REGULATION OF SERINE-ARGININE PROTEIN KINASE 1 FUNCTIONS BY HUMAN PAPILLOMAVIRUS

By

EMMA LOUISE PRESCOTT

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Cancer Sciences

College of Medicine and Dentistry

University of Birmingham

2011

ABSTRACT

The role of the E4 protein in the human papillomavirus (HPV) life cycle is an enigma even though it has varied effects on cell behaviour and organisation in overexpression studies. Full-length E4 proteins are derived from E1^E4 spliced RNA transcripts and E1^E4 proteins from diverse HPV types interact with serine-arginine (SR)-specific protein kinase SRPK1, that regulates diverse cellular functions including RNA splicing. This thesis has sought to address the hypothesis that E1^E4 alters SRPK1 activity and influences SRPK1 functions in the HPV life cycle. This study has uncovered the novel finding that E1^E4 protein of HPV1, but not HPV5, 16 and 18, is a potent inhibitor of SRPK1 activity *in vitro* and *in vivo* and inhibition is dependent upon E1^E4 binding to SRPK1. Whilst HPV1 E1^E4 inhibits SRPK1 phosphorylation of cellular (ASF/SF2, SRp20, SC35, 9G8 and SRp75) and viral (HPV E2) SR protein substrates, it has only weak effects on SR protein cellular localisation and on cellular and viral RNA splicing in minigene systems. Addition of the small molecule inhibitor of SRPK, SRPIN340 to organotypic raft cultures of HPV18 genome-containing keratinocytes enhances the morphological features of HPV viral replication suggesting that the HPV may modulate SRPK activity to facilitate the virus life cycle.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Sally Roberts, for her help and advice throughout my studies. I would also like to thank my previous co-supervisor Ian Bell and past members of the HPV group, especially Craig Delury, for their advice and support. I would like to thank Ben Willcox for his help with the Biacore experiments and all the other people in the School of Cancer Sciences who have given me help and support during the project. Additionally I would like to thank all the good friends I have made during my time in Cancer Sciences, especially the 2007 intake of PhD students, for helping me have fun during my studies and supporting me through the hard times.

I would like to thank the College of Medicine and Dentistry for their financial support during my studies and Cancer Research UK for funding the laboratory of Sally Roberts.

Finally I thank my family for putting up with me during this stressful time and for believing that I could finish this thesis.

LIST OF FIGURES

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Chapter 6

LIST OF TABLES

ABBREVIATIONS

APS	Ammonium persulphate
ASF/SF2	Alternative Splicing Factor/Splicing Factor 2
ATP	Adenosine triphosphate
BCC	Basal cell carcinoma
CCE	Cornified cell envelope
CDK	Cyclin-dependent kinase
CIN	Cervical intraepithelial neoplasia
CMV	Cytomegalovirus
DABCO	1,4-diazabicyclo[2.2.2]octane
DAPI	4',6-diamidion-2-phenylindole
DBD	DNA binding domain
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
ESE	Exonic splicing enhancers
ESS	Exonic splicing silencers
EV	Epidermodysplasia verruciformis
GFP	Green fluorescent protein
GST	Glutathione-S-transferase
H and E	Haematoxylin and eosin
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HFK	Human foreskin keratinocyte
HIV	Human immunodeficiency virus
hnRNP	Heterogeneous nuclear ribonucleoprotein
HPV	Human papillomavirus
HSV	Herpes simplex virus
IGC	Interchromatin granule clusters
IRES	Internal ribosome entry site
LCR	Long control region
MCM7	Minichromosome maintenance complex component 7
mRNA	Messenger RNA
mTOR	Mammalian target of rapamycin
NES	Nuclear export signal
NLS	Nuclear localisation signal

NMD	Nonsense mediated decay
NMSC	Non-melanoma skin cancer
NRE	Negative regulatory element
Nt	Nucleotide
ORF	Open reading frame
PBS	Phosphate buffered saline
PIAS1	Protein inhibitor of activated STAT-1
PKCI-r	Protein kinase C-interacting protein 1 related
PML	Promyelocytic leukaemia protein
PMSF	Phenylmethanesulfonyl fluoride
PP1	Protein phosphatase 1
RNA	Ribonucleic acid
RRM	RNA recognition motif
RS	Arginine serine
SCC	Squamous cell carcinoma
snRNP	Small nuclear ribonucleoprotein
SR	Serine arginine
SRSF	Serine arginine splicing factor
SRPK1	Serine arginine protein kinase 1
Ss	Splice site
SUMO	Small ubiquitin-related modifier
TCA	Trichloro acetic acid
UTR	Untranslated region
VLP	Virus like particle
VZV	Varicella-Zoster virus

CHAPTER 1 GENERAL INTRODUCTION

1.1 Small DNA tumour viruses

At the start of the twentieth century Rous (1911) identified that cells taken from a chicken tumour could cause a sarcoma to be induced when injected into another chicken and thus identified Rous sarcoma virus, a RNA virus capable of causing cancer in chickens. Since this initial discovery many other viruses have been identified as cancer causing agents and now viruses and other infectious agents are predicted to be causal agents in around 17.8% of human cancers (Parkin, 2006). Of the viruses which have been shown to cause cancer *in vivo* and/or *in vitro*, a number can be classified by genomic composition as DNA tumour viruses; these include papillomaviruses, polyomaviruses and adenoviruses. Other viruses shown to cause cancer include hepatitis B virus (HBV), hepatitis C virus (HCV), Epstein-Barr virus (EBV), human herpes virus 8 (HHV-8), human T-cell lymphotrophic virus, type-1 (HTLV-1) and Merkel cell polyomavirus (MCV) (Parkin, 2006). Cancer is an unintended consequence of the virus infection and results in disruption of the productive life cycle, preventing the production of viral progeny. Most DNA tumour viruses integrate into the host genome and only express part of the viral genome (Pagano et al., 2004). Normally infection with the virus is only a prerequisite for cancer development, with additional factors required for the progression to carcinogenesis (Liao, 2006).

Tumour viruses have become important tools in understanding not only the mechanisms of carcinogenesis but also general cellular biology. Studying these viruses has led to the discovery of multiple cellular oncoproteins and tumour suppressors, such as p53, along with an increased understanding of cellular pathways (Carrillo-Infante et al., 2007). Additionally, virally induced tumours can be treated with preventative therapeutics, such as the

prophylactic vaccines against HBV and HPV, giving new methods to help reduce the cancer burden.

1.2 Human papillomavirus

Papillomaviruses are a highly diverse group of small double-stranded DNA viruses, which can cause a range of lesions and diseases in a number of species. The most commonly studied host of papillomaviruses are humans, which are infected with 120 different types of human papillomavirus (HPV) (Bernard et al., 2010). HPV infects cutaneous and mucosal epithelia and causes a variety of lesions ranging from benign tumours (plantar, flat and common warts, genital condylomas and papillomas) to cervical neoplasia and cancer. Cutaneous and genital warts have been known to be infectious and caused by a virus since the early twentieth century, and in 1949 the viral particles isolated from skin papillomas were visualised by electron microscopy (Strauss et al., 1949, zur Hausen, 2009). The carcinogenic potential of papillomaviruses was initially demonstrated in rabbits, followed by investigations into the genital condylomata acuminata causing virus as a possible cause for cervical cancer (zur Hausen, 1976, 2009). HPV has since been shown to be a causative agent for cancers of the anogenital tract (cervical, vaginal, vulval, anal and penile) non-melanoma skin cancer and oropharyngeal cancer (Parkin, 2006).

1.3 Classification of papillomaviruses

Of the 120 types of HPV identified approximately one-third have been shown to target epithelial cells in the genital tract and these types can be divided into two groups of high and low risk viruses. The high risk HPV types include HPV16 and HPV18 and are found in 99.7% of cervical cancers (Walboomers et al., 1999). The low risk types, which include

HPV6 and HPV11, are more commonly associated with genital warts and are rarely present in cancers. HPV is the most common sexually transmitted disease and there is an 80% chance of contracting a genital infection of HPV in the sexually active population (Bekkers et al., 2004). High risk HPV infections usually persist for 12-18 months (Bodily and Laimins, 2011) and are cleared by the immune system, however about 10% to 20% of individuals do not clear the infection and develop a persistent infection (Stanley, 2008). Persistent infection by high risk HPV is a risk factor for the development of cervical intraepithelial neoplasia (CIN) and cancer.

The genomes of papillomaviruses are very static and mutations causing sequence changes are rare, occurring at a similar frequency as the genomes of the host (de Villiers et al., 2004). Within the genome the most conserved gene is the L1 ORF, this ORF is therefore used to classify different HPV types. HPV types are classified as a new type if the sequence of L1 differs by more than 10% to the closest known type; a subtype is therefore between 2% and 10% difference in homology and a variant if there is less than 2% variance (de Villiers et al., 2004). The HPV types can be further classified into genera and species, genera share between 60% and 70% nucleotide identity and species share between 71% and 89% (de Villiers et al., 2004). There are five HPV genera of which the main ones are alpha and beta which contain about 90% of known HPV types (Figure 1.1A). The largest genus is the alpha papillomaviruses, which contains both the high and low risk mucosal HPV types, along with some cutaneous viruses. The beta papillomaviruses are normally associated with insignificant cutaneous infections; however in individuals who are immunocompromised or have the genetic disease epidermodysplasia verruciformis (EV), the beta viruses can be associated with the development of non-melanoma skin cancer (Harwood and Proby, 2002). EV is a rare genodermatosis associated with a high-risk of non-melanoma skin cancer resulting from a high susceptibility to beta papillomavirus infection (Harwood and Proby, 2002).

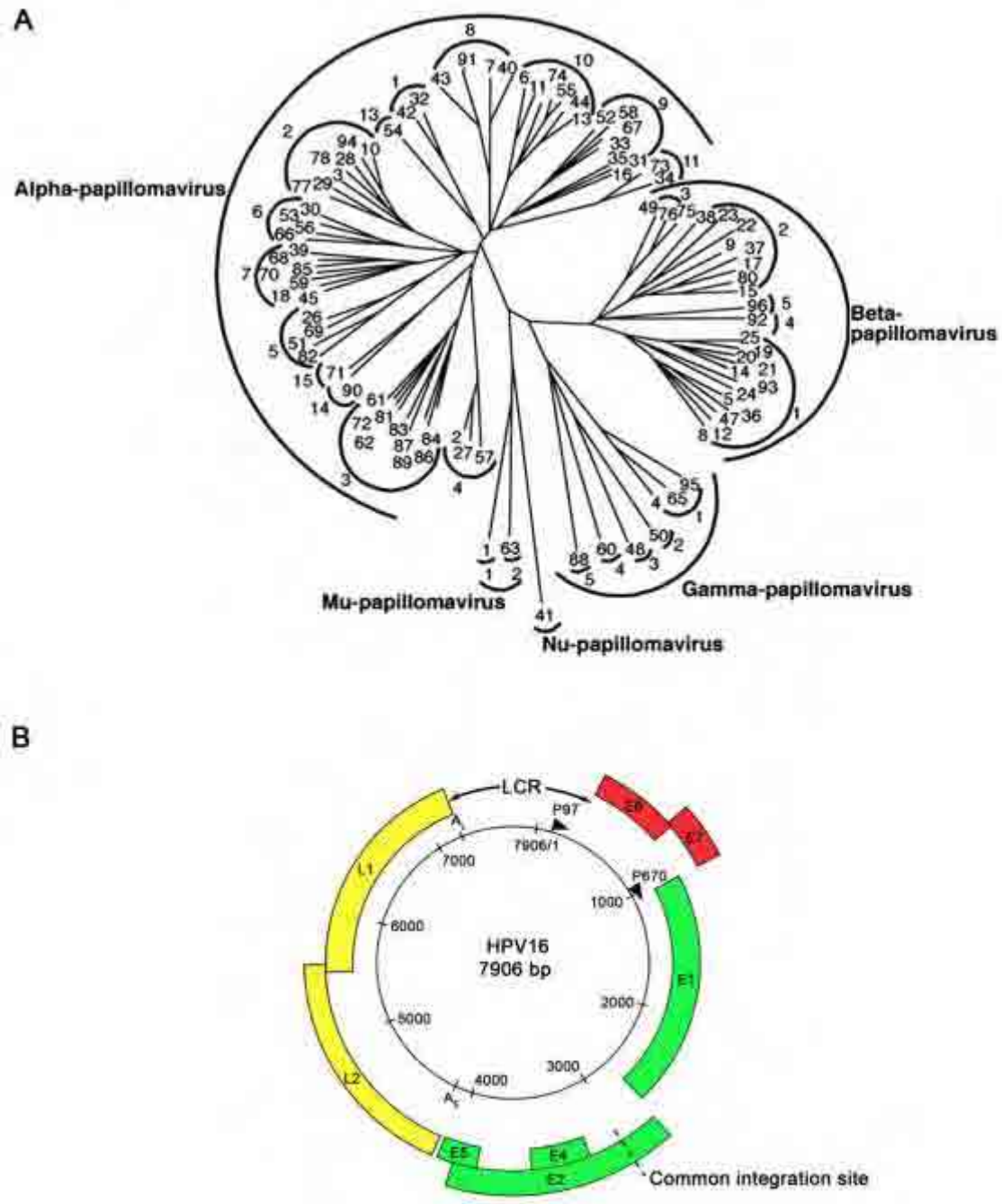

Figure 1.1 Phylogenetic tree of HPV and HPV genome map. (**A**) There are five evolutionary genera of HPV types and species within the genera. (**B**) HPV16 genome, early (p97) and late (p670) promoters are marked with arrows, polyadenylation sites A_E and A_L are marked, early proteins are in red and green, late proteins are in yellow. LCR: long control region. Integration often occurs within the E2 ORF, around the region indicated. Taken from (Doorbar, 2006)

The remainder of the HPVs belong to the genera gamma, mu and nu and are generally associated with cutaneous warts and verrucas which do not progress to cancer (de Villiers et al., 2004).

1.4 Warts and verrucas

Infection of the cornified stratified squamous epithelium of skin with HPV can cause mostly benign warts and papillomas to form; warts can vary in appearance and are classified based upon their appearance and epithelial surface infected. Different HPV types are more commonly found in the different wart types as shown in table 1. Warts have multiple histological differences, but are often elevated (acanthosis), with areas of hyperkeratosis (thickening of the cornified layer) and parakeratosis (retention of nuclei in cornified layer).

Table 1 HPV types associated with different warts

Clinical type	Distinguishing features	HPV type
Common warts (*Verruca vulgaris*)	Firm, rough keratotic papules and nodules on any skin surface	1, 2, 4, 57
Planar (flat) warts (*Verruca plana*)	2-4 mm in diameter, slightly elevated	3, 10
Intermediate warts (Butcher's warts)	Features of common and planar warts	2, 3, 7, 10, 28
Myrmecia (Palmoplantar warts)	Deep burrowing wart	1
Plantar warts (*Verruca plantaris*)	Keratotic surface with collar of thickened keratin	1, 2, 4, 57
Mosaic warts	Occur when palmar or plantar warts coalesce into large plaques	2
Anogenital warts (*Condyloma acuminata* or venereal warts)	Epidermal and dermal nodules and papules in the perineum and on the genitalia	6, 11
Oral warts	Small white of pink elevated papules on the oral mucosa	6, 11, 32

(Sterling et al., 2001)

Warts are most common in children and adolescents with the prevalence of visible warts in 11 and 16 year olds of between 3.9% and 4.9%, however most children clear warts easily and

spontaneously (Williams et al., 1993). In the United Kingdom there are regional variations as well as ethnic and socio-economic factors in the prevalence of warts in children (Williams et al., 1993). Warts can be painful, especially plantar warts and can be unsightly, leading to the need for treatment and removal, rather than wait for natural clearance. Treatment of cutaneous warts is usually by destructive methods such as treatment with salicylic acid or cryotherapy, however no treatment is 100% successful and most treatments require an immune response to completely clear the infection (Sterling et al., 2001). Generally there is not one therapy option that has been proven to be effective in all patients and thus multiple treatment options exist and are often used in combination.

1.5 Human papillomavirus and cancer

1.5.1 Non-genital

HPV is commonly associated with cancers of the anogenital tract; however HPV has been linked with a diverse range of malignancies, including non-melanoma skin cancer (NMSC). NMSC includes basal cell carcinomas (BCCs) and squamous cell carcinomas (SCCs) and is the most prevalent malignancy in the Caucasian population worldwide (Alam and Ratner, 2001). While BCCs are more common, accounting for around 80% of NMSC, and are locally aggressive they rarely metastasise, whereas SCCs can metastasise if not treated in a timely fashion, however still rarely cause death (Kiviat, 1999). Solar ultraviolet (UV) irradiation has long been established as the major risk factor for developing NMSC; however molecular studies have identified a role for HPV, acting as a co-factor with UV to cause skin cancer (Akgul et al., 2006). Skin cancers occur predominantly at sun-exposed sites and unlike anogenital cancers where at least one copy of the genome is present in each cell, only 1 in 20-50,000 cells contain HPV DNA (Pfister, 2003). The low copy numbers of viral DNA found

in NMSC suggests that cutaneous HPVs are involved in tumour initiation and progression, but are not required for the maintenance of the malignant phenotype (Pfister, 2003).

The immune status of the host is an important factor in the development of NMSC, with immunosuppressed organ-transplant recipients having a 10-fold increased risk of BCC and an up to 100-fold increased risk of SCC (Pfister, 2003). An increased risk of NMSC is also seen in individuals with the rare genodermatosis epidermodysplasia verruciformis (EV). EV is an inherited disorder of cell mediated immunity, meaning that individuals with this disease are unable to adequately control infection with certain HPV types (Kiviat, 1999). The mutations associated with EV are in the *EVER* genes, which negatively regulate zinc induced transcription factors and also regulate the zinc balance in keratinocytes (Lazarczyk et al., 2009). The knockdown of the *EVER* genes has been shown to correlate with increased activity of the AP-1 transcription factor, allowing the expression of transcription factors required for the expression and replication of the viral genome (Lazarczyk et al., 2009). The HPV types associated with EV are the beta papillomaviruses, with HPV5 and HPV8 being most commonly associated with SCC (Pfister, 2003). EV is characterised by widely spread, persistent, flat atypical warts and macular lesions which occur during childhood and can then metastasise later in life especially in sun-exposed areas of the skin (Pfister, 2003). The precise mechanism by which beta HPV types cause NMSC is unknown, as *in vitro* the E6 and E7 proteins of cutaneous HPVs have been shown to have very weak or no transforming properties (Akgul et al., 2006). The E6 protein of HPV38 is unable to induce p53 degradation, but has been shown to interact with the histone acetyltransferase p300 to inhibit p53 acetylation and therefore the inhibition of p53 mediated growth arrest and apoptosis (Muench et al., 2010). Additionally, several mechanisms by which HPV acts as a co-factor to UV have been proposed, these include the anti-apoptotic nature of E6 favouring the accumulation of somatic mutations, including p53 mutations which are common in NMSC

and abrogation of the p21 response to cellular stress by E6 (Alam and Ratner, 2001, Pfister, 2003). Also the immunomodulatory functions of E6 and E7 have been suggested to be involved in NMSC progression, as the secretion of anti-inflammatory cytokines by just a few HPV positive cells may impair the immunologic elimination of the cancer cells (Harwood and Proby, 2002, Pfister, 2003). The E6 proteins of HPV5, 10 and 77 have been shown to inhibit UV-induced apoptosis in a p53 independent manner, by abrogating the anti-apoptotic functions of Bak, by promoting its proteolytic degradation (Jackson et al., 2000). The E7 protein of HPV types 5, 8, 20 and 38 has been shown to upregulate the expression of lipocalin-2 in both monolayer and differentiated cells, a strong induction of lipocalin-2 is characterised by dysregulated epithelial differentiation. HPV-positive SCCs have increased levels of lipocalin-2 in the most differentiated layers and lipocalin-2 has been shown to have both pro- and anti-apoptotic functions. Although the precise biological effect of lipocalin-2 is unknown, the upregulation of lipocalin-2 by E7 may contribute to the progression of SCCs (Akgul et al., 2011).

The high-risk mucosal HPV types can infect the oral cavity and the oropharynx and have been etiologically linked with cancer at these sites. There are 5949 new cases of head and neck cancer each year in the UK (WHO/ICO, 2010) and the main causes of head and neck cancers are smoking and alcohol, however HPV has been found in up to 24% of oral SCCs and up to 60% of oropharyngeal SCCs in the USA, suggesting that HPV is a cause of these cancers (Kreimer et al., 2005, Marur et al., 2010). The main HPV type found in HPV-positive oropharyngeal SCCs and oral SCCs is HPV16, accounting for 87% and 68% respectively. Strangely HPV18 is rare in HPV-positive oropharyngeal SCCs, accounting for only 2.8%, but is more common in oral SCCs, accounting for 34% of the HPV-positive cancers (Kreimer et al., 2005). The incidence of oral HPV has been shown to increase with the number of sexual partners and is also more common in men, in HIV-infected individuals and in tobacco users

(Marur et al., 2010). HPV has been shown to be an important cause of oral cancers in both smokers and non-smokers but while the incidence of HPV-negative head and neck cancers, frequently associated with smoking and drinking, are decreasing, the incidence of HPV-positive head and neck cancers are increasing (Marur et al., 2010). The role of HPV in causing head and neck cancers is unknown, but it is predicted that HPV may play a similar role to that played in the aetiology of anogenital cancers.

1.5.2 Anogenital

Cervical cancer is the second most common cancer in women worldwide; in 2010 there were 529,828 cases of cervical cancer, which accounted for 275,128 deaths (WHO/ICO, 2010). The developing regions account for 76% of the cervical cancer cases worldwide and in 2010, 17% of the cervical cancer cases were in sub-Saharan Africa (Forouzanfar et al., 2011). In the United Kingdom the incidence of cervical cancer has more than halved since 1980, however in 2008 there were 2890 new cases of cervical cancer and 1111 deaths (Forouzanfar et al., 2011, WHO/ICO, 2010). Over 99.7% of cervical carcinomas have been shown to be positive for HPV DNA, allowing the conclusion to be reached that HPV infection is a necessary cause of cervical cancer (Walboomers et al., 1999).

Cervical cancer can be divided into two histological forms of cancer dependent upon the cell type which is infected. Squamous cell carcinomas (SCCs) of the cervix account for 80-85% of cervical cancer cases and are predominantly caused by HPV16 (62% HPV16, 8% HPV18) (de Sanjose et al., 2010). Adenocarcinomas, which are derived from adenomatous glandular cells, are less frequent, accounting for 15-20% of cases are commonly associated with HPV18 (32% HPV18, 50% HPV16) (de Sanjose et al., 2010, Green et al., 2003). However HPV types 31, 33, 35, 45, 52 and 58 have all been found in cervical cancers (de Sanjose et al., 2010). The prevalence of the different HPV types in cervical cancer varies with geography and age. Women with cervical cancers which are positive for HPV45 or HPV18 are generally

much younger than women with HPV16 positive cancers and who in turn are younger than women with cancers related to any other HPV type (de Sanjose et al., 2010). High risk HPV is also attributable to 40% of penile, vulvar and vaginal cancers and 90% of anal cancers (Parkin and Bray, 2006). The cancers of other anogenital regions are rare, however like cervical cancer they are normally SCCs, more common in the developing world and are usually caused by HPV16 and HPV18 (Parkin and Bray, 2006).

While infection with HPV is necessary for cervical cancer, HPV alone isn't sufficient to drive the progression to malignancy. A number of cofactors have been reported to be important environmental risk factors; these include smoking, sexual promiscuity, parity and oral contraceptive use (Green et al., 2003, Kjellberg et al., 2000). Sexual promiscuity increases the chances of contracting a HPV infection, whereas smoking has been shown to be the most significant non-HPV related environmental risk factor (Kjellberg et al., 2000). The chemical carcinogens found in cigarette smoke may influence progression to carcinogenesis by causing DNA damage and impairing the immune system, or may increase the risk of getting an infection with HPV and/or prolong that infection. Whatever the mechanism of action, smoking is an independent risk factor for cervical cancer and has been shown to be causal, with current smokers being twice as likely to be diagnosed with CIN as non-smokers (Collins et al., 2010, Kjellberg et al., 2000).

Cervical cancer has a well-characterised pre-malignant phase which can be analysed by cytological examination of exfoliated cervical cells and confirmed by histological examination of cervical tissue. The pre-malignant phases represent a range of histological abnormalities and are termed as cervical intraepithelial neoplasia (CIN) grades 1 to 3, with CIN1 representing mild dysplasia, CIN2, moderate dysplasia and CIN3, severe dysplasia or carcinoma *in situ* (Woodman et al., 2007). In low grade CIN, a productive life cycle of the

virus is supported, with viral progeny produced. However in high grade CIN, the HPV genome is often integrated into the host chromosomes, the site of integration is downstream of the E6 and E7 genes and often in the E1 or E2 region (Figure 1.1B). Integration leads to the disruption of E2 regulation of E6 and E7 expression and the subsequent upregulation of these oncoproteins (Woodman et al., 2007). When the genome is not integrated, the E2 binding sites in the early promoter are often methylated, preventing E2 binding and the repression of expression.

1.5.3 HPV screening and prophylactic vaccine

The rates of cervical cancer have been falling over the last few decades in developed countries, due to the implementation of screening programmes, which allow early detection and early treatment of pre-cancerous cervical lesions. In the UK the Papanicolaou (Pap) test of exfoliated cervical cells was introduced nationally in 1988 and since then the mortality rates of cervical cancer have decreased by about 80% (Peto et al., 2004). Cervical cancer screening is based upon cytological and colposcopic analyses, looking for the presence of abnormal cells rather than the presence of HPV DNA (zur Hausen, 2002). The screening programmes have reduced the incidence of cervical SCC since their introduction; however the rates of adenocarcinoma have been reported to have increased, this may be attributable to cervical screening being less efficient at the detection of adenocarcinoma (Parkin and Bray, 2006). The Pap test looks at the severity of cytological abnormalities; however adenocarcinoma, which is associated with HPV18, has only minor cytological abnormalities when compared to SCC. Therefore the exfoliated cells are less likely to have severe cytological abnormalities and be referred for colposcopy. This means that the screening programme is more likely to diagnose and interrupt disease progression with a HPV16 infection than with HPV18 (Woodman et al., 2007). The efficacy of cytological examinations has improved since the 1990s, with the introduction of liquid-based cytological screening and

new tests looking at viral load, viral genome integration and methylation of tumour suppressor genes are being developed, together these may aid in the detection of adenocarcinoma and reduce mortality rates further. As the number of cases of cervical cancer in the developed world decreases due to better understanding of the causes of cervical cancer and more improved screening protocols, the number of cases and mortality in the developing world is increasing or remaining constant (Forouzanfar et al., 2011, Parkin and Bray, 2006), in areas where screening is absent due to costs and expertise, different approaches are required to combat this virus.

Initial experiments with virus like particles (VLPs), which are the self-assembled papillomavirus structural proteins, in animals resulted in effective protection against primary papillomavirus infections (Breitburd et al., 1995, Suzich et al., 1995). There are now two prophylactic vaccines against HPV infection which have been licensed to date, the quadrivalent vaccine Gardasil® protects against HPV16, 18, 6 and 11 and is produced by Merck & Co., Inc. The bivalent vaccine Cervarix® is protective against HPV16 and 18 and is produced by GlaxoSmithKline plc. Both vaccines also show some level of cross-protection against infection with other HPV types; including the closely related types HPV31 and HPV45 (Bonanni et al., 2009). The vaccines are composed of HPV L1 proteins assembled into VLPs and are administered intramuscularly by injection in three doses at 0, 1 and 6 months (Harper et al., 2004, Villa et al., 2006). In the UK, both vaccines have been approved for use, but the bivalent vaccine Cervarix® has NHS funding for administration to girls aged 12-13, along with a catch up campaign for girls up to 18 years old, with the aim to vaccinate before sexual debut (Koulova et al., 2008). The uptake rate in England is 87.4% for the first injection, but only 76.6% for the third, meaning some girls are not getting the full protection. In some areas of England the uptake rate is less than 70%, this is often in areas with a large

ethnic population and low income families (DoHealth, 2011). In the UK males are not being vaccinated, as this is not considered to be cost-effective (Koulova et al., 2008).

Cervical cancer screening programmes are the most expensive preventative health care measure in the world, costing the NHS around £157 million in England (including treatment), the vaccination programme may help to combat some of this cost, however screening will still be required against the HPV types not targeted by the vaccines (Villa, 2011). The bivalent vaccine does not protect against HPV6 and 11 which cause genital warts and cost the NHS £22.4 million in 2003 (Brown et al., 2006), however this vaccine was chosen on purchase cost and not the potential savings. The vaccines are expensive and are administered in three injections which need to be cold, factors which make implementation of the vaccination programmes in developing countries unfeasible, therefore other methods of delivery are required for the countries in which the disease burden is greatest.

1.6 The HPV genome

Papillomaviruses have a circular double stranded DNA genome of about 8 kb contained within an icosahedral capsid of around 55 nm in diameter (Z. M. Zheng and Baker, 2006). The genome can be split into three regions: the long control region (LCR), the early (E) region and the late (L) region; they are separated by two polyadenylation (poly(A)) sites: early poly(A) (A_E) and late poly(A) (A_L) (Figure 1.1B) (Z. M. Zheng and Baker, 2006). The early region accounts for over 50% of the genome and commonly encodes for six early proteins: E1, E2, E4, E5, E6 and E7, which are responsible for transcription, viral genome replication and transformation of the host cell. In the case of genital HPV types an E8^E2 splice product is also coded for. The late region occupies nearly 40% of the genome and lies downstream from the early region, the late region encodes for two late viral proteins: L1 and

L2, which encode the structural major (L1) and minor (L2) capsid proteins (Z. M. Zheng and Baker, 2006). The LCR makes up the remaining 10% of the genome and has no-protein coding function, but has the origin of replication along with multiple transcription factor binding sites, which are involved in RNA polymerase II-initiated transcription regulation (Z. M. Zheng and Baker, 2006).

1.7 HPV life cycle

Papillomaviruses infect stratified squamous epithelia and link their productive life cycle to the differentiation of the host cell. In normal epithelia only cells in the basal and parabasal layers are undergoing cell division, as the process of cell differentiation involves the exit of cell cycle along with a series of complex changes in gene expression (Bodily and Laimins, 2011). The virus infects the dividing cells of the basal layer, most likely gaining entry to this layer through micro-abrasions of the tissue. Although the precise mechanism of viral uptake by the cells is unknown, it is believed that the heparin sulphate proteoglycans in the basement membrane are involved in the initial binding steps and/or virus uptake (Joyce et al., 1999). Following the initial binding steps, a conformational change of the capsid and furin cleavage of L2 occurs (Kines et al., 2009). Binding then occurs with a secondary receptor on the leading edge of the epithelial cells, which move to heal the wound and form the basal layer (Schiller et al., 2010). Entry into the cells is a relatively slow process and occurs via a keratinocyte specific receptor (Schiller et al., 2010). The virus enters the cell through clathrin-mediated endocytosis and travels through the endocytotic pathway where uncoating occurs. The uncoated genome remains bound to L2 and enters the nucleus where colocalisation with the ND10 body occurs and RNA transcription begins (Schiller et al., 2010). Following infection the life cycle of HPV can be divided into two stages, early and late (Figure 1.2).

1.7.1 Early virus life cycle

Following infection the viral genome is established as a stable episome in cells in the basal layer. Among the first viral proteins to be expressed are the replication factors E1 and E2, E2 acts to recruit E1 to the viral origin of replication, E1 then in turn recruits cellular replication factors to mediate viral replication (Longworth and Laimins, 2004b). In the basal layer the cells are undergoing cell cycle, therefore the viral genome can be replicated with the cellular DNA during S-phase. The replicated viral genomes are equally divided between the daughter cells, due to the anchoring of viral episomes to mitotic chromosomes by E2. In the basal layer the virus is maintained at approximately 20 to 100 episomal copies per cell (Longworth and Laimins, 2004b).

Following cell division, one of the daughter cells migrates upwards into the parabasal layer, where the process of terminal differentiation starts, while the other continues to divide in the basal layer. In uninfected epithelia, cells exit cell cycle once they leave the basal layer, however the virus requires the cellular replication machinery to replicate, therefore the early protein E7 drives the cells back into S-phase by binding to the retinoblastoma (Rb) family of tumour suppressors and other proteins involved in cell cycle regulation (Dyson et al., 1989). The re-entry into cell cycle induced by E7 activates pathways leading to growth arrest or apoptosis; however the early protein E6 has anti-apoptotic functions and associates with p53 to allow the cell to enter a pseudo-S-phase where the virus can replicate. During an infection caused by cutaneous viruses such as HPV1, only cells in the basal layer are proliferating, with the late phase of the life cycle occurring as soon as the cell leaves the basal layer (Peh et al., 2002). In infections with mucosal viruses, such as HPV6 and HPV16, a region of many layers of cells is retained in cell cycle prior to the onset of the late phase of viral genome amplification (Peh et al., 2002).

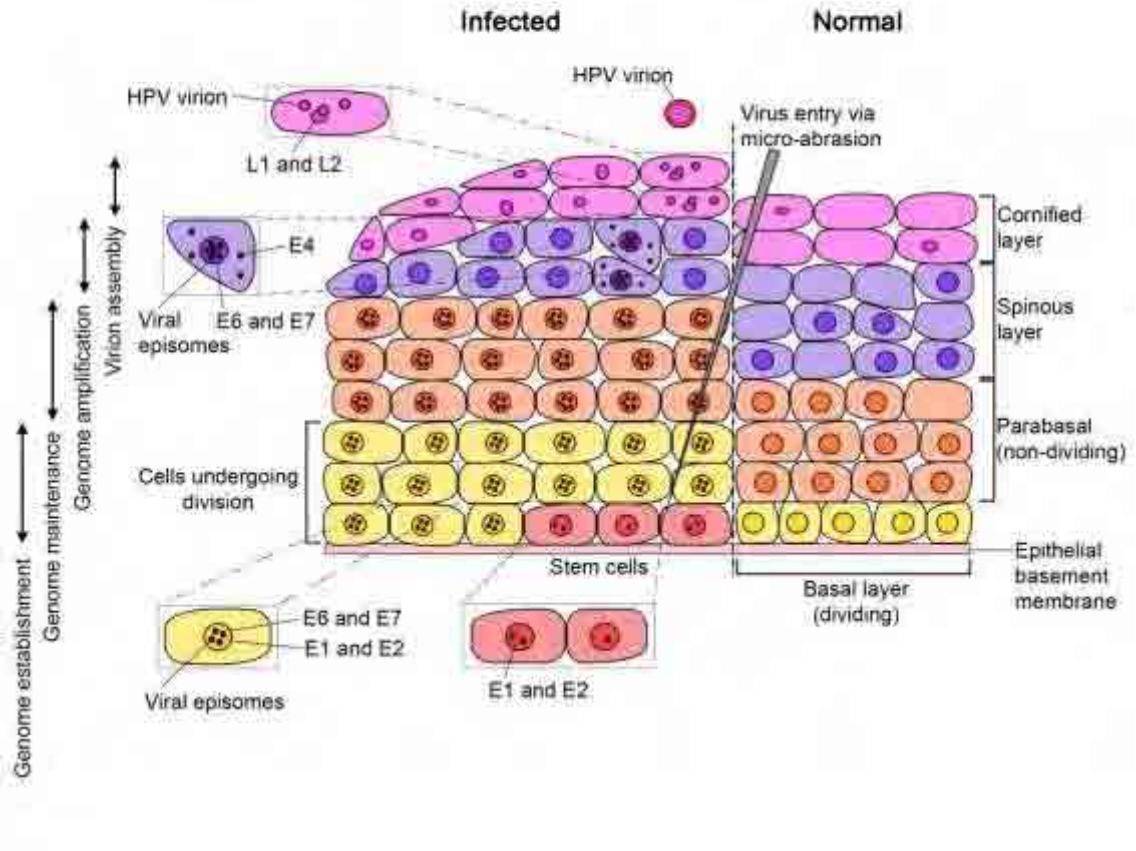

Figure 1.2 Life cycle of HPV in dividing epithelia. In normal epithelia (right side) cells exit cell cycle once they migrate upwards from the basal layer. In infected epithelia (left side), as the cells migrate from the basal layer they continue to undergo cell cycle. Viral proteins are expressed with differentiation as shown, with mature virions only produced in the most superficial layers of the epithelium. Adapted from (Frazer, 2004).

1.7.2 Late virus life cycle

The late stages of the virus life cycle are dependent upon cell differentiation and the up-regulation of the differentiation-dependent promoter. The trigger for the onset of the late events is not fully understood, however changes in cellular signalling are involved (Bernard, 2002). The activation of the late promoter leads to an increase in the levels of viral proteins required for replication, the proteins expressed are the E1, E2, E1^E4 and E5 early proteins and the L1 and L2 late proteins. The viral genome is amplified to thousands of copies per cell, before being packaged into infectious virions formed by the capsid proteins L1 and L2. The infectious virions are then shed with the desquamated cells (Bodily and Laimins, 2011).

1.8 Transcriptional control in HPV

1.8.1 Early promoter control

In high risk HPV types there are two major viral promoters, along with several other minor promoters, the main promoter which encodes early transcripts initiates upstream of the E6 ORF and is expressed prior to productive replication, in HPV16 this promoter is p97 in HPV18 it is p105 (Longworth and Laimins, 2004b). The cutaneous HPV types, such as HPV1 also contain a promoter upstream of the E6 ORF, which codes for the major early transcripts (Baker and Calef, 1997). The early promoter is tightly controlled by upstream cis-elements in the LCR, including four consensus E2-binding sites, (ACCGNNNNCGGT), which act to regulate transcription from the early promoter during differentiation (Bernard, 2002) (Figure 1.3A). In HPV1 and BPV1, E2 has been shown to activate the early promoter, however in mucosal HPVs, E2 acts primarily as a repressor (Bernard, 2002, Hirochika et al., 1987). In mucosal HPV types, the E2 binding sites overlap with a Sp1 binding site and the TATA box, the binding site for TFIID and general transcription factors, E2 binding to these sites displaces the transcription factors, first Sp1 and then TFIID as the concentration of E2

increases, this allows multiple levels of cell type specific repression of the promoter (Tan et al., 1994) (Figure 1.3B). In cutaneous HPV types, the distance between the E2 binding sites and the TATA box is greater, so E2 binding does not displace the transcription factors (Garrido-Guerrero et al., 1996). The repression of the early promoter by E2 allows feedback to control the level of transcription of the early genes and keep the episomal copy number of the genome stable. In over half of HPV16- and most of HPV18-positive malignant lesions (Moody and Laimins, 2010), the HPV genome is integrated into the host chromosomes, this integration generally occurs downstream of the E7 gene, disrupting the expression of E2. The loss of E2 regulation of transcription following integration leads to an increase in E6 and E7 expression and may impart a selective growth advantage over cells harbouring only episomal copies of the genome (Francis et al., 2000, Jeon et al., 1995).

The LCR contains an enhancer to stimulate the promoter; the enhancer is a 400 bp segment, 100 bp upstream of the promoter, which contains clusters of binding sites for cellular transcription factors that can activate the promoter over a long distance (Bernard, 2002). The HPV enhancer functions only in epithelial cells, restricting HPV exclusively to cells of epithelial origin. The epithelial specificity of the enhancers originates from transcription factors which are ubiquitously expressed, but made up from different subunits in epithelial cells to other cell types; these include activator protein-1 (AP-1), nuclear factor 1 (NFI) and transcriptional enhancer factor 1 (TEF-1) (Bernard, 2002). The 100 bp between the promoter and the enhancer contains binding sites for YY1 and CCAAT-displacement protein (CDP), which act as silencers for the promoter and repress transcription (Bernard, 2002). The cellular factors involved in repression, alter in concentration during differentiation, allowing the coupling of transcription and replication to epithelial differentiation (Bernard, 2002). The different activators and repressors which can bind to the enhancer and promoter are shown in Figure 1.3C.

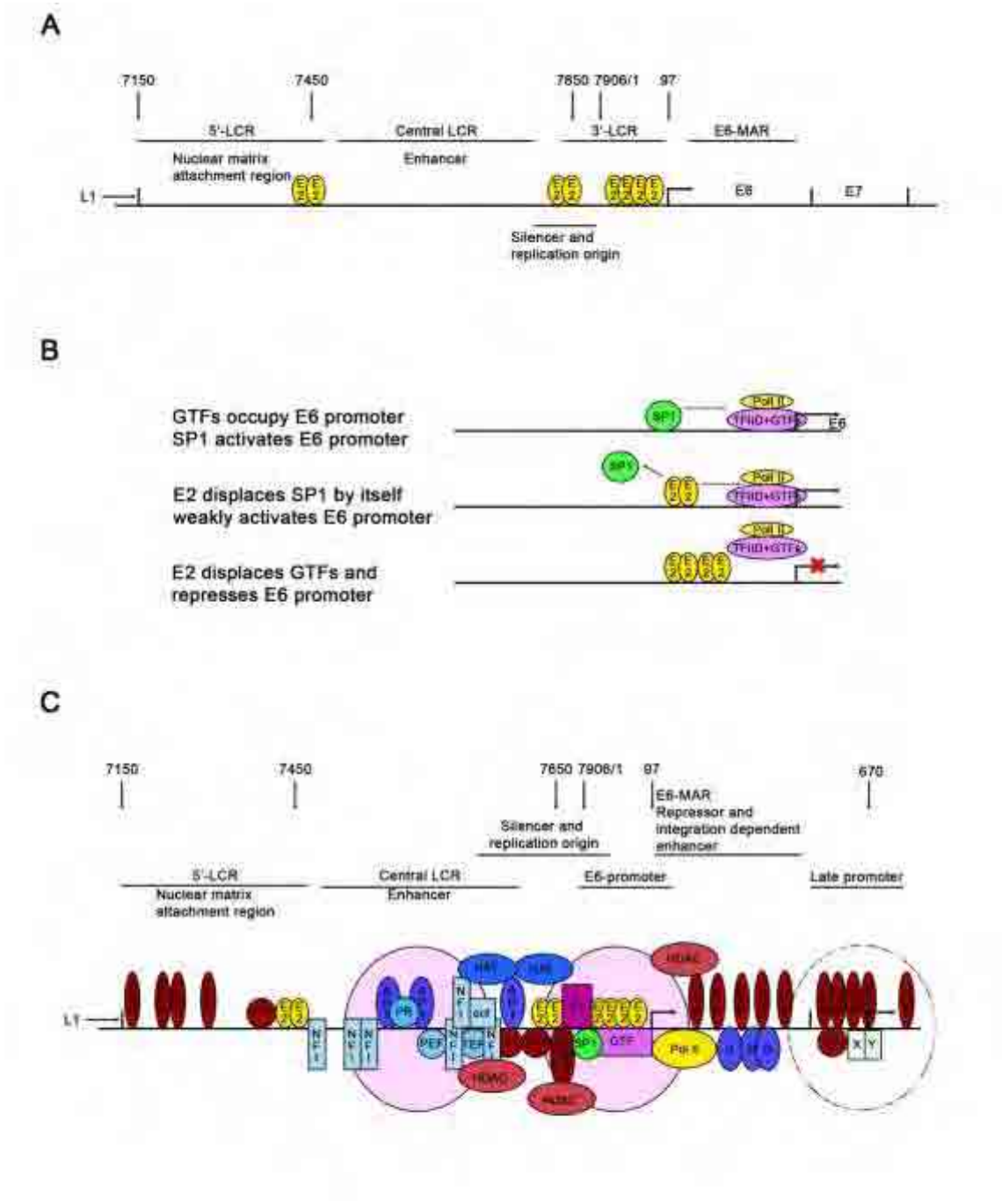

Figure 1.3 Control of transcription in HPV. (**A**) LCR of HPV16 showing the position of conserved E2 binding sites. (**B**) Modulation of the E6 promoter by E2 with increasing amounts of E2. (**C**) Summary of all currently known proteins bound to the HPV16 LCR and the E6 and E7 promoters. Activator proteins are in blue, repressor proteins are in red. Adapted from (Bernard, 2002)

1.8.2 Late promoter control

The late promoter is activated with the induction of the productive phase and in genital HPV types directs expression for a series of heterogeneous start sites located within the E7 ORF, in HPV16 this promoter is p670 and in HPV18 p811 (Grassmann et al., 1996, Longworth and Laimins, 2004b, X. Wang et al., 2011). However in non-genital HPV types such as HPV1 and HPV8, the late promoter is in the LCR upstream from the early promoter (Bernard, 2002), the promoter in the E7 region of HPV1 can also transcribe late genes (Baker and Calef, 1997). The location of the late promoter, results in the late mRNA transcripts containing all or most of the early region and the late region. The late mRNA can be processed into early region transcripts which are cleaved and polyadenylated at the early poly(A) site and late region transcripts which are polyadenylated at the late poly(A) site (Z. M. Zheng and Baker, 2006). The mechanisms of activation of the late promoter are unknown but several cellular transcription factors have been shown to repress the promoter, including CDP and YY1 (Bernard, 2002) (Figure 1.3C). The differentiation specific transcription factor Skn-1a can displace YY1 in a competitive manner and activate the late promoter (Kukimoto and Kanda, 2001), indicating that the late promoter could be controlled in a similar differentiation dependent manner to the early promoter (Bernard, 2002). Transcripts from the late promoter are only detectable in differentiated keratinocytes (Barksdale and Baker, 1995).

1.9 Post-transcriptional modifications

1.9.1 Genome splicing

The transcripts produced from the HPV promoters undergo a large amount of post-transcriptional modifications to produce multiple proteins from a single transcript. Alternative splicing allows the generation of multiple mRNAs from one single pre-mRNA

and is required for the differential expression of the HPV genes (Baker, 1997). In HPV1 there are three putative promoters located in the E7 ORF, upstream of the E6 ORF and in the upstream regulatory region (URR), the major promoter is located within the E7 ORF (Palermo-Dilts et al., 1990) (Figure 1.4). The early and late poly(A) sites have not been determined experimentally, but there are three poly(A) signals at 3985-3990, 7381-7386 and 7427-7432 (Baker, 1997). Within the genome there are four alternative 5' splice sites (ss), located at nt 827, 1231, 3592 and 7711 and three alternative 3' ss located at nt 2545, 3200 and 5432 (Baker and Calef, 1997, Palermo-Dilts et al., 1990). The most abundant mRNA in plantar warts is species A (Figure 1.4) which encodes the E1^E4 ORF and the most abundant L1 mRNA is species E (Figure 1.4), both transcripts are transcribed from the promoter in the E7 ORF, however species A is about 10 times more abundant than species E (Chow et al., 1987).

The early HPV16 pre-mRNAs produced from the p97 promoter have three exons and two introns, which undergo alternative RNA splicing and are polyadenylated at nt 4215 (Z. M. Zheng and Baker, 2006) (Figure 1.5). There is one alternative 5' ss at nt 226 and three alternative 3' ss in intron 1 (nt 409, 526 and 742), this allows the truncated species of E6 (E6*I, E6*II and E6*III) to be produced, when this intron is unspliced the full length oncogenic E6 is produced (Z. M. Zheng and Baker, 2006, Z. M. Zheng et al., 2004). Intron 2 contains three alternative 3' ss (nt 2582, 2709, 3358), which can lead to the production of another truncated E6 species, E6*IV, and E2, altogether the splices sites lead to the production of at least fourteen different mRNA transcripts with various coding potential (Z. M. Zheng and Baker, 2006). Interestingly the E6 coding regions for HPV18 and HPV31 have only one 3' ss and the low-risk HPV types such as HPV6 and HPV11 have none (Z. M. Zheng et al., 2004).

