for western blot analysis of M2 labelling of A549 lysate, SMEK1, PP4c and NFκB incubated samples after 10 minutes of exposure.

Figure 17(ii)
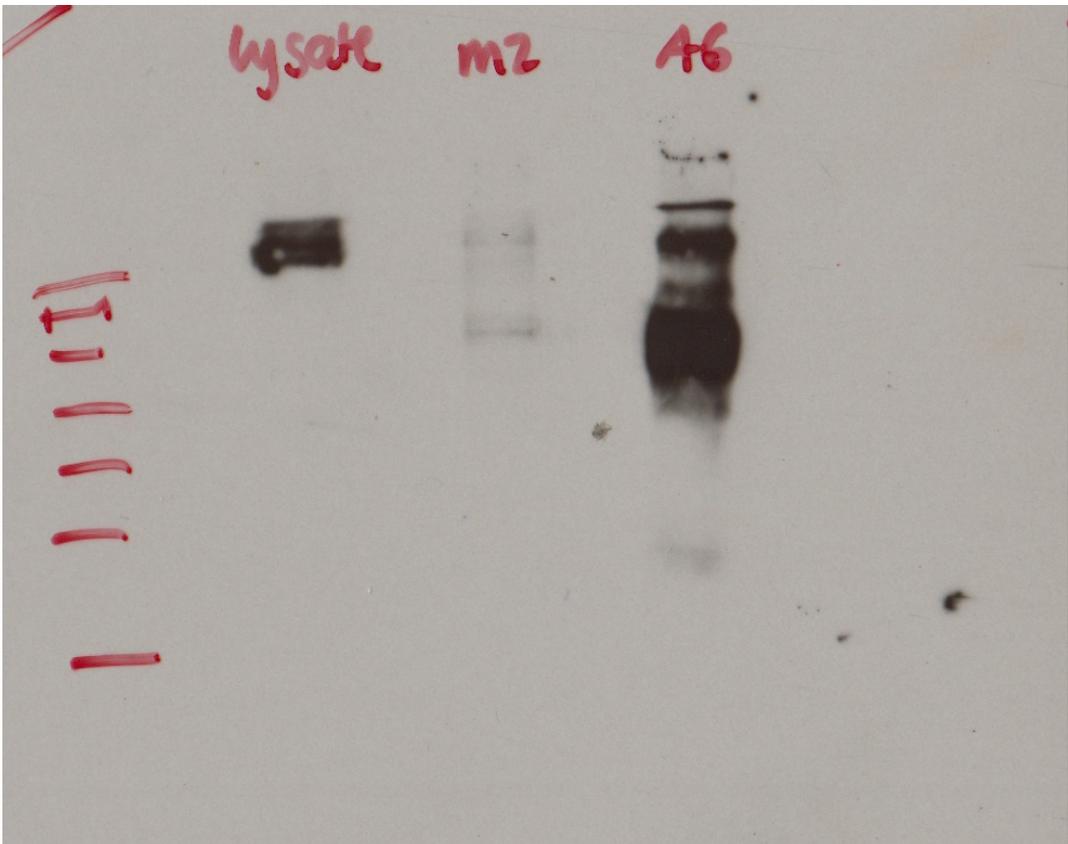
Full scanned X-ray film for western blot analysis of SMEK1 labelling of A549 lysate, M2 and Annexin A6 incubated samples after 10 minutes of exposure.