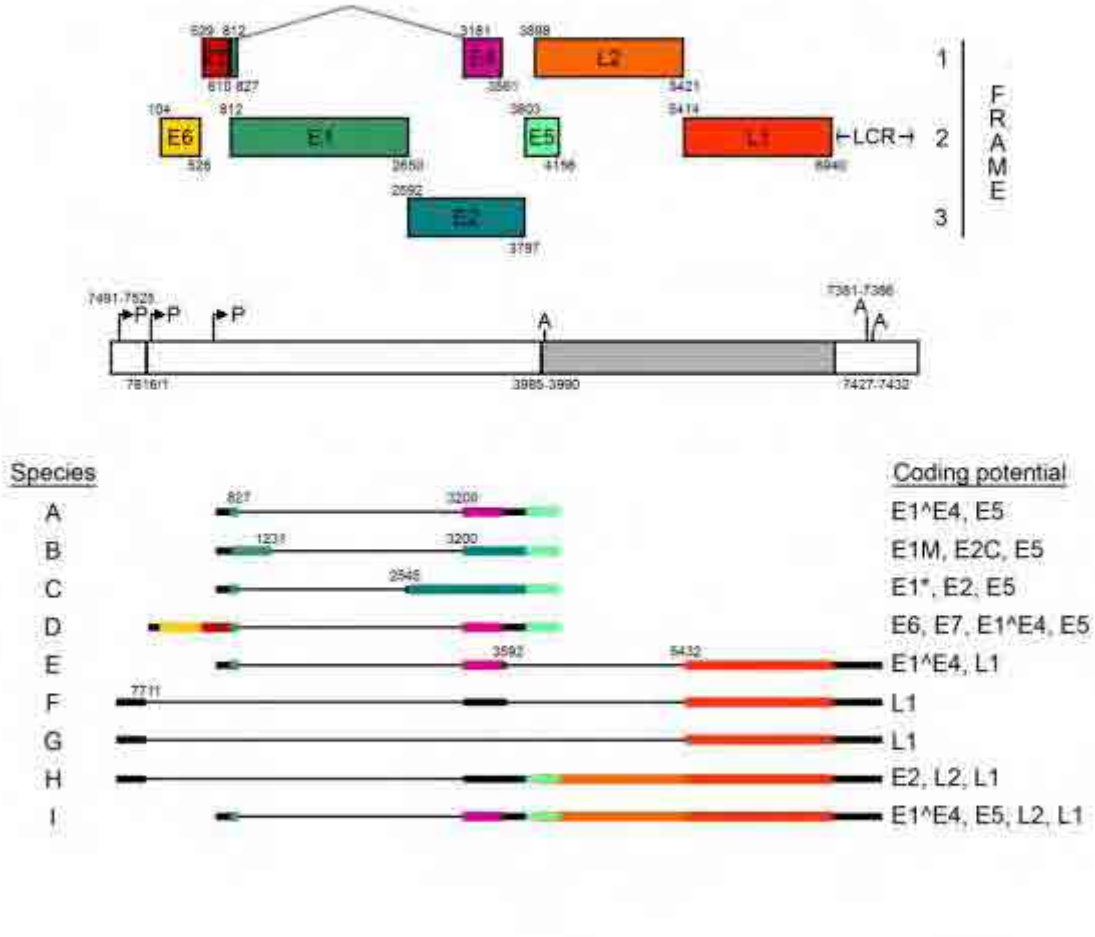

Figure 1.4 Genome structure and transcript map of HPV1. (Top) Open reading frames of HPV1 genes, numbers are positions of first nt of the start codon and last nt of stop codon. (Middle) Linear form of the HPV1 genome showing position of promoters and polyadenylation sites. (Bottom) Reported RNA species derived from alternative promoter usage and alternative splicing. Exons (thick lines) are coloured according to the protein produced, numbers are the splice site positions. Adapted from (Baker and Calef, 1997)

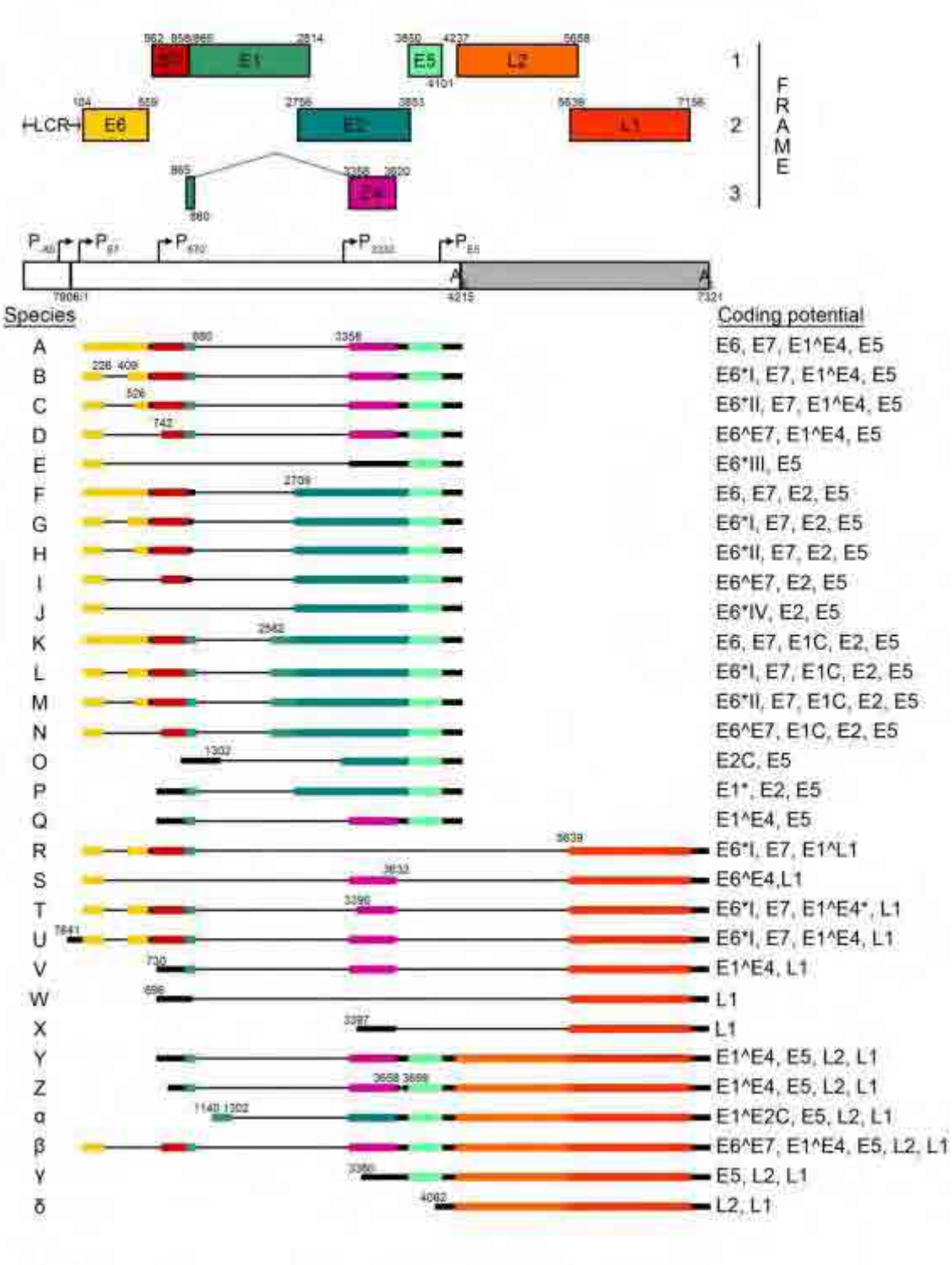

Figure 1.5 Genome structure and transcript map of HPV16. (Top) Open reading frames of HPV16 genes, numbers are positions of first nt of the start codon and last nt of stop codon. (Middle) Linear form of the HPV16 genome showing position of promoters and polyadenylation sites. (Bottom) Reported RNA species derived from alternative promoter usage and alternative splicing. Exons (thick lines) are coloured according to the protein produced, numbers are the splice site positions. Adapted from (Baker and Calef, 1997, Milligan et al., 2007, Z. M. Zheng and Baker, 2006)

The transcripts containing E6 and E7 are only efficiently spliced when capped, with cellular cap-binding factors involved in the splicing. The efficiency of the splicing of the capped RNA, depends on the distance of the cap-proximal intron from the RNA 5' cap (Z. M. Zheng et al., 2004). Splicing of the E6E7 mRNA has been shown to promote E7 translation, suggesting that the small cap-proximal exon allows HPV16 to confer efficient splicing of the transcripts to get better E7 expression. However splicing of the E6E7 transcript does not produce the full length E6 protein, so a balance in splicing needs to be maintained to produce the optimal ratio of E6 and E7 proteins throughout the virus life cycle (Z. M. Zheng et al., 2004).

1.9.2 Control of genome splicing

The late mRNAs of HPV16 are initiated and spliced at splice sites in the early region and are processed upstream of the early poly(A) site, a site which is used by both early and late mRNAs (Rush et al., 2005). Therefore the regulation of splicing of HPV16 late pre-mRNAs utilises exonic splicing enhancers (ESEs) and exonic splicing silencers (ESSs) to avoid premature late gene expression (Figure 1.6). The 3' splice site at nt 3358, which allows the production of E1^E4, is a suboptimal splice site, which is promoted through a 65 nucleotide AC-rich ESE located approximately 100 nucleotides downstream in exon 2 (Rush et al., 2005). The 3358 ss may also be promoted by terminal exon recognition, due to this splice site being the most 3' splice site to the early poly(A) signal. At the end of the E4 ORF is a 5' splice site, however, approximately 45 nucleotides upstream of the 3632 5' ss is an ESS, this may be present to counteract recognition of the late-specific 5' 3652 ss and enhance recognition of the early poly(A) site, to allow the production of early mRNAs terminating at the early poly(A) and block premature late mRNA expression (Rush et al., 2005). The late pre-mRNAs contain three exons and two introns along with two potential poly(A) sites, the second intron is an alternative intron/exon, retention of which allows the production of L2 (Z.

M. Zheng and Baker, 2006). The retention of the second intron may be enhanced by an ESS in the 3' terminal exon of the late transcripts, which binds hnRNP A1 (X. Zhao et al., 2004). Although the precise mechanisms of control of late gene expression are unknown, they are partly regulated by a family of splicing factors, serine arginine (SR) proteins, which will be covered in greater detail in a later section.

1.9.3 Polyadenylation

The maturation of the 3' end of most pre-mRNAs involves cleavage of the nascent transcript and addition of a poly(A) tail (polyadenylation), the polyadenylation of mRNA is another point where the expression of HPV proteins in controlled. In HPV16 the early and late transcription units overlap each other and the two poly(A) sites are in competition with each other (Z. M. Zheng and Baker, 2006). Cis-elements in the 3' untranslated regions (UTR) of late transcripts have been shown to be cryptic 5' splice sites which inhibit the usage of the late poly(A) site (Furth et al., 1994). In HPV16 the negatively regulatory element (NRE) is 79 nucleotides long and overlaps the 3' end of the L1 ORF and extends into the 3' UTR. The 5' end of the HPV16 NRE is composed of four cryptic 5' splice sites, which contribute together to make the NRE function and interact with a U1 snRNP-like complex (Cumming et al., 2003). The 5' end also contains a predicted stem-loop structure, whereas the 3' end is very GU rich and is a less potent inhibitor than the 5' end (Cumming et al., 2003). The NRE along with cis-acting elements which can up-regulate usage of the early poly(A) site means the early mRNAs are produced preferentially in undifferentiated cells, how the activation of the late poly(A) site occurs is unknown, but is probably due to changes in concentrations of the cellular proteins which interact with the cis-elements in the mRNA.

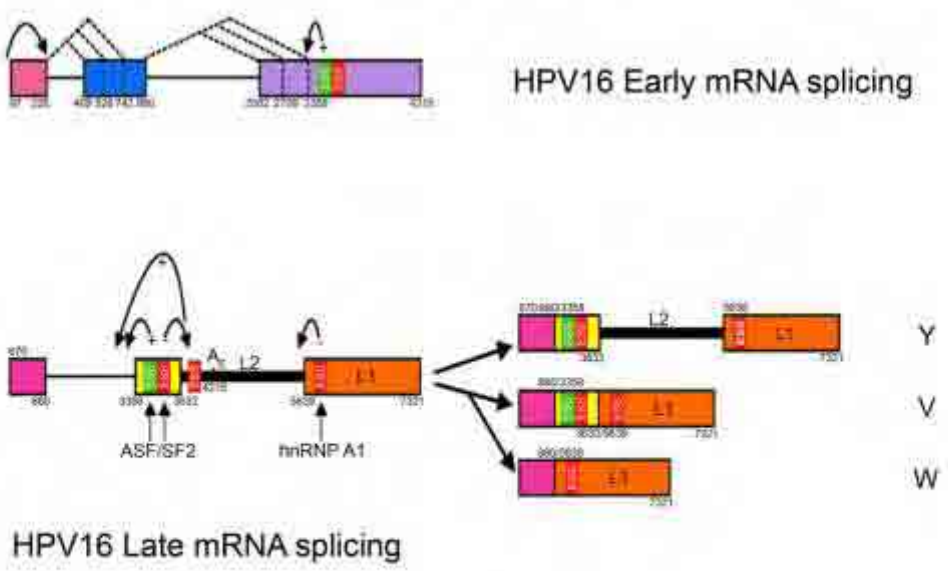

Figure 1.6 Regulation of HPV16 alternative splicing. (Top) Alternative 3' splice site selection in early transcripts. (Bottom) Production of L1 and L2 are regulated by multiple cis-acting elements. Adapted from (Z. M. Zheng and Baker, 2006)

1.9.4 Translation

The bicistronic and polycistronic transcripts produced by HPV which have tandemly arranged ORFs require specialised mechanisms for initiation of protein translation on a downstream ORF. Ribosomes scan mRNA in a linear fashion starting at the 5' AUG, how they bypass an upstream ORF to translate the next one is unknown. Several mechanisms have been proposed and have some experimental support in HPV, these include: translation re-initiation, leaky scanning and ribosome shunting (Remm et al., 1999, Stacey et al., 2000, Z. M. Zheng et al., 2004). In HPV16 the E6 and E7 ORFs are only separated by 2 nucleotides, in HPV18 this gap is 9 nucleotides, consequently the ribosome would not be able to efficiently re-initiate E7 translation after translating E6 in an unspliced full-length E6E7 mRNA, therefore the unspliced E6E7 mRNAs mainly serve for the translation of E6 (Tang et al., 2006). The splicing which produces E6*I mRNAs creates a premature stop codon after the splice junction and therefore increases the distance between the E6*I and the E7 ORFs to around 130 nucleotides, allowing a scanning ribosome to re-initiate E7 translation (Tang et al., 2006). Leaky scanning ribosomes have been shown to scan over the E6 ORF to then start translation at the E7 AUG (Stacey et al., 2000) and ribosomes have also been shown to discontinuously scan over the polycistronic transcripts containing E6, E7 and E1 ORFs to translate E1 (Remm et al., 1999). In practice the translation of the HPV genome may be down to a collection of several different mechanisms acting together to regulate the levels of the proteins throughout the virus life cycle.

1.10 HPV proteins

The early E1 and E2 proteins produced by HPV are the viral replication proteins, involved in initiation of replication, genome segregation and DNA replication. E5, E6 and E7 are viral oncoproteins, involved in the regulation of cell cycle, cellular transformation and immune

modulation (Kim et al., 2010). The main functions of the E1^E4 protein in the virus life cycle are largely unknown, however E1^E4 has been shown to be involved in the regulation of the late virus functions. The late proteins L1 and L2 spontaneously form the icosahedral capsids, in which the viral DNA is packaged.

1.10.1 HPV E1

The E1 protein is a phosphoprotein which is approximately 650 amino acids long (68 kDa) and belongs to the helicase superfamily III (SF3). E1 can be divided into three functional regions (Hickman and Dyda, 2005). The C-terminal domain has ATPase and helicase activity and can self assemble into hexamers, this domain is also used by E2 to recruit E1 to the AT-rich sequences at the viral origin of replication, which are located proximal to the start sites of early transcription (Hughes and Romanos, 1993, Longworth and Laimins, 2004b, Y. Sun et al., 1998, Titolo et al., 2000). The central region of E1 contains the DNA-binding domain (DBD), which binds to the origin and dimerises (Titolo et al., 2003). The N-terminal region of E1 contains a cyclin E/A-CDK2 binding motif, a nuclear localisation signal (NLS) and a CRM1-dependent nuclear export signal (NES), which act together to regulate the nucleocytoplasmic shuttling of the protein (Deng et al., 2004, Fradet-Turcotte et al., 2010). Alone, E1 only weakly binds origin sequences, but this binding is facilitated by the formation of a complex with E2 (Frattini and Laimins, 1994), once bound to the origin, the E1 proteins form hexamers with a high affinity for DNA (Sedman and Stenlund, 1998). The E1 complexes then efficiently unwind the supercoiled DNA with the help of chaperone proteins (Longworth and Laimins, 2004b). E1 also binds to the p68 subunit of DNA polymerase α-primase, to help recruit cellular replication complexes to the viral origin (Masterson et al., 1998). Phosphorylation of the N-terminal domain of E1 inhibits the NES, and results in E1 accumulation in the nucleus; the nuclear export of E1 is not essential for viral DNA replication but is required for long-term maintenance of episomal genomes (Fradet-Turcotte

et al., 2010). Continuous nuclear accumulation of E1 leads to reduced cellular proliferation, due to S phase progression being delayed (Fradet-Turcotte et al., 2010). The E1 protein is expressed at low levels in HPV-infected cells and an increase in E1 expression has a negative effect on replication efficiency, the ratio of E1 to E2 also has been shown to be important in determining replication efficiency (Deng et al., 2003).

1.10.2 HPV E2

The E2 proteins are approximately 365 amino acids long (50 kDa) and function as dimers; they are required for both viral DNA replication and transcriptional regulation (Longworth and Laimins, 2004b). The C-terminal region has a dimeric β-barrel structure and encodes a DNA-binding domain (Hegde et al., 1992); the DBD of E2 interacts with the DBD of E1 (Chen and Stenlund, 2000). The N-terminal region consists of a glutamine-rich α-helix packed against a β-sheet framework and contains a transactivation domain (Antson et al., 2000). The N- and C-terminal domains are linked by a variable length hinge region, which is unstructured (Gauthier et al., 1991). The E2 proteins bind to the consensus palindromic sequences in the LCR as dimers, three of the four E2-binding sites flank the E1 recognition sequences in the origin of replication, allowing E2 to load E1 onto the origin (Longworth and Laimins, 2004b). The transcription of early genes following infection is primarily activated by cellular transcription factors and at low concentrations E2 continues to activate early gene expression (Steger and Corbach, 1997). However, when the concentration of E2 increases, the binding of transcription factors to their recognition sequences is inhibited, as there is an overlap in the binding sites (Demeret et al., 1997). Upon differentiation, the late promoter, which is not repressed by E2, is activated, this leads to an increase in E1 and E2 expression and viral amplification occurs (Klumpp and Laimins, 1999). The E2 protein has also been shown to improve the efficiency of genome encapsidation, as the E2 brings the viral DNA to ND10 bodies where capsid formation occurs (K. N. Zhao et al., 2000).

Another role for the E2 protein is in the long-term episomal maintenance of viral genomes and in partitioning the viral DNA into daughter cells following cell division (Bastien and McBride, 2000, Lehman and Botchan, 1998). The E2 protein tethers the viral genome to the host chromosome during mitosis helping to maintain an equal division of nuclear located episomes between the cells (Bastien and McBride, 2000). The tethering of E2 to chromosomes is conserved among many HPV types; however the mitotic binding pattern varies between different types with the majority appearing as small speckles over the arms of all mitotic chromosomes (Oliveira et al., 2006). The E2 protein from the beta papillomavirus HPV8, bind as large foci to the pericentromeric regions of chromosomes (Oliveira et al., 2006). The E2 proteins from the beta papillomaviruses have longer hinge regions than those in the other genera, which are rich in RG and RS dipeptide motifs. The E2 protein of HPV16 has been shown to directly interact with the E7 protein and be involved in the recruitment of E7 to the mitotic chromosomes during the later stages of mitosis (Gammoh et al., 2006). The interaction between E2 and E7 results in a reduction in the ability of E7 to induce centrosome duplication abnormalities, thus reducing the chances of aberrant mitosis (Gammoh et al., 2009).

1.10.3 HPV E6 protein

The HPV E6 protein is approximately 150 amino acids long and has a high content of α-helical and β-sheet secondary structure; the protein has two CXXC zinc fingers and a PDZ binding domain but lacks any enzymatic activity (Bodily and Laimins, 2011, Nomine et al., 2006, Tungteakkhun and Duerksen-Hughes, 2008). The C-terminal region of HPV16 contains three NLSs, which lead to the E6 protein having a nuclear localisation, whereas the E6 proteins of low risk HPV types have a predominantly cytoplasmic localisation (Tao et al., 2003). The E6 gene is one of the earliest genes expressed following viral infection and is one of two principal oncogenes, along with E7. E6 expression can lead to hyperproliferation of

cells, loss of epithelial cell differentiation and tumour formation (Tungteakkhun and Duerksen-Hughes, 2008). Expression of high-risk E6 alone is sufficient to immortalise primary human mammary epithelial cells (Band et al., 1991) and together with E7 can transform primary human embryonic fibroblasts (Watanabe et al., 1989) and human foreskin keratinocytes in culture (Hawley-Nelson et al., 1989). E6 has been shown to interact with a large number of cellular proteins, as shown in figure 1.7A, but the most studied interaction of E6 is with the cellular E3 ubiquitin ligase, E6 associated protein (AP). E6 binding to E6AP recruits p53 to a trimeric complex, resulting in p53 ubiquitination and proteasomal degradation (Huibregtse et al., 1991, Scheffner et al., 1993). p53 is not a natural substrate for E6AP and is only targeted in the presence of high-risk E6 (Huibregtse et al., 1991), low-risk E6 proteins are able to bind E6AP but unable to degrade p53 (Bodily and Laimins, 2011). Not only does E6 target p53 for degradation, both high and low-risk E6 proteins can directly interact with p53 to inhibit p53 binding to DNA and cause aberrant p53 localisation (Lechner and Laimins, 1994, L. Sun et al., 2008). E6 also independently inactivates p53 through the binding of CBP/p300 and hADA3, which are histone acetyltransferases, leading to reduced acetylation of p53 (Bodily and Laimins, 2011, M. C. Thomas and Chiang, 2005). The degradation or loss of function of p53 leads to ablation of its ability to negatively regulate cell cycle progression, the G_1/S and G_2/M checkpoints and therefore facilitates the survival of the cells during differentiation (Tungteakkhun and Duerksen-Hughes, 2008). With many direct interactions and degradations via the E6AP (Bak, c-myc, E6TBP1, MCM7, NFX1-91), E6 is able to modulate G-protein signalling (E6TBP1), immune system recognition (IRF-3), chromosomal stability (MCM7) and induce telomerase activity (NFX1-91) (Howie et al., 2009). High-risk E6 upregulates hTERT transcription through a complex mechanism which involves the binding of SP1 and c-myc and the down-regulation of p300 (Pim and Banks, 2010).

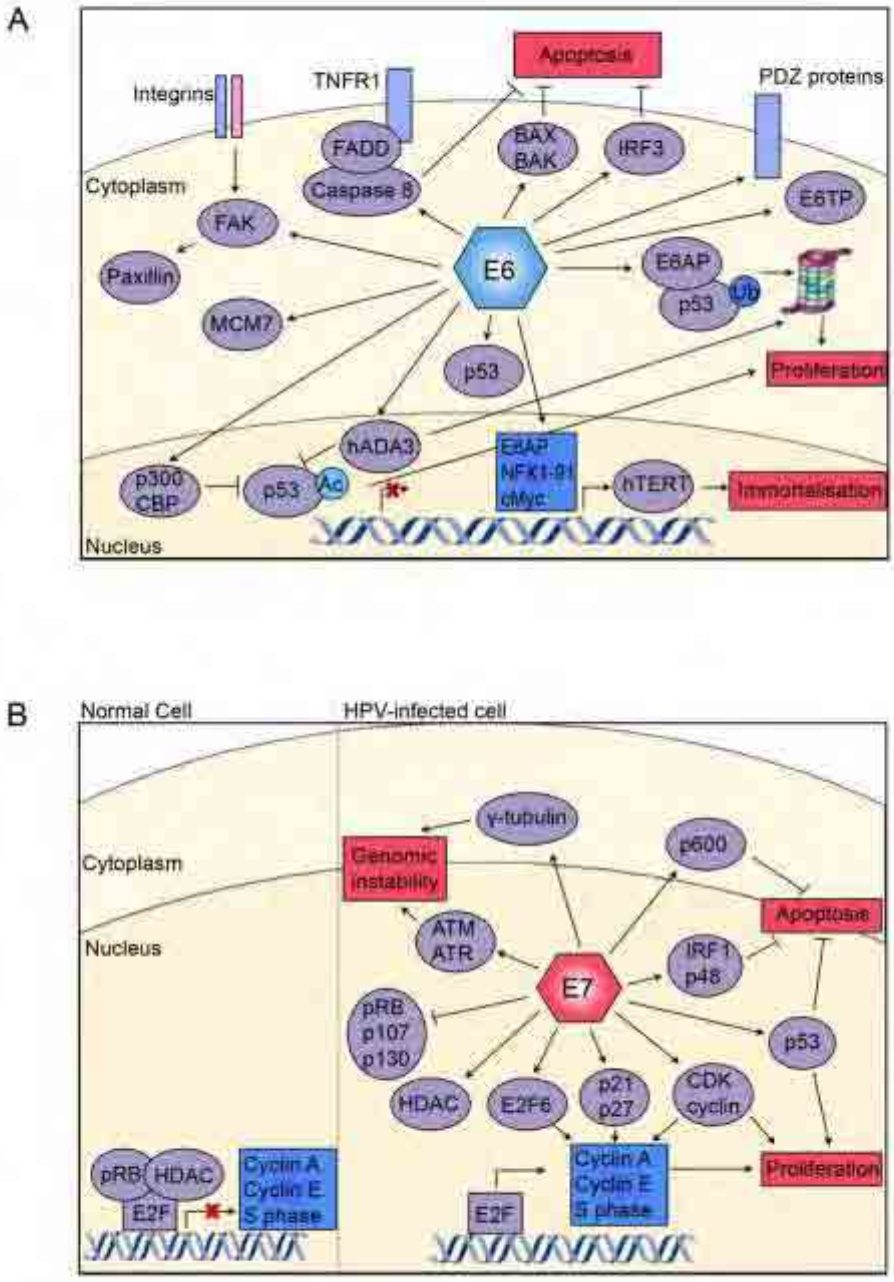

Figure 1.7 Cellular targets of E6 and E7 proteins. (**A**) Cellular proteins and pathways targeted by the E6 protein. High-risk E6 proteins inhibit p53 dependent growth arrest and apoptosis through several mechanisms. E6 is able to modulate G-protein signalling (E6TP), immune system recognition (IRF3), chromosomal stability (MCM7) and induce telomerase activity (NFX1-91). (**B**) Cellular pathways and proteins targeted by the E7 protein. High-risk E7 proteins subvert G_1-S arrest and induce hyperproliferation by interacting with RB family members and causing the constitutive activation of E2F-responsive genes. E7 also interacts with multiple proteins to cause genomic instability and inhibit apoptosis. Adapted from (Moody and Laimins, 2010)

Although the exact mechanism by which E6 activates hTERT is unknown, the prevention of telomere shortening is an important function of high-risk E6 in the immortalisation of cells (Pim and Banks, 2010).

At the C-terminus of high-risk E6 proteins there is a S/TXV motif, a PSD-95/Dlg/ZO-1 (PDZ) binding motif, which allows E6 to bind to PDZ proteins, these include hDlg, hScrib, MAGI-1, -2 and -3, MUPP1 and PTPN3 (Howie et al., 2009). Some of the PDZ proteins have been shown to be tumour suppressors and generally are involved in epithelial tight junctions, basal cell adhesion, cell:cell junctions and epithelial polarity (Howie et al., 2009). As negative regulators of cell proliferation the PDZ proteins are targeted for degradation by the E6 protein, although different HPV types have differing specificity for individual PDZ proteins (M. Thomas et al., 2001).

Together the functions of E6 allow the cell to bypass DNA damage checkpoints, resulting in cell proliferation in the presence of DNA damage, this enables mutations to be passed on to progeny cells and allows possible cancerous mutations to form.

1.10.4 HPV E7

The HPV E7 oncoprotein is approximately 100 amino acids in size and shares some sequence homology with adenovirus E1A and SV40 large T antigen (Phelps et al., 1988, Vousden and Jat, 1989). The C-terminus contains a CXXC zinc-binding domain, similar to that found in E6 (Barbosa et al., 1989) and like E6, E7 does not possess any intrinsic enzymatic activity and functions by binding to several cellular factors, as shown in figure 1.7B (Moody and Laimins, 2010). The N-terminal region of HPV16 E7 contains a NLS, whereas the C-terminal region codes for a second NLS and a NES, the subcellular localisation of E7 varies during the cell cycle (Knapp et al., 2009). Expression of high-risk E7 proteins alone can immortalise human keratinocytes at a low frequency (Moody and Laimins, 2010), however

E7 is unable to immortalise human mammary epithelial cells, highlighting that cell-specific factors may play a role in viral oncogene functions (Band et al., 1991). When expressed together with E6, E7 can efficiently immortalise most primary cells (Hawley-Nelson et al., 1989). The primary interaction of E7 is with members of the retinoblastoma (Rb) family, including p105 (pRB), p107 and p130; binding of E7 to these proteins targets them for degradation by a cullin 2 ubiquitin ligase complex (Dyson et al., 1989, Huh et al., 2007). The Rb family members control the G_1-S phase transition by regulating the activity of E2F transcription factors; pRB forms a complex with E2F, which recruits histone deacetylases (HDACs) and represses transcription from E2F-dependent promoters. The binding of E7 to pRB disrupts these complexes and E2F is released allowing the constitutive expression of E2F-responsive genes, such as cyclin A and cyclin E, promoting premature S phase entry and DNA synthesis, thus uncoupling the processes of differentiation and proliferation (Chellappan et al., 1992, Cheng et al., 1995, Moody and Laimins, 2010). E7 can also interact directly with the HDACs, independently of Rb, this interaction is required for stable maintenance of episomes (Longworth and Laimins, 2004a). E7 interacts with Rb family members via a conserved LXCXE motif in the E7 N-terminal region (Munger et al., 1989). The disruption of the pRB/E2F complex requires additional sequences found in the C-terminal region of E7 (E. W. Wu et al., 1993), which is highlighted in the case of HPV1 E7, which can bind to pRB as efficiently as HPV16 E7, but can not activate E2F-responsive promoters and transform primary cells (Schmitt et al., 1994). The binding of E7 to Rb family members targets them for proteasomal degradation through an ubiquitin dependent pathway, which abrogates other Rb activities such as DNA repair and maintenance of genomic integrity (Boyer et al., 1996). The HPV16 E2 protein can directly bind to E7 and inhibit the E7 induced degradation of pRB, demonstrating a viral mechanism by which the transforming activity of E7 can be modulated (Gammoh et al., 2009). The abrogation of Rb functions by E7 sensitises the cells

to p53-dependent apoptosis and stabilises p53, which is why the E6 degradation of p53 is important in the transformation of cells.

High-risk E7 proteins also bind to p21 and p27, which are CDK inhibitors, the binding of E7 to the inhibitors activates CDK2 and promotes G_1 to S phase entry and progression through cyclin A and cyclin E respectively (Moody and Laimins, 2010). Genomic instability is necessary for malignant progression to occur and high-risk E7 can induce this by rapidly inducing centrosome amplification, which occurs due to high levels of CDK2 activity and E7 binding directly to γ-tubulin, which results in the removal of γ-tubulin from the mitotic spindle (Moody and Laimins, 2010, Nguyen et al., 2007). The E2 protein of HPV16 can inhibit the centrosome amplification caused by HPV16 E7, by directly binding E7 and recruiting it to the mitotic chromosomes (Gammoh et al., 2009). High-risk E7 activates the ATM DNA damage pathway, this pathway is essential for differentiation-dependent viral genome amplification but not for stable episomal maintenance, however the DNA damage response could lead to chromosomal alterations and the induction of genomic instability (Moody and Laimins, 2009). HPV E7 also binds to a number of factors which help regulate the subversion of the interferon antiviral response (IRF1 and p48) and the inhibition of anoikis (p600) (Moody and Laimins, 2010).

Together the functions of E7 allow the deregulation of the link between cellular differentiation and proliferation in epithelia, allowing the virus to replicate in differentiating cells which would have normally withdrawn from the cell cycle.

1.10.5 HPV E5

Although E6 and E7 provide the primary transforming activities of high-risk HPV, the small oncoprotein E5 can augment their function and contribute to tumour progression. The HPV E5 proteins are small membrane bound, hydrophobic proteins, which are around 83-93 amino

acids long and localise to the endoplasmic reticulum, the Golgi apparatus and perinuclear membrane (Conrad et al., 1993). The bovine papillomavirus (BPV) E5 protein is smaller than HPV E5 and is the primary transforming protein, BPV E5 acts by associating with the platelet-derived growth factor (PDGF) receptor and constitutively activating it (Petti et al., 1991). The HPV E5 proteins have little homology with BPV E5 and do not associate with the PDGF receptor, however HPV16 E5 has been shown to transform rodent fibroblasts and human keratinocytes (Straight et al., 1993). The proposed mechanism by which HPV E5 transforms cells is via the epidermal growth factor receptor (EGFR) signalling pathway (Kim et al., 2010, Straight et al., 1993). The over expression of EGFR signalling is linked to the majority of cancers and activation of EGFR regulates gene transcription and modulates cell proliferation, apoptosis, angiogenesis, tumour invasion and metastasis (Kim et al., 2010). E5 acts to increase the number of EGF receptors, by interacting with and inhibiting vacuolar ATPase, interfering with the degradation of EGFR (Conrad et al., 1993), HPV16 E5 also increases EGFR phosphorylation in the presence of EGF (Straight et al., 1993). The E5 protein of HPV16 can also activate MAP kinase p38 and ERK1/2 in an EGF independent manner (Crusius et al., 1997) and down-regulate the tumour suppressors p21 and p27 (Pedroza-Saavedra et al., 2010, Tsao et al., 1996). E5 has also been shown to down-regulate the expression of surface HLA class I, reducing the recognition of infected cells by CD8 T cells (Campo et al., 2010). Although E5 has transforming properties, the protein is likely to be expressed primarily during the late phase of the life cycle in differentiated cells (Longworth and Laimins, 2004b). Mutation of HPV16 E5 in differentiated cells did not prevent viral DNA amplification; however the amount of suprabasal cells undergoing DNA synthesis was reduced compared to wild type (Genther et al., 2003).

These interactions suggest that the role played by E5 in the productive phase of the life cycle is subtle, but E5 helps modify the differentiation-dependent cell cycle exit of the cells to

support proliferative competence (Fehrmann et al., 2003). HPV E5 contributes to malignant transformation through the augmentation and supplementation of the roles of E6 and E7, also the upregulation of EGFR signalling pathway allows E5 to induce the inflammatory cell signalling, COX-2-PGE_2 pathway, which is involved in various oncogenic processes such as angiogenesis and anti-apoptosis (Kim et al., 2010). However, due to integration of the HPV genome during malignant progression, the expression of E5 is lost, whereas E6 and E7 are continually expressed, suggesting that the role played by E5 as an oncoprotein is only small compared to E6 and E7.

1.10.6 Capsid proteins

The major capsid protein, L1, is approximately 500 to 530 amino acids long (57 kDa) and is the primary structural element in the infectious virion. The virions contain 360 copies of the L1 protein organised into 72 pentavalent capsomeres, which make up an icosahedral shell using the C-terminal arms of the L1 proteins to link them together (Modis et al., 2002). The minor capsid protein, L2, is around 470 to 520 amino acids long (78 kDa) and unlike L1, the L2 protein is not highly conserved among HPV types, except for 50 amino acids in the N-terminal region (Doorbar and Gallimore, 1987). Within a capsid, the precise number of L2 proteins is unknown and the general consensus was 12 L2 molecules per capsid; however this can vary up to 72, with one L2 protein at the centre of every L1 pentavalent capsomere (Buck et al., 2008). A hydrophobic region in the C-terminal region of L2 is responsible for the binding of L2 to the L1 pentamers and is thought to insert into the central hole in the capsomere (Finnen et al., 2003). The L2 protein is expressed before the L1 protein upon differentiation of the epithelia (Florin et al., 2002). The L1 and L2 ORFs are transcribed in the undifferentiated epithelia but are not translated; the reason for this is that these ORFs contain codons which are not frequently found in mammalian cells. As the cells differentiate

the tRNAs for the rare codons may become more available, allowing translation of the late proteins (Zhou et al., 1999).

The L2 protein contains nuclear localisation sequences in its N- and C-terminal regions and is transported to the nuclear domain 10 (ND10 or PML) bodies once in the nucleus (Florin et al., 2002). The L2 protein then displaces the ND10 component SP100 and the viral DNA is recruited to the ND10 bodies by E2 (Florin et al., 2002). The L1 protein assembles into capsomeres in the cytoplasm, prior to nuclear relocation and then is recruited to ND10 bodies (Florin et al., 2002). The capsid particles can assemble in the absence of L2; however L2 contributes to efficient packaging (Stauffer et al., 1998) and enhances virus infectivity (Roden et al., 2001). The L2 protein of HPV16 can be stabilised by sumoylation, however sumoylation of L2 inhibits the interaction with L1, demonstrating a control mechanism, which modulates the formation of virions in cells (Marusic et al., 2010). The virion matures and stabilises as the infected cells approach the cell surface as a result of disulphide cross-linking (Modis et al., 2002). The C-terminal 'arm' which links the pentamers together is exposed on the surface of the virion and is a likely target of the immune system (Modis et al., 2002). The neutralising antibodies produced against L1 are HPV type specific, whereas those produced against L2 are broadly cross-neutralising. The production of a VLP with the N-terminal regions of eight different L2s displayed on the surface, delivers cross-protection against multiple HPV types from multiple genera, therefore it is a viable prophylactic alternative to the vaccines currently in production (Tumban et al., 2011).

1.11 HPV E1^E4

The E4 protein is the most highly expressed protein in HPV infections, representing up to 30% of the total protein in HPV1 warts (Doorbar et al., 1986), the expression is mainly cytoplasmic although nuclear forms are detected. However despite numerous studies of its

actions in overexpression studies the precise role of this protein in the HPV life-cycle is an enigma.

The E4 ORF lies within the early region of the genome and is present in early transcripts as a spliced E1^E4 product. HPV E4 is expressed as the fusion protein E1^E4 and is encoded for from spliced E1 and E4 ORFs, with the first five amino acids of E1 fused to E4 (Chow et al., 1987, Palermo-Dilts et al., 1990). The E4 ORF lies entirely within the E2 ORF and overlaps with the highly variable E2-hinge region, resulting in poor sequence homology in E4 between HPV types. Although the E4 ORF is in the early region, E4 is the third ORF on the polycistronic mRNAs (Figures 1.4 and 1.5). The precise mechanism of translation of HPV transcripts is unknown, however it is likely that little E1^E4 protein is expressed from these transcripts in the early stages of the virus life cycle (Longworth and Laimins, 2004b). Upon epithelial cell differentiation and activation of the late viral promoter, E4 is the first ORF on the late transcripts, leading to high expression of E1^E4 (Longworth and Laimins, 2004b). This protein only becomes detectable once cells have entered the productive phase, coinciding with the onset of viral genome amplification and persists in cells as they migrate to the surface of the wart or papilloma (Figure 1.8A); in cutaneous lesions caused by HPV1 this occurs as soon as cells have moved up from the basal cell layer (Breitburd et al., 1987) whilst in anogenital infection of HPV16 and HPV18, this phase is restricted to the upper spinous layers (Doorbar et al., 1997, Peh et al., 2002).

The size of the E1^E4 protein ranges from 73 (HPV34) to 304 (HPV47) amino acids long (the E1^E4 protein from HPV1 is 125 amino acids long and 17 kDa). E1^E4 contains several conserved regions including a leucine rich motif, LLXLL, a proline rich region in the N-terminus and a C-terminal domain which contains conserved hydrophobic residues (Figure 1.8B), these domains have been shown to be important in the functions of E4 (Ashmole et al.,

1998, C. E. Davy et al., 2002, Roberts et al., 1994, Roberts et al., 1993). Overexpression studies of E4 have shown E4 to affect diverse biological functions in cells including the suppression of chromosomal DNA synthesis (Roberts et al., 2008), inhibition of progression of cells through the mitotic cycle (C. E. Davy et al., 2002, Knight et al., 2004) and promotion of apoptosis (Raj et al., 2004). It also has a destabilising effect on structures associated with cellular integrity such as the keratin cytoskeleton (Doorbar et al., 1991, Roberts et al., 1993) and the corneocyte (Bryan and Brown, 2000) as well as sub-nuclear ND10 bodies (Roberts et al., 2003). The interactions of E1^E4 and the timing of E1^E4 expression suggest roles in the facilitation and support of viral genome amplification, regulation of late gene expression, control of virus maturation and mediation of virus release (Roberts, 2006).

Indeed studies knocking out expression of E1^E4 in which E2 expression is preserved in viral genome containing cells demonstrated that E4 is required for viral DNA amplification and late gene expression of HPV18 and HVP31 (Wilson et al., 2005, Wilson et al., 2007) and the cottontail rabbit papillomavirus (Peh et al., 2004), but not for the maintenance of viral episomes (Wilson et al., 2005, Wilson et al., 2007). Deletional analysis across the E4 ORF of HPV16 genomes indicated that only the C-terminus of E4 is required for activation of the late functions (Nakahara et al., 2005). Somewhat surprising was the finding that E4 is not required for viral DNA amplification and late gene expression of HPV11 genomes, which may reflect differences in experimental systems or that there are type specific functions of E4 (Fang et al., 2006).

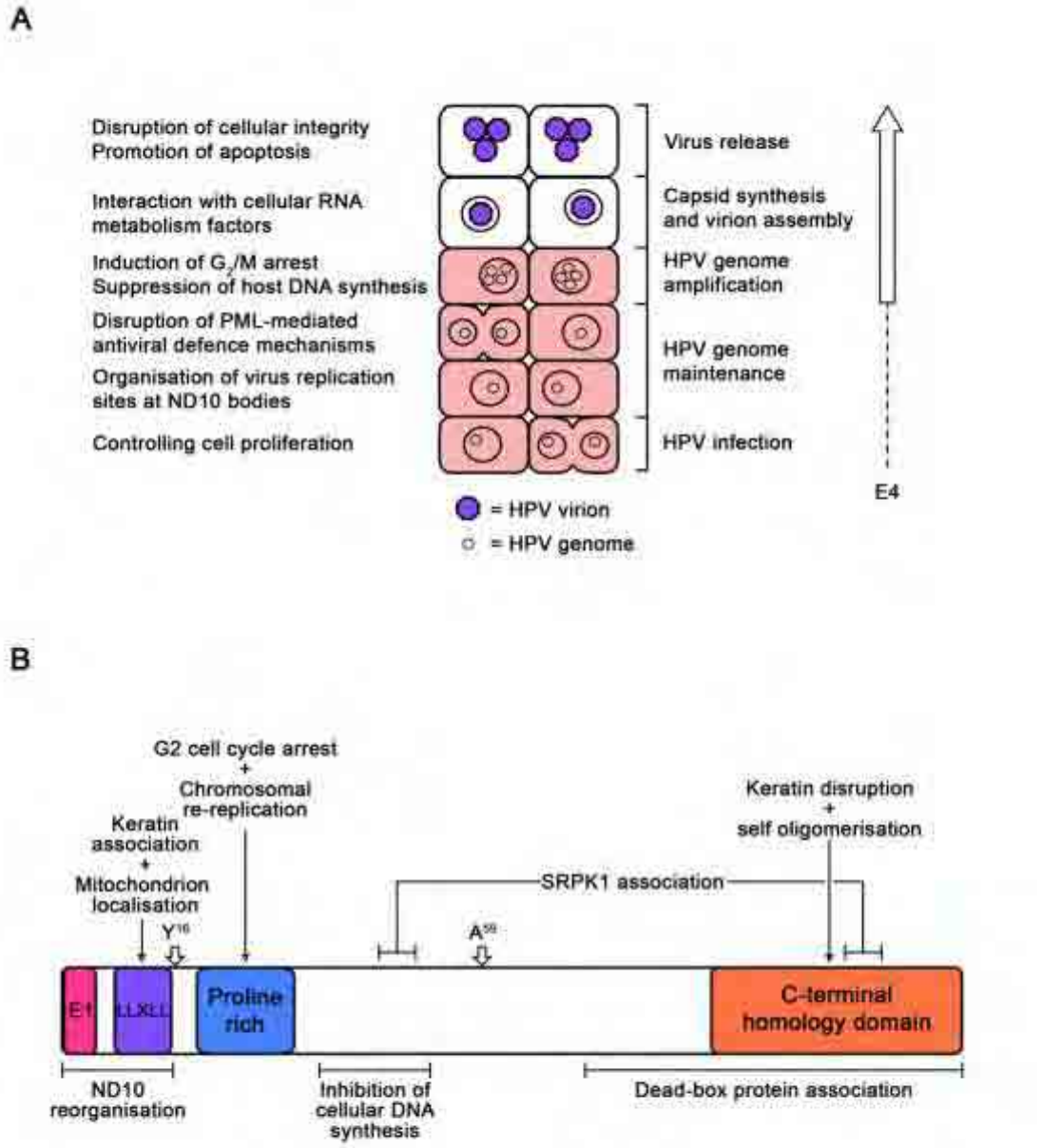

Figure 1.8 The HPV E1^E4 protein. (**A**) Schematic showing the position of E4 expression and E4 functions in the life cycle of alpha HPVs. (**B**) Schematic showing the regions involved in known biological functions and interactions with cellular proteins. Arrows indicate positions of proteolytic cleavage sites in HPV1. Adapted from figure provided by Sally Roberts and (Roberts, 2006)

1.11.1 Post-translational modifications of E4

The E4 proteins exist in natural infections as multiple species, arising from a combination of proteolysis, phosphorylation and oligomerisation and these post-translational modifications play a significant role in regulating E4 biological functions. The most studied E4 protein with regards to posttranslational modifications is the E4 protein of HPV1, the full-length E1^E4 protein of 17 kDa undergoes progressive N-terminal proteolytic cleavage to produce 16, 11 and 10 kDa forms (Doorbar et al., 1988). The first 15 N-terminal residues are cleaved off to produce the 16 kDa species, followed by a further 43 residues to produce the 11 and 10 kDa peptides (Figure 1.8B) (Roberts et al., 1994). In HPV1 warts the full-length 17 kDa species is expressed in the parabasal cell layers, with the other species accumulating in a progressive manner so that in the most superficial cells the major E4 species present are the most processed forms of 11/10 kDa (Breitburd et al., 1987). The E1^E4 proteins of other HPV types are subject to cleavage as well with calpain being the protease responsible for the cleavage of HPV16 and HPV18 E1^E4 in differentiating keratinocytes (Khan et al., 2011). The proteolysed peptides of E4 do not contain sequences necessary for some of the interactions the full-length protein is capable of, suggesting that some of the E4 functions are not required throughout the productive phase of the HPV life-cycle (Roberts et al., 1994, Roberts et al., 2003). Additionally the truncated E4 species have biological activities, not possessed by the full-length E1^E4 protein, indicating that E4 function is regulated by post-translational modifications (Knight et al., 2004).

The E1^E4 protein is subject to phosphorylation at multiple serine and threonine residues by several cellular kinases including cAMP dependent protein kinase (PKA), serine-arginine protein kinase 1 (SRPK1), cyclin dependent kinases (CDKs), extracellular signal-regulated kinase (ERK) and mitogen-activated protein kinase (MAPK) (Bell et al., 2007, Bryan et al., 2000, Grand et al., 1989, Knight et al., 2011, Q. Wang et al., 2009b). Interestingly, *in vivo* the

full length 17 kDa E4 protein of HPV1 is phosphorylated, but the 16 kDa proteolysed species is not, even though both proteins are substrates for PKA *in vitro* (Grand et al., 1989). This data suggests that the functions of E1^E4 may be modulated by phosphorylation in the parabasal layers of the epithelium but not in the more differentiated layers where cell death and virion assembly are occurring (Grand et al., 1989). Phosphorylation of the HPV16 E1^E4 protein by ERK at the onset of viral amplification causes the stabilisation of E1^E4 and enhanced binding of E1^E4 to the keratin network (Q. Wang et al., 2009b).

The E1^E4 protein has been shown to self-multimerise *in vivo* and exists as multiple stable oligomers (Doorbar et al., 1988). Hydrophobic residues in the C-terminus are important for this self-oligomerisation (Ashmole et al., 1998, Roberts et al., 1997). The deletion of the N-terminal region of E1^E4 leads the assembly of E4 into amyloid like fibrils, which facilitates the accumulation of E4 in the upper layers of the epithelia (McIntosh et al., 2008). The multimerisation of E4 proteins is important for inducing the collapse of the keratin network (Q. Wang et al., 2004).

1.11.2 Functions of E1^E4

Much of our understanding of the E4 functions and actions have been derived from overexpression studied in established epithelial cell lines and these are discussed in more detail below and summarised in Figure 1.8A.

1.11.2.1 Interaction with structural proteins

The intermediate filaments (IFs) in epithelial cells are composed of keratin and as the cells differentiate, the keratin network undergoes keratinisation, which is where the keratin filaments aggregate to form a tough impenetrable barrier to the cell. The E1^E4 protein from cutaneous and mucosal HPV types can interact with the keratin network, but only the E1^E4 proteins from some HPV types can cause the collapse of the keratin network into a

perinuclear fibrous clump (Doorbar et al., 1991, Roberts, 2006, Roberts et al., 1993). It has been proposed that the disruption of the keratin network by HPV16 E1^E4 may facilitate the release of infectious viral particles later in the virus life cycle (Doorbar et al., 1991). The highly conserved N-terminal leucine-rich motif (LLXLL) has been shown to be involved in the association between E1^E4 and the keratin filaments (Roberts et al., 1994), but in HPV16 a hydrophobic C-terminal domain, which is conserved amongst mucosal HPV types, is required for the induction of cytokeratin collapse (Roberts et al., 1997). The C-terminal region in HPV16 required for cytokeratin collapse is the same as the region required for oligomerisation, suggesting that the self-association of E1^E4 proteins may allow cross-linking of the keratin filaments and the subsequent collapse of the network (Roberts et al., 1997). The cross-linking of the keratin filaments prevents the keratins from freely moving between the soluble and insoluble compartments, thus disturbing the normal keratin dynamic reorganisation during the cellular life cycle (Q. Wang et al., 2004). Additionally the interaction between HPV16 E1^E4 and the keratin network, results in the keratin filaments being hyperphosphorylated and ubiquitinated, which targets them for proteosomal degradation (McIntosh et al., 2010). Interestingly the LLXLL motif required for association with the keratin filaments is lost after proteolysis of the E1^E4 protein (Roberts et al., 1994), and this suggests that interactions with the keratin cytoskeleton are modified during the virus life cycle. The mechanism of E4 induced keratin collapse is still controversial, as although E4 and keratin have been shown to colocalise in natural HPV16 infections, the number of cells with this colocalisation is very small (Q. Wang et al., 2004).

The cornified cell envelope (CCE) is an insoluble matrix which forms in terminally differentiated keratinocytes to provide a barrier against external factors. HPV11 E1^E4 has been shown to bind to the CCE and induce the degradation of two components of the CCE, loricrin and cytokeratin 10 (K10). The binding of E1^E4 to these CCE components results in

a CCE which has a thinner wall and increased fragility, which may facilitate virion release (Bryan and Brown, 2000).

1.11.2.2 Regulation of the cell cycle and cellular DNA synthesis

The expression of E4 proteins from several different HPV types (1, 11, 16 and 18) has been shown to modulate the progression of the cell cycle and induce a G_2 arrest (C. E. Davy et al., 2002, Knight et al., 2004, Nakahara et al., 2002). Studies of the different E4 proteins has however shown that these proteins induce G_2/M arrest using different mechanisms and in the case of HPV1, E4 mediates two distinct mechanisms of cell cycle arrest that are dependent on the processing of the viral protein.

The progression of the cell cycle past the G_2/M checkpoint is driven by the cyclin B1-CDK1 (cyclin-dependent kinase) complex, which is a core component of the mitosis-promoting factor. The E1^E4 protein of HPV16 induces the cytoplasmic retention of active cyclin B1-CDK1 complexes, preventing them from translocating to the nucleus and allowing the cells to proceed through G_2 and into mitosis (C. E. Davy et al., 2002). In contrast HPV1 E4 protein inactivates the complex by promoting the inhibitory phosphorylation of CDK1 or by reducing the levels of cyclin B1 (Knight et al., 2006).

Intriguingly unlike HPV16 or HPV18, the full length E1^E4 protein of HPV1 is not able to cause G_2 arrest; inhibition of cell division is mediated by the truncated 16 kDa form expressed either alone or together with the full-length polypeptide (Knight et al., 2004). The expression of the truncated 16 kDa form of E4 alone leads to a reduction in the levels of cyclin B1, whereas coexpression of the full length E1^E4 with the 16 kDa form leads to inactivated CDK1 due to phosphorylation by the kinase Wee1 (Knight et al., 2006).

Expression of full-length HPV18 E1^E4 and the 16 kDa form of HPV1 E4 was also associated with chromosomal rereplication suggesting that E4 promotes G_2 arrest and an

environment that is conducive to viral genome amplification through the access to the host replication machinery. However when the full length HPV1 E1^E4 and the 16 kDa E4 species are coexpressed, chromosomal rereplication was suppressed as was entry into S phase (Knight et al., 2004). Further investigations revealed that cellular DNA replication was suppressed by this E4 complex by E4 blocking replication initiation through the prevention of the loading of replication licensing factors, minichromosome maintenance proteins (MCMs) onto chromatin (Roberts et al., 2008). This E4 function may enhance viral genome amplification in replication activated cells by reducing competition of cellular replication factors by suppressing host cell DNA synthesis. Studies in organotypic raft cultures of HPV16 and HPV18 genome containing cells have shown that viral genome amplification is largely confined to cells which show little evidence of cellular DNA synthesis, demonstrating that viral and cellular DNA are not replicated concurrently (Nakahara et al., 2005, H. K. Wang et al., 2009a). Also, other unrelated DNA viruses such as EBV and CMV also have strategies to selectively inhibit host cell DNA replication (Roberts et al., 2008).

An important question is therefore what purpose this conserved function of cell cycle arrest serves within the HPV life cycle. A recent study from Roberts and co-workers has investigated this in the context of the HPV18 virus. Mutations within a CDK-bipartite consensus recognition motif that disrupt the G_2/M arrest function of the type 18 protein were inserted in to the E4 ORF of HPV18 genomes. Primary foreskin keratinocytes were transfected with the genomes and stable cell lines established. Following differentiation of the cells, viral genomes were efficiently amplified in the cells carrying mutant genomes indicating that this E4 function is not necessary for this vegetative phase of replication, even though loss of full-length E1^E4 expression does correlate with inefficient DNA amplification (Knight et al., 2011, Wilson et al., 2007). The G_2 arrest function might also be necessary to favour capsid expression (J. Ding et al., 2010), but there was no evidence from

organotypic raft cultures of a defect in the production of the major capsid L1 in the absence of the E4 G_2/M arrest function (Knight et al., 2011).

The study does not rule out a role for E4 G_2/M arrest in viral genome replication. The HPV18 E7 protein is able to maintain cells in an extended G_2 phase after the induction of S phase re-entry through the activation of CDC2 and cyclin B1 (Banerjee et al., 2011). This study shows the induction of cyclin B1 activity occurs before the expression of E1^E4 is upregulated, demonstrating that E1^E4 may be assisting this E7 function in the upper layers of the epithelia, to maintain the G_2 arrest for the maximum length of time. The combined activities of E7 and E1^E4 could stimulate the cells to be maintained in a pseudo S phase state, where replication enzymes are present, but the viral DNA is not competing with cellular DNA for these enzymes (Knight et al., 2004, Knight et al., 2011).

1.11.2.3 Disruption of ND10 bodies

The HPV1 E1^E4 protein forms cytoplasmic and nuclear inclusions in productively infected keratinocytes; these electron dense structures often contain relocalised cellular proteins. One protein which HPV1 E1^E4 induces the relocalisation of is promyelocytic leukaemia protein (PML) (Roberts et al., 2003). PML is a component of ND10 bodies, which are discrete nuclear foci, which contain multiple cellular proteins. The exact function of ND10 bodies is unknown, but they have been linked with ontogenesis, DNA damage repair, apoptosis, stress response, senescence, the ubiquitin pathway and regulation of gene expression (Tavalai and Stamminger, 2008). ND10 bodies have also been shown to constitute an interferon-mediated defence mechanism (Tavalai and Stamminger, 2008). Multiple viruses are known to target ND10 bodies, however the reasons for this are unclear; the targeting of these structures may be a viral strategy to evade a cellular regulatory process which would inhibit viral replication. Alternatively disruption of ND10 bodies may facilitate the formation of viral replication centres by sequestration of the cellular proteins required for replication (Tavalai and

Stamminger, 2008). The redistribution of PML to E1^E4 nuclear inclusions occurs in naturally infected cells which have switched to vegetative genome replication, suggesting that reorganisation of the ND10 bodies is only necessary when the virus changes its replication strategy (Roberts et al., 2003), this is similar to what occurs with viruses which have lytic replication cycles (Tavalai and Stamminger, 2008).

1.11.2.4 Induction of apoptosis

The E1^E4 protein of HPV16 has been shown to induce the redistribution of mitochondria from the microtubule network to perinuclear E4 aggregates (Raj et al., 2004). This redistribution of mitochondria is specific as the localisation of other organelles such as the Golgi body and lysosomes are unaffected. The region of E1^E4 required for the redistribution of mitochondria is the leucine cluster in the N-terminus of the protein, this is the same region required for the association of E1^E4 to the keratin network (Raj et al., 2004). The detachment of the mitochondria from the microtubules causes a severe reduction in the mitochondrial membrane potential and an induction of apoptosis. Together with the disturbance of the cytokeratin network, the relocalisation of mitochondria and subsequent apoptosis may aid with the exit of the infectious virus particles (Raj et al., 2004).

1.11.2.5 Interaction with components of RNA metabolism

The HPV16 E1^E4 protein has also been shown to associate with a novel member of the DEAD box protein family of RNA helicases, E4-DBP (Doorbar et al., 2000). This interaction is not conserved with HPV1 or HPV6 E1^E4 and occurs via sequences in the C-terminus of the protein (Doorbar et al., 2000). The interaction between E1^E4 and E4-DBP inhibits both the ATP binding and the RNA-independent ATPase activity of E4-DBP *in vitro*. E4-DBP has homology with bacterial and yeast DEAD box proteins involved in the

regulation of mRNA stability and degradation suggesting that E1^E4 has roles in regulating the stability of late viral transcripts (Doorbar et al., 2000).

The E1^E4 proteins from HPV types 1, 16 and 18 have been shown to directly interact with the mammalian kinase, serine-arginine (SR) specific protein kinase 1 (SRPK1). The HPV1 E1^E4 protein but not that of HPV16 or HPV18 is a substrate for this kinase *in vitro* (Bell et al., 2007). The study also found that SRPK1 is sequestered to E4 inclusion bodies in terminally differentiated cells in HPV1 warts and in cells expressing HPV1 E1^E4 alone (Bell et al., 2007). The function of this novel interaction in the HPV life cycle is unknown, but may be involved in the facilitation of late viral gene expression (Bell et al., 2007). The association between HPV E1^E4 and SRPK1 forms the subject of this thesis and is discussed in more detail later.

In conclusion, E4 has multiple effects upon cell behaviour, but little is understood about the physiological significance of these functions. It is likely that the protein has multiple actions within the life cycle and these are regulated by protein modifications. Therefore identifying the functions of E4 and understanding their significance is important since the protein may be a valuable target for the development of anti-viral therapies.

1.12 Serine-arginine specific protein kinases

Serine-arginine (SR) specific protein kinases (SRPKs) are a family of serine-threonine kinases which specifically phosphorylate serine residues in serine-arginine/arginine-serine (SR/RS) dipeptide motifs. SRPK1 was the first of the SRPK family to be indentified and was isolated following a search for a kinase which phosphorylated serine-arginine (SR) proteins (Gui et al., 1994a, Gui et al., 1994b). SRPK1 was shown to phosphorylate the serine residues in the arginine-serine (RS) domain of the SR proteins and release the SR proteins from

nuclear speckles (Gui et al., 1994a, Gui et al., 1994b). SRPK1 was found to be structurally similar to a fission yeast kinase, Dsk1, and both kinases have been shown to have a spacer domain separating the catalytic domain into two structural entities (Takeuchi and Yanagida, 1993). All members of the SRPK family contain highly conserved kinase domains separated by spacer sequences and these spacer sequences are unique to the SRPK family member (J. H. Ding et al., 2006). The spacer domains may confer specific functions and regulatory properties to individual SRPKs.

The second human SRPK family member SRPK2, is highly homologous to SRPK1, however SRPK2 has a proline rich tract at the N-terminus and an acidic region in the spacer domain (Kuroyanagi et al., 1998, H. Y. Wang et al., 1998). SRPK1 and SRPK2 are very similar with respect to their enzymatic activity and substrate specificity; however these kinases have levels of tissue specific expression, with SRPK1 being highly expressed in the pancreas and SRPK2 being highly expressed in the brain (H. Y. Wang et al., 1998). Both kinases are moderately expressed in other tissues including heart and skeletal muscle and expressed at low levels in lung, liver and kidney cells (H. Y. Wang et al., 1998). SRPK1 has been shown to associate with the U1-snRNP, which is involved in 5' splice site selection (Kamachi et al., 2002), whereas SRPK2 has been shown to be required for the formation of the U4/U6-U5 tri-snRNP, which is involved in 3' splice site selection (Mathew et al., 2008).

Another human member of the SRPK family is SRPK1a, which is a product of the SRPK1 gene and is produced by alternative splicing causing the insertion of an additional 171 amino acids at the N-terminus (Nikolakaki et al., 2001). The N-terminal insertion in SRPK1a is proline rich, which is reminiscent of SRPK2. Both SRPK1 and SRPK1a have similar substrate specificity and subcellular localisation (Nikolakaki et al., 2001).

SRPK3 has been identified in mouse and pig and is postulated to exist in humans; however no studies have shown SRPK3 expression and function in humans. SRPK3 has been shown to have very limited and tissue specific expression in both mouse and pig (Giannakouros et al., 2011). SRPK family members have also been identified in budding yeast (*Saccharomyces cerevisiae,* Sky1), nematodes (*Caenorhabditis elegans,* SPK-1), insects (*Drosophila melanogaster*, Srpk79D), protozoan parasites (*Trypanosoma cruzi,* TcSRPK) and slime moulds (*Physarum polycephalum,* PSRPK), showing that SRPKs occur in organisms with varying evolutionary scales (Giannakouros et al., 2011). The sequence conservation of the kinase domains is shown in Figure 1.9.

SRPKs have a predominantly cytoplasmic localisation, however they can be translocated into the nucleus following stress, or other signals and during the cell cycle (J. H. Ding et al., 2006, Giannakouros et al., 2011). Multiple elements in the spacer sequence in SRPKs control the cytoplasmic localisation, with deletion of the spacer leading to nuclear translocation (J. H. Ding et al., 2006). As SRPKs are constitutively active, they are likely to be regulated by cellular distribution as an increased amount of SRPK1 in the nucleus following osmotic shock leads to differential phosphorylation of SR proteins and altered splice site selection (Zhong et al., 2009a). In contrast to this, the kinase translocates into the nucleus in response to a cell cycle signal before the initiation of the mitotic phase, indicating that the kinase may play a role in cell cycle progression (J. H. Ding et al., 2006). SRPK1 has been shown to interact with two co-chaperones for the major heat shock proteins, Hsp70 and Hsp90, with the ATPase activity of Hsp90 acting to regulate the cellular distribution of SRPK1 (Zhong et al., 2009a). Stress signals induced by osmotic shock can cause the inhibition of Hsp90 ATPase activity and the subsequent dissociation of SRPK1 from the chaperone complexes and translocation of SRPK1 into the nucleus (Zhong et al., 2009a).

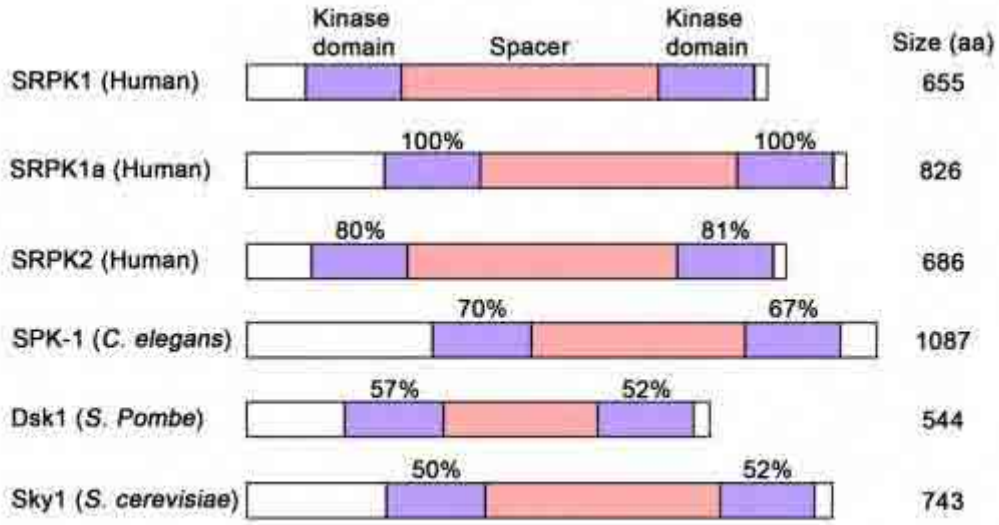

Figure 1.9 Sequence comparison of SRPK1 family members. Schematic showing the conservation of kinase domains between SRPK family members. Percentage shown is comparison to SRPK1. Adapted from (Siebel et al., 1999)

The heat shock proteins are involved in numerous signal transduction pathways suggesting that the regulation of SRPKs is complex and varies in response to a range of developmental cues and cellular signalling (Zhong et al., 2009a).

1.13 Serine arginine-rich proteins

Serine arginine-rich (SR) proteins are a family of RNA-binding proteins which are present in all metazoan organisms and in plants (Manley and Krainer, 2010). SR proteins contain an arginine serine-rich (RS) domain along with one or two RNA recognition motifs (RRM). The RS domain is involved in protein-protein and is extensively phosphorylated by several different kinases. There are currently twelve members of the human SR protein family, which have recently been given a new nomenclature based on the root SRSF (SR splicing factor). According to the new definitions, SR proteins contain one or two conserved RRMs followed by a C-terminal RS domain of at least 50 amino acids with >40% RS content (Manley and Krainer, 2010). The new and old names for the SR proteins along with their domain structures are shown in Figure 1.10A. However, for this thesis the original nomenclature will be used.

The first SR protein to be identified was alternative splicing factor/splicing factor 2 (ASF/SF2), which was discovered and characterised by two groups independently. ASF was identified as a protein which could influence the selection of 5' splice sites in an SV40 early pre-mRNA (Ge and Manley, 1990) whereas SF2 was identified as a protein required for the splicing of a β-globin pre-mRNA (Krainer and Maniatis, 1985), later SF2 was shown to be essential for splicing (Krainer et al., 1990). ASF/SF2 is considered the prototypical SR protein as it contains two RRMs and a classical RS domain, the sequence of the RS domain of ASF/SF2 is shown in Figure 1.10B.

The second human SR protein to be discovered was SC35, which was detected with a monoclonal antibody against purified spliceosomes, and was shown to be necessary for splicing and spliceosome assembly *in vitro* (Fu and Maniatis, 1990). Following the production of the monoclonal antibody, mAb104, the SR proteins SRp20, SRp75, SRp40 and SRp55 were identified (Zahler et al., 1992). The SR protein 9G8 was identified using a monoclonal antibody and unlike the other family members contains a zinc knuckle in-between the RRM and RS domain (Cavaloc et al., 1994). Other SR proteins including SRp30c and SRp54 were identified by oligonucleotide probes against the RNA recognition motif and auto-antiserum reactivity (Screaton et al., 1995, W. J. Zhang and Wu, 1996).

Along with the SR proteins there are a large number of proteins which contain RS or RS-like domains; these include the U1 snRNP-specific protein U1 70K, both subunits of the U2AF splicing factor and the splicing co-activators SRm 160/300. A study of RS domain containing proteins has revealed proteins with functions not only in pre-mRNA splicing but in chromatin remodelling, translation, transcription by RNA polymerase II and cell cycle progression (Boucher et al., 2001).

1.13.1 Localisation of SR proteins

Many proteins involved in pre-mRNA splicing, including SR proteins and small nuclear ribonucleoprotein (snRNPs) are found in nuclear speckles, which are distinct nuclear compartments which occur throughout the nucleus. There are two distinct structural types of speckles at the electron microscopic level, these are interchromatin granule clusters (IGCs) and perichromatin fibrils (Spector, 1993). Based on [^{3}H]uridine incorporation studies, transcription is thought to occur at the perichromatin fibrils which are approximately 3-5 nm in diameter (Spector, 1993, 1996). The IGCs are about 20-25 nm in diameter and are sites of storage and/or reassembly of splicing factors (Spector, 1993, 1996).

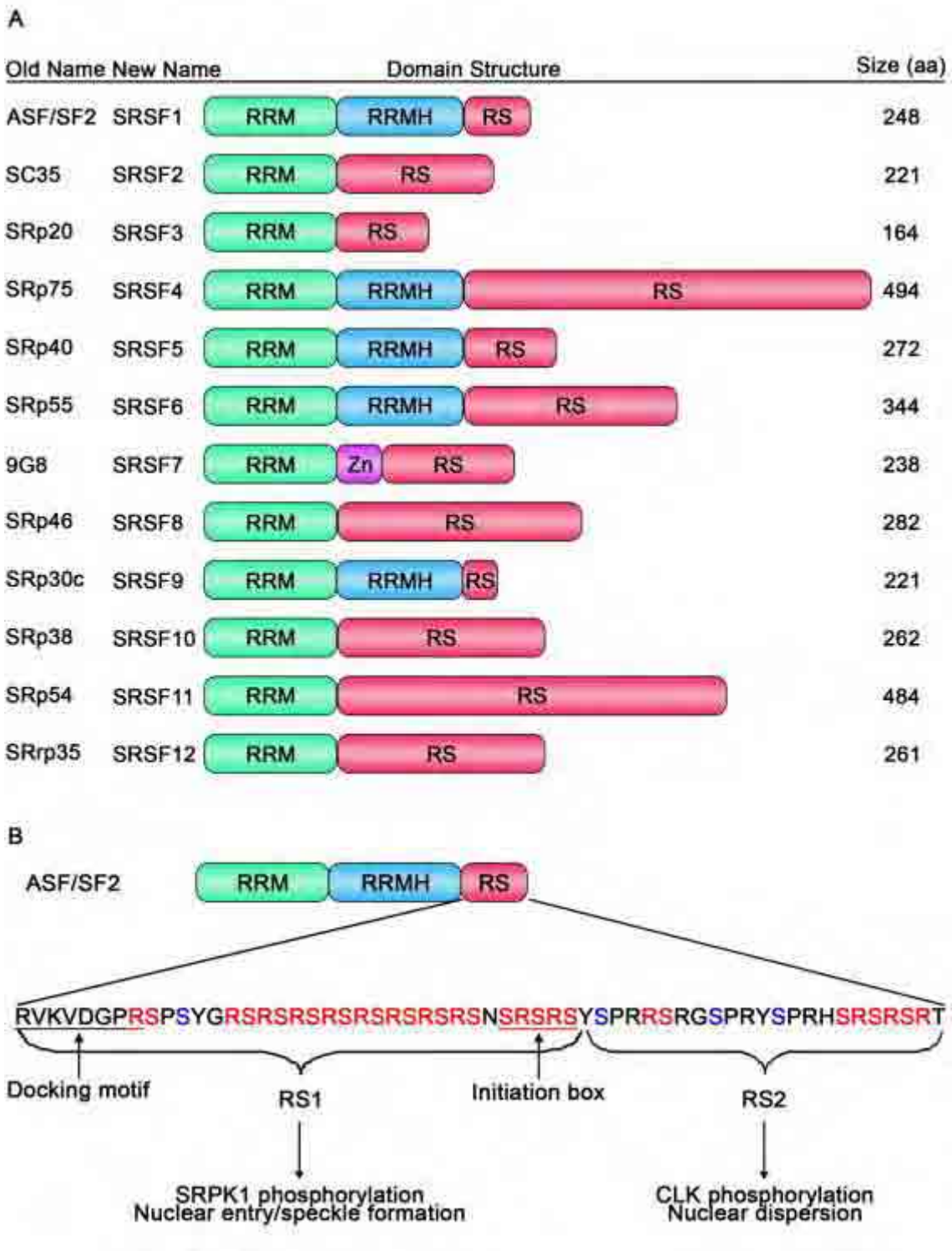

Figure 1.10 SR protein domain structure. (**A**) The old names along with the new nomenclature and the domain structure are shown for the 12 human SR proteins. RRM: RNA recognition motif, RRMH: RRM homology, RS: arginine serine-rich domain, Zn: Zinc knuckle. (**B**) Sequence of the RS domain of ASF/SF2, the RS domain can be divided into two subdomains (RS1 and RS2), RS/SR residues are shown in red, single serines are blue, docking motif and initiation box which are important for SRPK1 phosphorylation are identified. Adapted from (Manley and Krainer, 2010) and (Ghosh and Adams, 2011).

There are visible connections between the speckles and the SR proteins and snRNPs are recruited from the IGCs to the perichromatin fibrils which are the active sites of transcription and splicing. Inhibition of RNA polymerase II or splicing leads to the IGCs becoming larger and more uniform in shape, loss of the perichromatin fibrils and loss of the connections between speckles, confirming that the IGCs are sites of storage of splicing factors (O'Keefe et al., 1994, Spector, 1996). Both phosphorylation of the RS domain and the presence of the RRM are required for the recruitment of the SR proteins from the IGCs to the perichromatin fibrils (Misteli et al., 1998).

The RS domains of the SR proteins are partly responsible for targeting the SR proteins to nuclear speckles. In SR proteins which have a single RRM, such as SC35 and SRp20, the RS domain is both necessary and sufficient to target the SR protein to the speckles (Caceres et al., 1997). In contrast, in SR proteins which contain two RRMs, such as ASF/SF2, the RS domain is neither necessary nor sufficient to target the protein to nuclear speckles, with two of the three constituent domains of ASF/SF2 being required for subnuclear compartmentalisation (Caceres et al., 1997).

A subset of the SR proteins such as ASF/SF2, SRp20 and 9G8 shuttle rapidly and continuously between the nucleus and the cytoplasm, whereas SC35 and SRp40 do not, due to the presence of a dominant nuclear retention signal in the RS domain (Cazalla et al., 2002). The SR proteins SRp75 and SRp55 can shuttle between the nucleus and cytoplasm but with slower kinetics (Sapra et al., 2009). Most shuttling proteins are predominantly nuclear at steady state and present in the cytoplasm only transiently. SR proteins are exported from the nucleus as part of an RNA-protein complex, so require their RRM for cytoplasmic translocation. ASF/SF2 and SRp20 require active transcription for reimport (Caceres et al., 1998).

The reimport of SR proteins into the nucleus is mediated by transportin SR (TRN-SR), a member of the importin-β/transportin family (Kataoka et al., 1999). The interaction of the SR proteins with TRN-SR is via the RS domain and only occurs when the RS domain is phosphorylated (Lai et al., 2000).

1.14 Functions of SR proteins

1.14.1 pre-mRNA splicing

Pre-mRNA splicing is an essential process required for the expression of most eukaryotic protein-encoding genes. The primary transcripts of almost all protein-coding genes contain multiple non-coding intervening sequences (introns) in between the coding sequences (exons), which must be precisely removed to give translatable mRNAs. Alternative splicing allows single genes to encode multiple proteins, via the process of alternative exon inclusion and expands the proteomic complexity encoded by the genome. The process of intron excision is referred to as splicing and takes place in a large ribonucleoprotein complex known as the spliceosome. The spliceosome consists of five small nuclear ribonucleoproteins (snRNPs), designated U1, U2, U4, U5 and U6 along with a large number of non-snRNPs, including the SR proteins (Graveley, 2000). Spliceosome assembly is partly directed by the RNA sequence at the splice sites, with introns being bound by conserved sequences which define their 5' and 3' splice sites. In mammals the 5' splice site consensus sequence is AG/GURAGU (/ denotes the exon/intron boundary). The 3' splice site is defined by three distinct sequence elements, the branchpoint (YNYURAC), which is 20-40 nucleotides upstream of the splice site, a polypyrimidine tract and the 3' splice site (YAG/N) (Graveley, 2000). The spliceosome assembles de novo on each intron, in a sequential and coordinated pathway, although parts of the spliceosome may be pre-assembled (Stevens et al., 2002). The formation of the early (E) complex is initiated by the binding of the U1 snRNP to the 5' splice

site, splicing factor 1 (SF1) to the branchpoint sequence and the heterodimeric splicing factor U2 auxiliary factor (U2AF) to the pyrimidine tract and 3' splice site (Graveley, 2000). U2AF then recruits U2 snRNP in an ATP dependent manner to the branchpoint to form complex A, followed by the recruitment of the U4/U6•U5 tri-snRNP to form complex B. Finally the complex rearranges and U1 and U4 snRNPs are released, leading to the formation of the catalytically active complex C (Graveley, 2000).

1.14.2 SR protein control of pre-mRNA splicing

SR proteins participate at multiple steps in the assembly of the spliceosome; the selection of splice sites relies on the interactions of SR proteins and other RNA binding proteins with *cis*-acting elements within exonic or intronic sequences. The roles of the SR proteins in splicing can be divided into two categories: exon dependent and exon independent. The exon dependent functions of SR proteins are the best-characterised functions and involve the binding of SR proteins to exon sequences (Graveley, 2000). SR proteins can bind to exonic splicing enhancers (ESEs), where they function to enhance the splicing of the adjacent intron by stabilising the binding of U1 snRNP and U2AF to the 5' and 3 splice sites respectively, this process is known as exon definition (Figure 1.11A) (Boukis et al., 2004, Sanford et al., 2005). SR proteins are important in the regulation of alternative splicing and many alternative exons are designated by weak splicing signals. SR protein binding to ESEs can promote alternative 3' splice site selection by recruiting U2AF to suboptimal 3' splice sites with weak polypyrimidine tracts (Figure 1.11B). Exonic splicing silencers (ESSs) can recruit splicing repressor proteins such as hnRNP A1 and block 3' splice site selection, SR proteins can antagonise the effects of splicing repressors, thus promoting splice site selection (Figure 1.911) (Sanford et al., 2005). SR proteins can also promote alternative splice site usage by binding to intronic sequences close to the splice sites and inhibiting recruitment of

spliceosomal components through steric hindrance or non-productive spliceosomal assembly (Figure 1.11D) (Shepard and Hertel, 2009).

The exon independent functions of SR proteins are considered to be exon independent because it is not known whether the SR proteins must bind to exons to perform these functions or not. SR proteins facilitate splice site pairing by simultaneously interacting with U1-70K snRNP and $U2AF^{35}$ across the intron (Figure 1.11E) (J. Y. Wu and Maniatis, 1993), the SR proteins are also required for *trans*-splicing, where the 5' and 3' splice sites are contained on separate RNA molecules (Bruzik and Maniatis, 1995). SR proteins have been shown to play a role in the recruitment of the U4/U6•U5 tri-snRNP into the spliceosome (Figure 1.11E) (Roscigno and Garcia-Blanco, 1995), with the SR proteins interacting with either the U4/U6•U5-27K or the U5-100K (Graveley, 2000).

Multiple SR proteins have been shown to participate in the removal of each intron, suggesting that each splicing event requires at least one SR protein to perform a regulated exon dependent function and another SR protein to perform the exon independent functions. The regulated exon dependent functions of the SR proteins have phosphorylation requirements which are distinct from the other SR protein functions (Graveley, 2000). The splicing of strong introns only requires the exon independent functions of SR proteins and the SR proteins are all able to perform this activity (Hertel and Maniatis, 1999).

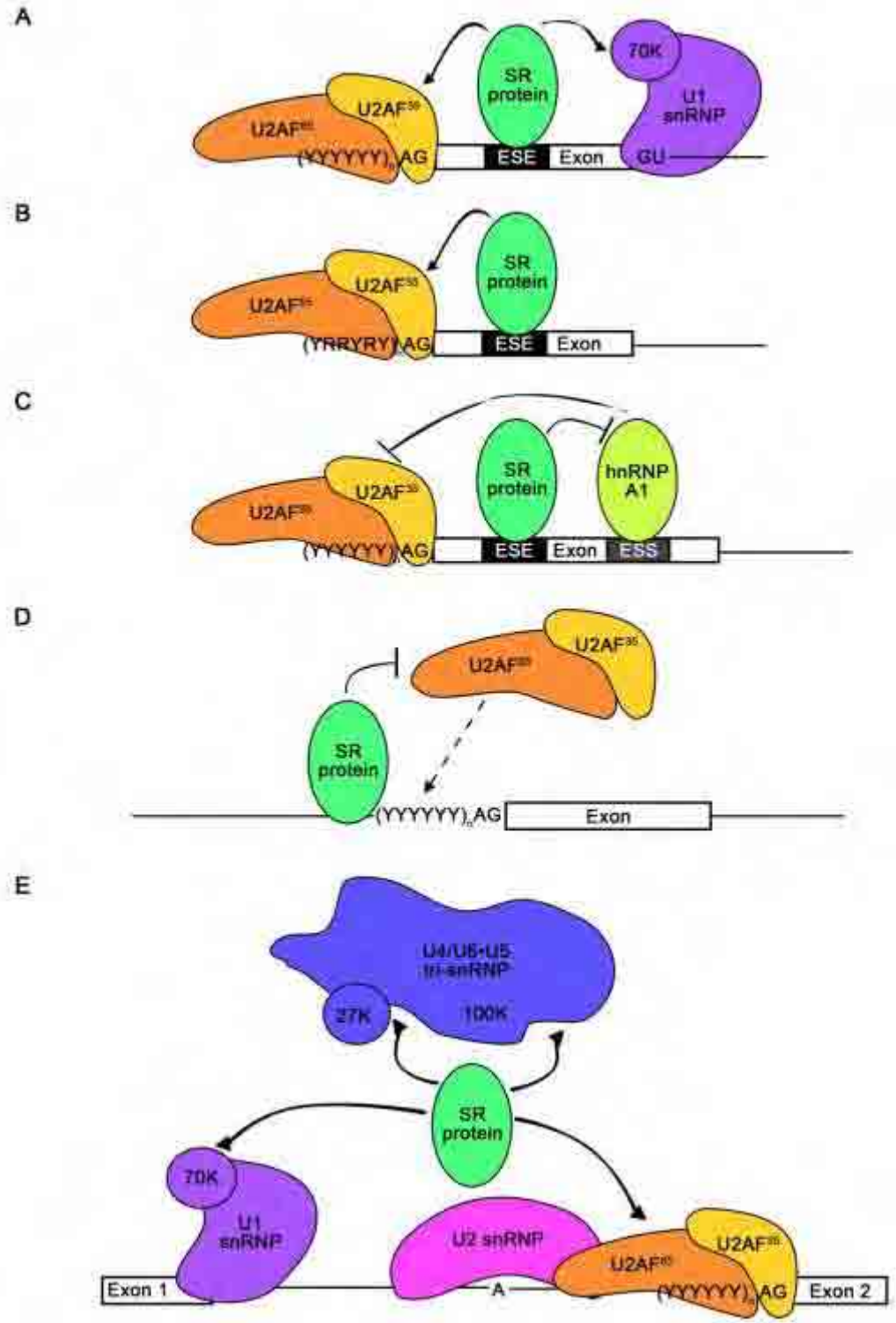

Figure 1.11 SR protein control of splicing. (**A**) SR proteins bound to an ESE can interact with U1 snRNP bound to the 5' site and U2AF bound at the 3' site. (**B**) SR proteins bound to ESEs can promote alterative 3' splice site selection by recruiting U2AF to a suboptimal 3' splice site (weak polypyrimidine tract). (**C**) Splicing repressor proteins such as hnRNP A1 can bind to ESSs and block 3' splice site selection by U2AF, SR proteins can antagonise these effects. (**D**) SR proteins can associate with intronic sequences and inhibit U2AF recruitment sterically. (**E**) SR proteins facilitate splice site pairing and assist in U4/U6•U5 tri-snRNP recruitment. $(YYYYYY)_n$ represents pyrimidine tract Adapted from (Shepard and Hertel, 2009).

1.14.3 mRNA export, stability and translation

Splicing not only removes introns from the mRNA, but also rearranges the mRNP particle to promote export from the nucleus. Some mRNAs however do not require splicing, but are still exported. The shuttling SR proteins, SRp20 and 9G8, have been shown to promote the nuclear export of intronless histone H2A mRNAs, by specifically binding to a 22 nt sequence within the H2A mRNA (Huang and Steitz, 2001). The same two SR proteins have been shown to be adapter proteins, which stimulate the export of mRNA by forming a ternary complex with Tip-associated protein (TAP), which is an essential nuclear export factor (Hautbergue et al., 2008). The binding of TAP to the SR proteins increases the RNA binding affinity of TAP and results in more efficient mRNA handover (Hautbergue et al., 2008).

SR proteins have been implicated in RNA stability and quality control. ASF/SF2 actively promotes the degradation of PKCI-r (protein kinase C-interacting protein 1 related) mRNA through the binding of a purine-rich element within the PKCI-r 3' untranslated region (Lemaire et al., 2002). The overexpression of various SR proteins strongly enhances nonsense-mediated decay (NMD) of mRNAs containing premature termination codons, suggesting that SR proteins can enhance the nuclear steps of NMD and target mRNAs to this pathway (Z. Zhang and Krainer, 2004).

SR proteins have been shown to influence translation either indirectly or directly. An indirect function of ASF/SF2 is by influencing the alternative splicing of the pre-mRNA for the protein kinase MNK2, a kinase which regulates translation initiation. High levels of ASF/SF2 promote the production of an MNK2 isoform which enhances cap-dependent translation, whereas low levels of ASF/SF2 promote the opposite (Karni et al., 2007). ASF/SF2 can also regulate translation directly, as ASF/SF2 can interact with polyribosome fractions in cytoplasmic extracts and enhance the translation efficiency of an ESE-containing luciferase reporter (Sanford et al., 2004). ASF/SF2 has been shown to affect translation through the

mediation of recruitment of components of the mTOR (mammalian target of rapamycin) signalling pathway. ASF/SF2 interacts with the protein kinase mTOR and the phosphatase PP2A and potentiates hyperphosphorylation of 4E-BP (eIF4E binding protein), blocking the inhibitory activity of 4E-BP on cap-dependent translation (Michlewski et al., 2008). SRp20 has been shown to be required for the stimulation of internal ribosome entry site (IRES)-mediated translation of poliovirus RNA (Bedard et al., 2007), while both SRp20 and 9G8 have been shown to increase the translation efficiency of unspliced mRNA containing a constitutive transport element (Swartz et al., 2007). The numerous roles of SR proteins in post-transcriptional gene expression are summarised in table 2.

Table 2 Roles of SR proteins in post-transcriptional gene expression

	Constitutive /alternative splicing	Splicing repression	Nucleo-cytoplasmic shuttling	mRNA export	mRNA stability	NMD	mRNA translation
ASF/SF2	+	+	+	-	+	+	+
SC35	+	+	-	-	ND	+	-
SRp20	+	-	+	+	ND	-	+
SRp75	+	-	+	ND	ND	ND	ND
SRp40	+	-	-	ND	ND	+	-
SRp55	+	-	+	ND	ND	+	ND
9G8	+	-	+	+	ND	-	+
SRp46	+	-	ND	ND	ND	ND	ND
SRp30c	+	-	ND	ND	ND	ND	ND
SRp38	-	+	ND	ND	ND	ND	ND
SRp54	+	-	ND	ND	ND	ND	ND

ND: not determined, +: experimentally determined, -: no demonstrated role (Sanford et al., 2005)

1.14.4 Other functions of SR proteins

The functions of SR proteins are not limited to post-transcriptional activities and there is growing evidence that the SR proteins act to couple transcriptional elongation to splicing. Depletion of ASF/SF2 and SC35 has been shown to dramatically attenuate the production of nascent RNA *in vivo* and SC35 has been shown to have an active role in transcriptional elongation (Lin et al., 2008). The SR proteins interact with the phosphorylated C-terminal

domain (CTD) of the largest subunit of PolII (Misteli and Spector, 1999) and *in vivo* depletion of SC35 diminishes the association of PolII with pTEFb, a kinase that phosphorylates CTD at Ser2 positions, which is critical for transcriptional elongation (Lin et al., 2008). This suggests that SR protein recruitment to PolII may help stabilise critical transcriptional elongation factors such as pTEFb and SR proteins may play an integral part in transcriptional elongation as well as reverse coupling transcription to splicing (Zhong et al., 2009b). The SR proteins SRp20 and ASF/SF2 can also bind directly to the histone H3 tail via the RS domain, the SR proteins may bind to the exposed H3 tail as a consequence of nucleosome rearrangement during transcriptional elongation (Loomis et al., 2009). SRp20 and ASF/SF2 associate with chromatin at interphase and are released from the hyperphosphorylated mitotic chromosomes and then reassociated with chromatin in the late M-phase. The phosphorylation of the H3 tail at serine 10 by Aurora kinase B causes the detachment of the SR proteins from the chromatin, also hyperphosphorylation of the SR proteins by SRPKs causes the interaction between SR proteins and the H3 tail to diminish. SR proteins have been shown to become hyperphosphorylated when cells enter M-phase. ASF/SF2 has also been shown to be required for the dismantling the heterochromatin mark HP1 from chromatin, suggesting that SR proteins may facilitate chromosome condensation and contribute to cell-cycle progression through M-phase (Loomis et al., 2009, Zhong et al., 2009b).

The SR protein ASF/SF2 has been shown to negatively autoregulate its expression to maintain homeostatic levels. ASF/SF2 can modulate alternative splicing to downregulate itself by decreasing the production of the protein-coding isoform and increasing the isoforms that are retained in the nucleus and are therefore not translated or are degraded by nonsense mediated decay (NMD) (S. Sun et al., 2010). ASF/SF2 can also form a negative feedback loop with a microRNA (miR-7), with ASF/SF2 promoting miR-7 maturation and mature miR-7 targeting the 3' UTR of ASF/SF2 to repress its translation (H. Wu et al., 2010). The SR

protein SRp20 has also been shown to negatively regulate the alternative splicing of its own mRNA, with overexpression of SRp20 leading to a transcript which encodes a truncated protein lacking the RS domain. Interestingly ASF/SF2 antagonises this auto-regulation and suppresses the formation of the truncated protein (Jumaa and Nielsen, 1997).

Another function of ASF/SF2 is the regulation of the sumoylation pathway, overexpression of ASF/SF2 stimulates SUMO (small ubiquitin-related modifier) conjugation, with ASF/SF2 knockdown inhibiting sumoylation. ASF/SF2 interacts with the SUMO E2 conjugating enzyme, Ubc9 and with specific substrates facilitates the transfer of SUMO from the thioester intermediate to the substrate, similar to a SUMO E3 ligase. In addition, ASF/SF2 interacts with the SUMO E3 ligase PIAS1 (protein inhibitor of activated STAT-1) and regulates PIAS1-induced protein sumoylation (Pelisch et al., 2010).

SR proteins are described as splicing factors; however they have diverse roles that couple the processes of translation, splicing, mRNA export and translation, along with many other roles which are unrelated to mRNA splicing.

1.14.5 SRPKs, SR proteins and cancer

The gene encoding the SR protein ASF/SF2 (*SFRS1*) is a proto-oncogene with roles in the establishment and maintenance of transformation. The *SFRS1* gene is upregulated in tumours of the colon, thyroid, small intestine, kidney, lung and cervix (Fay et al., 2009, Karni et al., 2007). ASF/SF2 controls the alternative splicing of the putative tumour suppressor BIN1, the transcription factor TEF-1, the kinase MNK2 and the ribosomal-protein kinase SK61. Overexpression of ASF/SF2 affects the splicing of these proteins and hence affects their control of transformation and apoptosis (Karni et al., 2007). Another regulatory protein which is differentially spliced by ASF/SF2 is vascular endothelial growth factor A (VEGFA),

which is a key regulator of physiological and pathological angiogenesis. Alternative splicing of VEGFA leads to two different isoforms which are either pro-angiogenic or anti-angiogenic. Overexpression of ASF/SF2 leads to the production of the pro-angiogenic isoform, whereas overexpression of SRp55 favours the productions of the anti-angiogenic isoform (Nowak et al., 2008). The phosphorylation of ASF/SF2 by SRPK1 has been shown to be involved in the production of the pro-angiogenic isoform of VEGFA (Nowak et al., 2010). Further evidence of the oncogenic functions of ASF/SF2 come from overexpression studies in which the overexpression of ASF/SF2 can cause colony formation in NIH 3T3 and Rat1 fibroblasts grown in soft agar; however the effect is not seen to the same extent with SC35 and SRp55 (Karni et al., 2007). Furthermore injection of ASF/SF2 overexpressing NIH 3T3 cells into nude mice formed large tumours. ASF/SF2 has also been shown cooperate with H-Ras to enhance cellular proliferation and appeared to protect cells from DNA damage and serum deprivation induced apoptosis (Karni et al., 2007).

SRPK1 has been shown to upregulated in breast and colonic tumours, with the expression of SRPK1 increasing with tumour grade (Hayes et al., 2007). The down-regulation of SRPK1 using siRNA, sensitised tumour cells to chemotherapeutic intervention with cisplatin and gemcitabine. Additionally alterations in the splice patterns of MAPK kinase 2 were observed upon SRPK1 inhibition, which resulted in a decrease in phosphorylation of MAPK2, MAPK3 and AKT (Hayes et al., 2007). SRPK2 can bind to and phosphorylate the SR-like protein acinus, causing its relocalisation from nuclear speckles to the nucleoplasm, resulting in the activation of cyclin A1. The overexpression of SRPK2 or acinus increases leukaemia cell proliferation and both of these proteins have been shown to be overexpressed in human acute myelogenous leukaemia patients (Jang et al., 2008). These studies demonstrate that SRPKs are involved in mediating the regulation of splicing of oncogenes.

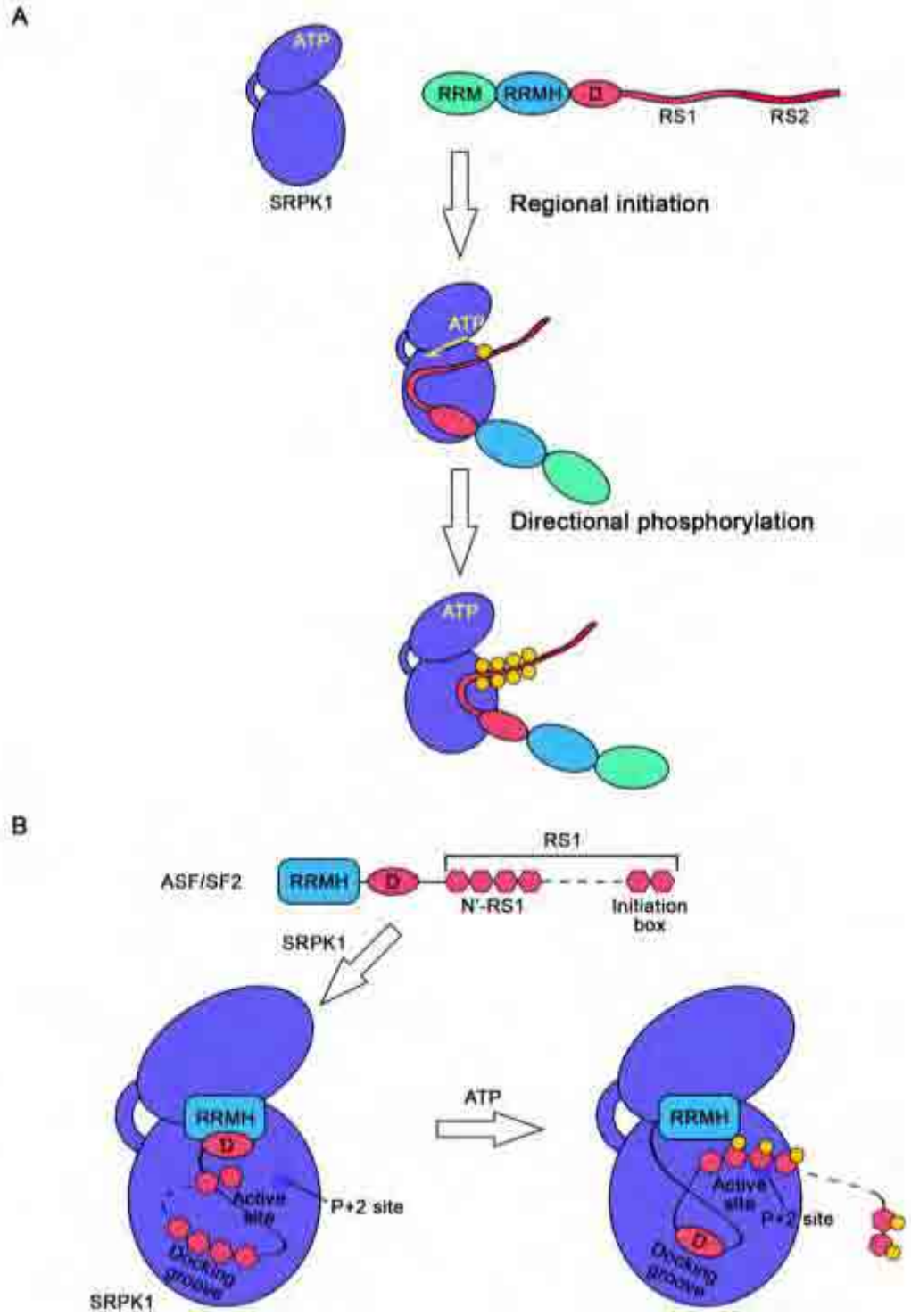

Figure 1.12 Mechanism of phosphorylation of ASF/SF2 by SRPK1. (**A**) Grab-and-pull mechanism for the phosphorylation of ASF/SF2. SRPK1 docks at the C-terminal end of RS1 and loosely initiates phosphorylation at the RS1/RS2 boundary (regional initiation). SRPK1 then pulls N-terminal serines into the active site for phosphorylation. (**B**) The N-terminal portion of RS1 (N'-RS1) binds in the docking groove while the initiation box occupies the active site. In the presence of ATP, RS1 is phosphorylated until the docking motif (D) occupies the docking groove. Basic side chains in the P+2 site stabilise the phosphates on RS1. Pink hexagons: RS pairs, dotted line: other RS pairs in RS1, yellow circles: phosphate groups. Adapted from (Ma et al., 2008) and (Ngo et al., 2008).

1.15 Regulation of SR proteins

1.15.1 SRPK1 phosphorylation

The RS domain of the SR proteins controls a large number of the SR protein interactions and functions; therefore one method of controlling these interactions is phosphorylation of this domain. A dynamic cycle of phosphorylation and dephosphorylation is required for pre-mRNA splicing, as upon recruitment of the SR proteins for splicing, the RS domain is hyperphosphorylated (Gui et al., 1994a); however the RS domain becomes partially dephosphorylated during splicing in a step which is necessary for SR protein nuclear export (Lai and Tarn, 2004). Multiple studies have been performed to identify the kinases and phosphatases responsible for regulating the SR proteins *in vivo*.

The first kinase to be identified which phosphorylated the RS domain was SRPK1, which specifically phosphorylates the serine residues in the RS dipeptides. SRPK1 is known to preferentially phosphorylate up to twelve serines in the N-terminal region of the RS domain (RS1) (Figure 1.10B) (Hagopian et al., 2008). A docking motif (RVKVDGPR) in ASF/SF2, located between the second RRM and the RS domain specifically interacts with a docking groove in SRPK1 (Lukasiewicz et al., 2007, Ngo et al., 2005). Start-trap experiments have shown that once bound SRPK1 phosphorylates ASF/SF2 in a processive manner (Aubol et al., 2003, Velazquez-Dones et al., 2005) in the direction of the N-terminus using a 'grab-and-pull' mechanism to control the regiospecific phosphorylation of the protein (Figure 1.12A) (Ma et al., 2008). SRPK1 initially binds to an initiation box (Ser221-Ser225) and rapidly phosphorylates the first eight serines, the remaining serines are then modified in a slower phase, suggesting a dual-track mechanism, incorporating both processive and distributive phosphorylation steps (Ghosh and Adams, 2011, Ma et al., 2008). Although SRPK1 prefers to bind to the initiation box at the C-terminal end of RS1, mutations in this region do not stop phosphorylation of the RS1 domain, indicating that SRPK1 can bind to different sites (Ma et

al., 2008). Different crystallisation studies have identified that both the initiation box and the docking motif of ASF/SF2 can bind to the docking groove in SRPK1, indicating that the structure of the SRPK1-ASF/SF2 complex changes as a function of phosphorylation. Therefore a proposed model of phosphorylation is that the RS repeat motif (residues 204-210) serves as a mobile docking element, placing the C-terminal serine from the initiation box in the active site. As each serine is phosphorylated the docking motif moves by two residues at a time towards the N-terminus, as the whole RS1 motif is fed through the active site of the kinase the N-terminal docking motif enters the docking groove (Figure 1.12B) (Ngo et al., 2008). The basic pocket at the P+2 site (prime site) interacts with a phosphoserine and stabilises the active conformation of the kinase, thus facilitating phosphorylation, after each round of phosphorylation a new phosphoserine replaces the old one with enhanced catalytic efficiency. The priming phosphorylation demonstrated by SRPK1 suggests that the rate-limiting step in the processive phosphorylation of ASF/SF2 is the phosphorylation of the first serine due to the lack of binding to the prime site (Ngo et al., 2008). However it has been shown that SRPK1 binds to unphosphorylated ASF/SF2 with high affinity, but this affinity weakens during subsequent phosphorylation steps, promoting dissociation over forward catalysis leading to a slower rate of phosphorylation (Ma et al., 2008). The fully phosphorylated ASF/SF2 protein binds to SRPK1 with five fold less affinity than the unphosphorylated protein (Koizumi et al., 1999).

1.15.2 Other kinases and regulatory systems

The second family of kinases which can phosphorylate SR proteins are the CLK/STY (CDC28/cdc2^{+}-like kinase/serine threonine tyrosine) family of dual specificity kinases, consisting of CLK1-4 (Colwill et al., 1996). The CLK family can phosphorylate both serine/threonine and tyrosine residues and CLK is the prototypical member of the LAMMER family of protein kinases, which are conserved throughout evolution and share a conserved

sub-domain (EHLAMMERILGPLP) within the kinase domain (Colwill et al., 1996). Unlike SRPK1, CLK can phosphorylate all of the serine residues in the RS domain of ASF/SF2 with the phosphorylation occurring by a processive mechanism (Velazquez-Dones et al., 2005). In contrast to SRPK1, X-ray structures of CLK reveal no deep docking groove; instead a shallow area with both acidic and basic charge patches is present. The charge distribution in CLK suggests that the hypophosphorylated RS domain could interact with CLK with higher efficiency than the unphosphorylated domain, leading to the suggestion that the product of SRPK1 phosphorylation is the substrate for CLK (Ghosh and Adams, 2011).

In addition to the SRPKs and CLKs, DNA topoisomerase I has been shown to phosphorylate the RS domain, DNA topoisomerase I has no obvious kinase domain, but can bind ATP with high affinity and can phosphorylate the serine residues in the RS domain of several SR proteins (Rossi et al., 1996). The phosphorylation of the RS domain by DNA topoisomerase I is dependent upon the number of RS repeats, with DNA topoisomerase I requiring at least five RS dipeptide repeats for efficient binding and phosphorylation of ASF/SF2 (Labourier et al., 1998). Another kinase shown to be able to phosphorylate the RS domain of SR proteins is mammalian PRP4 (pre-mRNA processing mutant 4). PRP4 contains an RS domain and putative nuclear localisation signals in the N-terminal region and has been shown to interact with and phosphorylate ASF/SF2 (Kojima et al., 2001). The protein kinases C and A (Colwill et al., 1996), along with CDC2 kinase (Okamoto et al., 1998) and AKT kinase (Blaustein et al., 2005) can also phosphorylate ASF/SF2 *in vitro*.

The phosphorylation of the RS domain has been shown to influence the localisation of the SR proteins as overexpression of both SRPKs and CLKs can cause the disassembly of nuclear speckles and a redistribution of splicing factors (Colwill et al., 1996, Gui et al., 1994a). Additionally, the phosphorylation of the SR proteins by SRPKs in the cytoplasm is required

for the interaction with the transportin-SR complex and nuclear reimport (Lai et al., 2000). The phosphorylation state of the RS domain controls the protein-protein interactions and the protein-RNA interactions, with phosphorylation required for the interaction between ASF/SF2 and U1-70K snRNP (Xiao and Manley, 1997), whereas dephosphorylation is required for the binding of ASF/SF2 to the RNA export factor Tap/NFX1 (Huang et al., 2004). The current model of the role of kinases in the control of SR proteins is that CLKs is responsible for recruiting the SR proteins to sites of active transcription and the SRPKs are responsible for nuclear reimport (Figure 1.13). However, as SRPKs can enter the nucleus and multiple phosphorylation states are required for different functions, the regulation of SR protein by kinases is likely to be a great deal more complex.

As dephosphorylation is required for the SR protein control of splicing, specifically spliceosome maturation, the export of mRNA and the release of the SR proteins from mRNA, phosphatases are required to remove the phosphate groups from the RS domain. Protein phosphatase 1 (PP1) has been identified as a phosphatase responsible for dephosphorylating SR proteins. PP1 removes the phosphate groups from the RS1 domain in an N to C direction, opposite to the direction which SRPK1 acts (Ma et al., 2010). The interaction of ASF/SF2 with ESEs negatively regulates dephosphorylation of the RS1 domain as does phosphorylation of the RS2 domain by CLK/STY, suggesting that the control of dephosphorylation is regiospecific (Ma et al., 2010).

In addition to phosphorylation the SR proteins are also controlled by methylation, with three residues in the linker region between the RRM and RRMH domain of ASF/SF2 being methylated by the human homolog of Hmt1p (PRMT1). This methylation influences the nucleo-cytoplasmic distribution as an ASF/SF2 mutant which could not be methylated was localised in the cytoplasm (Sinha et al., 2010).

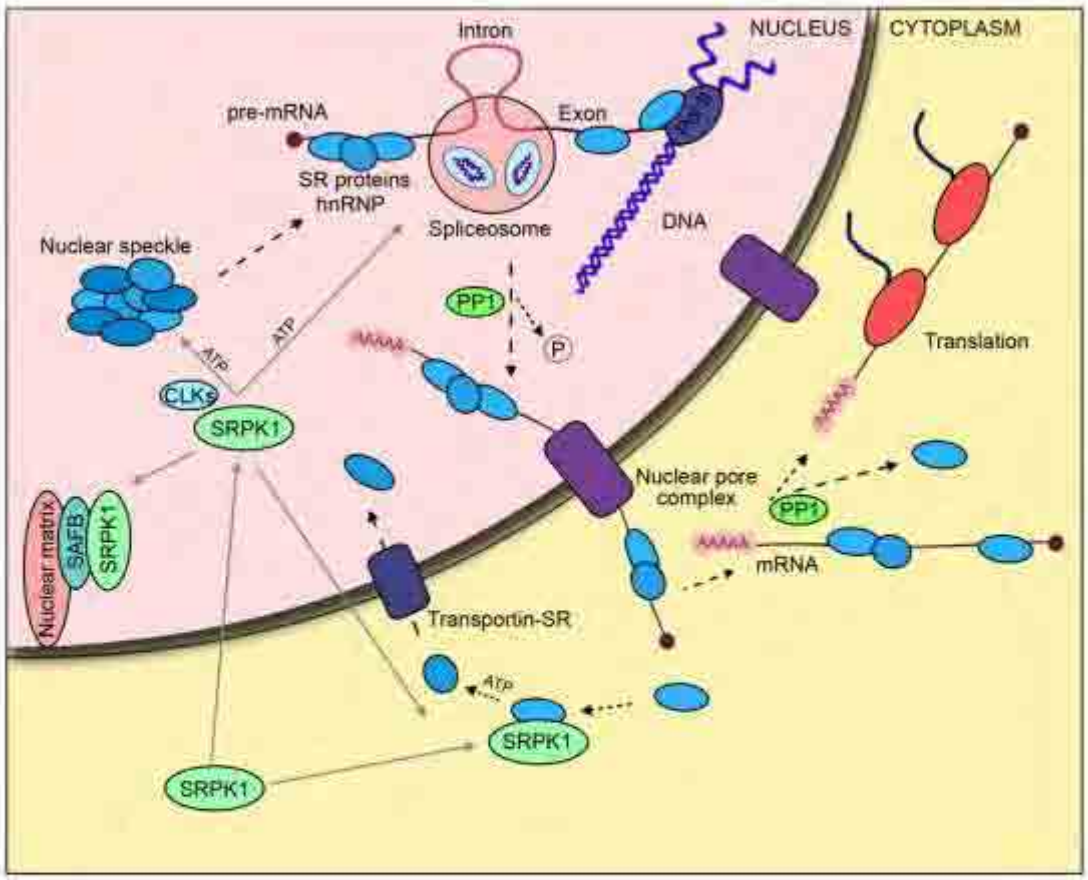

Figure 1.13 Regulation of SR proteins. SR proteins are usually in nuclear speckles; phosphorylation releases the SR proteins to participate in spliceosome assembly. Phosphorylation and dephosphorylation of the SR proteins is required during splicing. PP1 dephosphorylation of the SR proteins releases some post-splicing and the remainder after export from the nucleus. The SR proteins are phosphorylated by SRPK1 to facilitate re-import into the nucleus. Grey lines indicate involvement or movement of SRPK1; dashed lines indicate movement of SR proteins.

1.16 HPV regulation of splicing factors

As the genome of HPV requires extensive constitutive and alternative splicing, interactions between HPV proteins and splicing factors are to be expected in order for the virus to be preferentially spliced. In normal epithelium the expression of ASF/SF2 is typically restricted to the dividing cells in the basal layer; however infection of the epithelium with HPV16 causes ASF/SF2 to be expressed in the upper layers of the epithelium (McPhillips et al., 2004). The induction of ASF/SF2 expression by the virus is via the HPV16 E2 protein binding and *trans*-activating the promoter of the *SFRS1* gene, causing a 4-8 fold increase in ASF/SF2 levels in differentiated cells (McPhillips et al., 2004, Mole et al., 2009b). HPV16 E2 has also been shown to specifically up-regulate SRp20 and SC35 (Mole et al., 2009a) and HPV31 has also been shown to regulate ASF/SF2 (Mole et al., 2009b). The virally induced upregulation of the SR proteins is essential for the expression of some virus proteins, as HPV-infected HFKs expressing a mutant E2 which cannot stimulate the expression of SR proteins showed reduced levels of late virus RNAs, including the virus capsid proteins which are required for virion formation. This indicates that SR proteins are required for the expression of virus proteins in the mid-to-upper layers of the epithelium (McFarlane and Graham, 2010). The levels of SR proteins are also increased with cervical disease progression, with increased SR protein expression throughout the epithelial layers in higher grades of premalignant lesions (Mole et al., 2009a). The levels of ASF/SF2, SRp20 and SC35 increase in cells with integrated HPV genomes, therefore this increase cannot be due to E2 regulation as the expression of E2 is abrogated in integration (McFarlane and Graham, 2010). The hypothesised model of SR protein upregulation is that chromosomal rearrangements or changes in methylation patterns result in the amplification or activation of the genes (McFarlane and Graham, 2010).

The most frequently used splice site in HPV16 is splice acceptor 3358 (SA3358), which is used to produce E6, E7 and E1^E4 transcripts. SA3358 is followed by an ESE with 15 potential binding sites for ASF/SF2; mutational inactivation of the 15 sites redirected splicing to the late splice site SA5639 demonstrating that the SA3358 splice site is dependent on ASF/SF2 binding (Somberg and Schwartz, 2010). Therefore high levels of ASF/SF2 may be a requirement for the high expression of E6 and E7. The SA3358 ESE can also bind to SRp20 and this binding suppresses use of the late splice site SA5639 (Jia et al., 2009). High levels of SRp20 promote the expression of E6 and E7, whereas low levels promote the expression of L1, with L1 and SRp20 having an inverse expression in HPV infections (Jia et al., 2009).

The E2 proteins of the beta-HPV types (HPV5 and 8) contain RS dipeptide repeats in the hinge between the N- and C-terminal domains. The RS dipeptide repeats allow the E2 protein of HPV5 to interact with the SR proteins ASF/SF2, SRp20, SC35, SRp40, SRp55 and SRp75 and also with U1-70K and U1-100K. The E2 protein could not recruit the splicing factors and *trans*-activate splicing without activating gene transcription through binding to the promoter. This data indicates that the E2 proteins can play a dual role in gene expression with the RS-rich hinge of HPV5 E2 functioning to recruit splicing factors to splice pre-mRNA made via transactivation by E2 itself (Lai et al., 1999).

A number of different DNA viruses target splicing factors via various different mechanisms to control the balance between host and viral post-transcriptional gene regulation, demonstrating that the SR proteins are important targets for the viruses and as more functions for these proteins emerge they have become potential targets for anti viral therapy.

1.17 Aims and objectives

The E1^E4 protein of HPV types 1, 16 and 18 interacts with SRPK1 (Bell et al., 2007). In warts, SRPK1 is associated with E4 containing inclusion bodies in cells present in the upper cell layers of the lesion (Bell et al., 2007). SRPK1 cellular substrates are known to be involved in the regulation of HPV gene expression (Jia et al., 2009, Somberg and Schwartz, 2010). I hypothesise that the interaction between E1^E4 and SRPK1 alters SRPK1 activity and that this association influences the functions of SRPK1 in the HPV life cycle. This PhD study aims to investigate how E1^E4 affects SRPK1 functions and to understand the role of SRPK1 in the HPV life cycle and how the virus regulates SRPK1.

The specific objectives of this thesis are:

To investigate how E1^E4 proteins from alpha, beta and mu papillomaviruses influence SRPK1 kinase activity *in vitro* (Chapter 3).

To investigate the cellular and viral substrate specificity of any effect on SRPK1 activity caused by the interaction between SRPK1 and HPV E1^E4 (Chapter 4).

To determine the effect of the interaction between SRPK1 and HPV E1^E4 on the cellular localisation of SRPK1 cellular substrates and SRPK1 functions such as splicing regulation (Chapter 5).

To investigate the role of SRPK1 in the HPV life cycle (Chapter 6).

CHAPTER 2 MATERIALS AND METHODS

2.1 Bacterial cell culture

2.1.1 Bacterial hosts growth and storage

The Escherichia coli (*E. coli*) strain DH5α was used as a bacterial host for growth of plasmid vectors to prepare the plasmid DNA. The *E. coli* strains BL21 (Stratagene), BL21 pLys and BL21 CodonPlus® (Stratagene) were used for the expression of recombinant proteins. Bacteria were streaked onto agar plates and grown at 37 °C (1.5% w/v agar in Luria-Bertani (LB) medium with appropriate antibiotic). Plates were stored at 4°C for short periods of time or the bacteria were stored as a glycerol stock (25% glycerol, 50% overnight bacterial culture) at -80°C for longer periods of time.

Ampicillin was used at 50 μg/ml, kanomycin used at 20 μg/ml and chloramphenicol used at 34 μg/ml final concentration. Chloramphenicol was dissolved in ethanol at 34 mg/ml.

2.1.2 Preparation of chemically-competent bacteria

A streak plate was prepared from the bacterial stock on a LB agar plate with no selection antibiotic. Four millilitres of LB media was inoculated with a single colony from the agar plate and grown overnight at 37°C with shaking at 200 rpm. The following day 200 ml of LB media was inoculated with the starter culture and grown at 37°C with shaking until an $OD600_{nm}$ of 0.4-0.6. The bacteria were pelleted by centrifugation at 3,834 x g for 15 min at 4°C and re-suspended in 50 ml of cold TFB1 buffer (30 mM KOAc, 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$, 15% v/v glycerol) and pelleted again by centrifugation at 3,834 x g. The pellet was re-suspended in 8 ml of cold TFB2 buffer (10 mM Na-MOPS pH 7.0, 75 mM

$CaCl_2$, 10 mM KCl, 15% v/v glycerol), aliquoted, snap frozen in liquid nitrogen and stored at -80°C.

2.1.3 Transformation of competent bacteria with plasmid DNA

Plasmid DNA (0.5-1 µg) was added to 50 µl of competent bacteria in a pre-chilled polypropylene tube and incubated on ice for 10 min. The bacteria were heat shocked at 42°C for 90 s and returned to the ice for 2 min. Five hundred µl of LB media was added and the sample shaken at 200 rpm at 37°C for 45 min. Fifty µl of bacteria was spread onto an LB agar plate containing the relevant antibiotic and incubated overnight at 37°C.

2.1.4 Preparation of plasmid DNA

2.1.4.1 Small scale preparation

Four ml of LB media containing the relevant antibiotic was inoculated with a single colony and grown overnight at 37°C with shaking at 200 rpm. The following day 3 ml of the bacteria were pelleted by centrifugation at 16100 x g for 5 min at 4°C and the pellets processed using the QIAprep® Spin Miniprep kit (Qiagen) according to the manufacturer's instructions. DNA was eluted from the column into 50 µl of Tris-EDTA (TE) buffer and stored at 4°C.

2.1.4.2 Large scale preparation

A 4 ml starter culture of LB media with antibiotics was inoculated with a single colony and grown at 37°C with shaking at 200 rpm. After 6-8 h 200 ml of LB media was inoculated with the starter culture and grown at 37°C overnight with shaking. The bacteria were pelleted by centrifugation at 3,834 x g for 15 min at 4°C and the pellet processed using the Qiagen HiSpeed® Plasmid Maxi kit (Qiagen) or PureLink HiPure® Plasmid Filter Maxiprep kit (Invitrogen) according to the manufacturer's instructions. The DNA was re-suspended in 0.5-1 ml of TE buffer and stored at 4°C (short term) or at -20°C (long term). The concentration

of the plasmid DNA solution was determined using the Nanodrop™ spectrophotometer (Thermo Scientific).

2.1.5 Bacterial expression of recombinant GST and His-tagged fusion proteins

2.1.5.1 Small scale expression

To determine the optimum condition for the expression of glutathione-S-transferase (GST) and His-fusion (polyhistidine, His6) proteins a 20 ml starter culture of LB media with antibiotics was inoculated with a single bacterial colony and grown overnight at 37°C with shaking at 200 rpm. The following day 60 ml of LB media with antibiotics was inoculated with 6 ml of the starter culture and grown at 37°C with shaking. After 2 h, isopropyl β-D-1-thiogalactopyranoside (IPTG, Sigma) was added to a final concentration of 1 mM and a 6 ml aliquot was taken as a 0 h sample and the bacteria were pelleted by centrifugation at 16100 x g for 5 min. Cultures were grown at 37°C or 30°C with shaking and 6 ml samples taken at 1 h, 4 h and 24 h post IPTG-induction.

2.1.5.2 Large scale expression

A 20 ml starter culture of LB media was prepared as described in section 2.1.5.1 and used to inoculate 200 ml of LB media with antibiotics and 1% glucose and grown at 37°C with shaking. After 2 h, IPTG was added to a final concentration of 1 mM and the culture grown at 37°C or 30°C with shaking for an appropriate time that ensured maximum expression of the tagged fusion protein. The bacteria were pelleted by centrifugation at 3,834 x g for 15 min at 4°C.

2.1.5.3 Purification of recombinant GST or His-tagged fusion proteins

The bacterial pellet was re-suspended in 4-6 ml (400 µl for small scale) of *E. coli* lysis buffer (0.5% v/v Triton® X-100 and Complete™ protease inhibitors (Roche) in phosphate buffered saline (PBS, Dulbecco A)) and sonicated with three 30 s, 10 Watt pulses, using a microson ultrasonic cell disrupter (Misonix). The samples were centrifuged at 16100 x g for 5 min at 4°C and then 200 µl (40 µl for small scale) of a 50% slurry of glutathione (Sigma-Aldrich) or nickel-NTA (Qiagen) agarose beads, washed in *E. coli* lysis buffer, were added to the supernatant. The proteins were bound to the beads with slow speed rotation with inversion at 4°C for 1-2 h. The beads were pelleted by centrifugation at 200 x g for 1 min and the supernatant removed. The beads were then washed 6 times with 0.5 ml of *E. coli* lysis buffer and the proteins were eluted off the beads with 2 x 125 µl aliquots of reduced glutathione (50 mM Tris-HCl pH 8.0, 0.1 M glutathione) in *E. coli* lysis buffer or nickel elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole) for 30 min at RT. The proteins were dialysed using GeBAflex-tubes (MWCO 6-8 kDa, Gene Bio-Application Ltd) into 25 mM Tris-HCl pH 7.5, 200 mM NaCl overnight at 4°C, glycerol added to 5 %, aliquoted and stored at -80°C. Protein concentration was determined using the Bradford assay (Section 2.3.1).

2.1.5.4 Thrombin cleavage of GST-tag

The pGex-2T vector system (GE Healthcare) has a thrombin cleavage site to enable the removal of the GST tag. To remove the GST-tag from GST-HPV1 E1^E4, the fusion protein was expressed on a large scale (section 2.1.5.2), but not eluted from the glutathione agarose beads. The beads were incubated with 5 units of thrombin protease (GE Healthcare) in PBS per 100 µl bed volume of beads, overnight at 30°C. The cleaved protein was loaded onto a pre-equilibrated Benzamidine FF Hitrap column (GE Healthcare) and washed with 50 mM Tris-HCl pH 7.5, 0.5 M NaCl followed by 50 mM Tris-HCl pH 7.5, 1 M NaCl. The protein was collected in the high salt wash and the salt reduced to 200 mM by dialysis.

2.2 Molecular Cloning

Table 3 Plasmid DNA vectors

Plasmid	Source
pGex-2T	GE Healthcare
pcDNA 3 and 3.1	Invitrogen
pEGFP-C1	Clontech
pCMVFlag4-SRPK1 (codes for N-terminal flag tag (Sigma)	(Bell et al., 2007)
pcDNA HPV1 E1^E4	(Roberts et al., 1993)
pcDNA HPV E1^E4 Δ44-48	(Roberts et al., 1994)
pGex-2T HPV1 E1^E4	(Bell et al., 2007)
pGex-3X HPV16 E1^E4	(Bell et al., 2007)
pGex-3X HPV18 E1^E4	(Bell et al., 2007)
pGex-2T HPV1 E1^E4 Δ44-48	(Bell et al., 2007)
pGex-2T HPV1 E1^E4 Δ49-53	(Bell et al., 2007)
pEGFP-ASF/SF2	A gift from Bettina Heinrich, University of Erlangen,
pSG1aE2	A gift from Saleem Khan, University of Pittsburgh
HPV 5b complete genome	A gift from Ethel-Michele de Villiers, Deutches Krebsforshungszentrum
HPV 8 complete genome	A gift from Ethel-Michele de Villiers, Deutches Krebsforshungszentrum
pRSET B-SRPK1	A gift from Bai-Gong Yue, University of Leicester

2.2.1 Polymerase chain reaction (PCR)

DNA for cloning was prepared by PCR amplification with specific primers (Table 4) using the Expand High Fidelity PCR system (Roche) or Taq polymerase (Roche). The forward and reverse primers (25 pmol) (generated by AltaBiosciences) were mixed with 2 mM dNTPs (Roche), buffer (10 mM Tris-HCl pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl), 1 µg of DNA template, 1 µl DMSO and 3.5 units of Expand or 5 units of Taq polymerase in a total volume of 100 µl. Amplification was performed on a thermal cycler (2720 or 2700, Applied

Biosciences) using a 2 min 94°C hot start followed by 25 repeats of 94°C for 1 min, 55°C for 30 s and 72°C for 3 min, followed by 7 min at 72°C before cooling to 4°C. The PCR products were cleaned with High Pure PCR product Purification Kit (Roche) according to the manufacturer's instructions. DNA was eluted from the column in 50 µl dH_2O and stored at -20°C.

2.2.2 Restriction enzyme digestion

PCR product (20 µl) or plasmid DNA (2-4 µg) was digested with 10 units of relevant restriction enzyme (Roche), in the provided appropriate buffers in a total volume of 40 µl. After 1 h at 37°C another 10 units of enzyme was added and the reaction continued at 37°C for a further hour. Samples were purified using agarose gel electrophoresis and gel extraction.

2.2.3 Calf Intestinal Phosphatase (CIP) treatment of plasmid DNA

If necessary, the terminal 5' phosphate groups were removed from the digested products using CIP. Digested products were mixed with 1 unit of alkaline phosphatase (Roche) and the dephosphorylation buffer (0.5 M Tris-HCl pH 8.5, 1 mM EDTA) in a total volume of 100 µl and incubated at 37°C for 15 min. Samples were purified using agarose gel electrophoresis and gel extraction.

2.2.4 Agarose gel electrophoresis

For resolution of DNA samples, horizontal 1% (w/v) agarose gel (containing 0.25 mg/ml ethidium bromide) in TBE buffer (45 mM Tris, 45 mM Boric Acid, 1 mM EDTA). The agarose solution was dissolved by boiling and cast in a Mini Sub-Cell GT electrophoresis tank (Bio-Rad) or a Fisherbrand HU13 horizontal electrophoresis tank (Fisher Scientific). DNA samples were mixed with 6x Orange G Loading Buffer (4.4 mM Orange G, 10% v/v glycerol) prior to loading. Small gels were run at 80V and larger gels run at 120V for 1 h in TBE

buffer and DNA bands visualised with a Gene Flash UV light box (Syngene Bio Imaging). A 1 kb plus DNA ladder^{TM} (Invitrogen) was run on each gel for size determination.

To purify the DNA from the gel, bands of the expected sizes were excised using a clean scalpel upon UV illumination and purified from the gel using QIAquick Gel Extraction kit (Qiagen) according to the manufacturer's instructions. The DNA was eluted from the column into 30 µl dH_2O and stored at -20°C.

2.2.5 DNA ligation reactions

The prepared vector DNA and PCR product were mixed together in a ratio of 1:3, estimated via band intensity on an agarose gel, with 1 unit of T4 DNA ligase (Roche) and ligase buffer (660 mM Tris-HCl pH 7.5, 50 mM $MgCl_2$, 10 mM dithioerythritol, 10 mM ATP) in a total volume of 10 µl, and the reaction was carried out at 16°C overnight.

2.2.6 Cloning of DNA into a GST plasmid vector

The coding sequence of the protein shown in table 4 was amplified by PCR using the shown primers and following PCR purification digested using the shown restriction enzymes and ligated into the pGex-2T vector which had been digested with the same enzymes.

2.2.7 Generation of HA-tagged pcDNAs

To allow expression and identification of proteins in mammalian cells a HA-tagged pcDNA vector was created. A cassette containing the multiple cloning region and the sequence of the HA tag was created using the following primers: 5' primer 5'-AG CTT ATG TAC CCA TAC GAT GTT CCA GAT TAC GCT GGA TCC CCG GGA ATT CC-3' and 3' primer 5'-TCG AGG AAT TCC CGG GGA TCC AGC GTA ATC TGG AAC ATC GTA TGG GTA CAT A-3'. The primers were re-suspended to 100 µM and 10 µl of each primer was mixed together with 2 µl of Roche buffer M and heated to 95°C for 2 min then cooled very slowly to

room temperature. The cassette was diluted to 50 nM and then 5 µl of the mix was ligated into 100 ng of pcDNA3 vector digested with HindIII and XhoI.

The various cDNAs were removed from the pGex-2T constructs following digestion with the same enzymes used to clone. The inserts were then ligated into the pcDNA3-HA vector digested with the same restriction enzymes.

2.2.8 PCR sequencing

Plasmid DNA (1 µg) was mixed with 2.5 pmol primer, 1 µl BigDye® terminator v.3.1 ready reaction mix (Applied Biosystems) and relevant buffer in a total volume of 20 µl. Reaction was carried out with 25 cycles of 96°C for 10 s, 50°C for 5 s and 60°C for 4 min on a thermal cycler. The sequencing primers used are listed in Table 5, and were generated by AltaBiosciences. The DNA was precipitated with 5 µl 125 mM EDTA and 60 µl 100% ethanol at room temperature for 15 min, and the sample centrifuged at 16100 x g for 45 min at 4°C. The supernatant was removed and the pellet washed in 70 µl 70% ethanol, the sample was then centrifuged at 16100 x g for 15 min at 4°C. The supernatant was removed and the pellet air dried and re-suspended in 10 µl HiDi Formamide (Applied Biosystems). The DNA was run in a 3130xl ABI Prism™ DNA capillary sequencer (Applied Biosystems). Sequencing data was collected using the 3130 data collection software version 3.0 and was analysed using FinchTV Version 1.4.0 (Geospiza)

Table 4 Primers, template DNA and enzymes used for cloning

Protein		Primer	Template	Restriction enzyme
ASF/SF2	5’	5’ GCG CGG ATC CAT GTC GGG AGG TGG TGT GAT TCG TG 3’	pEGFP-ASF/SF2	BamHI
	3’	5’ GCG CGG ATC CTT ATG TAC GAG AGC GAG ATC TGC T 3’		
9G8	5’	5’ GCG CGG ATC CAT GTC GCG TTA CGG GCG GTA C 3’	IMAGE: 2967417	BamHI
	3’	5’ GCG CGG ATC CTC AGT CCA TTC TTT CAG GAC T 3’		
SRp20	5’	5’ GCG CGG ATC CAT GCA TCG TGA TTC CTG TCC 3’	IMAGE: 3049167	BamHI and EcoRI
	3’	5’ GCG CGA ATT CCT ATT TCC TTT CAT TTG ACC TAG 3’		
SRp75	5’	5’ GCG CGG ATC CAT GCC GCG GGT GTA CAT CGG C 3’	IMAGE: 3619538	BamHI and EcoRI
	3’	5’ GCG CGA ATT CTT AGG ACC TTG AGT GGG ACC 3’		
SC35	5’	5’ GCG CGG ATC CAT GAG CTA CGG CCG CCC CCC TC 3’	IMAGE: 5926371	BamHI and EcoRI
	3’	5’ GCG CGA ATT CTT AAG AGG ACA CCG CTC CTT C 3’		
HPV1 E2	5’	5’ GCG CGA ATT CAT ATG GAA AAC CTC AGC AGT CGC 3’	pSG1aE2	EcoRI
	3’	5’ GCG CGA ATT CTT AAG ACC CAT TAA ACT GTC C 3’		
HPV1 E2 H	5’	5’ GCG CGA ATT CAT GTT ATG TCT TCC ACT AGC TCC 3’	pSG1aE2	EcoRI
	3’	5’ GCG CGA ATT CTT ATA CAC AGA CCA CGG GTG G 3’		
HPV5 E2	5’	5’ GCG CGG ATC CAT GGA GAA TCT CAG CGA GCG 3’	HPV5b complete genome	BamHI and EcoRI
	3’	5’ GCG CGA ATT CTT AAA GAC TGT CCA GGT TGC C 3’		
HPV5 E2 H	5’	5’ GCG CGG ATC CGT CAC CAG CTC CAC GCC TCC 3’	HPV5b complete genome	BamHI and EcoRI
	3’	5’ GCG CGA ATT CTT ATT TGA CAA TGA TTA CTG G 3’		
HPV8 E2	5’	5’ GCG CGG ATC CAT GGA GAA TCT CAG CGA GCG TTT C 3’	HPV8 complete genome	BamHI and EcoRI
	3’	5’ GCG CGA ATT CTT ATA GAC TGT CCA GGT TAC C 3’		
HPV8 E2 H	5’	5’ GCG CGG ATC CAC TGT GTT TGC TCC TGT TAC C 3’	HPV8 complete genome	BamHI and EcoRI
	3’	5’ GCG CGA ATT CTT ATC GAA CCA ATA TTA CTG G 3’		
HPV5 E1^E4	5’	5’ GCG CGA ATT CAT ACG GAT CCT AAT CCT AAA GCT CCA CGC CTC CAG GGT C 3’	pGex-2T HPV5 E2	EcoRI
	3’	5’ GCG CGA ATT CTT ACT GGG GGG TCG CGA GCT TCT TCC 3’		

Table 5 Sequencing primers

Primer	Direction	Primer Sequence
pGex-2T	F	5' TTG AAG CTA TCC CAC AAA TTG 3'
	R	5' CAC CCG CCA ACA CCC GCT GAC 3'
T7	F	5' TAA TAC GAC TCA CTA TAG GG 3'
BGH	R	5' CTG TGA ATG CTG TGG CAG GGG 3'
ASF	261F	5' CTT CGT TGA GTT CGA GGA CC 3'
SRp75	371F	5' GGC CGA GAT AAA TAT GGC CC 3'
	1222R	5' CTC ACT CTT GCT GCG GCT GC 3'

2.3 Protein Biochemistry

2.3.1 Bradford Assay for protein concentration determination

Protein concentration was determined using the Bradford method using a Bio Photometer (Eppendorf). A calibration curve was made using 0, 4, 8, 16 and 20 µg of 1 mg/ml BSA in 1 ml of 1:4 diluted protein assay reagent (Bio-Rad). A known amount (1-5 µl) of the protein samples were then added to the reagent and the concentration determined.

2.3.2 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Protein samples were resolved on 10-12.5% polyacrylamide separating gels (375 mM Tris-HCl pH 8.8, 10-12.5% (w/v) acrylamide (Protogel, National diagnostics), 0.1% (w/v) SDS, polymerised with 0.04% (w/v) ammonium persulphate (APS) and 0.08% (v/v) TEMED) according to method published (Laemmli, 1970). To ensure a flat surface in order for the samples to enter the gel uniformly, a layer of isopropanol was added to the surface of the gel before it set. A stacking gel (125 mM Tris-HCl pH 6.8, 4.5% (w/v) acrylamide, 0.1% (w/v) SDS, polymerised with 0.05% (w/v) APS and 0.125% (v/v) TEMED) was then poured on top of the resolving gel and the comb inserted. The gels were run in running buffer (25 mM Tris-

HCl, 192 mM glycine, 0.1% (w/v) SDS). With small gels (Mini Protean® 3 cell, Bio-Rad) run at 150 V for 1-2 h and large gels (the Sturdier system, Amersham Biosciences) run at 53 V overnight.

Protein samples were mixed 1:1 with 2X Laemmli sample loading buffer (Bio-Rad or 125 mM Tris-HCl pH 6.8, 2% (w/v) SDS, 9% (v/v) glycerol, 0.01% (w/v) bromophenol blue) supplemented with 5% (v/v) β-mercaptoethanol, and heated at 100°C for 5 min before loading onto the gel along with a protein marker ladder (Bio-Rad Precision Plus Protein Dual Colour or NEB Prestained protein marker, Broad range (7-175 kDa)). Protein bands were stained using Bio-Safe Coomassie (Bio-Rad).

2.3.3 Western blot analysis

Resolved proteins were transferred to Biotrace® NT nitrocellulose membrane (Pall Life Sciences) using a Trans-Blot® Cell transfer tank (Bio-Rad) in transfer buffer (25 mM Tris-HCl, 192 mM glycine, 20% (v/v) methanol) at 350 mA for 3.5 h or 15 V overnight. The total protein transfer was visualised using ponceau stain (1% (w/v) Ponceau Red, 3% (w/v) trichloroacetic acid), which was removed with repeated washing with 0.2% (v/v) Tween-20 in PBS. The membrane was blocked in 2% (w/v) dried skimmed milk in PBS for 1 h at room temperature or overnight at 4°C, for phospho-specific antibodies the membrane was blocked in 5% BSA in Tris-buffered saline (20 mM Tris-HCl pH 7.6, 138 mM NaCl) (TBS) for 1 h at room temperature. Antibodies were diluted in 2% (w/v) dried skimmed milk in PBS or 1% BSA in TBS (dilutions shown in table 6). The primary antibody was incubated for 1 h at room temperature or overnight at 4°C and the secondary antibody for 1 h at room temperature. The membranes were washed with 0.2% Tween (v/v) in PBS for 1 h with 4 changes of buffer after incubation with the primary and secondary antibodies. Membranes were developed by enhanced chemiluminescence (Amersham™ ECL™ Western blotting detection reagents, GE Healthcare) and signal detected using autoradiography film.

Table 6 Antibody dilutions used in Western blotting

Antibody Name	Manufacturer	Species	Dilution
Human anti SRPK1 (Clone 12)	BD Transduction Labs	Mouse	1:1000
Anti-Splicing Factor 2 (SF2/ASF) (Clone 103)	Zymed Labs	Mouse	1:1000
HA.II Monoclonal Antibody (Anti-HA) (Clone 16B12)	Covance	Mouse	1:1000
Anti-FLAG M2 Monoclonal Antibody	Sigma	Mouse	1:5000
Anti-FLAG Polyclonal	Sigma	Rabbit	1:1000
Monoclonal anti-polyHistidine	Sigma	Mouse	1:1000
Anti-SR Proteins (1H4) (Clone 1H4G7)	Zymed Labs	Mouse	1:1000
Anti-SRPK2 (Clone 23)	BD Transduction Labs	Mouse	1:1000
Anti-SFRS4 (SRp75)	Sigma	Rabbit	1:1000
Anti-SFRS3 (SRp20)	Sigma	Rabbit	1:1000
Splicing Factor SC-35	Sigma	Mouse	1:1000
Anti-GST	Pharmacia Biotech	Goat	1:1000
4.37 (Anti-HPV1 E4)	In house (Doorbar et al., 1988)	Mouse	1:150
GAPDH	Santa Cruz	Mouse	1:1000
Lamin A/C	Santa Cruz	Mouse	1:1000
mAb104 (Anti-phospho RS domain)	A gift from Sheila Graham, University of Glasgow	Mouse	Neat
Anti-mouse IgG (whole molecule)-Peroxidase	Sigma	Goat	1:3000
Polyclonal Swine Anti-Rabbit Immunoglobulins/HRP	DAKO	Swine	1:3000
Anti-goat IgG (whole molecule)-Peroxidase	Sigma	Rabbit	1:3000

2.3.4 Densitometry

Protein levels from Western blot films were analysed using a Scanning Densitometer GS-800 (Bio-Rad) and quantified using the Quantity One 4.6.7 software (Bio-Rad)

2.3.5 *In vitro* kinase assay

In vitro kinase assays were carried out following the previously described method of Papoutsopoulou et al. (1999). The reactions were performed in a buffer containing 25 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$ and 20 μM ATP. Between 1 and 2 μg of recombinant His-SRPK1 along with 5 μCi [γ-^{32}P] ATP (Perkin Elmer) was added to 1-25 μg of substrate protein (GST-SR, -E2 and -E1^E4 proteins). For inhibition studies 1-10 μg of the inhibitory protein (GST and GST-E1^E4 proteins) or 10 or 50 μM of SRPIN340 (N4-[2-Piperidino-5-(trifluoromethyl)phenyl]isonicotinamide, Maybridge) in DMSO was added to the reaction. The reaction was incubated for 10-30 min at 30°C and the reactions stopped by the addition of 2X Laemmli loading buffer. The samples were resolved by SDS-PAGE gel electrophoresis on gels containing 12.5% acrylamide and stained using Bio-Safe Coomassie (Bio-Rad). The stained gels were dried down and exposed to autoradiography film. Band intensities were quantified using a phosphorimager; gels were exposed to a Molecular Dynamics storage phosphor screen, then analysed using a Typhoon 8600 Variable mode imager and Molecular Dynamics ImageQuant Version 5.1.

The amount of substrate protein and the length of incubation were optimised experimentally.

2.3.6 Trichloroacetic acid (TCA) protein precipitation

To quantify the amount of phosphate incorporated into the substrate proteins, trichloroacetic acid (TCA) protein precipitation was used to remove the unincorporated [γ-^{32}P] ATP from the sample. The kinase assay was carried out as described in section 2.3.4, but at the end of the reaction the sample was split into two aliquots. One aliquot was prepared as described for

analysis by SDS-PAGE (section 2.3.4). To the other aliquot, 20% (w/v) TCA and 0.2 mg bovine serum albumin (BSA) were added; the sample was vortexed and spun at 16100 x g for 5 min. The pellet was washed with 10% (w/v) TCA three times and then re-suspended in acetone and centrifuged at 16100 x g for 5 min. The pellet was allowed to air dry and re-suspended in scintillation cocktail (OptiPhase 'HiSafe' 3, Perkin Elmer). The amount of phosphate incorporated was determined using a scintillation counter (Packard Tri-Carb liquid scintillation counter) and calculated using the following equations

$$\text{pmol incorporated} = \frac{\text{Scintillation count}}{\text{Count 1 pmol gives}}$$

$$\text{Count 1 pmol gives} = \frac{\text{1 µCi count}}{[^{32}\text{P}] \text{ / specific activity}}$$

1 µCi count determined by measuring the scintillation count of 1 µCi of radioisotope

$[^{32}P]$ is concentration of radioisotope determined using Perkin Elmer decay calculator

Specific activity of batch of radioisotope determined using Perkin Elmer decay calculator

http://www.perkinelmer.co.uk/tools/RadCalculator

2.3.7 *In vitro* co-precipitations

2.3.7.1 GST "pull down" from lysates

Co-precipitations were carried out using recombinantly expressed GST-proteins and His-SRPK1, in *E. coli* lysis buffer (0.5% v/v Triton® X-100 and Complete™ protease inhibitors (Roche) in PBS). The proteins were recombinantly expressed as section 2.1.5, and the GST-tagged proteins were bound to glutathione agarose (Sigma) (washed in *E. coli* lysis buffer) for 1 h at 4°C. After washing in *E. coli* lysis buffer the glutathione agarose bound proteins were mixed with the His-SRPK1 lysate for 1 h at 4°C. Following washing, the beads were re-suspended in 2X Laemmli loading buffer, resolved by SDS-PAGE and analysed by Western blotting.

2.3.7.2 GST "pull-down" using purified proteins

Co-precipitations were carried out using recombinantly expressed GST proteins and His-SRPK1, in a buffer containing 25 mM Tris-HCl pH 7.5, 200 mM NaCl, 10 mM $MgCl_2$ and 20 μM ATP. The GST-tagged proteins (10-50 μg) were purified as described in section 2.1.5.3 and then bound to glutathione agarose beads, washed and incubated with His-tagged SRPK1 for 2 h at 4°C. To "pull-downs" containing GST-ASF/SF2; HPV1 E1^E4 (10 μg), SRPIN340 (5 μM) and Staurosporine (5 μM) were added. Following washing the beads were re-suspended in 2X Laemmli loading buffer, resolved by SDS-PAGE and analysed by Western blotting.

2.3.8 Biacore® Surface Plasmon Resonance

Surface Plasmon resonance was performed on a Biacore® 3000 and analysed using BIAevaluation version 3.1.

2.3.8.1 His-SRPK1 on a NTA chip

Two micrograms of purified His-tagged SRPK1 was bound to a nickel primer nitrilotriacetic acid (NTA) chip (GE Healthcare) by nickel chelation in HBS-EP (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA, 0.005% Surfactant P20) using a flow rate of 5 μl/min. A His-tagged control protein was obtained from Ben Wilcox (School of Cancer Sciences), and 2 μg was bound. The quick-inject programme was used to flow 1 μg of GST-ASF/SF2 and 1 μg of GST-HPV1 E1^E4 over the relevant chambers. The NTA surface was regenerated using EDTA regeneration solution (GE Healthcare).

2.3.8.2 GST-chip

A goat anti-GST antibody (GE Healthcare) was diluted to 30 μg/ml in immobilisation buffer (10 mM sodium acetate pH 5.0), and then following EDC/NHS (1-Ethyl-3-[3-dimethylaminopropyl]carbodiimide hydrochloride/*N*-hydroxysulfosuccinimide) activation of

the CM5 chip, the goat anti-GST antibody was amine-coupled onto the activated surface. The surface was then deactivated with 1 M ethanolamine HCl pH 8.5 and regenerated with 10 mM glycine-HCl pH 2.2. Once the anti-GST surface was ready, 3.5 µg of GST, 1.2 µg of GST-ASF/SF2 or 1.4 µg of GST-HPV1 E1^E4 were bound to the surface in 25 mM Tris-HCl pH 7.5, 200 mM NaCl with a flow rate of 5 µl/min. His-SRPK1 was then flowed over the chambers in 25 mM Tris-HCl pH 7.5, 200 mM NaCl, or in 25 mM Tris-HCl pH 7.5, 200 mM NaCl, 2 mM $MgCl_2$ and 1 mM ATP, for 5 min and 15 min at 5 µl/min.

2.3.8.3 Amine-coupling of the proteins directly to the CM5 chip

To determine the binding conditions, the proteins were flowed over the chamber until a characteristic binding pattern was observed and this was performed at a pH lower than the pI of the protein. GST-ASF/SF2 was diluted to 50 µg/ml in 10 mM sodium borate pH 7.9, 10 mM sodium acetate pH 5.5 or 10 mM sodium acetate pH 5.0 and passed through the chamber until binding was observed at pH 5.0. His-SRPK1 was diluted to 10 µg/ml in 10 mM sodium acetate pH 5.0, 10 mM sodium acetate pH 4.5 or 10 mM sodium acetate pH 4.0 and passed through a different chamber until binding was observed at pH 4.0. His-SRPK1 was also diluted in the presence of 1 mM ATP and 2 mM $MgCl_2$ in 10 mM sodium acetate pH 5.0 or 10 mM sodium acetate pH 5.5 and passed through a chamber until binding was observed at pH 5.0. The chip was then regenerated with 10 mM glycine-HCl pH 2.2 before further use.

The CM5 chip was activated by EDC/NHS and 50 µg/ml GST-ASF/SF2 in 10 mM sodium acetate pH 5.0 was amine coupled onto the chip, followed by 1 M ethanolamine deactivation. For His-SRPK1, the protein was diluted to 10 µg/ml in 10 mM sodium acetate pH 4.0, or 10 mM sodium acetate pH 5.0, 2 mM $MgCl_2$, 1 mM ATP for the amine coupling step.

GST-ASF/SF2 (1 µg) was passed through the chambers in 10 mM HEPES pH 7.5, 150 mM NaCl, followed by 1 µg of His-SRPK1, and then by 1, 2 and 5 µg of GST-HPV1 E1^E4. For

the His-SRPK1 bound in the presence of $MgCl_2$ and ATP, 5 µg of GST-ASF/SF2 was used, and 5 and 20 µg of GST-HPV1 E1^E4 were passed through the chamber. The flow rate was 5 µl/min and the binding interactions carried out at 4°C.

2.3.9 Statistical analysis

The significance of the results was calculated using a two-tailed Student's t-test, when the raw data was compared or a two-tailed Welch's t-test when the data was normalised. A result was considered significant when p was less than 0.05.

2.4 Mammalian cell culture

2.4.1 Cell lines and growth media

Table 7 Cell lines used in this study

Cell line	Cell type
HeLa	Cervical adenocarcinoma derived cell line, contains integrated HPV18 genomes
COS-1	Simian virus 40 (SV40)-immortalised African green monkey epithelial cell line
SVJD	SV40-immortalised human epidermal keratinocytes (Roberts et al., 1993)
HEK 293	Adenovirus 5 transformed human embryonic kidney cell line
H1299	Non-small lung carcinoma derived cell line

The cell lines COS-1, HeLa, HEK 293 and H1299 were all cultured in Dulbecco's modified Eagle medium (DMEM) HEPES modification (Sigma-Aldrich) supplemented with 4 mM glutamine and 10% (v/v) foetal calf serum. SV40-immortalized human epidermal keratinocytes (SVJD) were grown in Minimum Essential Medium Jocklik modification (Sigma-Aldrich), supplemented with 4 mM glutamine and 10% (v/v) foetal calf serum. Cells were kept at 37°C, 5.0% CO_2 in an incubator and were routinely grown in 75 cm^3 flasks (Iwaki) or 10 cm tissue culture dishes (Iwaki).

2.4.2 Passaging of cells

The cells were harvested once they reached about 80% confluency, the media was removed and the cells washed with PBS. The cells were removed from the dish using 0.05% trypsin EDTA (Gibco-Invitrogen) with gentle tapping over a 5 min incubation at 37°C, the adhesion of the cells was monitored frequently under a microscope. The trypsin was neutralised by addition of the appropriate growth media. The cells were counted in a haemocytometer and seeded out at the required density for experimentation.

2.4.3 Freezing and thawing of cells

Cells were grown to 70-80% confluency before freezing and were harvested as described in section 2.3.2. Following counting the cells were re-suspended to a concentration of 2 x 10^6 cells/ml in 85% (v/v) FBS, 15% DMSO. Cells were transferred in 1 ml aliquots to Nunc cryovials (Nalgene-Nunc) and stored at -80°C in Nalgene® Mr Frosty, a slow-cool freezing chamber with isopropanol (Nalgene Nunc) for 16 h before being transferred to liquid nitrogen for long term storage.

To defrost the cells, the cryovials were removed from liquid nitrogen and thawed at 37°C for 1 min. The cells were transferred using a transfer pipette into a 22.5 ml universal tube and 10 ml of growth media added drop-wise, before centrifugation at 538 x g for 5 min. The cells were then re-suspended in 10 ml of the appropriate media and grown in the appropriate tissue culture vessel.

2.4.4 Transfection in 6 well plates

Cells were seeded out at 3 x 10^5 cells per well of a 6-well plate (Iwaki) and grown to 80% confluence. The cells were transfected with 3 µg DNA and Lipofectamine™ 2000 (2 µl/µg DNA transfected) (Invitrogen) reagent or Lipofectamine™ LTX (1.5 µl/µg DNA) (Invitrogen) with PLUS reagent (1 µl/µg DNA) in 500 µl Opti-mem® (Invitrogen) according

to the manufacturer's instructions. The cells were flooded after 4 h with the appropriate growth media and harvested after 48 h with various methods.

2.4.5 Harvesting of cells and protein extraction

Cells were harvested from dishes using trypsin treatment. Following neutralisation of the trypsin, the cells were pelleted by centrifugation at 538 x g for 5 min before being washed in PBS and re-pelleted twice. The pellet was then re-suspended in UTB (8 M Urea, 25 mM Tris-HCl pH 8.0, 0.15 M β-mercaptoethanol) and incubated on ice for 20 min. The lysate was sonicated for 20 s at 2 watts using a microson ultrasonic cell disrupter (Misonix) and the insoluble material removed by high speed centrifugation at 16100 x g for 10 min. The proteins were analysed using a Bradford assay (2.3.1) before resolution by SDS PAGE (2.3.2).

2.4.6 Immunofluorescence

2.4.6.1 Transfection on slides

Cells were seeded out at 3-5 x 10^4 cells per spot on frosted multispot glass slides (Hendley-Essex) and grown to 80% confluence. The cells were transfected with 0.5-1 µg DNA per spot and Lipofectamine™ 2000 (2 µl/µg DNA) (Invitrogen) or Lipofectamine™ LTX (1.5 µl/µg DNA) (Invitrogen) with PLUS reagent (1 µl/µg DNA) in 100 µl Opti-mem®, according to the manufacturer's instructions. After 4 h the slides were flooded with 15 ml of the appropriate growth media.

2.4.6.2 Fixation of cells

Cells were fixed 48 h after transfection; various fixation methods were used following a wash in ice-cold saline. (a) Fixation with 4% paraformaldehyde in PBS for 5 min followed by two washes in PBS and permeabilisation in acetone at -20°C for 10 min. (b) A 30 s pre-wash in 0.1% Triton® X-100 (Sigma) in PBS followed by fixation with 4% paraformaldehyde for 5

min and permeabilisation in acetone (-20°C) for 10 min. For both fixation methods, slides were then air-dried and stored wrapped in aluminium foil at -20 C.

2.4.6.3 Immune-labelling

Slides were immersed in blocking buffer (20% (v/v) HINGS, 0.1% (w/v) BSA in PBS) at room temperature for 30 min, the primary antibody, diluted in blocking buffer (dilutions shown in table 8) was applied and incubated at 37°C in a humidity chamber for 90 min. After washing twice in PBS for 15 min, the secondary AlexaFluor® dye conjugate (Molecular Probes) were diluted 1:500 in blocking buffer and incubated for 45 min at 37°C. The slides were washed twice in PBS for 15 min and then for a further 10 min with 4',6-diamidion-2-phenylindole (DAPI, a nuclear stain) (Sigma-Aldrich) added to the wash buffer. Slides were mounted on coverslips using a mounting media containing 2.5 % (w/v) 1,4-Diazabicyclo[2.2.2]octane (DABCO, an anti-fade reagent) (Sigma-Aldrich) in 90% (v/v) glycerol and PBS, pH 6.8.

2.4.6.4 Microscopy

Epiflourescence microscopy was performed on a Nikon Eclipse E600 microscope and images were captured using Nikon DXM1200F digital camera. Confocal microscopy was performed on a LSM 510 META confocal microscope (Zeiss) using multi-track imaging. Images were assembled in Adobe Photoshop® CS2.

Table 8 Dilutions for antibodies used in immunofluorescence

Antibody Name	Manufacturer	Species	Dilution
Human anti-SRPK1 (Clone 12)	BD Transduction Labs	Mouse IgG1	1:500
Anti-Splicing Factor 2 (SF2/ASF) (Clone 103)	Zymed Labs	Mouse IgG1	1:100
HA.II Monoclonal Antibody anti-HA tag (Clone 16B12)	Covance	Mouse IgG1	1:1000
Anti-FLAG M2 Monoclonal Antibody	Sigma-Aldrich	Mouse IgG1	1:200
Anti-FLAG Polyclonal	Sigma-Aldrich	Rabbit	1:200
Anti-SR Proteins (1H4) (Clone 1H4G7)	Zymed Labs	Mouse IgG1	1:100
Anti-β-Galactosidase, purified monoclonal	Promega	Mouse IgG2a	1:2000
Splicing Factor SC-35	Sigma-Aldrich	Mouse IgG1	1:2000
4.37 (Anti-HPV1 E4)	In house (Doorbar et al., 1988)	Mouse IgG2a	1:200
R424 (Anti-HPV18 E4)	In house (Wilson et al., 2007)	Rabbit	1:1000
mAb104 (Anti-phospho RS domain)	A gift from Sheila Graham, University of Glasgow	Mouse IgM	Neat
AlexaFluor® 488 Goat anti-mouse IgG	Molecular Probes (Invitrogen)		1:500
AlexaFluor® 488 Goat anti-mouse IgG1	Molecular Probes (Invitrogen)		1:500
AlexaFluor® 594 Goat anti-mouse IgG2a	Molecular Probes (Invitrogen)		1:500
AlexaFluor® 594 Goat anti-rabbit IgG	Molecular Probes (Invitrogen)		1:500

2.4.7 Immunoprecipitation

Cells grown in six-well plates, transfected with the appropriate plasmids were washed with ice cold PBS and lysed in NP-40 buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% (v/v) NP40, Complete™ protease inhibitors (Roche)) or RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% (w/v) SDS, 0.5% (w/v) sodium deoxycholate, 1% (v/v) Triton® X-100, Complete™ protease inhibitors) cells were harvested using a cell scraper. Following incubation on ice for 10 min the lysate was cleared by centrifugation at 16100 x g for 10 min and an aliquot taken as a control sample. The lysate was mixed with 0.5 - 2 µg of the relevant antibody for 1 h at 4°C and then 30 µl of a 50% slurry of protein G/A Sepharose was added and rotated at 4°C for 1 h. The beads were washed 5 times with lysis buffer and re-suspended in 30 µl Laemmli buffer and boiled for 5 min.

2.4.8 *In vivo* kinase assay

Cells grown in six-well plates, transfected with the appropriate plasmids were washed with ice cold PBS and lysed in kinase lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% (v/v) Triton® X-100, 1 mM PMSF, 5 mM DTT) cells were harvested using a cell scraper. Following incubation on ice for 30 min the lysate was cleared by centrifugation at 16100 x g for 30 min. The supernatant was analysed by Bradford assay (2.3.1) and 1 mg of lysate was mixed with 1 µg of M2 (mouse anti-Flag) antibody or 200 µl of 4.37 (mouse anti-HPV1 E4) for 30 min at 4°C, then 50 µl of 50% slurry protein G Sepharose was added and rotated at 4°C for 1 h. Beads were washed 2 times with kinase lysis buffer, then 2 times with 50 mM Tris-HCl pH 7.5, 150 mM NaCl. The beads were resuspended in 50 µl of 50 mM Tris-HCl pH 7.5, 150 mM NaCl and had 1 µg GST-ASF/SF2 added, followed by 5 µCi [γ-^{32}P] ATP (Perkin Elmer) in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$ and 20 µM ATP.

Following 30 min at 30°C, the reaction was stopped by the addition of 50 µl 2 X Laemmli loading buffer and resolved by SDS PAGE. Analysis was carried out as with 2.3.5.

2.5 *In vitro* splicing assay

2.5.1 Minigene constructs

The minigene constructs were generously donated from other laboratories: The pCRXH construct which contains the 3' region of exon 3, exon 4 and 5' region of exon 5 of X16 (Mouse SRp20) was donated by Peter Nielsen, Max Planck Institute for Immunobiology, Germany. The pMTE1A construct which contains the Adenovirus E1A coding sequence was donated by Adrian Krainer, Cold Spring Harbor Laboratory, USA. The primers used to amplify the minigenes are shown in table 9.

Table 9 Minigene primers

Minigene	Primer Name	Sequence
pCRXH (X16)	X16 Exon 3 F	GAG ATG ATT ACC GCA GGA GGA GTC
	X16 Exon 5 R	CTT CTC CTT CTT GGG GAT
pMTE1A (E1A)	E1A Exon 1 F	GTT TTC TCC TCC GAG CCG CTC CGA
	E1A Exon 2 R	CTC AGG CTC AGG TTC AGA CAC AGG

2.5.2 Transfection, harvesting of DNA and reverse transcription

Cells were transfected with 2 µg of the minigene plasmid and 1 µg of each of the possible effector plasmids, with 1.5 µl Lipofectamine™ LTX per µg of DNA, and 1 µl PLUS reagent per µg of DNA. After 48 h the cells were washed with PBS and harvested using the Qiagen RNeasy® mini kit (Qiagen) according to the manufacturer's instructions. Total RNA was eluted into 50 µl of RNase-free water (Qiagen) and the RNA concentration measured using a Nanodrop™ 2000 spectrophotometer (Thermo Scientific). One microgram of RNA was

reverse transcribed using the Qiagen QuantiTect® Reverse Transcription kit (Qiagen) according to the manufacturer's instructions (genomic DNA elimination: 2 min at 42°C, reverse transcription: 15 min at 42°C, then 3 min at 95°C) using a Thermal cycler (2720 or 2700, Applied Biosystems).

2.5.3 Radioactive PCR

Two microlitres of the transcribed DNA was used as a template, using the primers in table 9, Taq polymerase, and 1 μCi of [α-^{32}P] dCTP (Perkin Elmer) to amplify the minigene construct in a total volume of 50 μl. The X16 minigene was amplified with an initial degradation at 94°C for 2 min then 30 cycles of 94°C for 30 s, 55°C for 1 min and 72°C for 1 min followed by 72°C for 20 min, before cooling to 4°C. The E1A minigene was amplified with an initial degradation at 94°C for 2 min, then 20 cycles of 94°C for 30 s, 65°C for 1 min, decreasing by 0.5°C each cycle, and 72°C for 2 min, then 10 cycles of 94°C for 30 s, 55°C for 1 min and 72°C for 2 min, followed by 72°C for 20 min before cooling to 4°C.

2.5.4 Native polyacrylamide electrophoresis (PAGE)

Ten microlitres of the PCR reaction was mixed with 5 μl FICOLL buffer (15% (w/v) FICOLL, 0.25% Bromophenol blue, 0.25% Orange G) and run on a native PAGE gel (10% (v/v) Protogel, 90 mM Tris, 90 mM boric Acid, 2 mM EDTA, 0.07% (w/v) APS, 0.08% TEMED) in TBE (90 mM Tris, 90 mM boric Acid, 2 mM EDTA) for 1 h at 125 V, a 1 kb plus DNA ladder (Invitrogen) was run on each gel for size determination. The gels were washed in ethidium bromide solution (0.25 mg/ml) for 15 min and visualised on Gene Flash UV light box (Syngene Bio Imaging). The gel was then dried down and exposed to autoradiography film. Band intensities were quantified using a phosphorimager as described in section 2.3.5.

2.6 Human Foreskin Keratinocytes (HFKs)

2.6.1 Maintenance of untransfected HFKs

Primary human foreskin keratinocytes (HFKs) isolated "in house" from neonatal foreskin by Sally Roberts were grown in serum-free keratinocyte growth medium (Gibco-Invitrogen), with fresh media every 2 days. Cells were grown in an incubator at 37°C and 5.0% CO_2 and grown to no more than 80% confluency. Cells were passaged by removing the media and washing with PBS, 2 ml of 0.05% trypsin-EDTA was then added and after 5 min at 37°C the trypsin was inactivated by 4 ml of trypsin neutralising solution (TNS, 0.25 mg/ml soybean trypsin inhibitor (Invitrogen) in PBS (magnesium and calcium free) pH 7.2). The cells were transferred to a tube and the plate washed with 6 ml serum-free media, then centrifuged at 538 x g for 5 min. The cells were re-suspended in 10 ml serum-free media and counted before being seeded out at 2 x 10^5 cells in a 10 cm tissue culture dish.

2.6.2 Maintenance of HFKs containing HPV18 genomes

HFKs containing HPV18 genomes were grown on irradiated J2-3T3 mouse fibroblasts in E-medium.

2.6.2.1 E-medium

To make 2 litres of E-medium the following components were mixed: 1200 ml DMEM HEPES modification, 640 ml Ham's F-12, 20 ml of 100 X cocktail (described below), 10 ml of 100 X Penicillin/Streptomycin (PAA Laboratories), 100 ml Foetal Bovine Serum (FBS, 5% (v/v)), 2 ml 1000 X Cholera toxin (ICN Biomedical) and 2 ml 1000 X Hydrocortisone (Sigma-Aldrich).

To make 200 ml of 100 X cocktail the following components were mixed: 20 ml 0.18 M Adenine (Sigma-Aldrich, 0.486 g in 15 ml H_2O with addition of 10 M HCl until dissolved, then addition of 5 ml H_2O), 20 ml 5 mg/ml Insulin (Sigma-Aldrich, 0.1 g added to 20 ml 0.1

M HCl), 20 ml 5 mg/ml transferin (Sigma-Aldrich, 0.1 g added to 20 ml PBS), 20 ml 2 x 10^{-8} M 3,3',5-Triiodo-L-thyronine (T_3) in PBS (Sigma-Aldrich, 13.6 g added to 100 ml 0.02 M NaOH to make 2 x 10^{-4} M T_3 further diluted in PBS to make 2 x 10^{-8} M). The 100 X cocktail mix was filter sterilised and frozen prior to use.

Following mixing, E-medium was filter sterilised and stored at 4°C in the dark. Immediately prior to use the media was supplemented with 4 mM L-glutamine and 5 ng/ml epidermal growth factor (EGF, BD Biosciences). E-medium was prepared by Craig Delury.

2.6.2.2 Maintenance and irradiation of J2-3T3 mouse fibroblasts

J2-3T3 mouse fibroblasts were routinely grown in Dulbecco's Modified Eagle Medium (DMEM) HEPES modification (Sigma-Aldrich) supplemented with 10% (v/v) new born bovine serum and 4 mM L-glutamine, in 10 cm tissue culture dishes (Iwaki). When the J2-3T3 cells were required as feeder cells, they were harvested and re-suspended in E-medium at 2 x 10^6 cells/ml in a 50 ml tube. Cells were irradiated using 50 Gray of a caesium-137 source. Irradiated J2-3T3s were plated out at 2 x 10^6 cells per 10 cm dish in E-medium and allowed to settle for at least 2 h before addition of HFKs. If not required within 24 h of irradiation, J2-3T3 cells were stored in E-medium at 4°C for up to 3 days.

2.6.2.3 Maintenance of transfected HFKs containing HPV18 genomes

Cell lines established from HFKs transfected with HPV18 genomes were generated by Sally Roberts and Gillian Knight. The cells were routinely grown on a feeder layer of irradiated J2-3T3 fibroblasts in E-medium at 37°C and 5.0% CO_2. The media was replaced every 2 days and cells not allowed to reach more than 80% confluency. The HPV18 genome containing cells were passaged by first removing the J2-3T3 feeder layer with 2 ml of warm 0.5 mM EDTA in PBS. Once the feeder layers began to detach, a transfer pipette was used to aspirate the EDTA and remove the J2-3T3 cells. The cells were washed with PBS and then removed

using 2 ml 0.05% Trypsin EDTA and incubation at 37°C for 5-10 min. The trypsin was neutralised with 8 ml E-medium and a further 5 ml used to wash the dish. The cells were centrifuged at 538 x g for 5 min and re-suspended in 10 ml E-medium. The cells were plated out at 2 x 10^5 cells on top of the irradiated J2-3T3 cells.

2.6.3 SRPIN340 toxicology study

Keratinocytes containing HPV18 genomes were grown on J2-3T3s seeded out in 6 well plates. Once cells reached 50% confluency SRPIN340 (in DMSO) was added to the cells, using SRPIN340 concentrations of 0, 1, 5, 10, 25 and 40 μM. Equivalent volumes of DMSO were added to the control wells. After 48 h cells were harvested as described in 2.6.2.3 except the original media was kept and added back to the trypsinised cells. After centrifugation at 538 x g for 5 min, the cells were re-suspended in 1 ml of fresh media and a 250 μl aliquot taken. To the aliquot of cells, 0.4% trypan blue stain (Sigma-Aldrich) was added and incubated at room temperature for 5 min. The total cells and the non-viable cells were counted using a haemocytometer. The remaining cells were used for DNA analysis.

2.6.4 Southern blot analysis

2.6.4.1 Isolation of DNA

The pelleted cells were washed in PBS and then re-suspended in 3 ml lysis buffer (400 mM NaCl, 10 mM Tris pH 7.4, 10 mM EDTA). RNase A was added to a final concentration of 400 μg/ml, the cell lysates were vortexed and put at 37°C for 30 min. Proteinase K (Roche) was then added to a final concentration of 50 μg/ml, and SDS to 0.2% (w/v), before vortexing and incubation at 37°C overnight. The cellular DNA was sheared by passing the lysate through a 20-gauge needle 10 times. Phenol-chloroform extraction was performed by the addition of 6 ml phenol/chloroform/isoamyl alcohol (25:24:1) (Sigma-Aldrich) before inversion of the tube 8 times and centrifugation at 537 x g for 5 min. The upper aqueous layer

was retained and the extraction repeated, before extraction with 6 ml chloroform/isoamyl alcohol (24:1). The DNA was then precipitated by the addition of 2 volumes of ethanol and 1/10th volume of 3 M sodium acetate, followed by incubation at -80°C for 1 h, or -20°C overnight. Following precipitation the samples were centrifuged at 537 x g for 30 min and the DNA pellet was washed twice with 70% ethanol. The DNA was allowed to air dry for a few minutes before re-suspension in 20-100 µl TE buffer and incubated at 37°C for 5 min. The DNA yield was determined using the Nanodrop^{TM} spectrophotometer (Thermo Scientific) and samples stored at 4°C.

2.6.4.2 Digestion of DNA for Southern blot analysis

The extracted DNA was subjected to digestion with BglII (Biolabs) and DpnI (Roche), with 5 µg of DNA being digested overnight at 37°C; this digests the expression plasmid and digests host genetic material without cutting the HPV18 genome. The DNA was resolved on a 0.8% (w/v) agarose gel in TBE (90 mM Tris-HCl, 90 mM boric Acid, 2 mM EDTA) buffer overnight at 50 V in a Fisherbrand HU20 Horizontal electrophoresis tank. To enable quantification of genome copy numbers, copy number standards were generated by digestion of pGEMII-HPV18 with *EcoRI* to release the HPV18 genome from the vector. The equivalent of 5 and 50 genome copies per cell were used as copy number standards on the gel.

2.6.4.3 Transfer of DNA from agarose gel to nylon membrane

Once the agarose gel had run, the loading and digestion of the DNA was checked using a Gene Flash UV light box (Syngene Bio Imaging). The gel was washed twice in 250 mM HCl for 30 min at room temperature with gentle agitation, before washing twice in 400 mM NaOH for 30 min. To transfer the DNA to the nylon membrane, a tray was set up with 1 L of 400 mM NaOH with a glass plate resting across the tray. A 24 x 33 cm single sheet of Whatman^{TM} 3 mm paper was soaked in 400 mM NaOH and laid across the glass plate with

both ends submerged in the NaOH. The bubbles were removed by rolling a pipette across and 3 more layers of 24 x 33 cm Whatman™ paper were placed on top in the same manner. The agarose gel was placed on top of the paper layers with the loading wells facing downwards and the bubbles removed with a pipette. A 20 x 22.5 cm sheet of Gene Screen Plus nylon membrane (Perkin Elmer) was soaked in 400 mM NaOH and placed on top of the gel, followed by 4 layers of 21 x 23.5 cm Whatman™ paper, placed on top one at a time with care. Two stacks of absorbent paper towels were placed on top, covering the paper, ensuring there were no gaps between the stacks. Another glass plate was placed on top of the towels and a weight centred on top. Saran wrap was then placed between the top and bottom layers of Whatman™ paper to ensure they were not in contact and the transfer was left overnight. The following day the position of the wells was marked onto the membrane and the DNA crosslinked to the membrane using a UV crosslinker (Stratalinker, Stratagene) on auto-crosslink mode. The membrane was soaked in SSC (600 mM sodium citrate, 300 mM NaCl pH 7.0) and stored at -20°C wrapped in saran wrap.

2.6.4.4 Preparation of radiolabelled DNA probe

To make the DNA probe the pGEMII-HPV18 plasmid was linearised by digestion with EcoRI and resolved on a 1% agarose gel. The ~ 8 kb fragment HPV18 genome fragment was excised and purified by using the QIAquick Gel Extraction kit (Qiagen). Fifty ng of the purified HPV18 linearised genome was diluted into 45 µl TE buffer and denatured by boiling at 95°C for 5 min followed by incubation on ice for 5 min. The denatured DNA was then bound to Amersham™ Ready-To-Go™ DNA labelling beads (GE Healthcare) and 50 µCi [α-^{32}P] dCTP (Perkin Elmer) added. Following incubation at 37°C for 30 min, the radiolabelled probe was loaded onto a pre-prepared Illustra Probe Quant G-50 microcolumn (Amersham) and centrifuged at 801 x g for 2 min to remove unincorporated nucleotide.

2.6.4.5 Hybridisation of radiolabelled probe to immobilised DNA

To make the hybridisation buffer: A 2 X hybridisation solution (20% (w/v) Dextran sulphate (USB), 10 X Denhardts (0.2% (w/v) Ficoll 400, 0.2% (w/v) Polyvinypyrolilone, 0.2% (w/v) BSA fraction V (Sigma-Aldrich)) and 5 X SSC (1.5 M sodium citrate, 750 mM NaCl pH 7.0)) was diluted 1:1 with deionised formamide and SDS (1% final concentration (w/v)) added. Salmon sperm DNA (200 µl of 10 mg/ml, Invitrogen) was denatured by boiling at 100°C for 10 min followed by cooling in an ice/water slurry for 5 min and subsequently added to 10 ml of the hybridisation buffer to make the pre-hybridisation buffer. The nylon membrane on which the DNA was immobilised was rolled with a sheet of mesh and placed into a hybridisation canister with the DNA side facing inwards. The pre-hybridisation buffer was added to the canister and put in the hybridisation oven with rotation at 42°C for 1 h. Following the pre-hybridisation, hybridisation buffer containing the radiolabelled DNA probe was prepared by addition of the probe to 200 µl of 10 mg/ml salmon sperm DNA, followed by boiling at 100°C for 10 min, cooling in ice/water for 5 min and subsequent addition to 10 ml of hybridisation buffer. The pre-hybridisation buffer was poured out of the canister and the hybridisation buffer containing the probe added the canister was incubated at 42°C with rotation overnight.

2.6.4.6 Stringency washes

Following removal from the hybridisation canister, the membrane was rinsed in buffer 1 (600 mM sodium citrate, 300 mM NaCl, pH 7.0, 0.1% (w/v) SDS), using a sponge to wipe the surface of the membrane, before 2 x 15 min washes with agitation in the same buffer. Further washes were 2 x 15 min with buffer 2 (150 mM sodium citrate, 75 mM NaCl, pH 7.0, 0.1% (w/v) SDS), then 2 x 15 min with buffer 3 (30 mM sodium citrate, 15 mM NaCl, pH 7.0, 0.1% (w/v) SDS) and finally 15 min at 55°C with buffer 4 (30 mM sodium citrate, 15 mM

NaCl, pH 7.0, 1% (w/v) SDS). The membrane was then wrapped in saran wrap and exposed to autoradiography film or a phosphorimager screen.

2.6.5 Organotypic raft cultures

Organotypic raft cultures were prepared on my behalf by Sally Roberts as described in Wilson and Laimins (2005) Briefly 2 x 10^6 keratinocytes (HFK or HPV18 genome containing keratinocytes) were grown on a collagen plug containing 2 x 10^6 J2-3T3 fibroblasts. After 24-48 h the collagen plug was placed on top of a wire mesh to create an air/liquid interface with E-media –EGF. The media was changed every 2 days and the rafts grown for 13 days. SRPIN340 (10 μM or 25 μM), or an equivalent amount of DMSO was added for the final 7 days and added fresh when the media was changed. The nucleotide analogue BrdU was added 12-15 h before harvest to identify nuclei containing synthesizing cellular DNA. Rafts were harvested by flooding the dish with 10% (v/v) formaldehyde (36.5- 38%, Sigma) in DMEM and the rafts were embedded in paraffin prior to sectioning (Propath UK Ltd).

CHAPTER 3 INVESTIGATING THE EFFECT OF THE INTERACTION BETWEEN HPV E1^E4 AND SRPK1 ON THE PROTOTYPICAL SR PROTEIN ASF/SF2

3.1 Introduction

The phosphorylation of ASF/SF2 by SRPK1 controls protein-protein and protein-RNA interactions along with the localisation of ASF/SF2 and its functions in splicing. This cellular kinase is a target of several unrelated viruses, including herpes simplex virus and Varicella-Zoster virus (Ote et al., 2009, Sciabica et al., 2003), indicating that the modulation of SRPK1 is beneficial for the life cycle of these viruses.

GST-HPV1 E1^E4 was used to coprecipitate cellular proteins from lysates prepared from human keratinocytes and mass spectroscopical analysis identified SRPK1 as an interacting partner, further analysis confirmed that there is a direct interaction between the two proteins (Bell et al., 2007). In keratinocytes SRPK1 colocalises to the cytoplasmic inclusion bodies formed by HPV1 E4. This novel association has physiological significance since in natural HPV1 infections, SRPK1 is found to be present in the E4 inclusions formed in cells of the more superficial layers where the virus is amplifying the viral DNA and producing viral capsids. The SRPK1 interaction is not confined to HPV1; the E1^E4 proteins of the mucosa-tropic HPV types 16 and 18 interacted with the cellular kinase as well. Two regions within HPV1 E1^E4 were identified as critical determinants in the interaction, the C-terminal E4 oligomerisation domain and an arginine-rich sequence (GRPRR) that bears resemblance to the arginine-rich sequences present within SRPK1 substrates (R-X-R/K-X-X-X-R) and indeed HPV1 E1^E4 was a substrate for SRPK1 *in vitro*, although the phosphorylated sequences were not confirmed (Bell et al., 2007). Of interest was the observation that in *in vitro* kinase

assays the interaction between HPV1 E1^E4 and SRPK1 also modulates SRPK1 autophosphorylation and this was dependent upon SRPK1 binding by E1^E4 (Bell et al., 2007).

The inhibition, modulation or sequestration of SRPK1 by E1^E4 may affect the functions of some of the substrates of SRPK1. SRPK1 has been shown to phosphorylate the lamin-B receptor (LBR) which interacts with B-type lamins and chromatin, the interaction with chromatin is dependent upon phosphorylation of RS repeats at the N-terminus (Takano et al., 2002). Inhibition of the phosphorylation of LBR may inhibit the attachment of LBR to chromatin and result in the loss of nuclear envelope integrity. The sequestration of SRPK1 to E4 inclusion bodies is restricted to the upper layers of the warts, where virion assembly occurs; therefore destabilisation of the nuclear envelope in these cells may facilitate the egress of virions (Bell et al., 2007). The regulation of HPV late gene expression is controlled by multiple *cis*-acting elements, including the major substrates of SRPK1, the SR proteins (McPhillips et al., 2004). The modulation of SRPK1 by E1^E4 may impact on SR protein regulation of host and/or virus gene expression and splicing (Bell et al., 2007). Additionally, there are viral proteins which are substrates for SRPK1, such as the HPV5 E2 protein which has multiple RS repeats in the hinge region (Lai et al., 1999). SRPK1 may influence the functions of HPV5 E2 by modulating E2s ability to activate transcription and/or replication. Interestingly the E1^E4 protein of HPV16 has been shown to stabilise HPV16 E2 and also to promote E2s translocation to the nucleus, this demonstrates that E1^E4 may be acting to regulate the functions of E2 by controlling how much active E2 there is present in the nucleus (C. Davy et al., 2009).

Besides HPV, other viruses interfere with SRPK1 and SR proteins, the ICP27 protein of herpes simplex virus (HSV)-1 plays multiple roles in post-transcriptional regulation and has been shown to interact with SR proteins, such as SRp20, this interaction occurs via the RRM

and not the RS domain (Sciabica et al., 2003). Additionally, ICP27 can interact with SRPK1 and cause its relocalisation to the nucleus. *In vitro,* ICP27 is a substrate for SRPK1 and the interaction between ICP27 and the kinase causes the hypophosphorylation of SR proteins and consequently the downregulation of cellular pre-mRNA splicing (Sciabica et al., 2003). ICP27 has been shown to inhibit splicing before the conversion of the intermediate B complex into the catalytically active C complex (Bryant et al., 2001). The result of ICP27 interacting with SRPK1 is to shut-off the splicing machinery and facilitate the export of the HSV-1 intronless RNAs, along with host-protein synthesis shut-off.

The Varicella-Zoster virus (VZV) immediate-early (IE) 4 protein acts as an important regulator of VZV and cellular gene expression and can directly interact with the SR proteins ASF/SF2, 9G8 and SRp20. The VZV IE4 protein can also interact with and is a substrate for SRPK1 *in vitro*, however this interaction does not seem to cause any relocalisation of SRPK1 nor alteration of the kinase activity (Ote et al., 2009). Additionally, the IE4 protein has been shown to interact with the TAP/NXF1 export receptor and bind to RNA, together this data suggests that IE4 acts with the SR proteins to export intronless mRNA via TAP/NXF1 and the reimport of IE4 into the nucleus is dependent upon phosphorylation by SRPK1 in a similar manner to the shuttling SR proteins (Ote et al., 2009).

The early region 4 (E4)-ORF4 protein of Adenovirus 5 can activate the dephosphorylation of SR proteins by interacting with the phosphatase PP2A, leading to the depression of IIIa splicing (Estmer Nilsson et al., 2001). E4-ORF4 also interacts with the SR proteins ASF/SF2 and SRp30c, which can bind to an intronic element of IIIa pre-mRNA and inhibit exon IIIa inclusion (Estmer Nilsson et al., 2001). Interestingly, the interaction between E4-ORF4 and PP2A can induce HPV16 late gene expression and the production of L1 mRNA, this switch in splice site selection is dependent upon the phosphorylation state of ASF/SF2 (Somberg et al., 2009).

The utilisation of host SR protein kinases during viral replication has led to the search for specific inhibitors of these kinases, the isonicotinamide compound SRPIN340 has been developed by Fukuhara et al. (2006), which specifically inhibits SRPK1 and SRPK2. Treatment of cells with this inhibitor has been shown to suppress the propagation of Sindbis virus, cytomegalovirus (CMV) and hepatitis C virus (HCV) (Fukuhara et al., 2006, Karakama et al., 2010). SRPIN340 has been shown to affect the replication of a wide range of DNA and RNA viruses, suggesting that the SRPKs play an important role in the replication cycles of many viruses.

Several different viruses have been shown to interact with or be dependent upon the activity of SRPK1, demonstrating the importance of SRPK1 in the life cycles of these infectious agents. Additionally proteins from HSV-1 and Adenovirus have been shown to affect the phosphorylation status of the SR proteins. This chapter examines the effect of the interaction between SRPK1 and the E1^E4 protein from multiple HPV types on SRPK1 kinase activity. The kinase activity of SRPK1 has been examined with *in vitro* kinase assays, using the prototypical SR protein ASF/SF2 as a substrate.

3.2 Results

3.2.1 Expression of GST-ASF/SF2

In order to investigate how the interaction between SRPK1 and HPV E4 proteins affects the kinase activity of SRPK1 in vit*ro,* the prototypical SR protein Alternative Splicing Factor/Splicing Factor 2 (ASF/SF2) was expressed recombinantly as a GST fusion protein. The coding sequence of ASF/SF2 was excised from a plasmid containing the ASF/SF2 cDNA (pEGFP-ASF/SF2) by BamHI digestion and ligated into the complementarily digested GST vector pGex-2T to form pGex-2T-ASF/SF2. *E. coli* BL21-CodonPlus® cells were transformed with pGex-2T-ASF/SF2 and used to express the GST-ASF/SF2 protein.

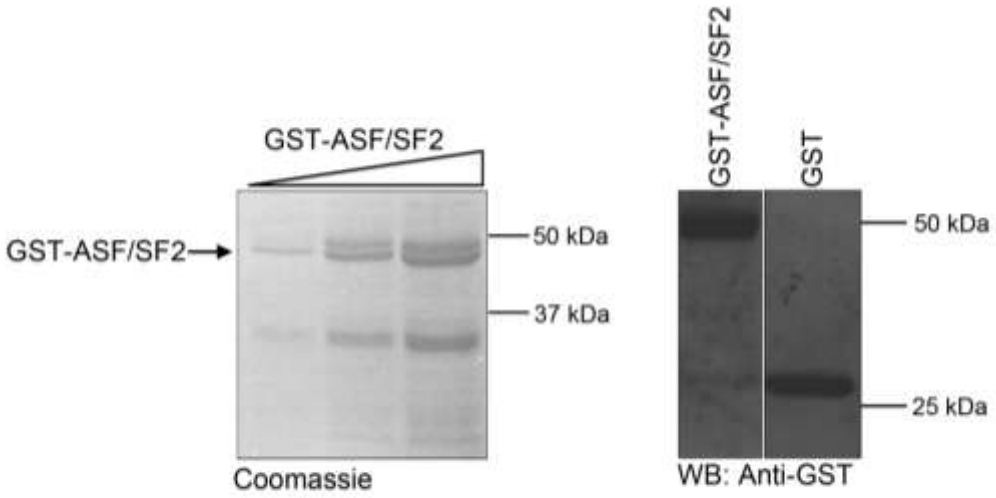

Figure 3.1 Expression of GST-ASF/SF2. Coomassie stained gel of increasing amounts of purified protein (0.5, 2.5 and 5 µg) and an anti-GST Western blot with 2.5 µg protein loaded, showing a doublet of approximately 50 kDa (expected size of GST-ASF/SF2 is 56 kDa).

The product migrated as a doublet of approximately 50 kDa (expected molecular mass of 56 kDa); this doublet was recognised by an anti-GST antibody (Figure 3.1) and an anti-ASF/SF2 antibody (data not shown).

3.2.2 GST-ASF/SF2 is a substrate for His-SRPK1

Initial experiments were carried out to show that the GST-ASF/SF2 could be phosphorylated by the recombinantly expressed His-tagged SRPK1 *in vitro*. *In vitro* kinase assays were carried out in the presence and absence of ^{32}P labelled ATP. GST-ASF/SF2 was added to the *in vitro* kinase assay with or without His-SRPK1 and GST was also included with His-SRPK1 as a control. The resolved gel containing ^{32}P labelled ATP was exposed to X-ray film (Figure 3.2A), and the gel without ^{32}P was Western blotted with a phospho-specific antibody against the phosphorylated RS domain of SR proteins, 1H4 (Figure 3.2B). Both methods identified a band corresponding to phosphorylated GST-ASF/SF2 only in the presence of the SRPK1 kinase, indicating that the recombinantly expressed GST-ASF/SF2 was a substrate *in vitro* for His-SRPK1. GST was not a substrate for His-SRPK1.

Once GST-ASF/SF2 had been established as a substrate for His-SRPK1 *in vitro*, the kinase assay conditions were optimised. In the first instance the optimum incubation time was determined. Experiments were carried out for increasing lengths of time, ranging from 0.5 to 60 min, with all other conditions remaining constant (Figure 3.3). The phosphorylation levels were determined by phosphorimager analysis of the radiolabelled gel. Consistently 50% of the total phosphorylation of GST-ASF/SF2 occurred after 10 min. Subsequent *in vitro* kinase assays were therefore carried out for either 10 or 30 min.

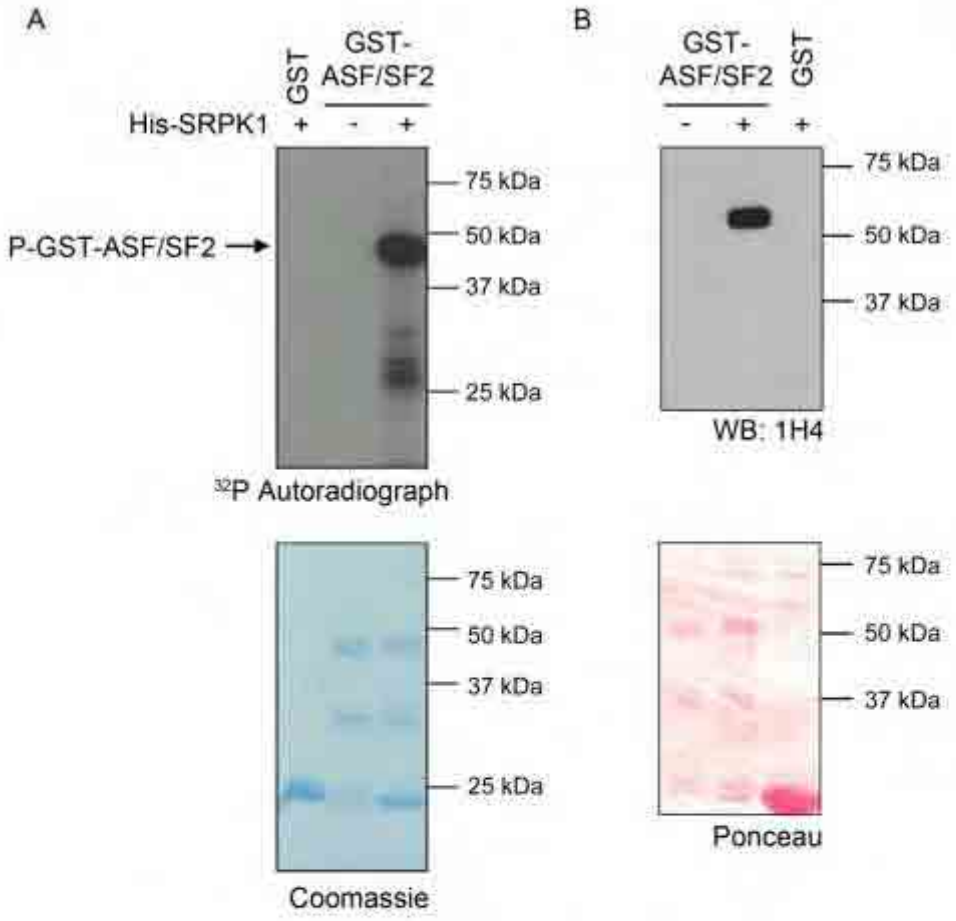

Figure 3.2 Phosphorylation of GST-ASF/SF2 by His-SRPK1. An *in vitro* kinase assay with 8 µg GST-ASF/SF2 in the presence or absence of 1 µg His-SRPK1. (**A**) Radiolabelled ^{32}P autoradiograph with band corresponding to phosphorylated GST-ASF/SF2 only in the presence of His-SRPK1. (**B**) Using the antibody 1H4 specific for the phosphorylated RS domain of SR proteins in a Western blot, a band corresponding to GST-ASF/SF2 was detected only in the presence of His-SRPK1.

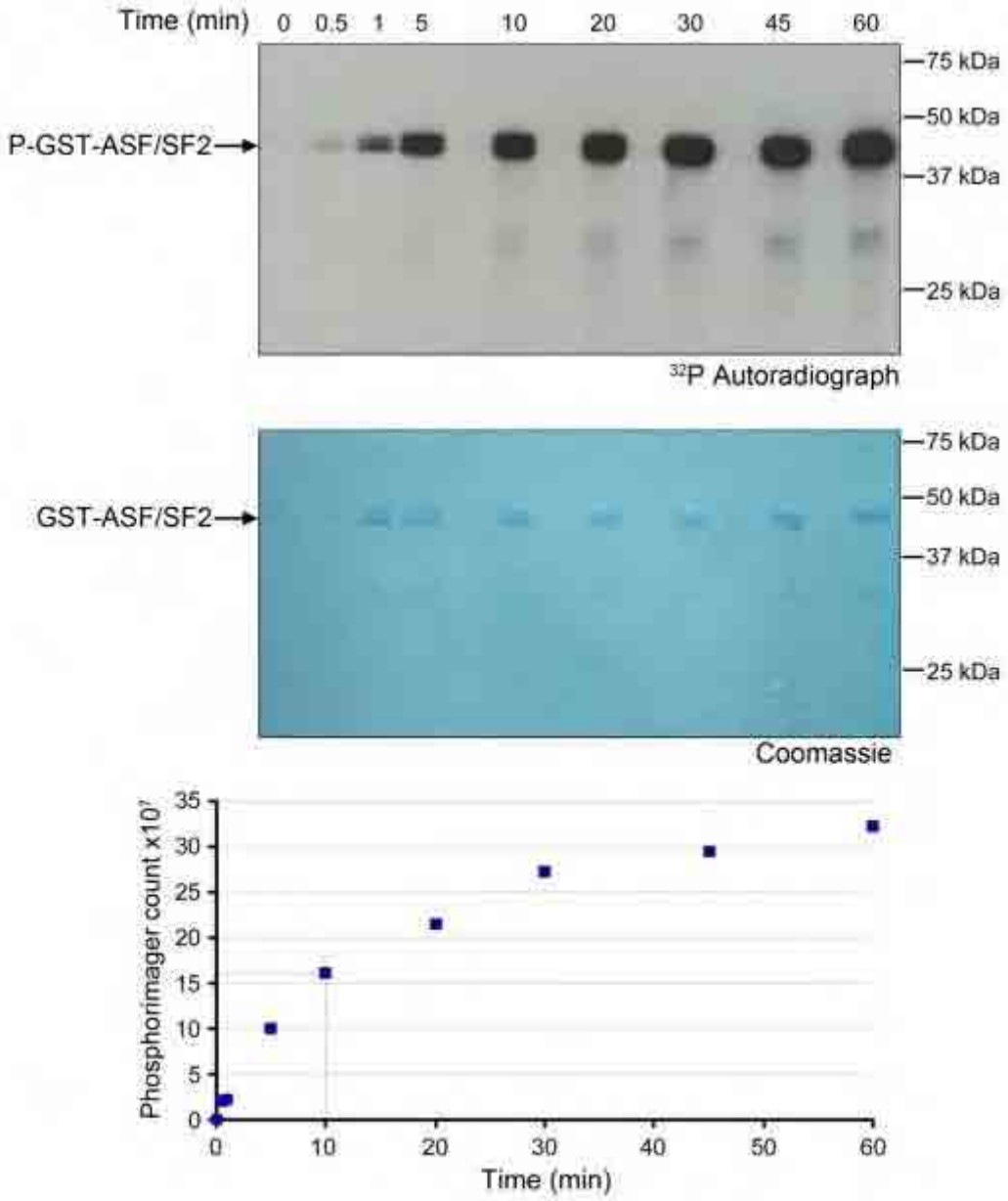

Figure 3.3 SRPK1 kinase assay time course. The kinase assays were carried out for increasing amounts of time and the level of labelled ATP incorporated into GST-ASF/SF2 (5 µg) was measured by phosphorimager analysis, 1 µg of His-SRPK1 was used. The data shown is representative of three data sets.

In order to ensure that the amount of His-SRPK1 was not limiting the level of phosphorylation of GST-ASF/SF2 in the assay, an *in vitro* kinase assay was carried in which the amount of GST-ASF/SF2 added was titrated from 17.5 to 0.007 μg (Figure 3.4). The level of phosphorylation was very similar with 17.5 and 14 μg of the substrate but decreased to 50% with 10.5 μg indicating that His-SRPK1 is in excess in this assay. Thus, between 10 and 12 μg of GST-ASF/SF2 will be used in successive kinase assays, when 2 μg of His-SRPK1 is used.

3.3 HPV1 E1^E4 inhibits SRPK1 phosphorylation of ASF/SF2

3.3.1 HPV1 E1^E4 interacts with SRPK1

To confirm that the recombinantly expressed proteins prepared for these studies interact as previously reported by this laboratory (Bell et al., 2007), GST-HPV1 E1^E4 bound to glutathione agarose beads was mixed with a bacterial lysate containing His-SRPK1. Following a brief incubation at 4°C for 1 h and washing of the beads, the proteins bound to the beads were resolved by SDS-PAGE and Western blotted with an anti-His antibody. GST-ASF/SF2 was used as a positive control and GST used to control for non-specific interactions. Figure 3.5 shows that GST-HPV1 E1^E4 coprecipitates His-SRPK1 to a similar level as the cellular substrate GST-ASF/SF2 and no interaction is seen with the GST protein.

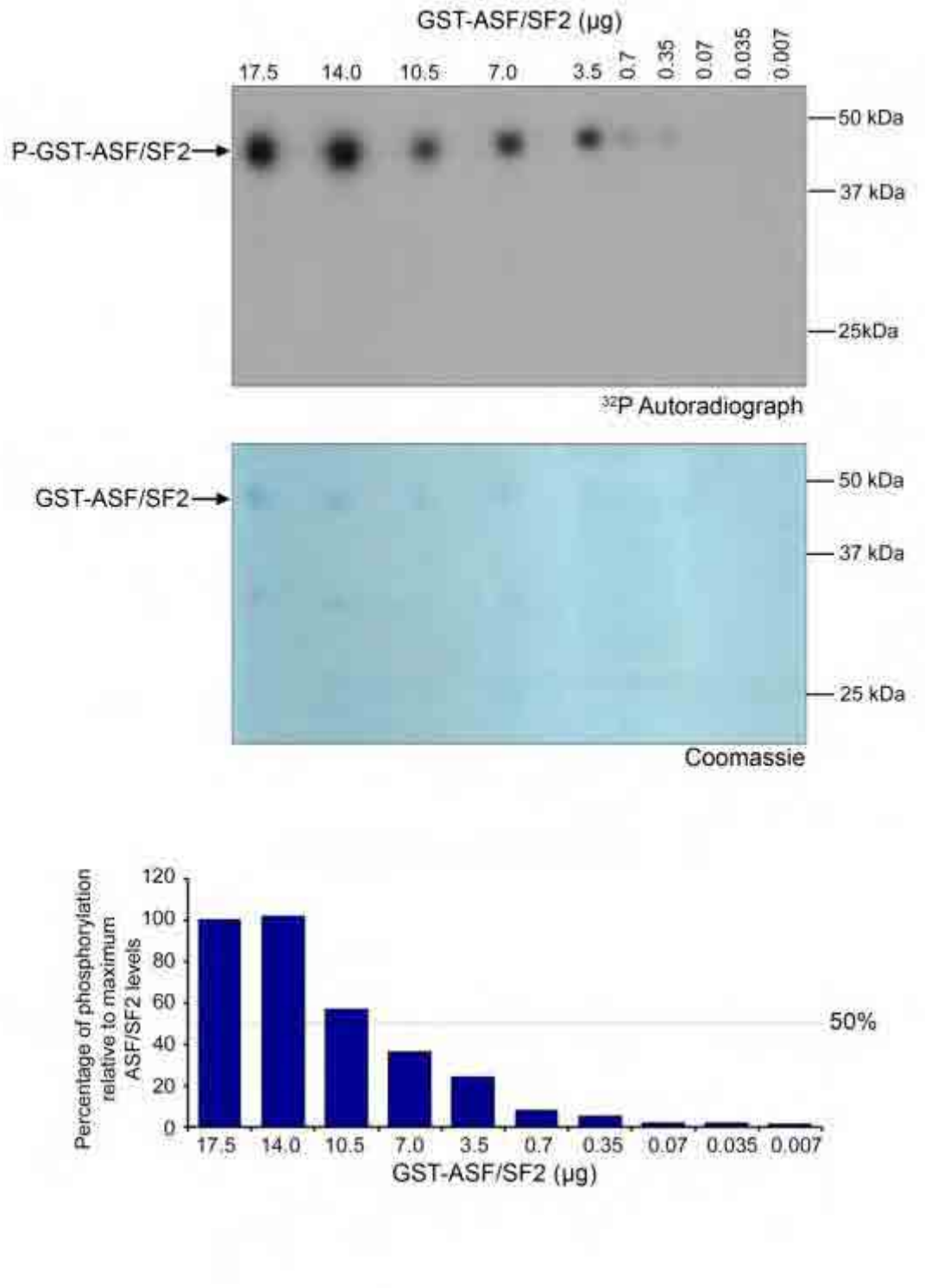

Figure 3.4 Titration of GST-ASF/SF2 into an *in vitro* kinase assay. An *in vitro* kinase assay with decreasing amounts of GST-ASF/SF2 (17.5, 14, 10.5, 7, 3.5, 0.7, 0.35, 0.07. 0.035 and 0.007 µg) and 2 µg His-SRPK1. The level of ^{32}P-ATP incorporated into the gel was measured by phosphorimager analysis and used to calculate the percentage of phosphorylation normalised to the maximum amount of GST-ASF/SF2. The graph shown contains the mean from two independent experiments.

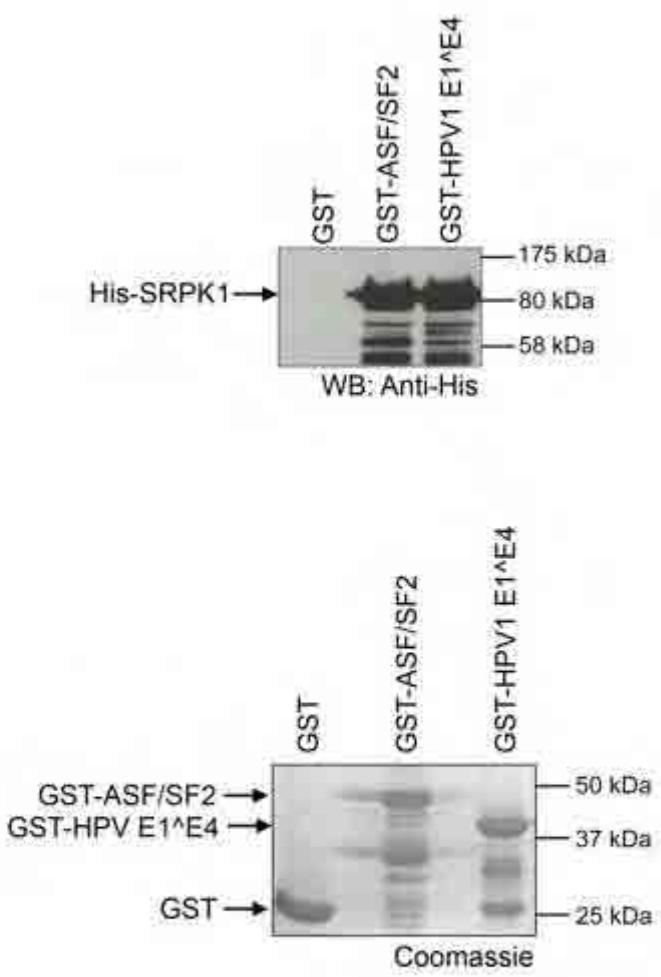

Figure 3.5 GST-HPV1 E1^E4 binds to His-SRPK1. Approximately 10 µg of GST-HPV1 E1^E4, GST-ASF/SF2 and GST were bound to glutathione agarose beads and mixed with bacterial lysate containing His-SRPK1. Western blotting for the His-tag revealed that GST-HPV1 E1^E4 and GST-ASF/SF2 coprecipitated equivalent amounts of His-SRPK1. GST proteins shown by Coomassie stain.

3.3.2 Inhibition of SRPK1 phosphorylation of ASF/SF2

Comparable levels of SRPK1 were coprecipitated by HPV1 E1^E4 and ASF/SF2. Therefore, to determine if the interaction with the E1^E4 protein has any effect on SRPK1 kinase activity, GST-HPV1 E1^E4 or GST were added in increasing amounts into an *in vitro* kinase assay in the presence of GST-ASF/SF2 and His-SRPK1. The level of phosphate incorporated into the ASF/SF2 protein was measured by using trichloroacetic acid (TCA) protein precipitation and scintillation counting of the precipitate, as described in Chapter 2 (Section 2.3.6). With increasing amounts of GST-HPV1 E1^E4 added to the *in vitro* kinase assay the level of phosphate incorporated into GST-ASF/SF2 decreased (Figure 3.6). No variation of phosphate incorporation occurred in the presence of the GST protein. At the maximum amount of GST-HPV1 E1^E4 added the phosphate incorporated into ASF/SF2 is reduced by 65% compared to an equivalent amount of GST ($p = 0.0002$). This data indicated that the GST-tagged HPV1 E1^E4 is a potent inhibitor of GST-ASF/SF2 phosphorylation by His-SRPK1.

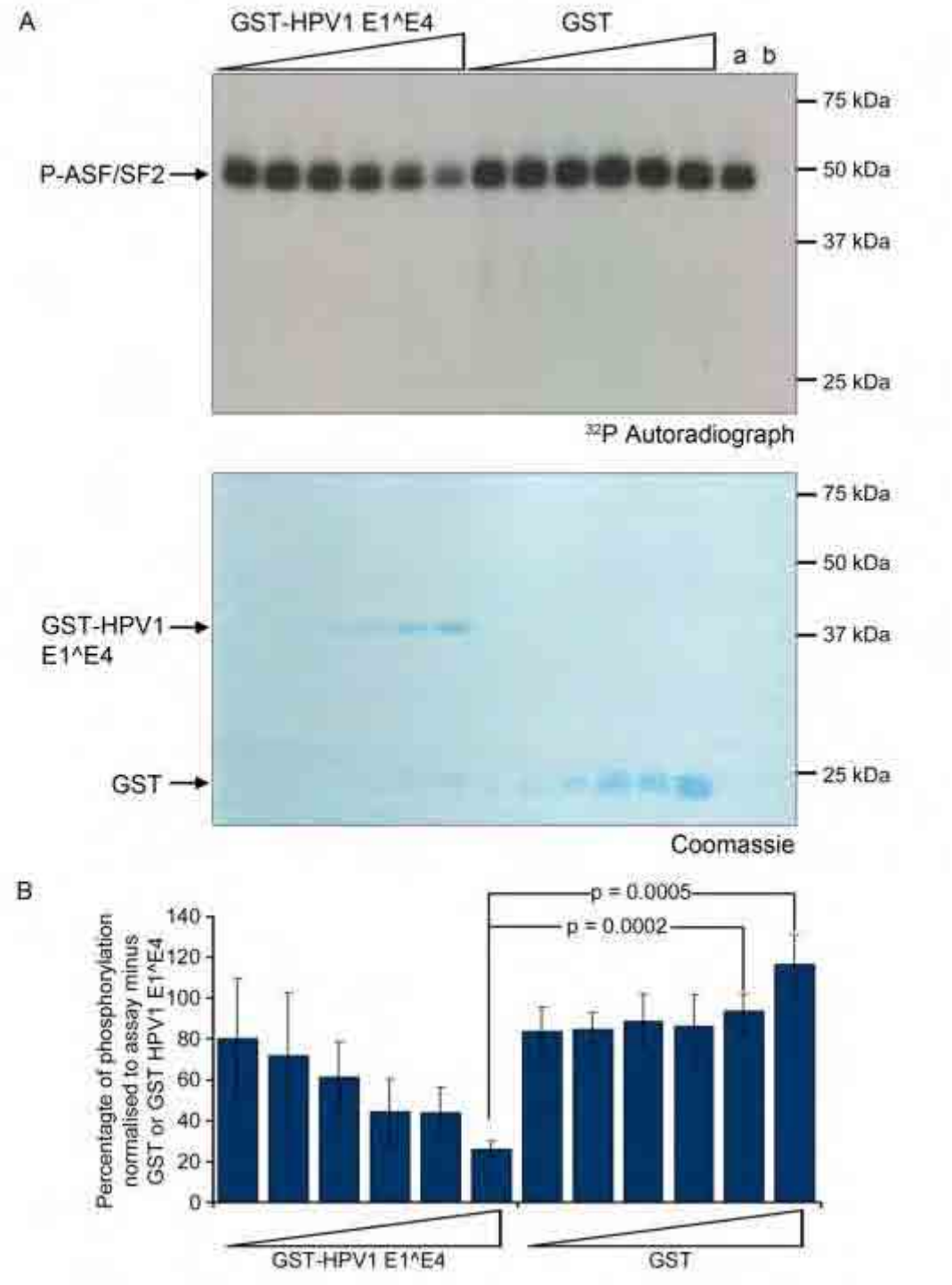

Figure 3.6 GST-HPV1 E1^E4 inhibits GST-ASF/SF2 phosphorylation by His-SRPK1. GST (0.3, 0.6, 1.2, 2.5, 5 or 10 µg) or GST-HPV1 E1^E4 (0.15, 0.3, 0.6, 1.2, 2.5 or 5 µg) was added in increasing amounts to an *in vitro* kinase assay containing 1 µg GST-ASF/SF2 and 1 µg His-SRPK1. (**A**) ^{32}P autoradiograph and Coomassie stained gel. Lane (a) contains only GST-ASF/SF2 and His-SRPK1; lane (b) contains only GST-ASF/SF2. (**B**) Bar graph showing the percentage of phosphorylation normalised to assay minus GST or GST HPV1 E1^E4 (lane (a)). The data shown were collected from three independent experimental replicates.

3.3.3 Inhibition is dependent on the interaction between SRPK1 and HPV1 E1^E4

To confirm that the inhibition of SRPK1 kinase activity by HPV1 E1^E4 is dependent on an interaction between the two proteins, the inhibitory actions of two HPV1 deletion E1^E4 mutants were tested. The mutant protein Δ44-48 which has been shown to be a non-binder of SRPK1 and the deletion mutant Δ49-53 which retains SRPK1 binding activity (Bell et al., 2007). These mutants, expressed as GST-fusion proteins were included in an *in vitro* kinase assay with GST-ASF/SF2 and His-SRPK1 (Figure 3.7). The non-binding deletion mutant Δ44-48 did not inhibit phosphorylation of GST-ASF/SF2 ($p=0.18$), whereas the binding mutant, Δ49-53, significantly ($p=0.048$) reduced the phosphorylation of ASF/SF2 when compared to phosphorylation in the presence of GST alone.

3.3.4 Determination of the mode of inhibition of HPV1 E1^E4

A selective inhibitor, *N*-[2-(1-piperidinyl)-5-(trifluoromethyl)phenyl]isonicotinamide (SRPIN340) (Fukuhara et al., 2006) (Figure 3.8A) has been shown to inhibit two members of the SRPK family (SRPK1 and SRPK2). To confirm that the recombinantly expressed His-SRPK1 was able to respond to this inhibitor, SRPIN340 was added to an *in vitro* kinase assay with GST-ASF/SF2 and His-SRPK1. As shown in Figure 3.8B, SRPIN340 is an effective inhibitor of His-SRPK1 mediated phosphorylation of the ASF/SF2 fusion protein, inhibiting the phosphorylation of this substrate by up to 70% when added at a concentration of 50 μM (Figure 3.8C).

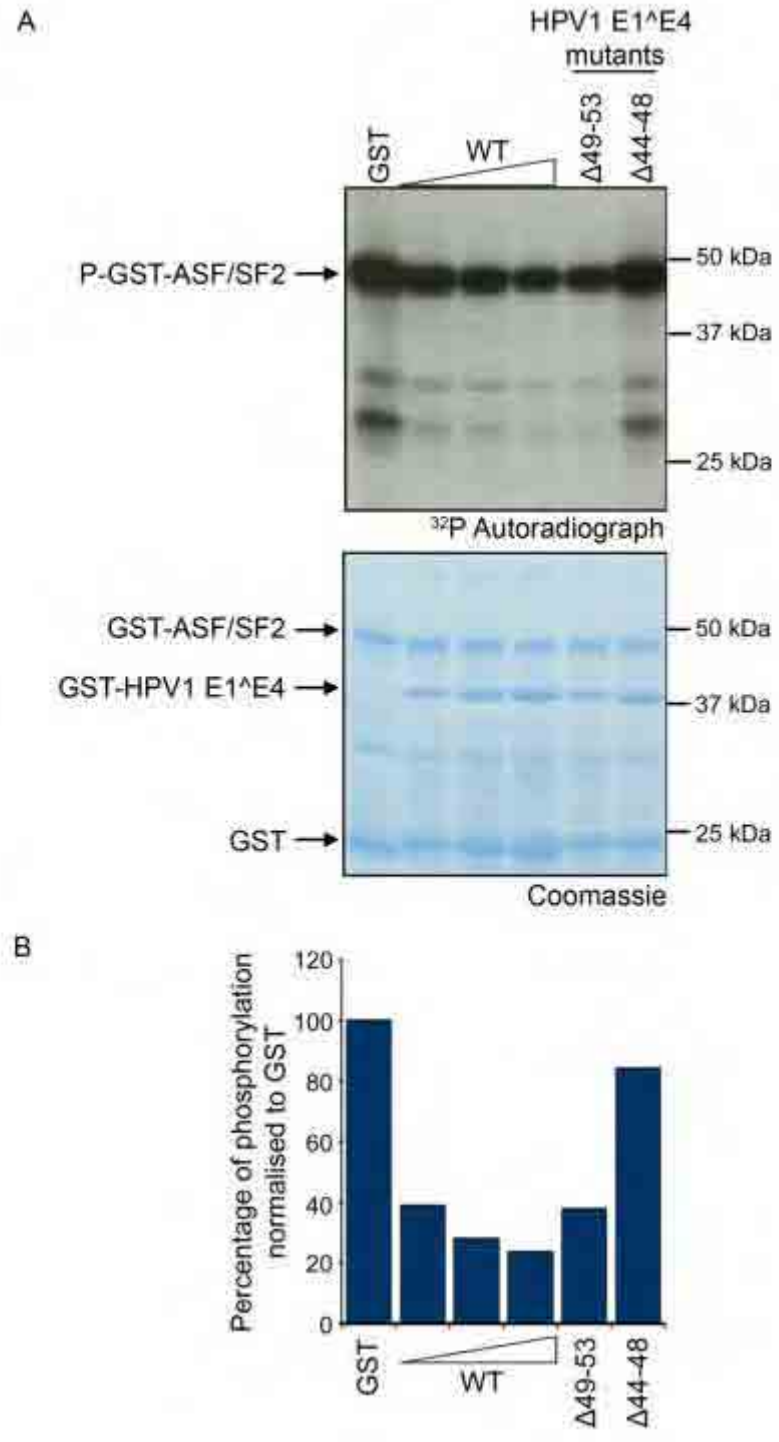

Figure 3.7 HPV1 E1^E4 mediated inhibition of SRPK1 activity is dependent on an interaction with E1^E4. GST-HPV1 E1^E4 deletion mutants, Δ44-48 (non-binding mutant, 3 µg) and Δ49-53 (a binding mutant, 5 µg) were added in an *in vitro* kinase assay containing 5 µg GST-ASF/SF2 and 1 µg His-SRPK1. GST-HPV1 E1^E4 (2.5, 5 and 10 µg) and GST (10 µg) were added as positive and negative controls respectively. (**A**) ^{32}P autoradiograph and Coomassie stained gel. (**B**) The data in the graph is the mean taken from two independent experiments.

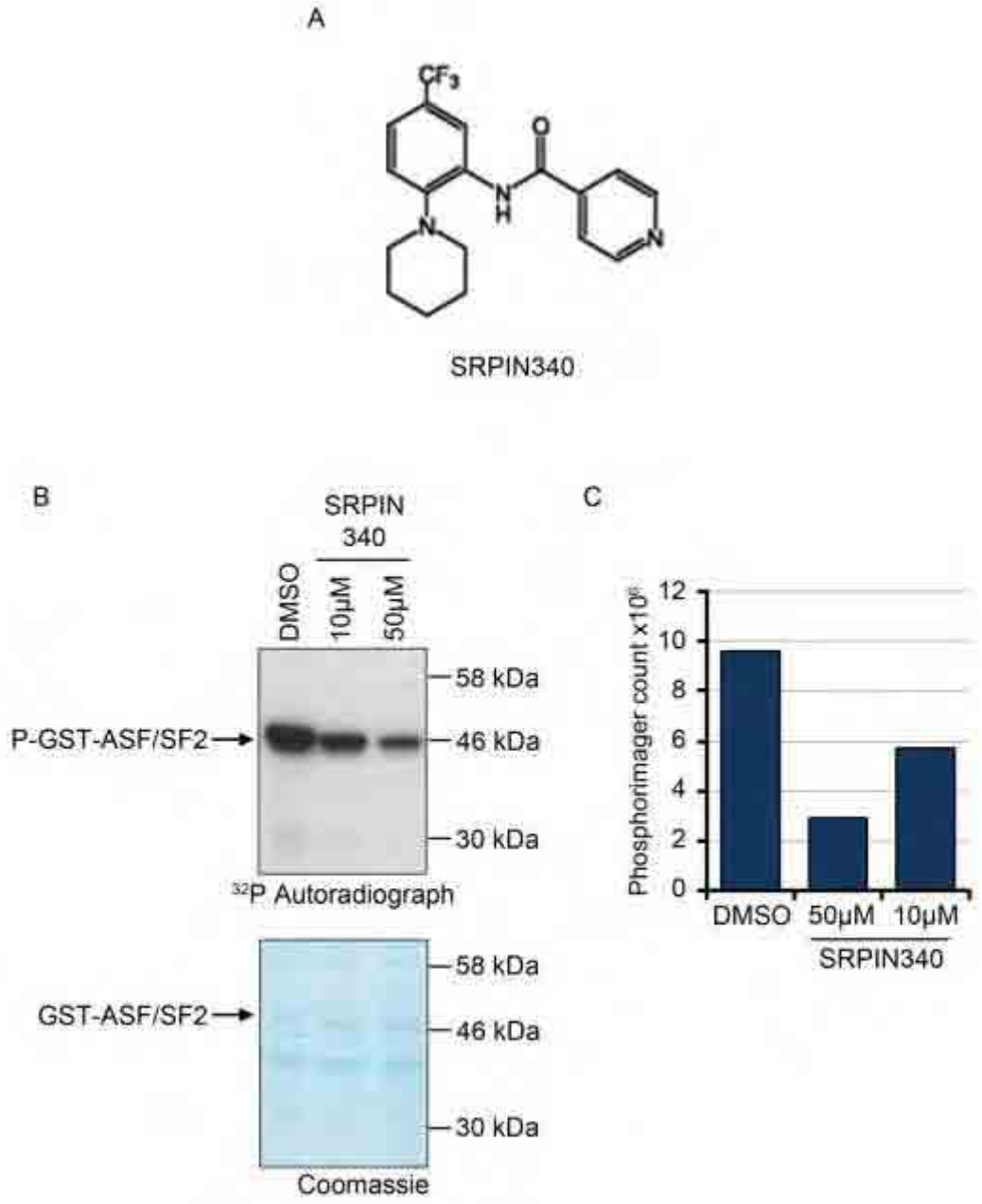

Figure 3.8 Inhibition of SRPK1 phosphorylation of ASF/SF2 by SRPIN340. (**A**) Structure of selective inhibitor SRPIN340. (**B**) ^{32}P autoradiograph and Coomassie stained gel showing the addition of 10 or 50 μM SRPIN340 to an *in vitro* kinase assay containing 2 μg GST-ASF/SF2 and 1 μg His-SRPK1. (**C**) Graph showing the phosphorimager count from gel B. The data shown is representative of two independent experiments.

SRPIN340 inhibits SRPK1 activity by binding to the active site and is referred to as a competitive inhibitor (Fukuhara et al., 2006). To determine if the mode of inhibition by HPV1 E1^E4 is competitive like SRPIN340, or if it is a non-competitive or uncompetitive inhibitor, an *in vitro* kinase assay was carried out in which the amount of GST-ASF/SF2 was titrated in the presence of constant amounts of His-SRPK1 and GST-HPV1 E1^E4, or GST as a control. The level of phosphate incorporated into the ASF/SF2 protein was measured by TCA protein precipitation (Section 2.3.6). The level of phosphate incorporation into ASF/SF2 in the presence of GST-HPV1 E1^E4 was consistently lower than with the GST control, with a reduction of 56% at the highest amount of GST-ASF/SF2 ($p=0.001$). To determine the mode of inhibition, a Lineweaver-Burk plot was produced (Figure 3.9C), where the K_m and V_{max} of the reaction can be determined from the x and y intercepts respectively. In the presence of a competitive inhibitor such as SRPIN340 the V_{max} of the reaction does not vary with increasing amounts of inhibitor. With a non-competitive inhibitor, which binds to an alternative site on the enzyme, the K_m does not vary and for an uncompetitive inhibitor the K_m/V_{max} does not vary in the presence of increasing amounts of substrate (Mathews et al., 2000). However, with increasing amounts of the substrate GST-ASF/SF2 in the presence of GST-HPV1 E1^E4, the K_m, V_{max} and K_m/V_{max} are all changed in comparison to the assay containing the GST control protein, and this response can be indicative of mixed inhibition. The Lineweaver-Burk plot assumes simple enzyme kinetics, with E+S → ES → E+P, however the kinase activity of SRPK1 is processive (Aubol et al., 2003), with the reaction kinetics varying between the initial phosphorylation and the following multisite phosphorylation (Aubol and Adams, 2011, Ngo et al., 2008). The measure of total phosphate incorporated into ASF/SF2 does not account for the changing reaction rate and therefore does not allow for accurate rate determination.

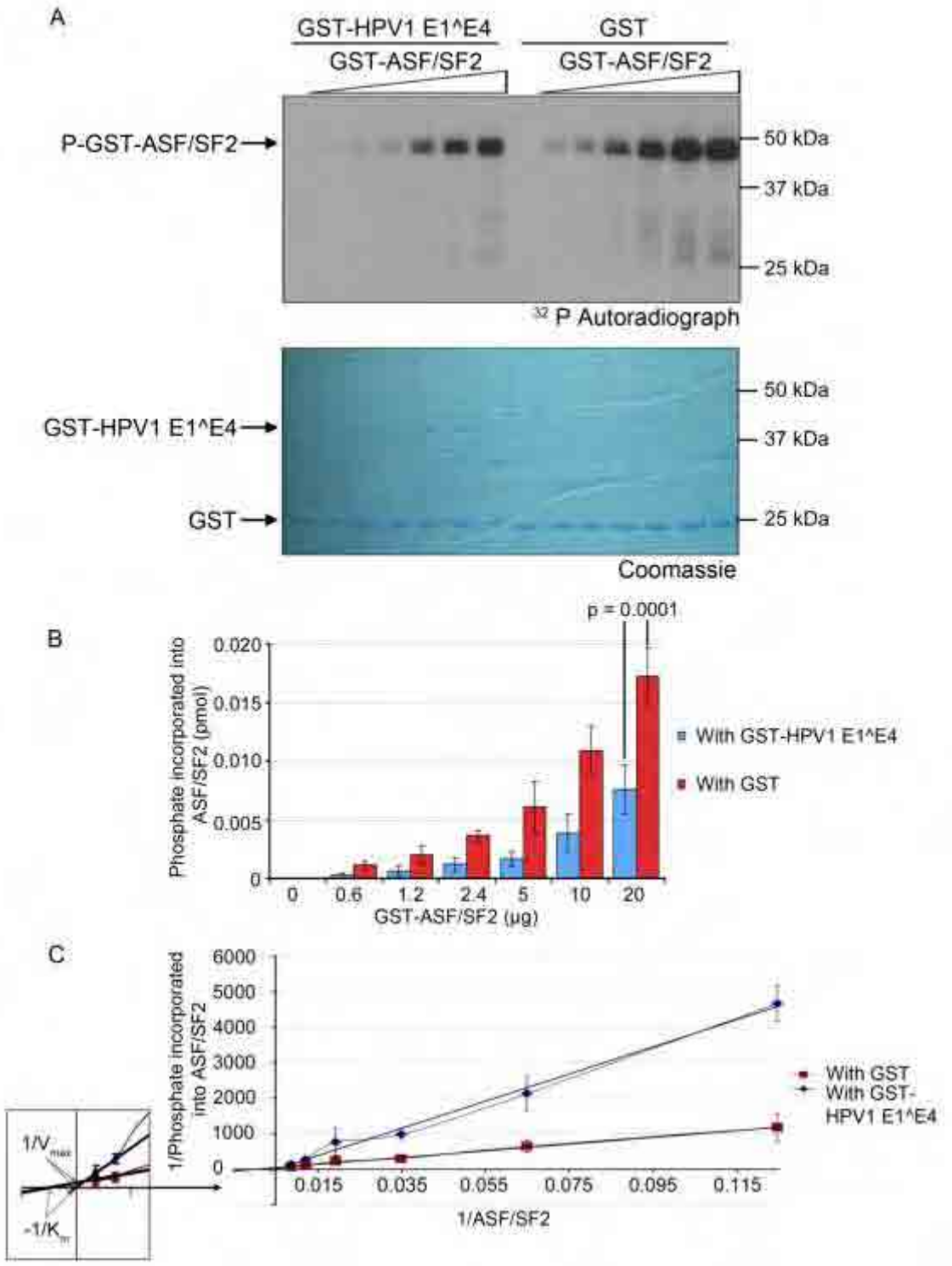

Figure 3.9 Determination of the mode of inhibition by HPV1 E1^E4 using Michaelis-Menten kinetics. (**A**) ^{32}P autoradiograph and Coomassie stained gel showing the titration of 0.6, 1.2, 2.4, 5, 10 or 20 µg GST-ASF/SF2 into an *in vitro* kinase assay containing 1 µg of His-SRPK1, in the presence of 5 µg GST-HPV1 E1^E4 or 7 µg GST. (**B**) The level of phosphate incorporated into the ASF/SF2 with increasing amounts of GST-ASF/SF2 is shown. The data in the graph shown is the mean (± standard deviation) from five experiments. (**C**) Lineweaver-Burk plot of the data in graph B, with linear regression trendlines.

This means that to accurately determine the mode of inhibition of SRPK1 by HPV1 E1^E4 using Michaelis-Menten kinetics, the use of the Lineweaver-Burk plot was inadequate.

3.3.5 HPV1 E1^E4 inhibits binding of SRPK1 to ASF/SF2

Initial attempts to establish the mode of inhibition of HPV1 E1^E4 on SRPK1 by determination of Michaelis-Menten kinetics were unsuccessful. This was possibly due to the limitations of the approach used as mentioned previously. Another method to investigate the binding of proteins is using surface plasmon resonance (SPR) to measure protein-protein interactions, as used by Biacore® technology. Biacore® can be used to accurately determine the specificity of interactions, the affinity of interactions and the kinetics of binding (k_a and k_d). The approach was therefore to use Biacore® technology to determine whether SRPK1 bound the viral E1^E4 protein with the same affinity as its cellular substrate ASF/SF2 and to determine the kinetics of this interaction. Initially, purified His-SRPK1 was bound to a nitrilotriacetic acid (NTA) chip, using the His-tag. Whilst the His-SRPK1 bound successfully to the chip (as determined by an increase in response units), neither GST-ASF/SF2 nor GST-HPV1 E1^E4 would bind to the bound SRPK1 (Figure 3.10A). To determine whether this was due to steric interactions interfering with the binding, His-SRPK1 was bound to nickel agarose beads and used to coprecipitate GST-ASF/SF2 or GST-HPV1 E1^E4. No binding of the GST-tagged proteins to SRPK1 was detectable by Western blotting (data not shown); suggesting that binding of the His-tag to a support induced steric hindrance to prevent SRPK1 binding to its substrate or to the viral inhibitor.

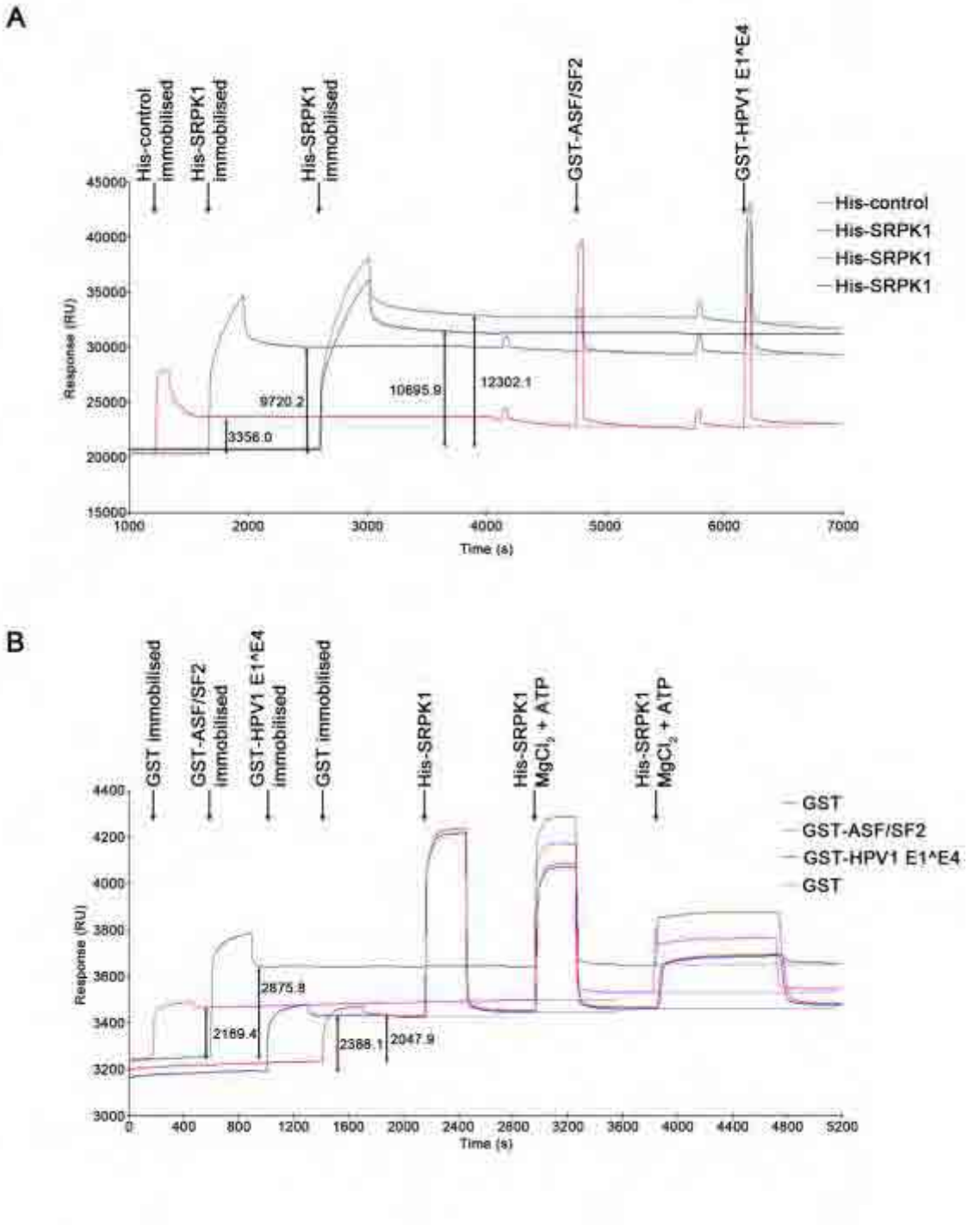

Figure 3.10 Biacore® interaction study using NTA and GST-antibody chips. (**A**) Two micrograms of His-SRPK1 was bound to a nitrilotriacetic acid (NTA) chip on chambers 2, 3 and 4. One microgram of GST-ASF/SF2 was passed through chambers 1 and 2 at 5 µl/min and 1 µg of GST-HPV1 E1^E4 was passed through chambers 1, 2 and 3 at 5 µl/min. (**B**) A goat anti-GST antibody was bound to a CM5 chip by amine coupling, then 3.5 µg of GST were bound to chambers 1 and 4, 1.2 µg of GST-ASF/SF2 bound to chamber 2 and 1.4 µg of GST-HPV1 E1^E4 bound to chamber 3. His-SRPK1 was passed through chambers 3 and 4 at 5 µl/min, then chambers 1-4 in the presence of MgCl2 and ATP at 5 µl/min for 5 min then 15 min.

A second approach therefore was to determine if the substrate GST-ASF/SF2 could retain binding to SRPK1 when bound to a solid support and whether the interaction could be blocked by addition of E1^E4 or small molecule inhibitors. Thus, GST or GST-ASF/SF2 were bound to glutathione agarose beads and mixed with His-SRPK1 in the presence of the various inhibitors; HPV1 E1^E4 that had been cleaved from the GST-tag using thrombin (Section 2.1.5.4) (Figure 3.11A), or the small molecule inhibitors SRPIN340 or staurosporine, an ATP analogue (a schematic of binding is shown in Figure 3.11B). After repeated washings, the proteins bound to GST and GST-ASF/SF2 were resolved by SDS PAGE and Western blotted with an anti-polyHis antibody (Figure 3.11C). As expected, the addition of the inhibitors, SRPIN340 and staurosporine, which are known to bind in the active site reduced binding of His-SRPK1 to the substrate GST-ASF/SF2. Addition of the cleaved HPV1 E1^E4 protein also reduced the binding of the kinase to GST-ASF/SF2, and to a similar extent as the two competitive inhibitors SRPIN340 and staurosporine.

This data shows that the bound substrate ASF/SF2 was recognised by SRPK1 in solution and that binding of the kinase is inhibited by E1^E4. Therefore a CM5 Biacore® chip was generated with an anti-GST antibody immobilised on its surface. However, whilst the GST proteins were all able to bind to the anti-GST chip, His-SRPK1 did not bind to any of the immobilised proteins (Figure 3.10B).

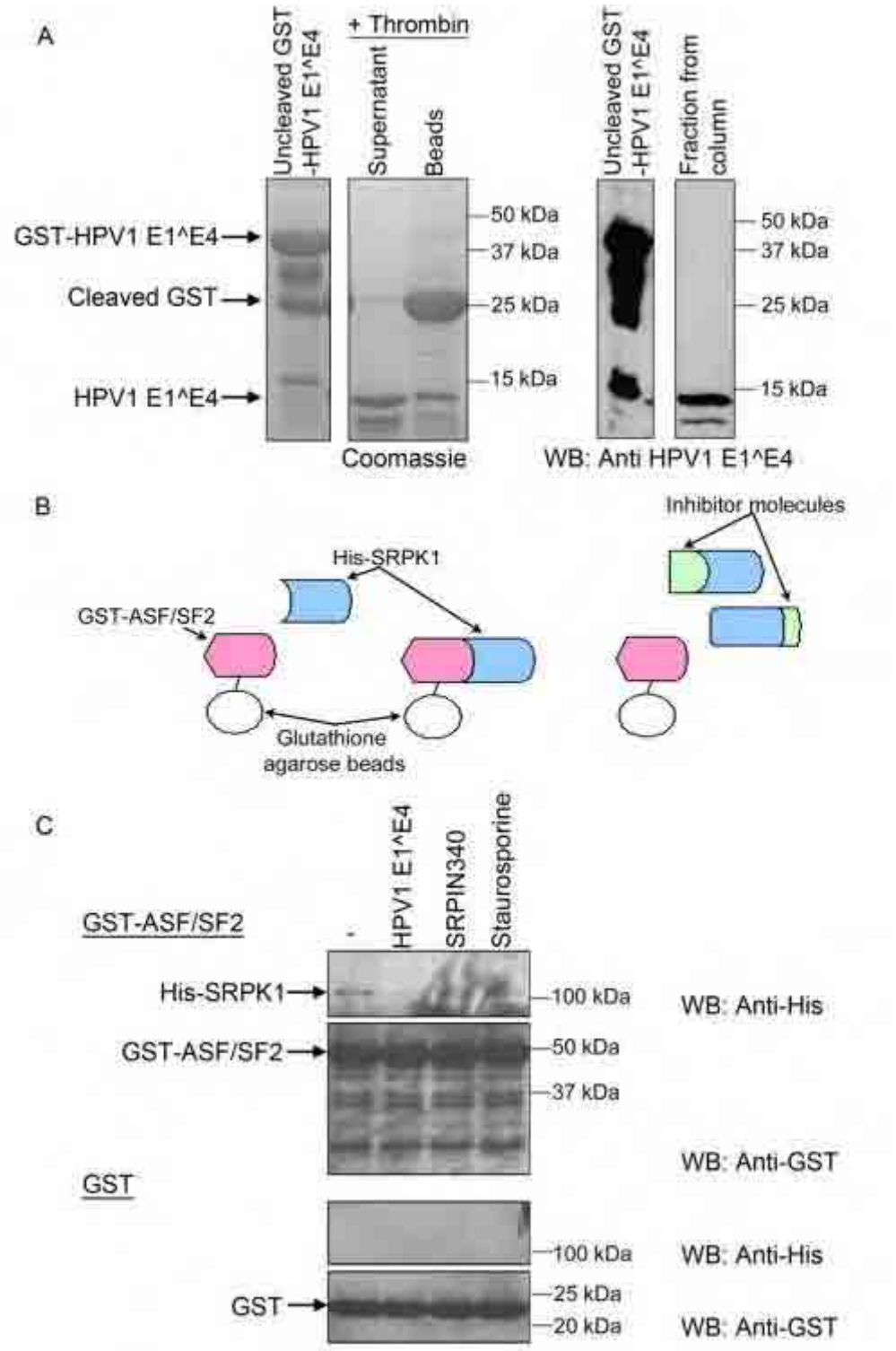

Figure 3.11 HPV1 E1^E4 inhibits the binding of His-SRPK1 to GST-ASF/SF2. (**A**) Coomassie gel and anti-HPV1 E1^E4 Western blot showing efficient thrombin cleavage of the GST-tag from GST-HPV1 E1^E4 and elution of the purified protein from a benzamidine column. (**B**) Schematic showing the coprecipitation of His-SRPK1 by GST-ASF/SF2 and the addition of inhibitors. (**C**) GST-ASF/SF2 or GST was bound to glutathione agarose beads and incubated with His-SRPK1 alone or in the presence of HPV1 E1^E4, SRPIN340 or Staurosporine. Following resolution by SDS PAGE the gels were Western blotted with anti-His and anti-GST antibodies. This data is representative of two experimental replicates.

The next approach was to couple the GST proteins and His-SRPK1 directly to the CM5 Biacore® chip via amine coupling. The coupling of the proteins was performed at a pH lower than the pI of the proteins, and this was determined experimentally by changing the pH until 'sticking' occurred between the protein and the chip (section 2.3.8.3). The GST-ASF/SF2 and His-SRPK1 were coupled to separate chambers of the chip at pH 5.0 and pH 4.0 respectively (Figure 3.12A). The His-SRPK1 was also coupled in the presence of ATP and $MgCl_2$ at pH 5.0 (Figure 3.12B), as this can help retain the ability of the kinase to bind to its substrate (Nordin et al., 2005). When GST-ASF/SF2 was flowed into the chamber containing immobilised His-SRPK1, binding occurred. However when GST-ASF/SF2 was added as a control to the chamber containing immobilised GST-ASF/SF2, binding also occurred to a similar level (Figure 3.12C). Injection of His-SRPK1 into the chambers also did not bind to the immobilised GST-ASF/SF2 or to the immobilised His-SRPK1. A small increase in response units (< 370 units) was seen when GST-HPV1 E1^E4 was added to the immobilised SRPK1, however normally a change of over 1000 response units is expected when binding occurs. Also, the small change in response units with GST-HPV1 E1^E4 was not a reproducible observation between chips as shown in figure 3.12D.

An explanation for the lack of binding may be due to the His-SRPK1 being bound in an inaccessible configuration. Therefore to test this, an anti-SRPK1 antibody was used to determine whether the His-SRPK1 immobilised on the chip was accessible. No binding of the antibody was observed, even though His-SRPK1 bound to the chip with 12,000 response units, leading to the conclusion that the SRPK1 was bound in an inaccessible configuration and the "binding" that was observed, was due to non-specific interactions with the chip.

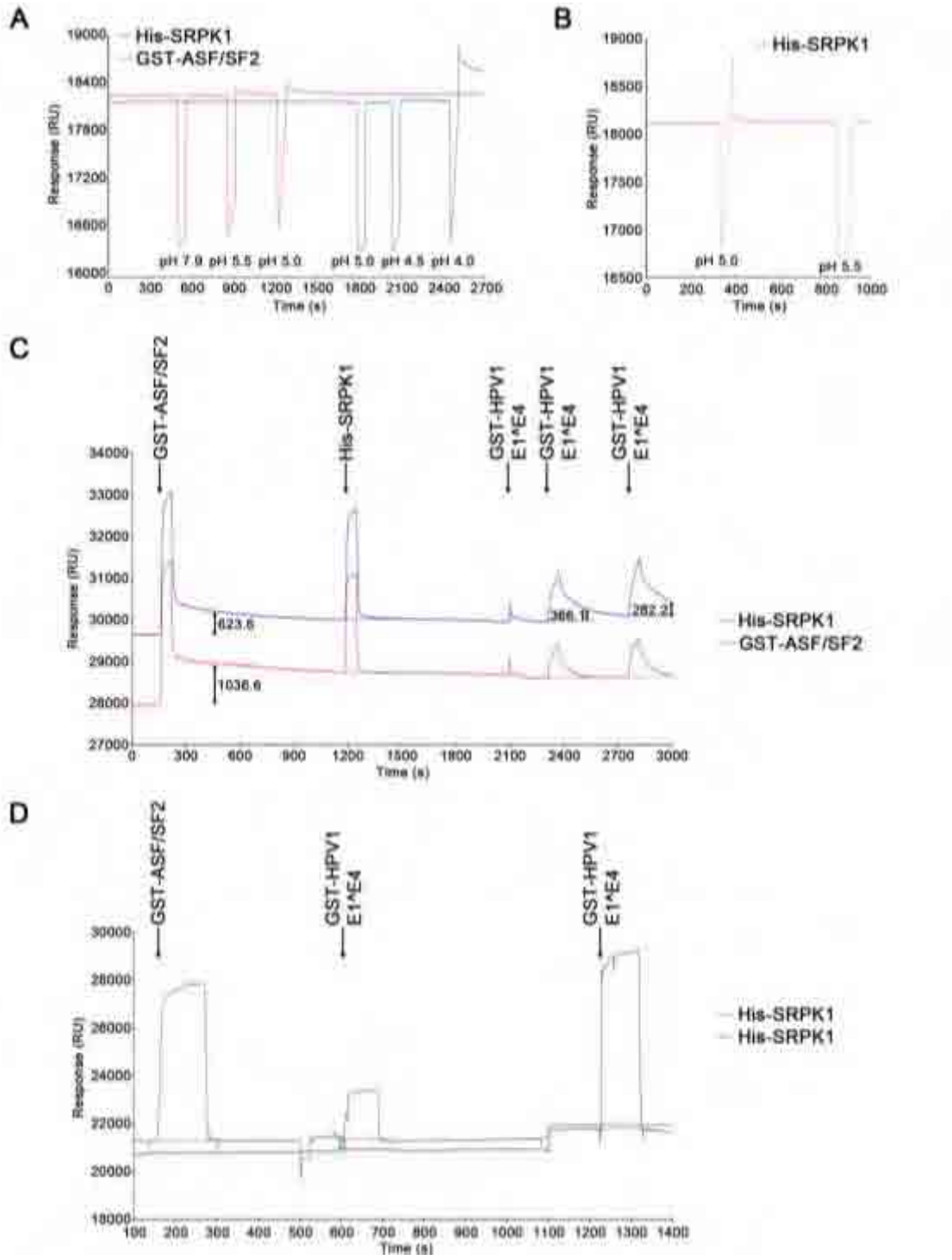

Figure 3.12 Biacore® interaction study using amine coupled proteins on a CM5 chip. (**A**) GST-ASF/SF2 was flowed over chamber 1 at pH 7.9, 5.5 and 5.0; His-SRPK1 was flowed over chamber 2 at pH 5.0, 4.5 and 4.0. (**B**) His-SRPK1 was passed through chamber 3 in the presence of 2 mM MgCl2 and 1 mM ATP at pH 5.0 and 5.5. (**C**) His-SRPK1 was bound at pH 4.0 and GST-ASF/SF2 bound at pH 5.0. One microgram of GST-ASF/SF2 was passed through the chambers, followed by 1 µg His-SRPK1, then 1 µg, 2 µg and 5 µg of GST-HPV1 E1^E4. (**D**) His-SRPK1 was bound at pH 5.0 in the presence of 2 mM MgCl2 and 1 mM ATP. Five micrograms of GST-ASF/SF2 was passed through chamber 1, followed by 5 µg of GST-HPV1 E1^E4. Twenty micrograms of GST-HPV1 E1^E4 was passed through chamber 2.

The failure of a range of strategies to use surface plasmon resonance analysis to determine kinetic information of the interaction between SRPK1 and the viral inhibitor HPV1 E1^E4 probably reflects a combination of complex kinetics of SRPK1 mediated phosphorylation and possibly a mix in the mode of inhibition exhibited. Therefore, due to time constraints the aim of obtaining information on the mode of binding was not pursued further.

3.4 HPV16 and HPV18 E1^E4 do not inhibit SRPK1 kinase activity

In addition to the cutaneous mu HPV1 virus, the E1^E4 proteins of the mucosatropic alpha HPV types HPV16 and HPV18 have also been shown to interact with SRPK1 (Bell et al., 2007). The interaction between His-SRPK1 and GST-tagged HPV16 and HPV18 E1^E4 was confirmed by "pull down" of the His-tagged SRPK1 from bacterial lysate. Western blotting with an anti-His or anti-SRPK1 antibody (Figure 3.13) showed that both GST proteins were able to interact with His-SRPK1. As noted previously (Bell et al., 2007), the GST-tagged HPV16 and HPV18 proteins were less effective at precipitation of His-SRPK1 than the type 1 protein.

To determine whether HPV16 and 18 E1^E4 proteins affect kinase activity, GST-tagged HPV16 and HPV18 E1^E4 proteins were added to an *in vitro* kinase assay containing GST-ASF/SF2 and His-SRPK1 (Figure 3.14). Whilst in the reactions with GST-HPV1 E1^E4, phosphorylation of ASF/SF2 decreased with increasing amounts of fusion protein, the level of phosphorylation in the presence of increasing amounts of the HPV16 or HPV18 E1^E4 proteins remained constant, suggesting that binding to SRPK1 does not inhibit its activity. Although phosphorylation increased slightly with HPV16 and HPV18 E1^E4 compared to the addition of GST alone, the difference was not significant ($p=0.36$ for HPV16 and $p=0.69$ for HPV18).

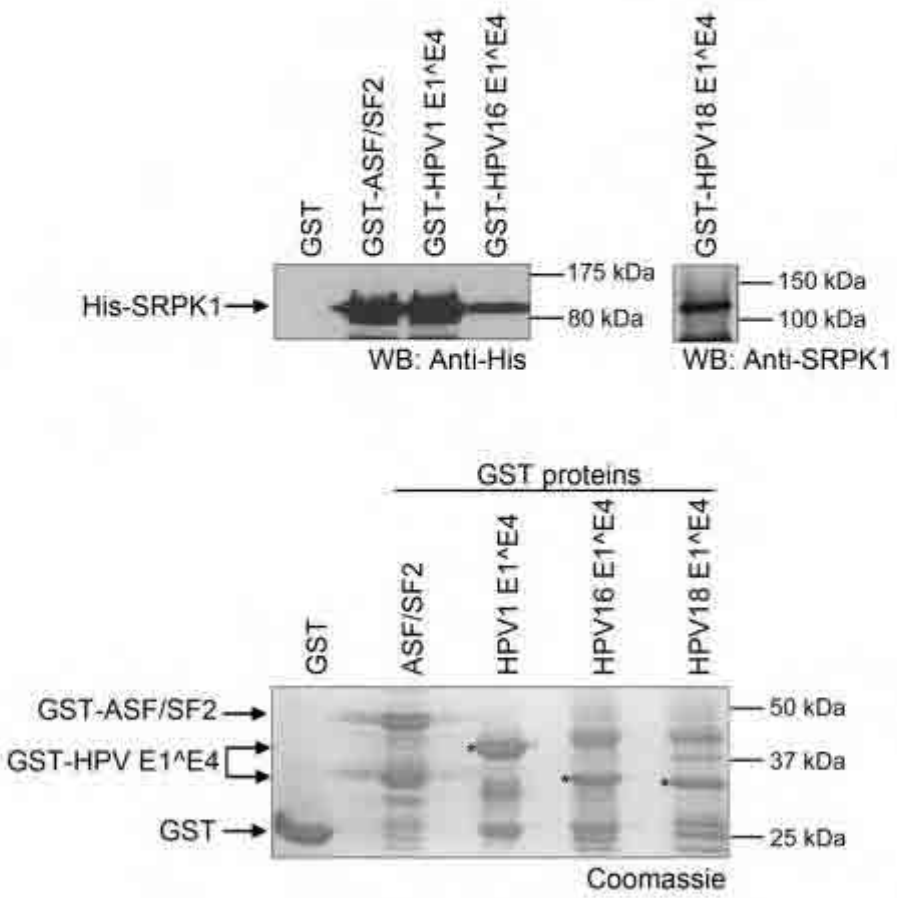

Figure 3.13 His-SRPK1 interacts with HPV type 16 and 18 E1^E4. Approximately 10 µg of GST and GST-tagged HPV1, 16 and 18 E1^E4 and ASF/SF2 proteins bound to glutathione agarose beads were mixed with bacterial lysate expressing His-SRPK1. Western blotting of the washed precipitates with anti-His or anti-SRPK1 antibodies showed that GST-HPV1, 16 and 18 E1^E4 proteins interacted with His-SRPK1 (GST-E1^E4 proteins marked with * in Coomassie). The data is representative of two independent experiments.

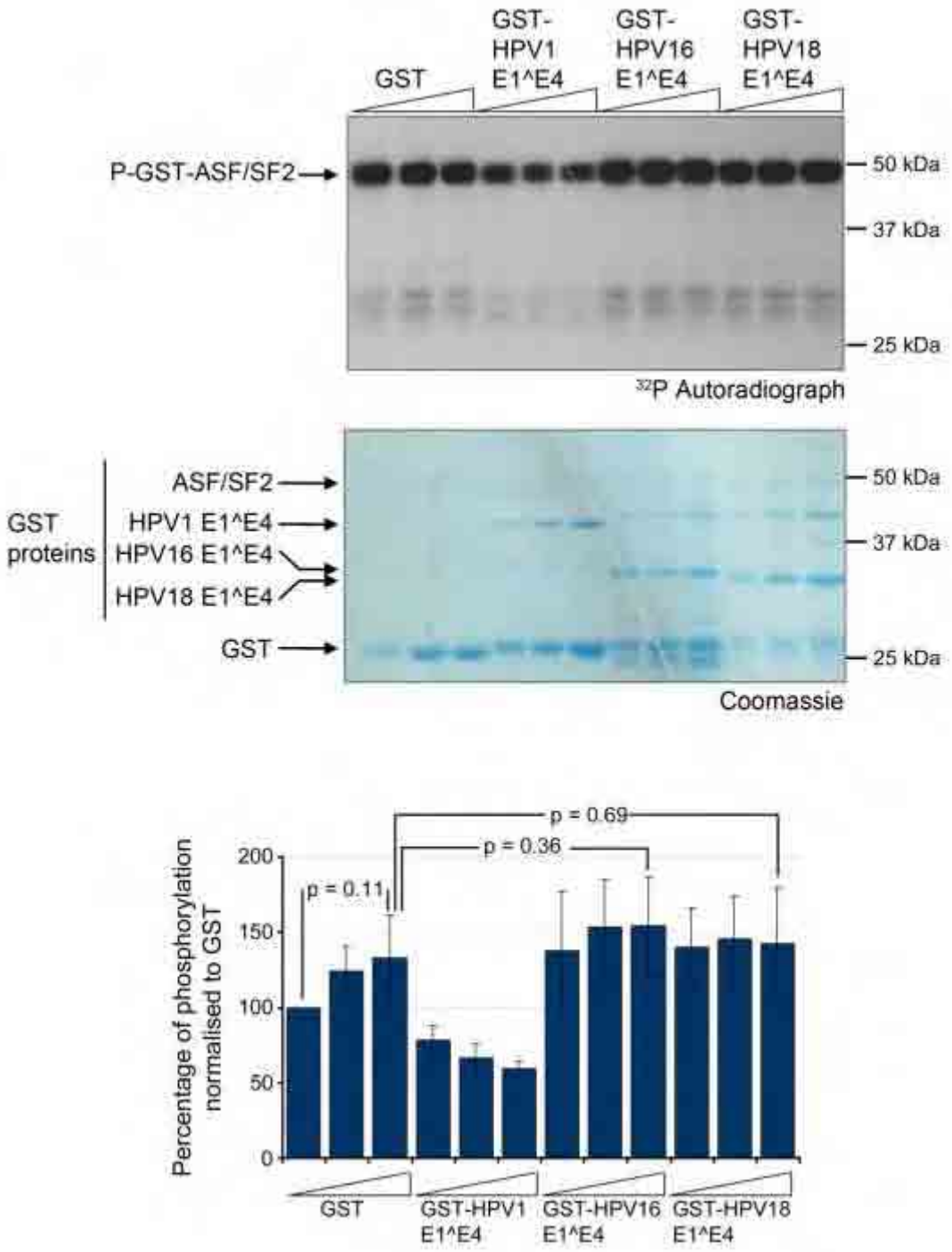

Figure 3.14 GST-HPV16 and GST-HPV18 E1^E4 proteins do not inhibit His-SRPK1 phosphorylation of GST-ASF/SF2. GST-tagged HPV1, 16 and 18 E1^E4 proteins were added in increasing amounts (2.5, 5 and 10 µg) to an *in vitro* kinase assay containing 3 µg GST-ASF/SF2 and 1 µg His-SRPK1. In comparison to the GST control, GST-HPV16 and 18 E1^E4 did not reduce the level of phosphorylation of GST-ASF/SF2. The data in the graph is the mean (± standard deviation) from four experimental replicates.

3.5 HPV5 E1^E4

The binding of SRPK1 to E1^E4 proteins of anogenital viruses HPV16 and HPV18 (α group) and of the cutaneous benign virus HPV1 (μ group), suggests that SRPK1 might be a conserved partner to E1^E4 proteins. To test whether SRPK1 interacts with E1^E4 from β-viruses the E1^E4 coding sequence of HPV5 was derived from the total HPV5 genome by PCR (see Chapter 2, Table 4). The cDNA of E1^E4 was ligated into the EcoRI restriction enzyme site of pGex-2T and following transformation into *E. coli* BL21 CodonPlus® cells, the fusion protein GST-HPV5 E1^E4 was produced. Following optimisation of expression conditions (Figure 3.15), a protein was produced of the expected size of 53 kDa and this was recognised by an anti-GST antibody. DNA sequencing of the E1^E4 cDNA showed that the E1^E4 sequence was most similar to HPV5b with only two amino acid changes to the published sequence (see Appendix 1A). For the purpose of this thesis, the HPV5b E1^E4 protein will be referred to as HPV5 E1^E4.

To determine whether HPV5 E1^E4 interacts with SRPK1, GST-HPV5 E1^E4 was bound to glutathione agarose beads, and mixed with His-SRPK1 for 10 min at 4°C. Western blotting of the washed precipitate with an anti-SRPK1 antibody (Figure 3.16) showed that GST-HPV5 E1^E4 did interact with the His-SRPK1. Binding to SRPK1 was observed in multiple experiments and consistently the HPV5 E1^E4 protein bound less SRPK1 than GST-HPV1 E1^E4, up to 50% less.

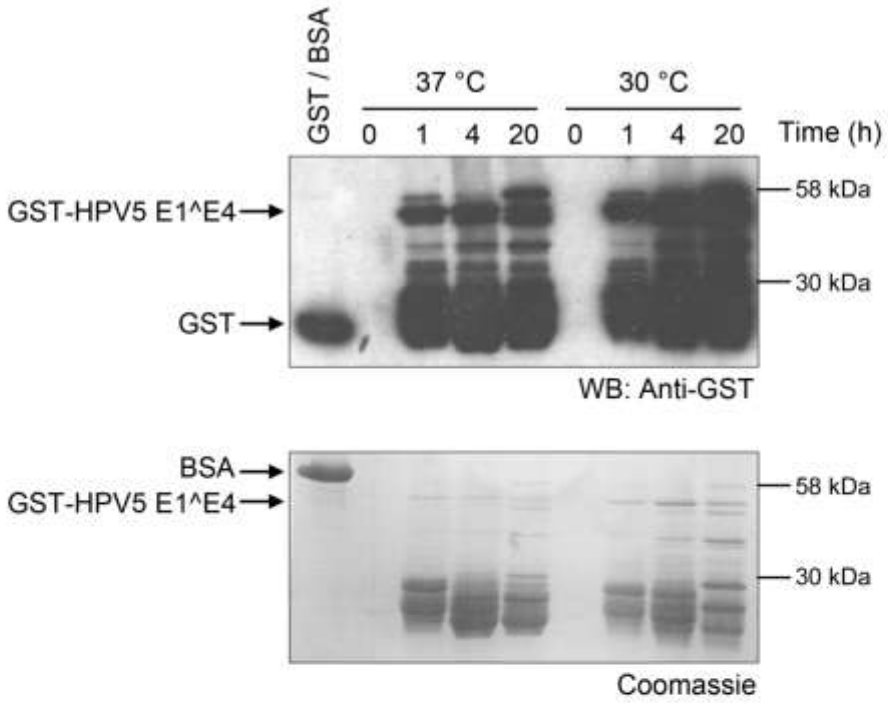

Figure 3.15 Expression of GST-HPV5 E1^E4. For optimisation of expression, E.coli BL21 CodonPlus® cells transformed with pGex-2T-HPV5 E1^E4 were grown at 37°C and 30°C for up to 20 h. The glutathione purified lysates were resolved by SDS PAGE and Coomassie stained or Western blotted using an anti-GST antibody.

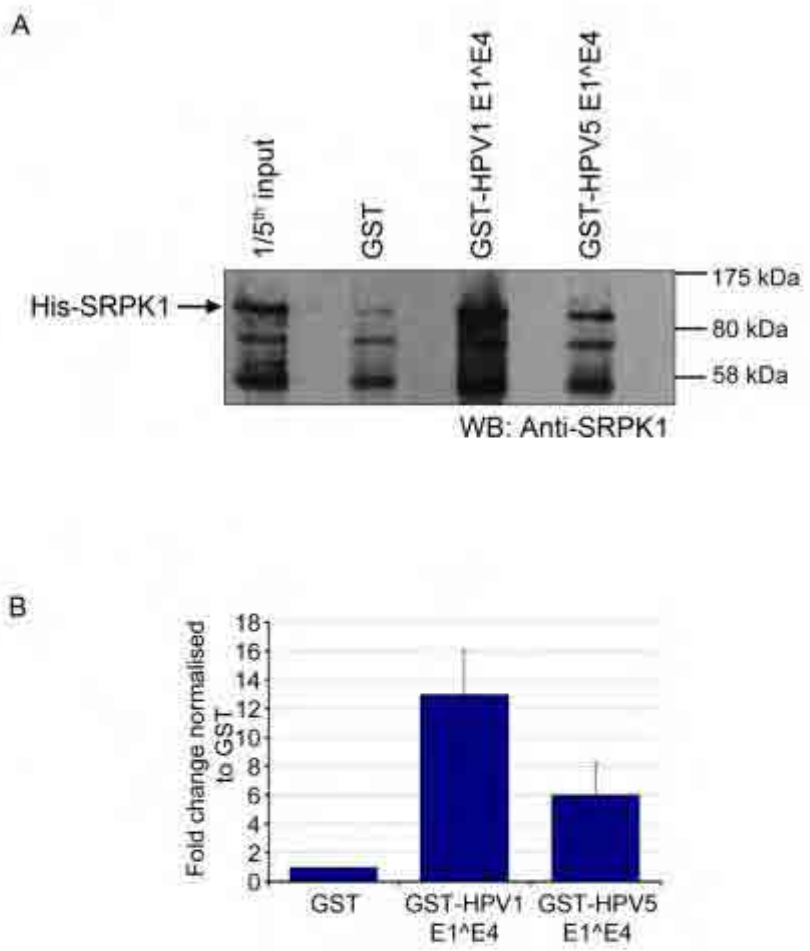

Figure 3.16 GST-HPV5 E1^E4 interacts with His-SRPK1. Approximately 10 µg of GST-HPV5 E1^E4 was bound to glutathione agarose beads, along with 10 µg of GST-HPV1 E1^E4 and GST, and mixed with 1 µg His-SRPK1 for 10 min. (**A**) Western blot of washed precipitates blotted with anti-SRPK1 antibody. (**B**) Densitometry plot of coprecipitated His-SRPK1 in western blot A, data is the average of three repeats.

3.5.1 HPV5 E1^E4 is a substrate for SRPK1 but does not affect kinase activity

To determine whether HPV5 E1^E4 affects SRPK1 kinase activity, GST-HPV5 E1^E4 protein was added to an *in vitro* kinase assay in increasing amounts with GST-ASF/SF2 as a substrate. Figure 3.17A shows that in comparison to the assays containing the GST protein the level of phosphorylation of GST-ASF/SF2 does not change significantly with increasing amounts of the GST-HPV5 E1^E4 protein. The observed small increase is not significant, ($p=0.095$). This data indicates that GST-HPV5 E1^E4 is not a potent inhibitor or activator of SRPK1 kinase activity.

HPV5 E1^E4 is however a potential substrate for SRPK1 as there are 2 RS repeats within the E1^E4 coding sequence (Appendix 1B). When GST-HPV5 E1^E4 is added only in the presence of His-SRPK1 a phospho-species of approximately 40 kDa is present (Figure 3.17B). This phospho-species is also present in the inhibitor assays in reactions containing the GST-HPV5 E1^E4 protein but not in the reactions containing the HPV1 fusion protein or GST control (Figure 3.17A). This species is smaller than the full-length GST-HPV5 E1^E4 of 53 kDa, but could represent phosphorylated E1^E4 that has arisen from proteolytic breakdown of the full-length species and is more susceptible to phosphorylation or indicates that that phospho-species exhibits increased migration upon SDS-PAGE. Interestingly, a similar increase in migration of the phospho-species was observed following *in vitro* SRPK1 phosphorylation of the GST-HPV1 E1^E4 protein (Bell et al., 2007).

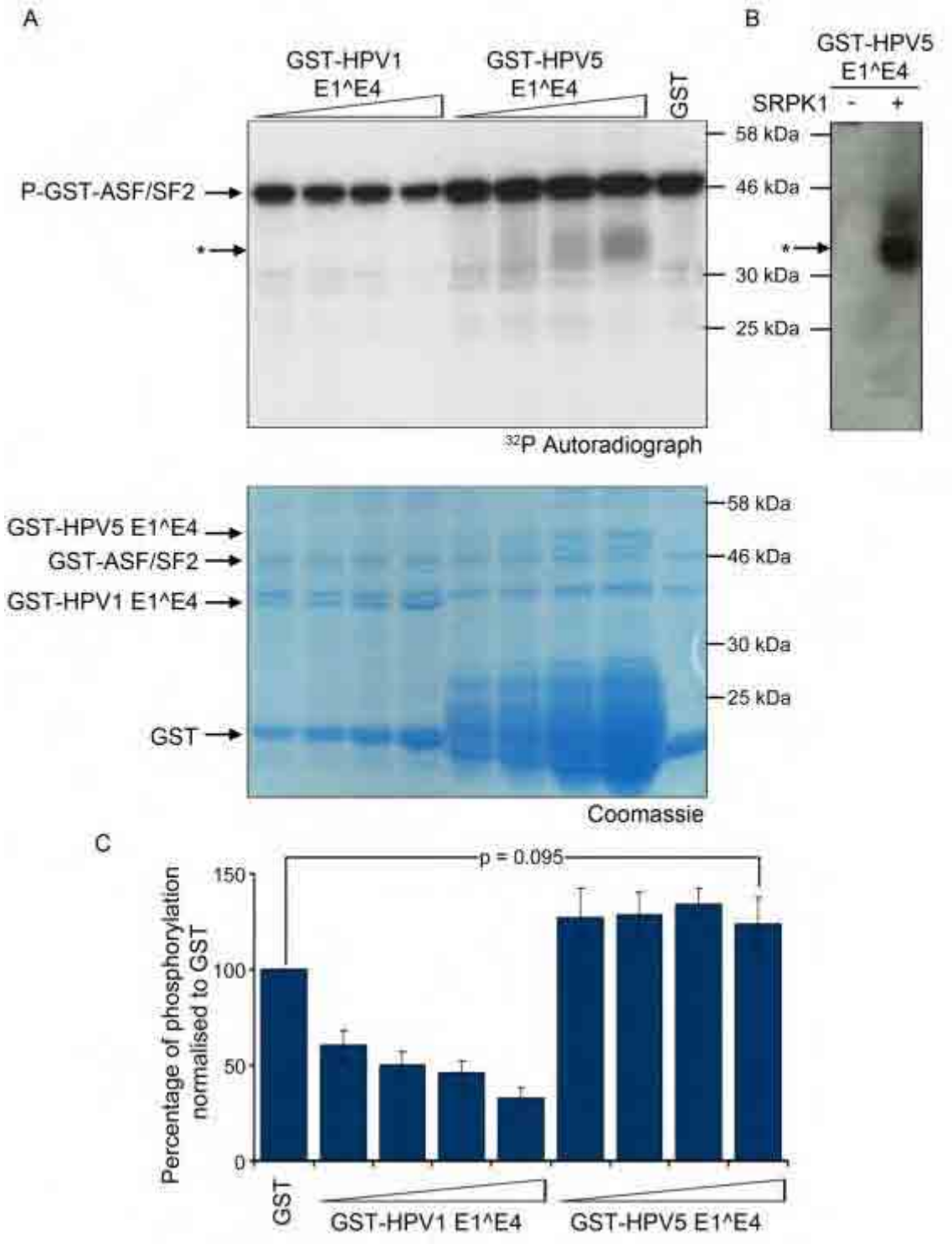

Figure 3.17 HPV5 E1^E4 is a substrate but not an inhibitor for SRPK1. (**A**) GST-HPV5 E1^E4 (1.25, 2.5, 5 and 10 µg) was added to an *in vitro* kinase assay containing 2 µg GST-ASF/SF2 and 1 µg His-SRPK1, GST-HPV1 E1^E4 (1.25, 2.5, 5 and 10 µg) and GST (10 µg) were added as controls. (**B**) GST-HPV5 E1^E4 (2.5 µg) in the presence and absence of 1 µg His-SRPK1. The band marked with the * is a phospho-species which is present only in the presence of GST-HPV5 E1^E4 and His-SRPK1, and also increases in intensity with increasing amounts of GST-HPV5 E1^E4. (**C**) Bar graph showing the mean (± standard deviation) percentage phosphorylation of GST-ASF/SF2 normalised to GST from three independent experiments.

3.6 Discussion

This study has shown for the first time that the interaction between HPV1 E1^E4 and SRPK1 causes the inhibition of phosphorylation of ASF/SF2 by SRPK1 *in vitro*. The maximum level of inhibition observed was 65%, the lack of total inhibition may be due to HPV1 E1^E4 sometimes binding to SRPK1 in a manner which does not inhibit SRPK1 activity. HPV1 E1^E4 must bind to SRPK1 for inhibition to occur, as a non-binding deletion mutant of HPV1 E1^E4 is unable to inhibit SRPK1 activity. The inhibition of SRPK1 activity was not however a conserved function of the other E1^E4 proteins of HPV5, HPV16 and HPV18 that bind to SRPK1. HPV1 and HPV5 are specific for cutaneous epithelia, whereas HPV16 and HPV18 are mucosaspecific, in cutaneous infections the late stages of the life cycle are induced as soon as the cell leaves the basal layer, whereas in mucosal infections the vegetative stages occur in the upper layers, due to the activity of E6 and E7. In the cutaneous infections the modulation of SRPK1 activity by E1^E4 may be required to alter cellular factors which are no longer present in the upper layers when the mucosal viruses are undergoing viral amplification. Alternatively the E1^E4 proteins of HPV16 and HPV18 may require additional cellular factors or post-translational modifications to modulate SRPK1 activity. Interestingly in the *in vitro* kinase assays, HPV5, HPV16 and HPV18 increase the activity of SRPK1; although this is not significant it may indicate that *in vivo* these proteins influence SRPK1 function by an alternative method.

The mechanism of inhibition of SRPK1 kinase activity by HPV1 E1^E4 was investigated using both kinetic and binding studies, however due to the processitivity of the kinase reaction and possible steric hindrances this was not successful. The sterically specific nature of the interaction between SRPK1 and other proteins suggests that the two domains of SRPK1 must be able to freely move and change conformation during the binding process. It may be possible to determine whether HPV1 E1^E4 is a competitive or non-competitive inhibitor by

varying the amount of ATP added to the *in vitro* kinase assay. This experiment would allow for a kinetic determination that was not affected by the processitivity of the reaction and has been utilised by Fukuhara et al., (2006) to demonstrate that SRPIN340 is a competitive inhibitor of SRPKs.

Whilst the mechanism of inhibition was not resolved here the interaction between HPV1 E1^E4 and SRPK1 was sufficient to prevent ASF/SF2 binding and therefore HPV1 E1^E4 could be binding to a region of SRPK1 such as the docking groove (Figure 1.12). HPV1 E1^E4 contains a bipartate arginine-rich motif and deletion of residues 44-48 (^{44}GRPRR^{48}) abrogates the binding of HPV1 E1^E4 to SRPK1 (Bell et al., 2007). This arginine-rich motif is similar to the ASF/SF2 docking motif and the consensus motif R-X-R/K-X-X-X-R identified by Ngo et al., (2005). Huynh et al., (2009) have shown that if the RS domain of ASF/SF2 does not initially bind to the docking groove, the RRMH domain cannot bind to SRPK1, demonstrating that ASF/SF2 requires multiple interactions with SRPK1 to be in the correct configuration for phosphorylation. The strict requirements for binding and phosphorylation may account for some of the difficulties encountered when investigating the interaction using Biacore^{TM}.

The HPV1 E1^E4 protein has been shown to be a substrate for SRPK1 *in vitro* by Bell et al., (2007), however in the experiments performed in this thesis, this phosphorylation was not observed. The amounts of SRPK1 and HPV1 E1^E4 used in the assays reported here were much lower than those used by Bell et al., (2007) due to ASF/SF2 being a superior substrate for SRPK1 than HPV1 E1^E4. In addition to HPV1 E1^E4 being a substrate for SRPK1, the E1^E4 protein of HPV5 has been shown to be a substrate for SRPK1 *in vitro* in this study; this phosphorylation may be required to modulate the activity of these proteins. SRPK1 and SRPK2 have been shown to phosphorylate the core protein of hepatitis B virus (HBV) (Daub et al., 2002) and this phosphorylation is required for multiple steps in the DNA synthesis of

HBV (Lewellyn and Loeb, 2011). Phosphorylation of three serines in the C-terminal domain of the core protein is required for efficient minus-strand DNA elongation, primer translocation, circularisation and plus-strand DNA elongation (Lewellyn and Loeb, 2011). In contrast, SRPK1 and SRPK2 have been shown to suppress viral replication of HBV by interfering with pre-genomic RNA packaging without affecting the formation of the viral core particles, this suppressive effect was not dependent upon the kinase activity of SRPK1/2 (Y. Zheng et al., 2005), demonstrating that SRPKs have multiple roles in HBV replication and this may be true for HPV too. Identification of the site of phosphorylation within the E1^E4 proteins might be helpful in understanding the consequence of SRPK1 mediated phosphorylation. Perhaps phosphorylation facilitates the interaction between E4 and cellular and/or viral partners and thus modulates their behaviour, for example the interaction between HPV16 E1^E4 and E2, modulates the functions of E2 (C. Davy et al., 2009).

CHAPTER 4 INTERACTION OF SRPK1 WITH CELLULAR AND VIRAL SR PROTEINS

4.1 Introduction

The SR proteins are a family of related proteins and although all members of the family contain one or two conserved RRMs and a C-terminal RS domain there are differences between family members in their functions and localisation (Chapter 1, Table 2). The SR proteins ASF/SF2, SRp20 and 9G8 have been shown to shuttle rapidly and continuously between the cytoplasm and nucleus, whereas SRp75 can shuttle with slower kinetics and SC35 does not shuttle (Cazalla et al., 2002, Sapra et al., 2009).

SRp20 is the smallest SR protein and contains a single RRM, as a shuttling SR protein, SRp20 has been shown to be involved in mRNA export and has been shown to promote mRNA translation (Hautbergue et al., 2008, Swartz et al., 2007). In HPV16 and BPV infection, SRp20 has been shown to interact with an A/C-rich element located upstream of the late 3' splice site, and regulate the selection of late splice sites, thus regulating the expression of late proteins (Jia et al., 2009). Interestingly in BPV, the A/C-rich element (SE4) only binds SRp20, whereas in HPV16, the A/C-rich ESE binds multiple SR proteins including SRp20, ASF/SF2 and SRp75 (Jia et al., 2009). SC35 is a non-shuttling SR protein which contains a single RRM and SC35 has been shown to be involved in transcriptional elongation in a gene-specific manner, allowing the coupling of transcription and splicing (Lin et al., 2008). 9G8 is a shuttling SR protein, with a single RRM, however 9G8 also contains a zinc knuckle. 9G8 has been shown to be involved in mRNA export and the promotion of mRNA translation (Hautbergue et al., 2008, Swartz et al., 2007). SRp75 is the largest SR protein, with two

RRMs and a long RS domain. Hyperphosphorylation of SRp75 by CLKs has been shown to modulate the alternative splicing of Adenovirus E1A (Yomoda et al., 2008).

The prototypical SR protein, ASF/SF2, has two RRMs, with the RRMH domain interacting with SRPK1 during phosphorylation to stabilise the complex (Ngo et al., 2008). Whether the RRM of SRp20 and SC35 interacts with SRPK1 is undetermined, but SRPK1 may bind to these proteins in a different configuration and therefore may be more or less susceptible to inhibition of binding.

There are many proteins, including multiple snRNP components and splicing regulators and coactivators which contain RS/SR dipeptides and can interact with the SR proteins via their RS domains and/or be phosphorylated by SRPK1. Included in this class of proteins are the E2 proteins of the beta HPV types associated with EV. They contain a relatively long hinge region, of around 200 amino acids, between the conserved transactivation and DNA binding domains. Within the hinge region of these proteins are multiple RS/SR dipeptides and HPV5 and HPV8 E2 have been shown to interact with several SR proteins and to colocalise with SC35 in nuclear speckles, suggesting that HPV5 E2 may have a role in pre-mRNA splicing (Lai et al., 1999, Sekhar et al., 2009). The accumulation of HPV5 and HPV8 E2 in nuclear speckles is dependent upon the hinge region of the proteins, but the hinge alone is not sufficient for this localisation (Lai et al., 1999, Sekhar et al., 2009). The hinge region of HPV5 E2 is extensively phosphorylated *in vivo* by SRPK1 and the overexpression of SRPK1 leads to accumulation of HPV5 E2 in the cytoplasm similar to the relocalisation observed with ASF/SF2 (Lai et al., 2000). The hinge region of HPV5 E2 can interact with the SR transport protein TRN-SR2 and this interaction is dependent upon phosphorylation of the RS dipeptides (Lai et al., 2000).

The binding of SRPK1 to the SR proteins with a single RRM may differ to those with two, therefore the inhibition caused by HPV1 E1^E4 may be substrate specific. Additionally as

certain SR proteins have been shown to have specific functions during viral infections, HPV1 may be purposely targeting one or more of these proteins. This chapter examines the substrate specificity of HPV1 E1^E4 inhibition of SRPK1 kinase activity, using cellular and viral SR proteins as substrates in *in vitro* kinase assays.

4.2 Results

4.2.1 Investigation of HPV E1^E4 action on SRPK1 phosphorylation of cellular SR proteins other than ASF/SF2

4.2.1.1 *In vitro* SRPK1 phosphorylation of cellular SR proteins

The effect on kinase activity of the interaction between SRPK1 and E1^E4 have been examined with respect to the prototypical SR protein, ASF/SF2. However, the effects observed could be substrate specific and therefore it is necessary to examine the influence of E1^E4 on SRPK1 phosphorylation of other SR proteins. To investigate this, the following SR proteins were chosen to express as GST-fusion proteins; SC35, 9G8, SRp75 and SRp20.

The intention was to express the SR proteins as GST fusion proteins, therefore the IMAGE clones were obtained from Source BioScience LifeSciences and following amplification and restriction digestion the cDNAs were ligated into complementarily digested pGex-2T (see Chapter 2, Table 4 for IMAGE clone details and primers). The GST-tagged SR proteins were expressed in *E.coli* BL21 CodonPlus® cells and following optimisation of expression conditions they were resolved by SDS PAGE (Figure 4.1).

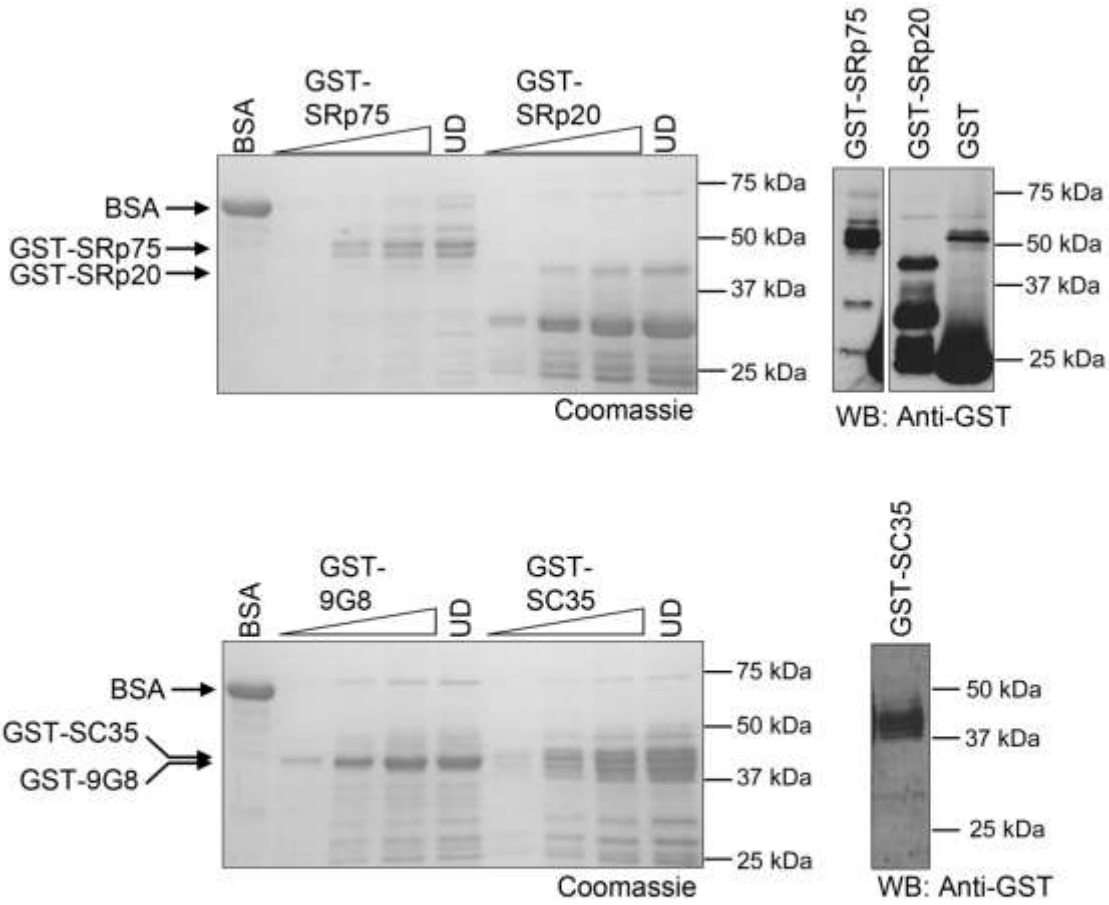

Figure 4.1 Expression of GST-tagged SR proteins. Coomassie stained gel showing the purified GST-tagged SR proteins: SC35, 9G8, SRp20 and SRp75 in increasing amounts. UD, sample taken before dialysis. BSA, aliquot of bovine serum albumin 5 µg, allowing for relative amounts of the SR proteins to be calculated. Western blots show that GST-tagged SC35, SRp20 and SRp75 were recognised by an anti-GST antibody.

All of the GST-SR fusion proteins migrated with masses smaller than the predicted sizes. GST-SC35, GST-9G8 and GST-SRp20 all produced products with masses of between 40 and 45 kDa, which were smaller than the expected sizes of 53 kDa, 55 kDa and 47 kDa respectively. The product from the expression of GST-SRp75 was of a mass of 50 kDa (expected 85 kDa), but this product was recognised by an anti-SRp75 (SFRS4) antibody (data not shown). Cleavage of some of the RS domain at the C-terminus may have occurred, as the GST-tag is at the N-terminus and the anti-SRp75 antibody recognises an epitope at the start of the RS domain.

To ensure that the GST-tagged SR proteins were therefore functional each one was added to an *in vitro* kinase assay in the presence and absence of His-SRPK1. In the autoradiograph shown in Figure 4.2 each of the GST-tagged SR proteins were phosphorylated by His-SRPK1 *in vitro*.

4.2.1.2 Effect of HPV1, 5, 16 and 18 E1^E4 on SRPK1 phosphorylation of cellular SR proteins

To test whether the inhibition of SRPK1 kinase activity *in vitro* by HPV1 E1^E4 is substrate specific, GST-HPV1 E1^E4 was added into an *in vitro* kinase assay with the GST-tagged SR proteins and His-SRPK1. As Figure 4.3 shows, GST-HPV1 E1^E4 inhibits His-SRPK1 phosphorylation of all the GST-tagged SR proteins. In comparison to phosphorylation of the SR proteins in the presence of GST, the level of inhibition by HPV1 E1^E4 was significant in all cases; GST-ASF/SF2 $p = 0.006$; GST-SRp75 $p = 0.0002$; GST-SRp20 $p = 0.02$ and GST-9G8 $p = 0.002$. It can be concluded that HPV1 E1^E4 inhibits SRPK1 phosphorylation of multiple substrates. However the potency of inhibition varied between substrates, ranging from 58% inhibition with GST-SRp20 to 86% inhibition with GST-SRp75.

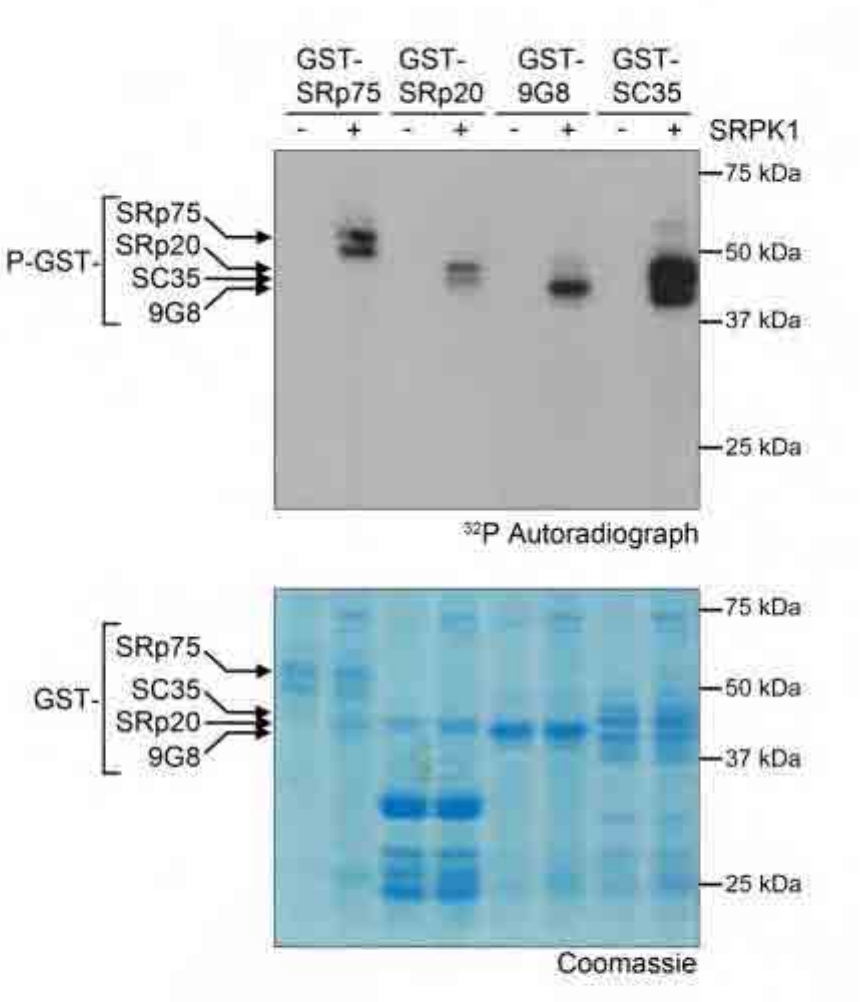

Figure 4.2 Phosphorylation of GST-SR proteins by His-SRPK1. GST-tagged SRp75 (2 µg), SRp20 (2 µg), 9G8 (4 µg) and SC35 (4 µg) were added to an *in vitro* kinase assay in the presence and absence of (1 µg) His-SRPK1.

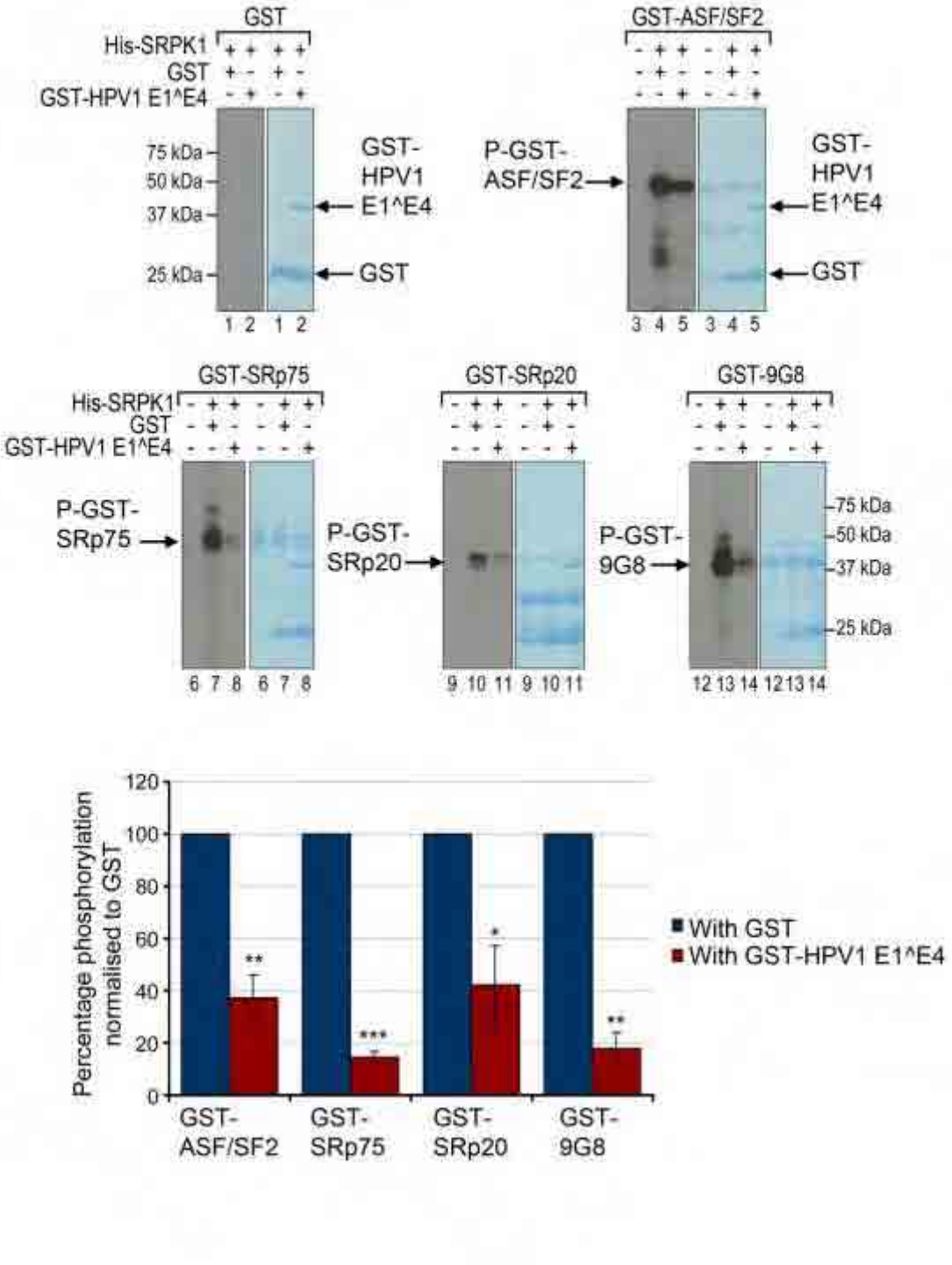

Figure 4.3 Inhibition of SRPK1 kinase activity by HPV1 E1^E4. Five micrograms of GST (lanes 1-2), GST-ASF/SF2 (lanes 3-5), GST-SRp75 (lanes 6-8), GST-SRp20 (lanes 9-11) and GST-9G8 (lanes 12-14) were added to an *in vitro* kinase assay with 1 µg His-SRPK1. Five micrograms of GST-HPV1 E1^E4 (lanes 2, 5, 8, 11 and 14) was added as an inhibitor and 5 µg of GST (lanes 1, 4, 7, 10 and 13) added as a control. The data in the graph shown is the mean ± standard deviation of three experimental replicates. P values: * <0.05, ** < 0.01, *** < 0.0005.

SRPK1 phosphorylation of several different SR protein substrates was inhibited by HPV1 E1^E4, indicating that there is no substrate specificity with regards to the inhibitory function of HPV1 E1^E4. However, whilst HPV16 and HPV18 E1^E4 do not inhibit SRPK1 phosphorylation of the shuttling SR protein ASF/SF2, it is possible that their binding to SRPK1 may inhibit the phosphorylation of other substrates. Therefore the effect of HPV16 and HPV18 E1^E4 proteins on SRPK1 phosphorylation of SC35 was investigated; SC35 was used in this assay as it is a good substrate *in vitro* for His-SRPK1 and as SC35 only contains a single RRM, the binding to SRPK1 may be different to that with ASF/SF2. GST-tagged HPV16 and HPV18 E1^E4 were added in increasing amounts to an *in vitro* kinase assay with GST-SC35 and His-SRPK1. Figure 4.4 shows that the GST-tagged E1^E4 proteins of HPV16 (lanes 5 to 7) and HPV18 (lanes 8 to 10) do not inhibit His-SRPK1 phosphorylation of GST-SC35. There was a small increase in the level of phosphorylation when compared to the reaction containing the GST control but this was not significant ($p = 0.41$ for HPV16; $p = 0.50$ for HPV18).

To confirm that the binding of GST-HPV1 E1^E4 to His-SRPK1 is required for the inhibition of phosphorylation of SC35, the deletion mutants of HPV1 E1^E4, Δ44-48 (an SRPK1 non-binding mutant) and Δ49-53 (an SRPK1 binding mutant) were added to the *in vitro* kinase assay with GST-SC35. As shown in Figure 4.4, the wild type HPV1 E1^E4 inhibited the phosphorylation of GST-SC35 ($p = 0.05$) (lanes 2 to 4), whereas the deletion mutant of HPV1 E1^E4, Δ44-48 which does not bind to SRPK1 did not (lane 12). As expected, addition of the deletion mutant Δ49-53 which retains SRPK1-binding function, significantly reduced the level of phosphorylation of GST-SC35 ($p = 0.005$) (lane 11).

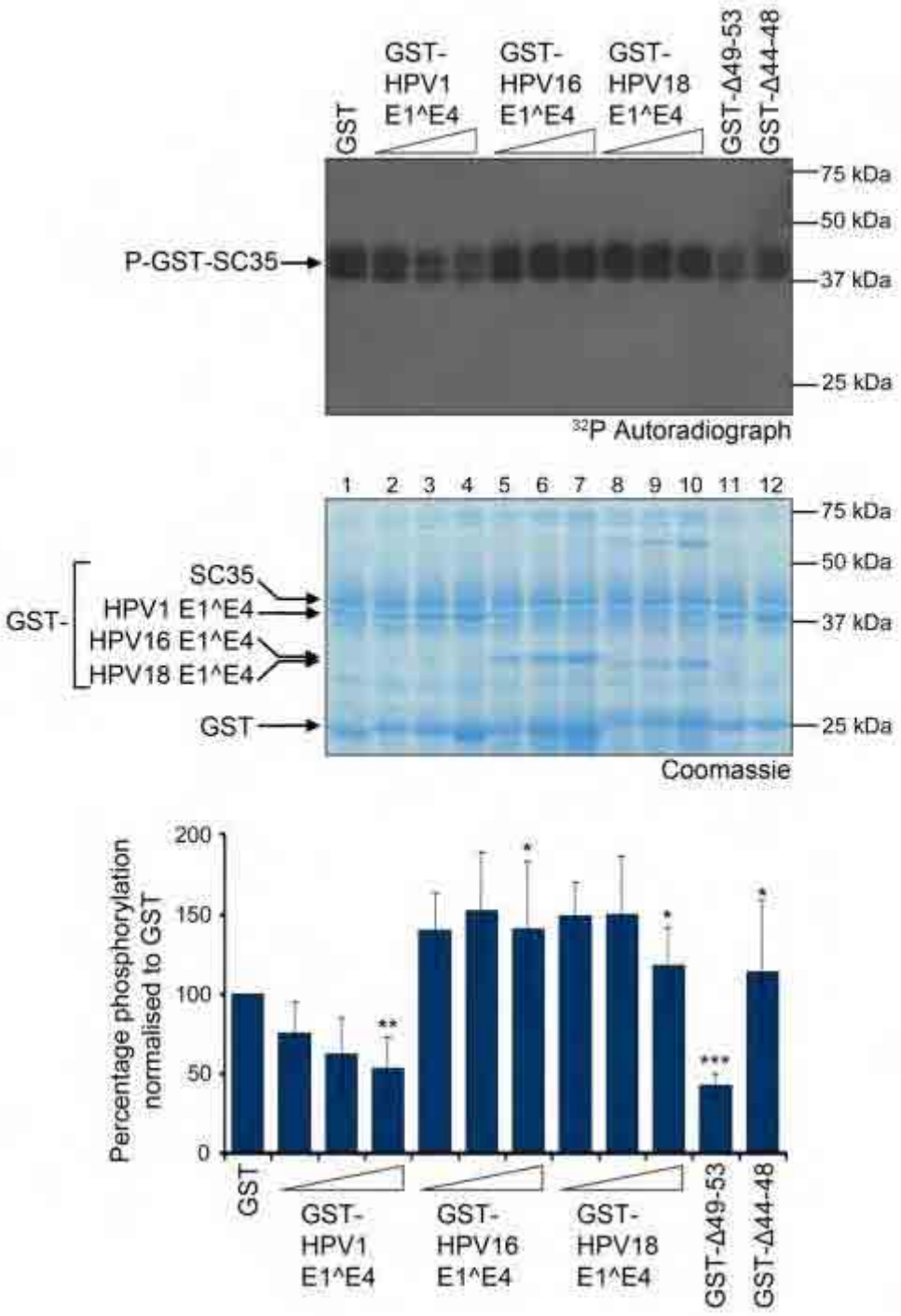

Figure 4.4 HPV16 and HPV18 E1^E4 do not inhibit SRPK1 phosphorylation of SC35. GST-tagged HPV16 and HPV18 E1^E4 along with HPV1 E1^E4 were added in increasing amounts (2.5, 5 and 10 µg) to an *in vitro* kinase assay with 1 µg His-SRPK1 and 5 µg GST-SC35, 10 µg of GST-tagged deletion mutants of HPV1 E1^E4 Δ44-48 and Δ49-53 were added along with 10 µg GST. The data in the bar graph is the mean ± standard deviation of three experimental replicates. P values: * > 0.1, ** < 0.05, *** < 0.01

GST-HPV5 E1^E4 was added to an *in vitro* kinase assay with GST-SRp20 and GST-SC35 as substrates and no detectable inhibition of SRPK1 kinase activity was observed (data not shown). GST-HPV5 E1^E4 produces a phospho-species (Chapter 3, section 3.5.1) which migrates at around 40 kDa, this phospho-species migrates at around the same size as the phosphorylated products produced by GST-SRp20 and GST-SC35 making analysis difficult, therefore no further experiments were performed using these substrates with HPV5 E1^E4.

4.2.2 Effect of E1^E4 on SRPK1 phosphorylation of viral SR proteins

The E2 proteins of the beta papillomaviruses have a large number of RS or SR dipeptide repeats in the hinge region (Figure 4.5A), HPV5 E2 has 27 repeats and HPV8 E2 has 24 (Figure 4.5B), making these viral proteins substrates for SRPK1 (Lai et al., 2000, W. S. Wang et al., 2009c). The interaction between HPV E1^E4 proteins and SRPK1 could modulate the phosphorylation of other HPV proteins by impinging on SRPK1 activity in differentiating cells.

To determine whether these E2 proteins were substrates for His-SRPK1 *in vitro*, GST-fusions of HPV5 and HPV8 E2 proteins were produced. In addition, the hinge regions (E2H) of both HPV types were also made into GST-fusion proteins (Chapter 2, Table 4). Since HPV1 E2 has 4 dipeptide repeats within the hinge region, the fusion proteins for HPV1 E2 and E2H were also produced (Figure 4.5B). Following ligation of the E2 and E2H cDNAs into pGex-2T, transformation into *E. coli* BL21 CodonPlus® cells and optimisation of expression conditions, the GST-fusion proteins were produced and resolved by SDS PAGE (Figure 4.5C). The HPV1 E2 and E2H constructs produced proteins of the expected sizes of 72 kDa and 42 kDa respectively. The expressed fusion proteins of HPV5 and HPV8 full length E2 were of smaller mass (55 kDa and 65 kDa) than the expected sizes of 86 kDa and 84 kDa respectively.

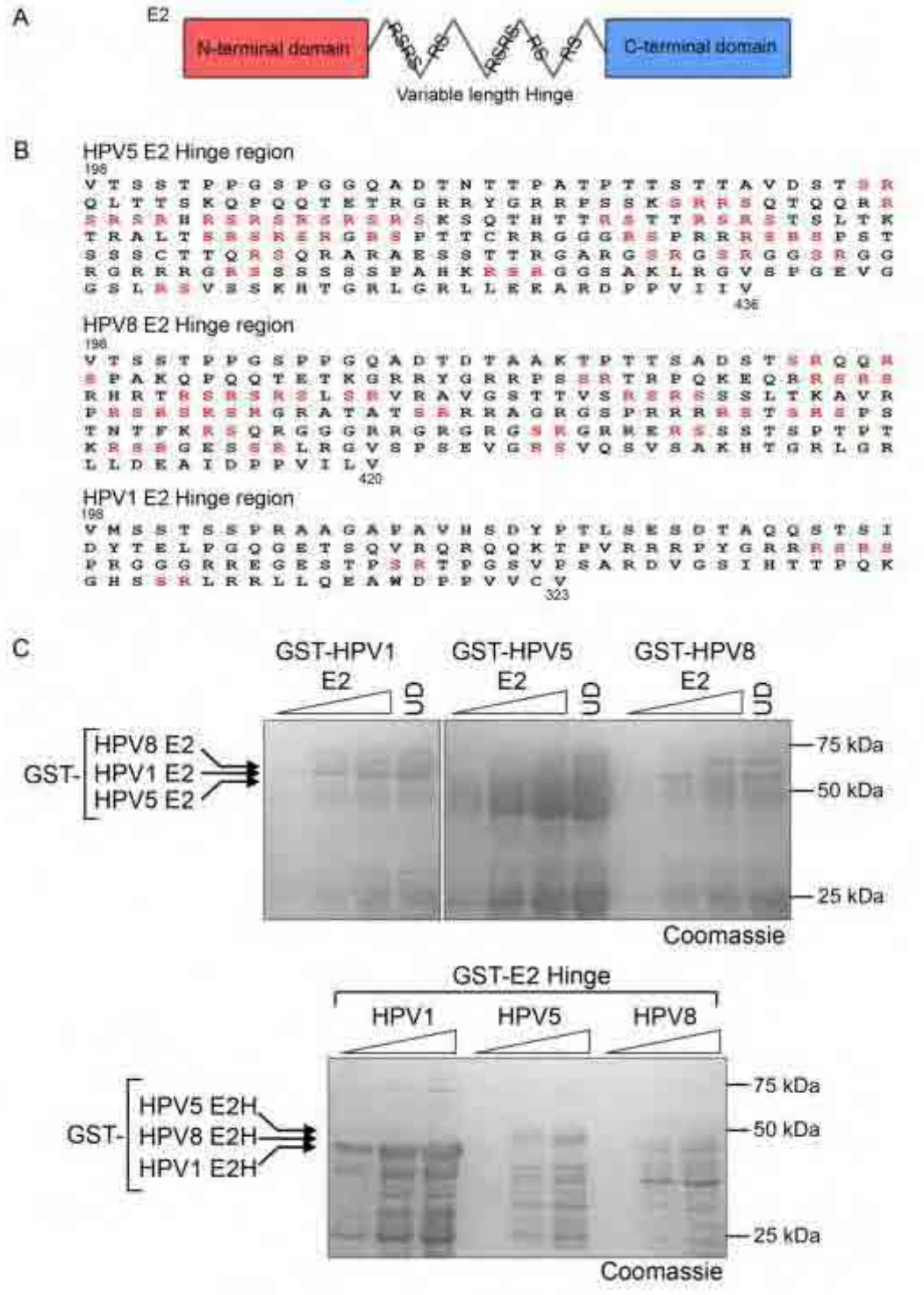

Figure 4.5 Expression of GST-tagged HPV1, HPV5 and HPV8 E2 and E2 Hinge. (**A**) Schematic of the HPV E2 protein of β viruses. (**B**) Amino acid sequence of HPV5 E2 hinge region, amino acids 198 to 436; HPV8 E2 hinge region, 198 to 420 and HPV1 E2 hinge region, 198 to 323. SR and RS dipeptides are shown in red. (**C**) Coomassie gels showing resolved GST-tagged E2 and E2H proteins of HPV1, HPV5 and HPV8 in increasing amounts and a sample taken before dialysis (UD).

The fusion proteins produced by the E2H constructs were of around the correct sizes of 54 kDa and 53 kDa. Protein species considerably smaller than the expected proteins were observed with all of the GST-E2 and E2H constructs, this may be due to proteolytic cleavage of the proteins or loss of the GST-tag during the purification steps.

To demonstrate that the E2 proteins and E2 hinge regions were substrates for His-SRPK1, the GST-tagged E2 and E2H proteins were added to an *in vitro* kinase assay with His-SRPK1. Figure 4.6 shows that GST-tagged HPV5 and HPV8 E2H were substrates for His-SRPK1 *in vitro*, GST-HPV1 E2H was phosphorylated by SRPK1, but only to a small extent, so could only be seen after a long exposure. These levels of phosphorylation are as predicted by the variation in the number of RS repeats in the proteins (Figure 4.5B). The GST-tagged full length E2 proteins of HPV1 and HPV5 were phosphorylated by SRPK1 to a similar extent as the corresponding E2H proteins. GST-HPV8 E2 however does not appear to be phosphorylated very efficiently by SRPK1 *in vitro* when compared to the E2H protein; this was possibly due to the folding of the protein being incorrect.

To test whether HPV1 E1^E4 inhibits SRPK1 kinase activity with viral substrates, GST-HPV1 E1^E4 was added along with the E2 and E2H GST-fusion proteins (Figure 4.7). The addition of GST-HPV1 E1^E4 reduced the level of phosphorylation of all of the GST-tagged E2 and E2H proteins compared to GST. The reduction in phosphorylation was significant with HPV1 and HPV5 E2H ($p = 0.007$ and 0.006 respectively) but was not significant with HPV8 E2H ($p = 0.14$). The viral substrates, like the cellular SR protein substrates, show variation in the potency of HPV1 E1^E4 as an inhibitor of SRPK1, with inhibition varying from 19% to 51%.

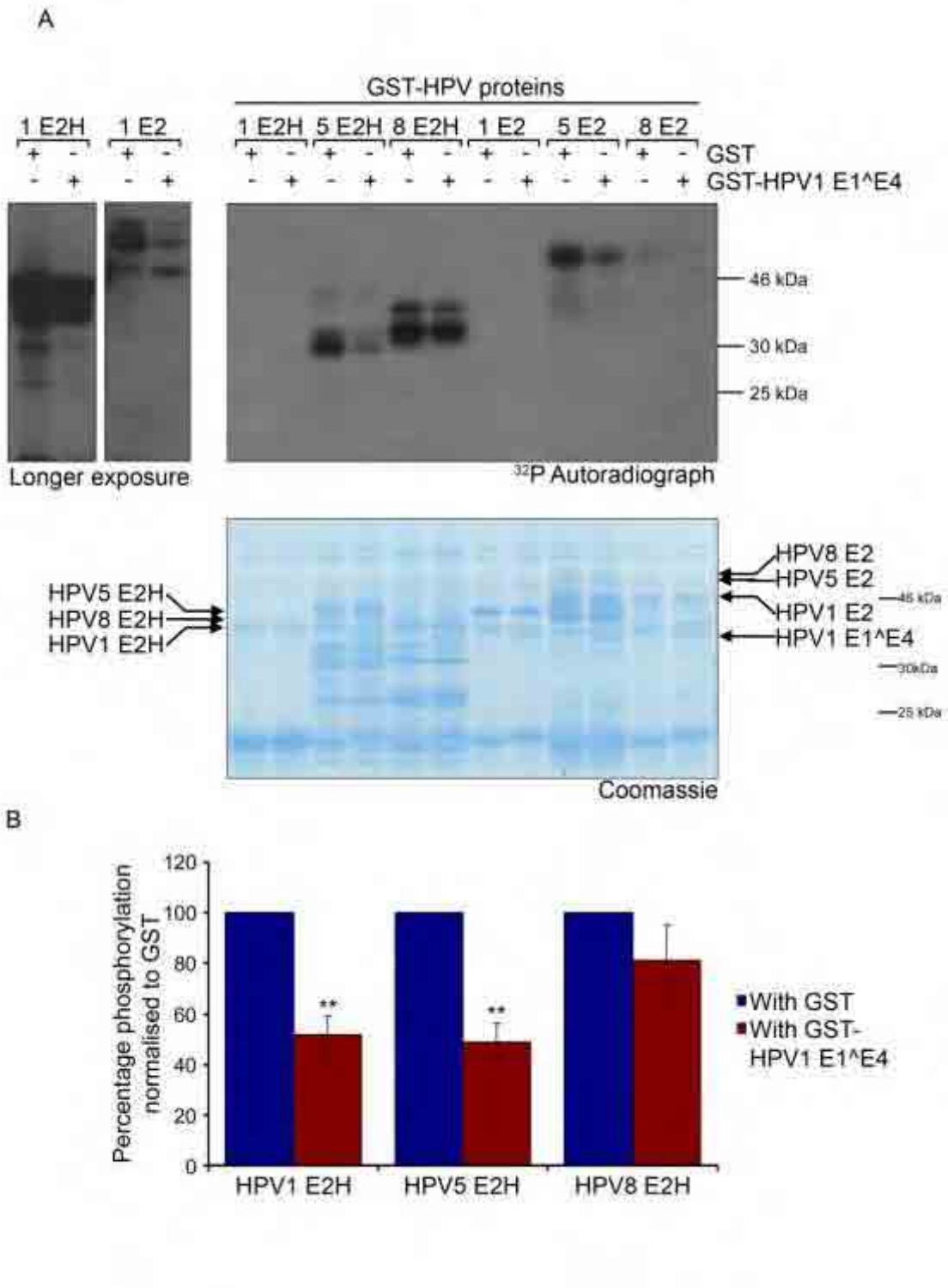

Figure 4.6 ***In vitro*** **kinase assay with viral E2 and E2H proteins.** (**A**) ^{32}P autoradiograph and Coomassie stained gel from *in vitro* kinase assay with 3 µg of GST-tagged HPV1, HPV5 and HPV8 E2 and E2H proteins, and 1 µg of GST-HPV1 E1^E4. A longer exposure autoradiograph allows the phosphorylation of the HPV1 proteins to be seen. (**B**) Graph showing the normalised data for the E2H proteins, ** = $p < 0.01$, graph contains the mean ± standard deviation from three experimental replicates.

To determine whether HPV5 E1^E4 altered SRPK1 mediated phosphorylation of HPV5 E2, GST-HPV5 E1^E4 was added in increasing amounts to an *in vitro* kinase assay in the presence of GST-tagged HPV5 E2 and E2H. GST-HPV5 E1^E4 did not inhibit SRPK1 phosphorylation of the GST-tagged proteins of HPV5 full length E2 or the E2 hinge region (data not shown). The level of phosphorylation of the full length E2 protein did increase with increasing amounts of HPV5 E1^E4 in comparison to GST, this may be due to an interaction between the E2 and E1^E4 tagged proteins, which may alter the conformation of the E2 protein allowing phosphorylation to occur more efficiently. This effect was not observed upon the phosphorylation of the E2 hinge region, however the phosphorylated forms of GST-HPV5 E1^E4 and E2H migrate at around the same molecular mass, making analysis of the effect of HPV5 E1^E4 on the phosphorylation of HPV5 E2H difficult.

4.3 Discussion

The interaction between HPV1 E1^E4 and SRPK1 caused the inhibition of SRPK1 phosphorylation of a number of different SR protein substrates; however the potency of inhibition did vary. As shown for ASF/SF2 the inhibition of SRPK1 activity towards these other SR substrates required the binding of HPV1 E1^E4 to SRPK1, as a non-binding mutant did not cause any inhibition, suggesting that the mechanism of inhibition was similar and the variation in potency is perhaps due to differences in SRPK1 substrate binding. The phosphorylation of SRp75 and 9G8 by SRPK1 was more susceptible to inhibition than the phosphorylation of SRp20. This data suggests that the larger SR proteins with more than one RRM, or in the case of 9G8 additional function domains may be more prone to E1^E4 mediated inhibition than those proteins with a single RRM.

The variation in the potency of inhibition by HPV1 E1^E4 on different cellular substrates, may allow for the differential regulation of the diverse SR protein functions.

As with ASF/SF2, HPV16 and HPV18 E1^E4 did not cause any significant change in the levels of phosphorylation of SC35, suggesting that the interaction between SRPK1 and HPV16 and HPV18 E1^E4 occurs via a different region of SRPK1 and does not interfere with substrate binding.

The hinge regions of the E2 proteins of the beta papillomaviruses HPV5 and HPV8, were good substrates for SRPK1 *in vitro* and this concurs with the findings of others (Lai et al., 1999, Sekhar et al., 2009). HPV5 E1^E4 did not inhibit SRPK1 phosphorylation of HPV5 E2; however the level of E2 phosphorylation observed did increase slightly and this may be due to the E2 and E1^E4 proteins interacting directly and causing a change in the conformation of E2, or the interaction between E1^E4 and SRPK1 may enable SRPK1 to bind E2 more efficiently. The E2 protein of the beta papillomaviruses has been shown to be extensively phosphorylated *in vivo*, with PKA phosphorylation of an RXXS motif in the hinge region being required for the association between E2 and mitotic chromosomes (Sekhar et al., 2009). Additionally phosphorylation of the RS dipeptides in the hinge region has been shown to be required for the localisation of the E2 proteins to nuclear speckles (Lai et al., 1999, Sekhar et al., 2009). Therefore SRPK1 phosphorylation of E2 may control both the localisation of E2 and some of the functions of E2 in the life cycle of the virus, such as the promotion the efficient packaging of viral genomes into virions (K. N. Zhao et al., 2000).

Additionally, the hinge region of HPV1 E2 was a substrate for SRPK1 and HPV1 E1^E4 inhibited SRPK1 phosphorylation of this viral SR protein. Whether SRPK1 can phosphorylate the hinge region of HPV1 E2 *in vivo* is unknown and the significance of the inhibition of SRPK1 phosphorylation of E2 by HPV1 E1^E4 on the life cycle of HPV1 is yet to be determined. In HPV16 the E1^E4 protein has been shown to stabilise the E2 protein

and control its subcellular localisation, thus regulating some of E2s functions (Davy 2009). The regulation of SRPK1 phosphorylation of HPV1 E2 by HPV1 E1^E4 may be a method of controlling the stability and functions of E2.

In summary, HPV1 E1^E4 can inhibit SRPK1 kinase activity with all tested substrates, (although the inhibition of HPV8 E2 hinge was not significant) but the variation in inhibition between the different substrates demonstrates that the potency of E1^E4 mediated SRPK1 inhibition is substrate specific.

CHAPTER 5 BIOLOGICAL FUNCTIONS OF THE INTERACTION BETWEEN HPV1 E1^E4 AND SRPK1

5.1 Introduction

HPV1 E1^E4 is found in numerous inclusions in a natural infection and these inclusions have been shown to be associated with several cellular proteins including the ND10 body defining protein PML, and the cellular kinase SRPK1 (Bell et al., 2007, Breitburd et al., 1987, Roberts et al., 2003). SRPK1 is relocalised to E4 inclusions in the granular layer of warts but not in the spinous layer, even though E4 inclusions are present in these layers (Bell et al., 2007). Indicating that SRPK1 sequestration to inclusions is required in the later stages of the HPV1 life cycle, where the capsid proteins are being expressed (Bell et al., 2007, Peh et al., 2002). Overexpression of HPV1 E1^E4 proteins in established epithelial cell lines gives rise to the formation of similar "*in vivo*-like" inclusions (Roberts et al., 2003, Rogel-Gaillard et al., 1992). Coexpression of HPV1 E1^E4 and SRPK1 causes the recruitment of SRPK1 to the cytoplasmic E4 inclusions, indicating that redistribution of SRPK1 is E4 dependent (Bell et al., 2007).

The phosphorylation state of the RS domain of the SR proteins is involved in mediating the subcellular localisation of the SR proteins as well as the interactions of the SR proteins with other proteins and RNA. SR proteins are predominantly concentrated in nuclear speckles which are sites of storage for transcription factors and also sites of active transcription. Overexpression of SRPK1, SRPK2 and CLK/STY has been shown to lead to the disassembly of nuclear speckles and cytoplasmic accumulation of the SR proteins (Colwill et al., 1996, Gui et al., 1994a, H. Y. Wang et al., 1998). The disassembly of nuclear speckles requires the

kinase to be active and leads to the hyperphosphorylation of the SR proteins (Koizumi et al., 1999, H. Y. Wang et al., 1998).

Pre-mRNA splicing is an essential step for gene expression in higher eukaryotes and is controlled by multiple factors. The minigene system allows the identification of *cis*-acting regulatory elements and *trans*-acting factors which bind to these elements and establish splicing efficiency and regulate alternative splicing. The expression of a minigene by transient transfection allows for the easy analysis of splicing regulation (Cooper, 2005). Use of a minigene coding for an alternatively spliced exon of mouse SRp20 has shown that SRp20 regulates the alternative splicing of its own mRNA (Jumaa and Nielsen, 1997). Overexpression of SRp20 leads to the inclusion of the alternatively spliced exon, which produces a truncated protein, lacking the RS domain that is inactive for many SR protein functions. In contrast, expression of ASF/SF2, promotes the skipping of the exon and therefore the production of the functional full length SRp20 protein (Jumaa and Nielsen, 1997). Interestingly the expression of the SR proteins SRp40 and 9G8 has no effect on the alternative exon inclusion, whereas SC35 leads to a novel splicing of the SRp20 minigene, demonstrating the range of effects the different SR proteins have on a single splice site (Jumaa and Nielsen, 2000). The autoregulation through splicing demonstrated by SRp20 has been observed with several other SR proteins including ASF/SF2 and SC35, with alternative ASF/SF2 transcripts being contained in the nucleus and SC35 producing unstable mRNAs (S. Sun et al., 2010, Sureau et al., 2001). The hyperphosphorylation of SR proteins by CLK2, has been shown to promote the skipping of the alternative exon in SRp20, to increase the amount of active protein produced (Stoss et al., 1999).

The minigene system can be used to analyse the cellular factors utilised by viruses to alternatively splice viral genes. The Adenovirus E1A gene has multiple splice isoforms and using an E1A minigene, the overexpression of ASF/SF2 has been shown to activate the

proximal 5' ss, prevent abnormal exon skipping and promote exon inclusion, leading to the production of the largest isoform of E1A (Caceres et al., 1994). Overexpression of the CLK kinases has been shown to promote the usage of the distal 5' ss and the production of the smaller isoforms (Yomoda et al., 2008). The same study showed that the overexpression of SRPK1 and other SR protein kinases had no effect on splice site selection in the E1A minigene (Yomoda et al., 2008).

In this chapter the effect of the interaction between HPV1 E1^E4 and SRPK1 on the kinase activity has been examined *in vivo*. In addition the effect of relocalisation of SRPK1 to E4 inclusion bodies will be examined with regards to cellular distribution of the SR proteins. Finally the effect of the interaction between SRPK1 and HPV1 E1^E4 on splice site selection will also be examined using the minigene system.

5.2 Results

5.2.1 *In vivo* inhibition of SRPK1 by HPV1 E1^E4

The inhibition studies carried out with His-SRPK1 and GST-HPV1 E1^E4 showed that HPV1 E1^E4 is a potent inhibitor of SRPK1 *in vitro*. However the effects of the interaction between SRPK1 and HPV E1^E4 in cells is unknown. In order to determine whether HPV1 E1^E4 can inhibit SRPK1 *in vivo*, Flag-tagged SRPK1 was transfected into H1299 cells along with plasmids expressing HPV1 E1^E4 and HPV1 E1^E4 Δ44-48 (a non-binding mutant). Flag-tagged MCM7 (minichromosome maintenance complex component 7), which has been shown to interact with HPV1 E1^E4 (Roberts et al., 2008), was also added to control for any non-specific effects. Following transfection the cells were harvested and immunoprecipitated with an anti Flag antibody or an anti-HPV1 E4 antibody. The immunoprecipitated proteins were then added to an *in vitro* kinase assay containing GST-ASF/SF2 and ^{32}P labelled γ-ATP. Following resolution by SDS PAGE, the gel was exposed to X-ray film (Figure 5.1A) and

quantified by phosphorimager analysis (Figure 5.1B). The HPV1 E4 immunoprecipitation was performed to ensure that the HPV1 E1^E4 deletion mutant Δ44-48 did not bind to SRPK1 (Figure 5.1C). Protein expression was analysed in the cell lysates by Western blot using anti-Flag and anti-HPV1 E4 antibodies (Figure 5.1D).

In the presence of the HPV1 E1^E4 protein, the immunoprecipitated Flag-SRPK1 (Flag IP) showed a significant reduction in phosphorylation of the substrate GST-ASF/SF2 in the *in vitro* kinase assay compared to the Flag-SRPK1 immunoprecipitated from cells co-transfected with Flag-SRPK1 and the HPV1 E1^E4 deletion mutant Δ44-48 which was unable to bind to SRPK1 (Figure 5.1A). The phosphorylation of GST-ASF/SF2 observed with immunoprecipitated Flag-MCM7 was due to non-specific sticking of kinases to the Sepharose beads and was considered the background level of phosphorylation. The proteins immunoprecipitated with HPV1 E1^E4 (HPV1 E4 IP) were able to phosphorylate GST-ASF/SF2 to a small extent (Figure 5.1C), this may be due to HPV1 E1^E4 sometimes binding to SRPK1 in a manner which does not cause inhibition of kinase activity.

The kinase assay showed that HPV1 E1^E4 can inhibit SRPK1 *in vivo*, with the level of inhibition of HPV1 E1^E4 wild type compared to the non-binding mutant Δ44-48 at 36% this inhibition is significant ($p = 0.030$). The level of inhibition *in vivo* however is not as great as the inhibition observed *in vitro* of 65% (Chapter 3, Figure 3.6). This was most likely due to the presence of SRPK1 that was not in complex with HPV1 E1^E4 or was bound in a conformation which does not cause inhibition.

This data indicates that HPV1 E1^E4 alters SRPK1 kinase activity *in vivo* and HPV1 E1^E4 may therefore affect SRPK1 functions.

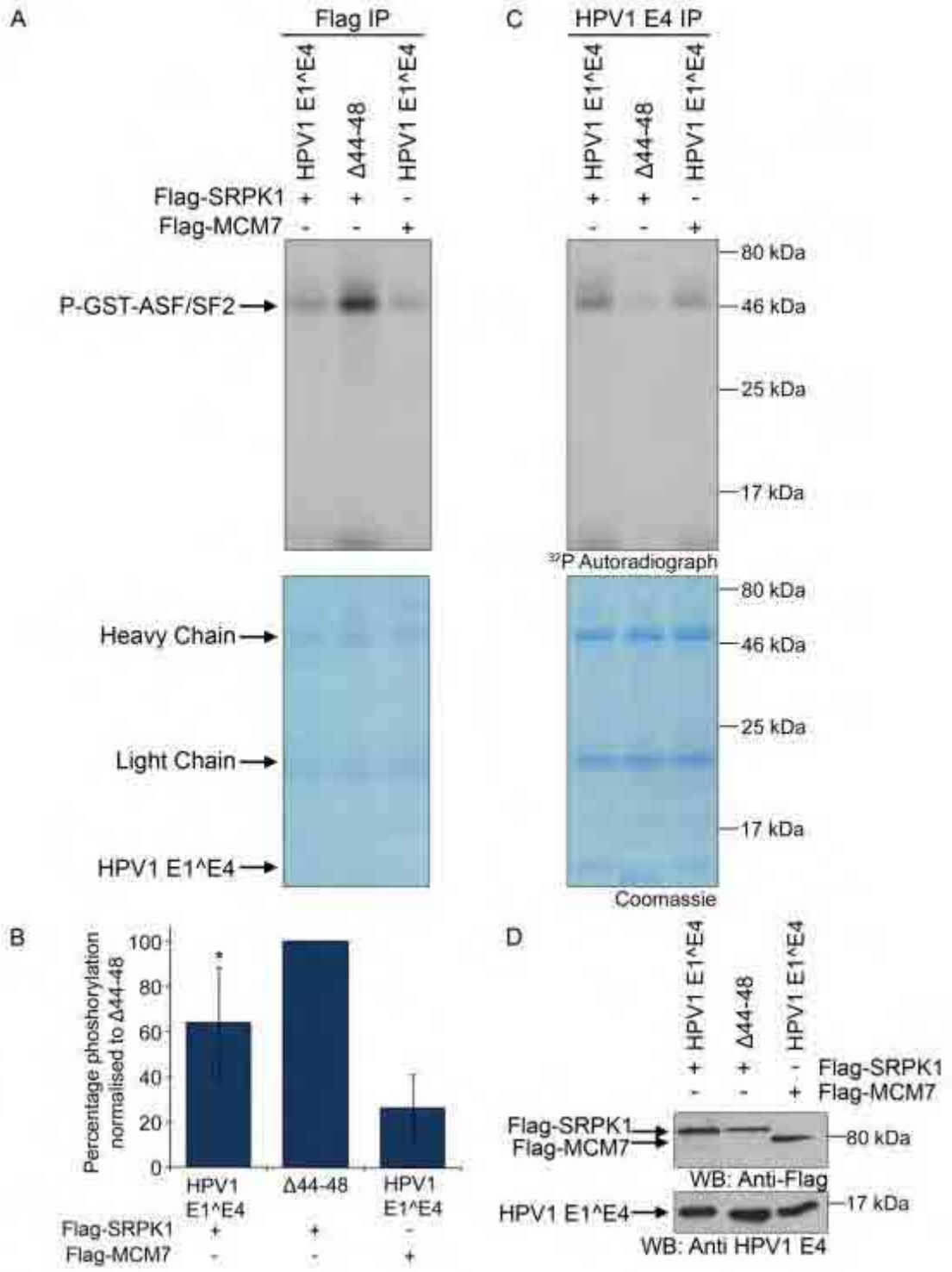

Figure 5.1 ***In vivo*** **inhibition of SRPK1 by HPV1 E1^E4.** Flag-tagged SRPK1 and MCM7 were co-transfected into cells with WT and Δ44-48 HPV1 E1^E4. The proteins were subject to IP with anti-Flag and anti-HPV1 E4 antibodies and added to an *in vitro* kinase assay with GST-ASF/SF2. (**A**) ^{32}P autoradiograph and Coomassie stained gel from *in vitro* kinase assay with Flag IP proteins. (**B**) Graph showing the normalised data for the Flag IP, * = $p < 0.05$, graph contains the mean ± standard deviation from five experimental replicates. (**C**) ^{32}P autoradiograph and Coomassie stained gel from *in vitro* kinase assay with proteins from HPV1 E4 IP. (**D**) Western blot with 1/20th of the input lysate blotted with rabbit anti-Flag and mouse anti-HPV1 E4 (4.37).

5.3 SRPK1 colocalises with HPV1 and HPV5 E1^E4 in cells

To understand the effect of HPV E4 proteins on SRPK1 functions the localisation of SRPK1 in the presence of HPV1 E1^E4 and HPV5 E1^E4 was examined. HPV1 E1^E4 is found in numerous nuclear and cytoplasmic inclusions in natural infections and when expressed in established epithelial cell lines (Breitburd et al., 1987, Roberts et al., 2003, Rogel-Gaillard et al., 1992). The distribution of E4 in naturally-occurring HPV5 infections has not been reported.

In immortalised human keratinocytes, Flag-SRPK1 co-localized to HPV1 E1^E4 inclusions (Bell et al., 2007), although the distribution of endogenous SRPK1 was not examined in this study. To confirm and extend these findings immortalised keratinocytes (SVJD), HPV containing cervical cancer keratinocytes (HeLa) and the non small cell lung carcinoma cells H1299 were each transfected with a plasmid expressing HPV1 E1^E4, fixed 48 h following transfection and stained for endogenous SRPK1 and HPV1 E4 (Figure 5.2). In cells transfected with the empty plasmid SRPK1 has a mainly cytoplasmic staining pattern, however there is some nuclear staining, and this is consistent with the literature (Figure 5.2A) (J. H. Ding et al., 2006). For all three cell lines, in cells expressing HPV1 E1^E4 a proportion of endogenous SRPK1 was present in E4 inclusions in the cytoplasm and in the nucleus (Figure 5.2B – H1299 cells, Figure 5.2C – HeLa cells and data not shown).

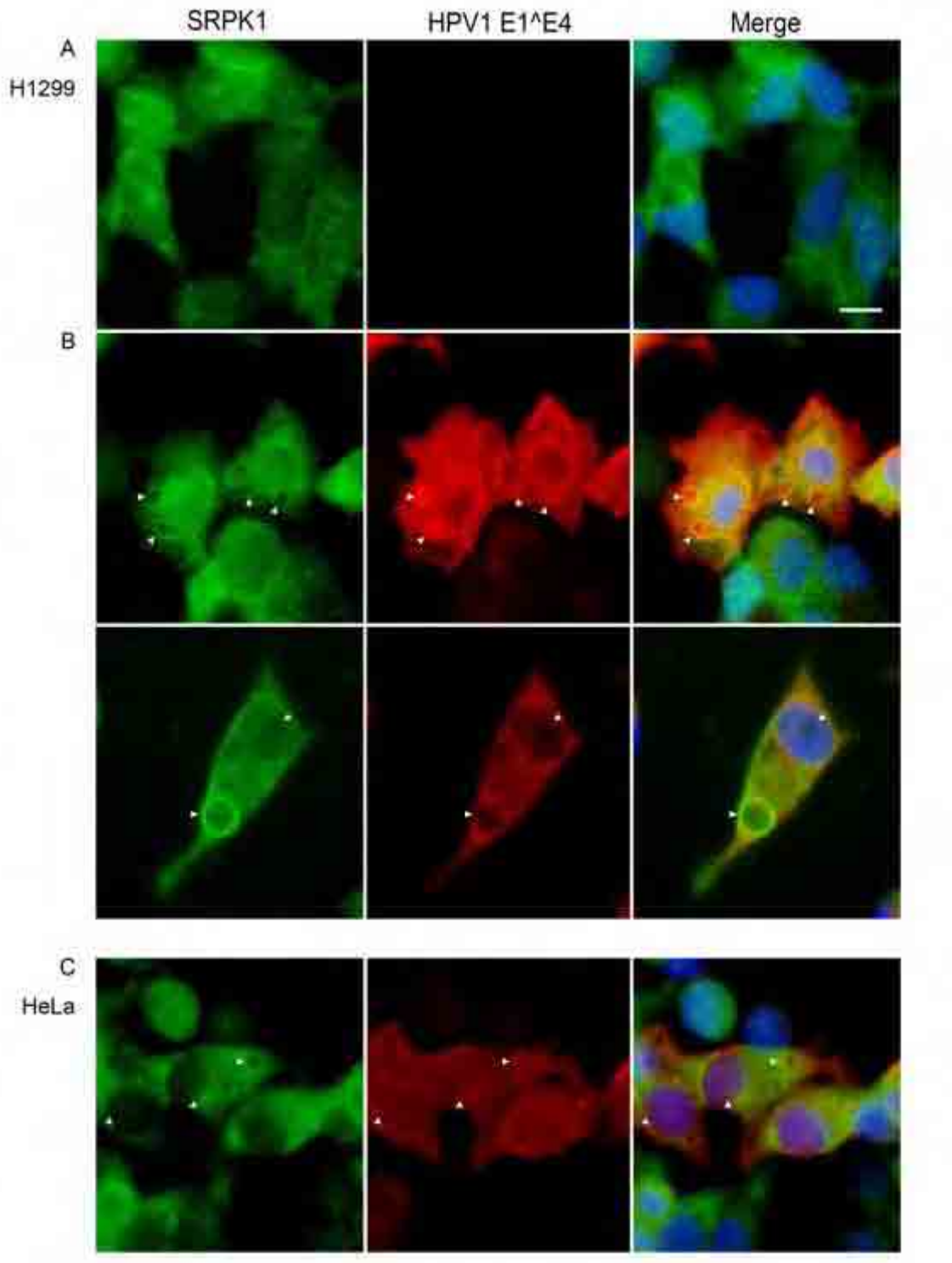

Figure 5.2 Localisation of SRPK1 with HPV1 E1^E4 in H1299 and HeLa cells. (**A**) H1299 cells were transfected with an empty plasmid. (**B**) H1299 cells were transfected with a plasmid expressing HPV1 E1^E4, alternative images under the same conditions shown. (**C**) HeLa cells were transfected with a plasmid expressing HPV1 E1^E4. The cells were fixed with a 0.1% Triton® X-100 pre treatment, then 4% paraformaldehyde followed by acetone permeabilisation. Cells were stained with mouse anti-SRPK1 (green) and mouse anti-HPV1 E4 (4.37) (red), nuclei were indentified with DAPI (blue). Arrowheads indicate HPV1 E4 inclusions. Scale bar indicates 10 μm.

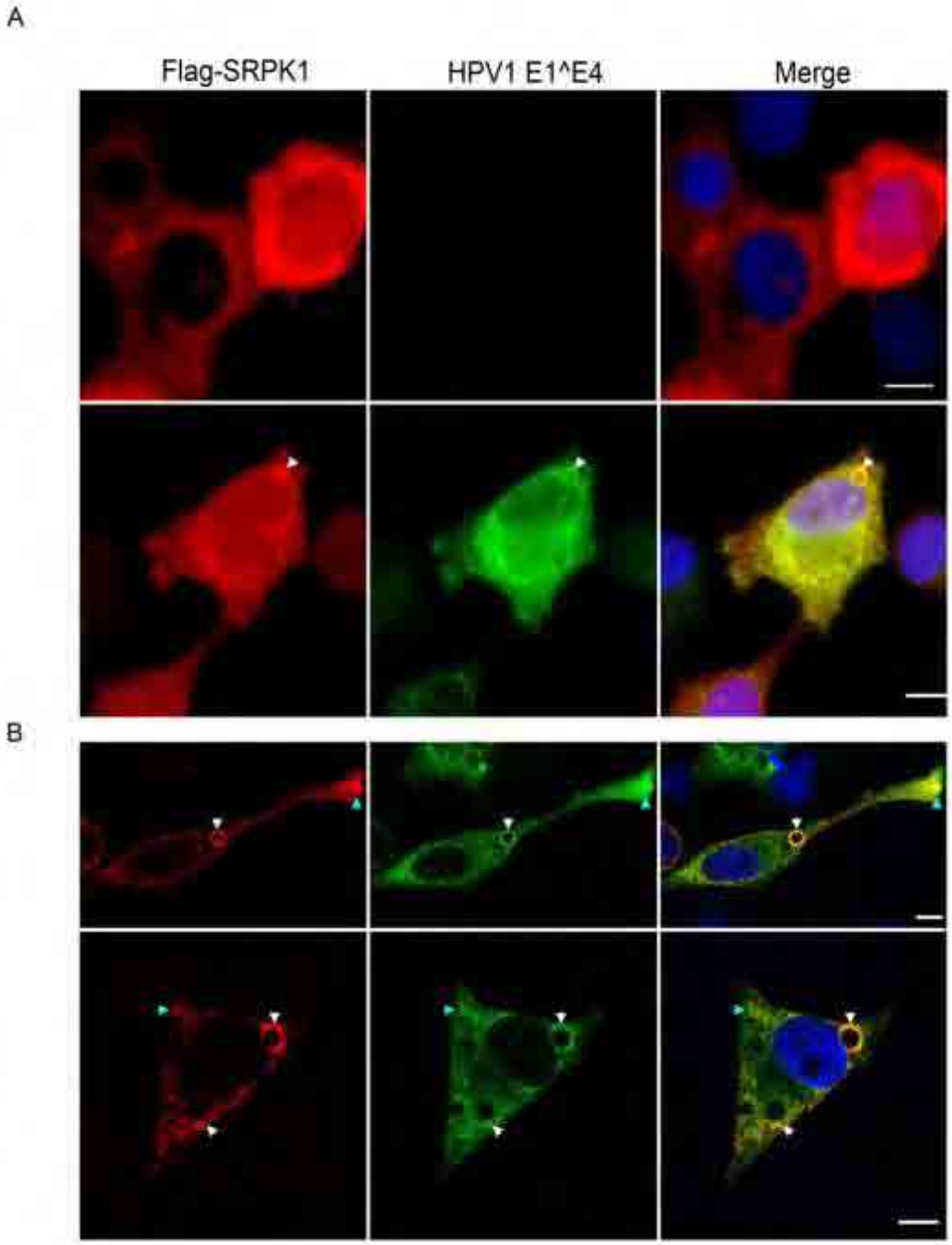

Figure 5.3 Localisation of Flag-SRPK1 with HPV1 E1^E4 in H1299 cells. H1299 cells were transfected with Flag-SRPK1 and a plasmid expressing HPV1 E1^E4. Cells were stained with rabbit anti-Flag (red) and mouse anti-HPV1 E4 (4.37) (green), nuclei were indentified with DAPI (blue). (**A**) Epifluorescent microscopy, top image: cells transfected with Flag-SRPK1 and empty plasmid, bottom image: cells transfected with Flag-SRPK1 and pcDNA HPV1 E1^E4. (**B**) Confocal microscopy showing two alternative images under the same conditions. White arrowheads indicate HPV1 E4 inclusions, blue arrowheads indicate colocalisation not related to inclusions. Scale bar indicates 10 μm.

Colocalisation between SRPK1 and HPV1 E1^E4 is also seen with exogenous Flag-tagged SRPK1 as previously reported (Bell et al., 2007). Flag-SRPK1 shows the same staining pattern as endogenous SRPK1, with mainly cytoplasmic staining; however as with endogenous SRPK1 a proportion of cells have intense nuclear staining (Figure 5.3A). The Flag-SRPK1 is sequestered to the HPV1 E4 inclusions and this is confirmed with confocal microscopy (Figure 5.3B). The Flag-SRPK1 is also colocalised with HPV1 E1^E4 not in inclusions, demonstrating that the colocalisation between SRPK1 and HPV1 E1^E4 is not dependent upon the E4 protein being in the inclusions, shown with blue arrowheads in Figure 5.3B.

Since HPV5 E1^E4 has been shown to be a substrate *in vitro* (Chapter 3, section 3.5.1) for SRPK1 it was important to determine if HPV5 E1^E4 altered the localisation of SRPK1 in cells. Therefore to determine whether HPV5 E1^E4 could cause a re-localisation of SRPK1 *in vivo*, a HA-tagged construct containing the HPV5 E1^E4 sequence was constructed using a HA cassette, (Chapter 2, section 2.2.7) and the pGex-2T HPV5 E1^E4 plasmid. The HA-tagged HPV5 E1^E4 was transfected into HeLa, H1299 and SVJD cells and fixed 48 h following transfection. Cells were pre-washed with 0.1% Triton® X-100 and then fixed with 4% paraformaldehyde followed by acetone permeabilisation. The HA-HPV5 E1^E4 was detected with an antibody against the HA tag. HPV5 E1^E4 was found present in the cytoplasm and nucleus, as a diffuse staining but also present as numerous distinct cytoplasmic aggregates, as shown in Figure 5.4 (A and B) in HeLa cells. In around 5% of the transfected population of cells, intense nuclear and cytoplasmic staining of HPV5 E1^E4 was observed (Figure 5.4C). The localisation of the HA-HPV5 E1^E4 was the same in the three cell lines used. In cells expressing HA-HPV5 E1^E4, a proportion of the endogenous SRPK1 formed aggregates and co-stained with the HA-HPV5 E1^E4 in some of these aggregates and around the nuclear membrane (Figure 5.5).

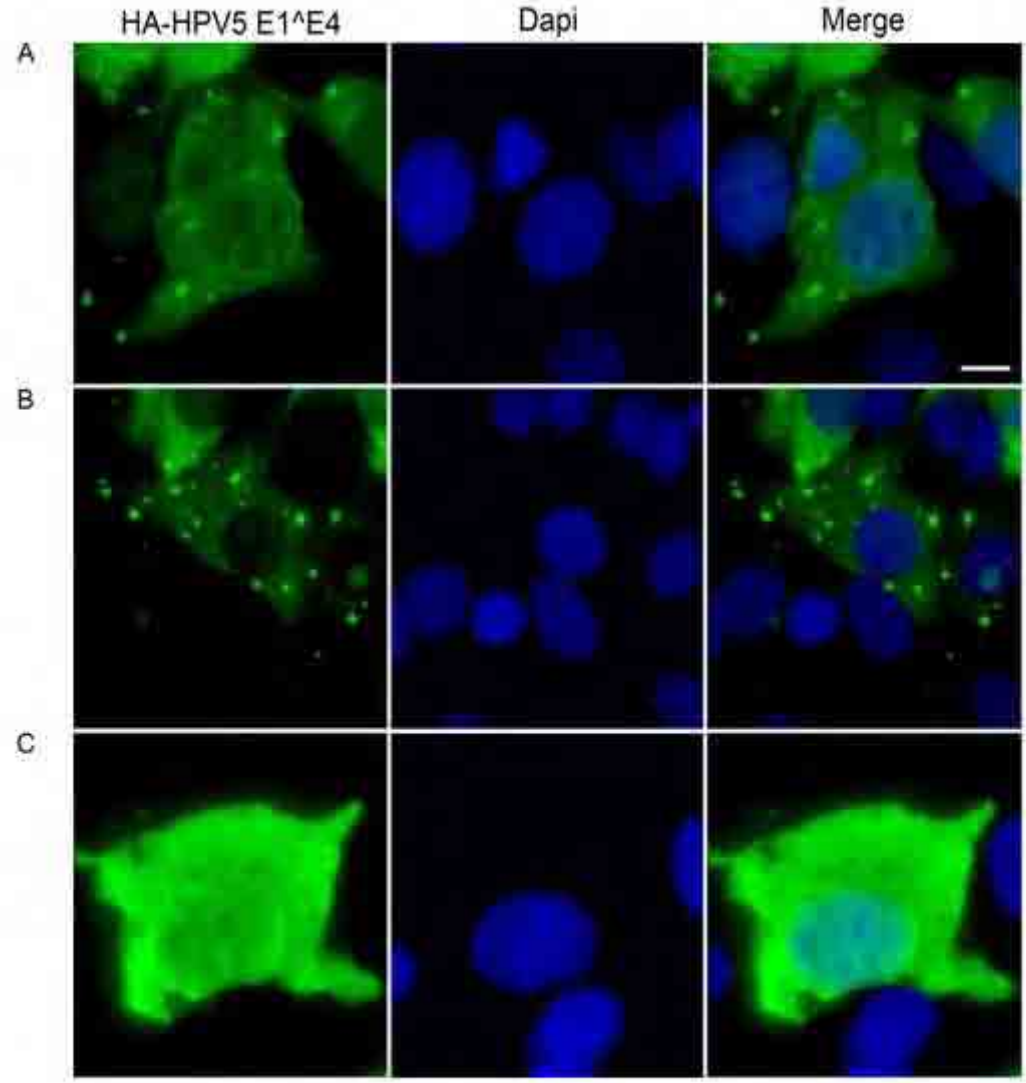

Figure 5.4 Localisation of HA-HPV5 E1^E4 in HeLa cells. HeLa cells were transfected with HA-HPV5 E1^E4 and showed three different staining patterns. (**A**) Diffuse cytoplasmic and nuclear staining with cytoplasmic aggregates. (**B**) Diffuse cytoplasmic staining with distinct cytoplasmic aggregates. (**C**) Intense cytoplasmic and nuclear staining. Cells were stained with mouse anti-HA (green) and nuclei were indentified with DAPI (blue). Scale bar indicates 10 μm.

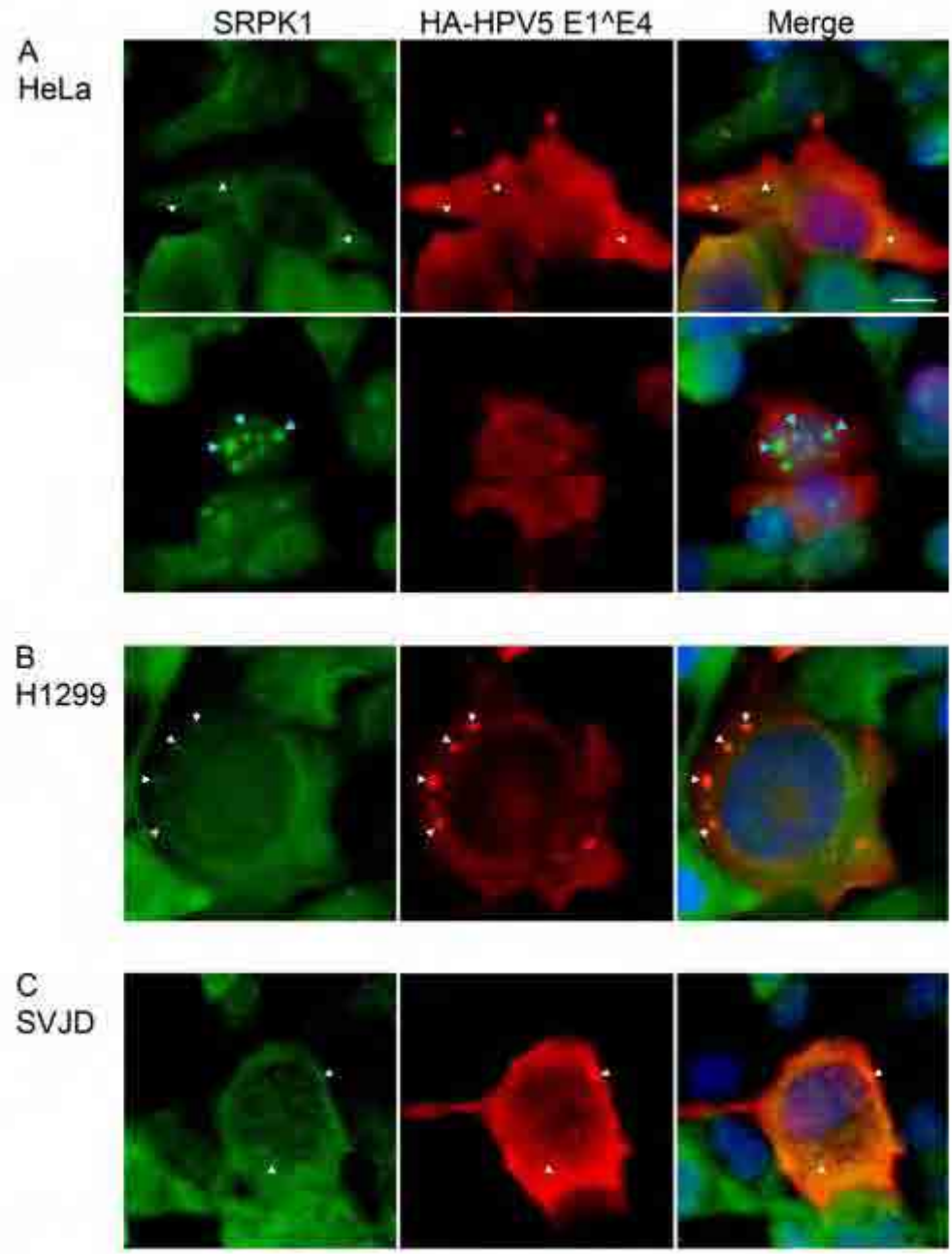

Figure 5.5 Localisation of HA-HPV5 E1^E4 and SRPK1 in HeLa, H1299 and SVJD cells. (**A**) HeLa (two alternative images under the same conditions), (**B**) H1299 and (**C**) SVJD cells were transfected with HA-HPV5 E1^E4. Cells were stained with mouse anti-SRPK1 (green) and rabbit anti-HA (red), nuclei were identified with DAPI (blue). White arrowheads indicate some areas of co-staining, blue arrowheads indicate nuclear SRPK1 aggregates which do not co-stain with HPV5 E1^E4. Scale bar indicates 10 μm.

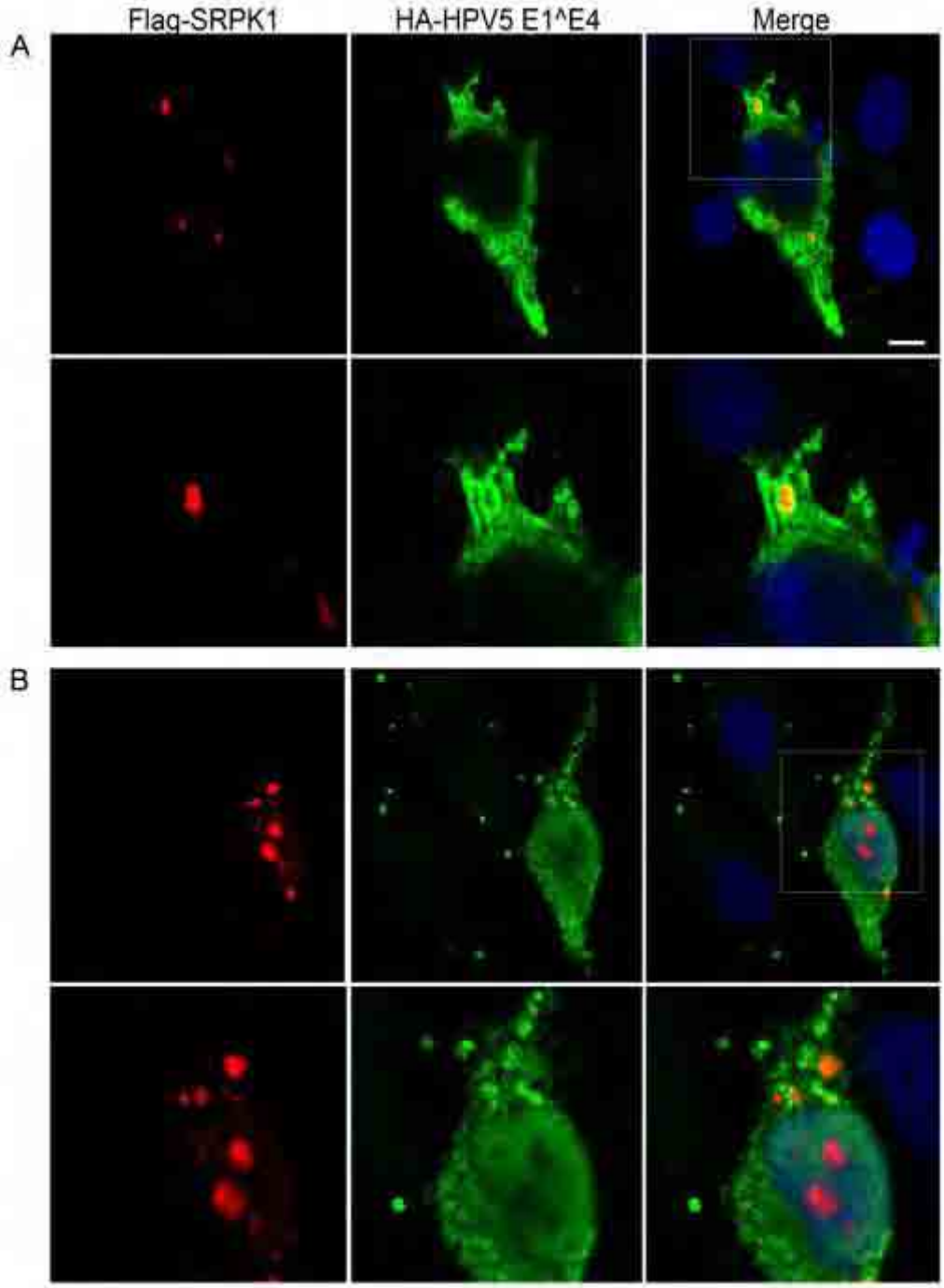

Figure 5.6 Localisation of Flag-SRPK1 with HA-HPV5 E1^E4 in HeLa cells. (**A** and **B**) HeLa cells were transfected with Flag-SRPK1 and HA-HPV5 E1^E4, two alternative images under the same conditions. Cells were stained with rabbit anti-Flag (red) and mouse anti-HA (green), nuclei were indentified with DAPI (blue) and analysed by confocal microscopy. Box indicates area enlarged below. Scale bar indicates 10 μm.

To extend the finding that HA-HPV5 E1^E4 caused a change in the staining pattern of endogenous SRPK1, co-transfection of HA-HPV5 E1^E4 with Flag-SRPK1 expression plasmids into cells was performed. As with the endogenous SRPK1, HA-HPV5 E1^E4 altered the localisation of the Flag-SRPK1 into aggregates, confocal analysis of these cells showed that SRPK1 was colocalised to the cytoplasmic E1^E4 aggregates (Figure 5.6A). Interestingly, in a subset of E1^E4 transfected cells, SRPK1 formed nuclear aggregates which were not colocalised with E1^E4 (Figure 5.6B).

This data suggests that HPV5 E1^E4 induces a relocalisation of SRPK1 in the cytoplasm and nucleus and that a proportion of SRPK1 is colocalised with the HPV5 protein, similar to that seen with HPV1 E1^E4.

5.4 HA tagged SR proteins

To aid with the visualisation of SR proteins *in vivo* and to investigate the localisation and functions of the SR proteins in overexpression and splicing studies, HA-tagged constructs of the SR proteins ASF/SF2, SC35 and SRp20 were generated (Chapter 2, 2.2.7). The expression and distribution of these tagged proteins was analysed following transfection into a panel of established cells lines including COS-1 epithelial cells, HeLa cervical cancer cells and H1299 lung carcinoma cells. After 48 h post transfection cellular lysates (COS-1 cells) were solubilised in RIPA buffer, resolved by SDS-PAGE and Western blotted using an anti-HA antibody. Bands corresponding to the expected sizes of 28.9 kDa for HA-ASF/SF2, 20.4 kDa for HA-SRp20 and 26.9 kDa for HA-SC35 were detected (Figure 5.7A). The HA-tagged SR proteins are resolved as doublets as they can be both hypo- and hyper-phosphorylated *in vivo*.

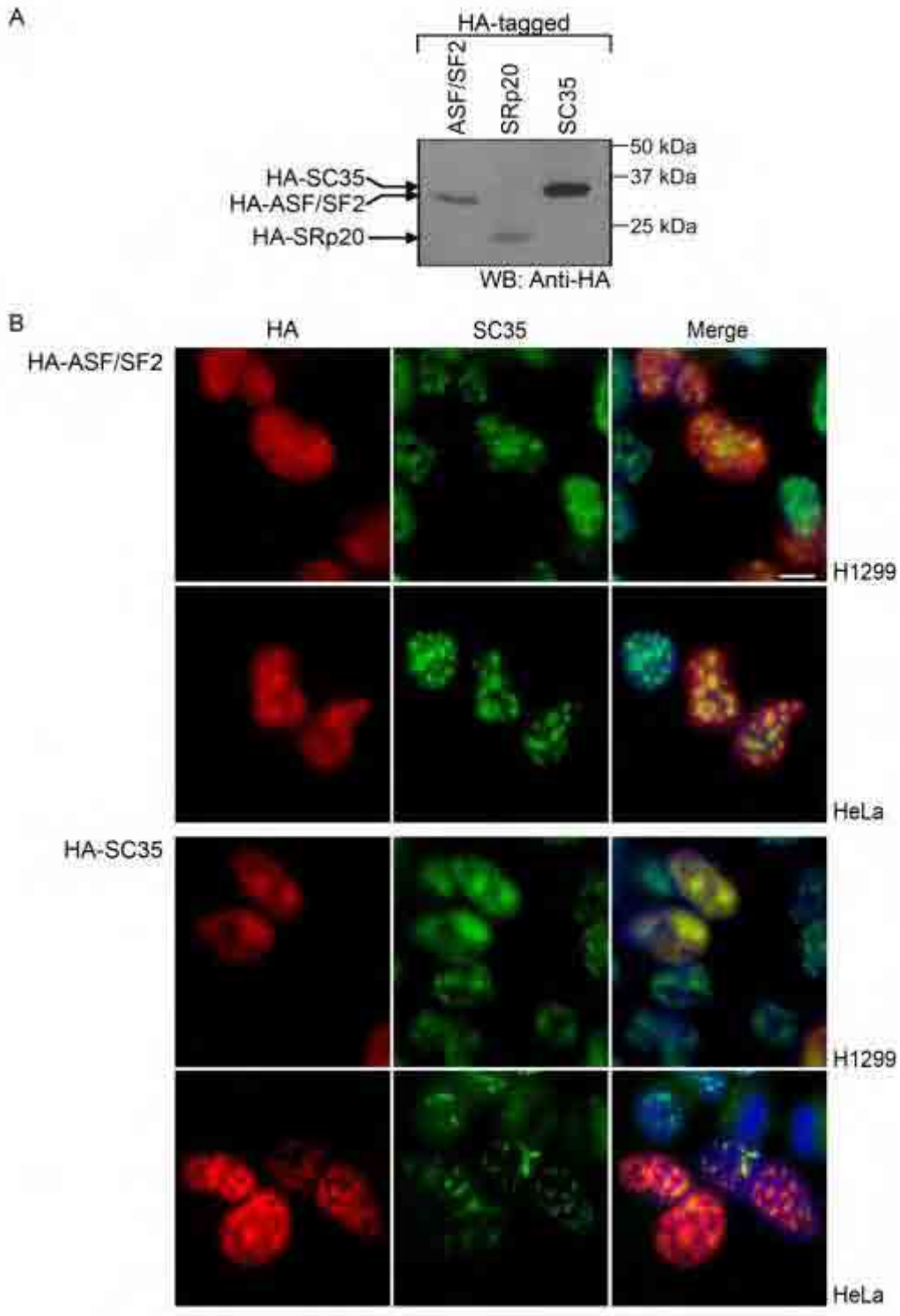

Figure 5.7 Expression and localisation of HA-tagged SR proteins. (**A**) Cos cells were transfected with HA-tagged constructs, lysed in RIPA buffer and the gel western blotted with a mouse anti-HA antibody. (**B**) H1299 and HeLa cells were transfected with HA-ASF/SF2 or HA-SC35 and stained with rabbit anti-HA (red) and mouse anti-SC35 (green), nuclei were stained with DAPI (blue). Scale bar indicates 10 μm.

The HA-tagged SR proteins were transfected into H1299 and HeLa cells on slides and after 48 h washed with 0.1% Triton® X-100 and then fixed with 4% paraformaldehyde followed by acetone permeabilisation. The transfection showed that the HA-tagged ASF/SF2, SRp20 and SC35 had a mainly nuclear localisation and could often be visualised in nuclear speckles. Co-staining with endogenous SC35 showed colocalisation between the endogenous SC35 speckles and a proportion of the HA-tagged ASF/SF2 and SC35 proteins (Figure 5.7B). The antibody used to identify SC35 speckles is for a phosphorylated form of the protein and showed that the majority of the HA-SC35 is not phosphorylated in the cells. The expression of the HA-SRp20 plasmid was poor in the cell lines used, with very low levels of staining seen in the transfected cells; therefore no further analysis was performed using this construct.

5.5 Overexpression of SRPK1 causes the relocalisation of SR proteins

The phosphorylation of SR proteins is required for multiple functions during pre-mRNA splicing and several groups have reported previously that overexpression of SRPK1 leads to the release of SR proteins from nuclear speckles (Gui et al., 1994a, H. Y. Wang et al., 1998). To demonstrate that this was observed in the cell lines used for this study, Flag-tagged SRPK1 was transfected into H1299 cells and 48 h post transfection the cells were fixed with various fixation methods. The cells were stained using antibodies against the endogenous phosphorylated SR proteins (Figure 5.8), endogenous ASF/SF2 (Figure 5.9), and endogenous SC35 (Figures 5.10 and 5.11).

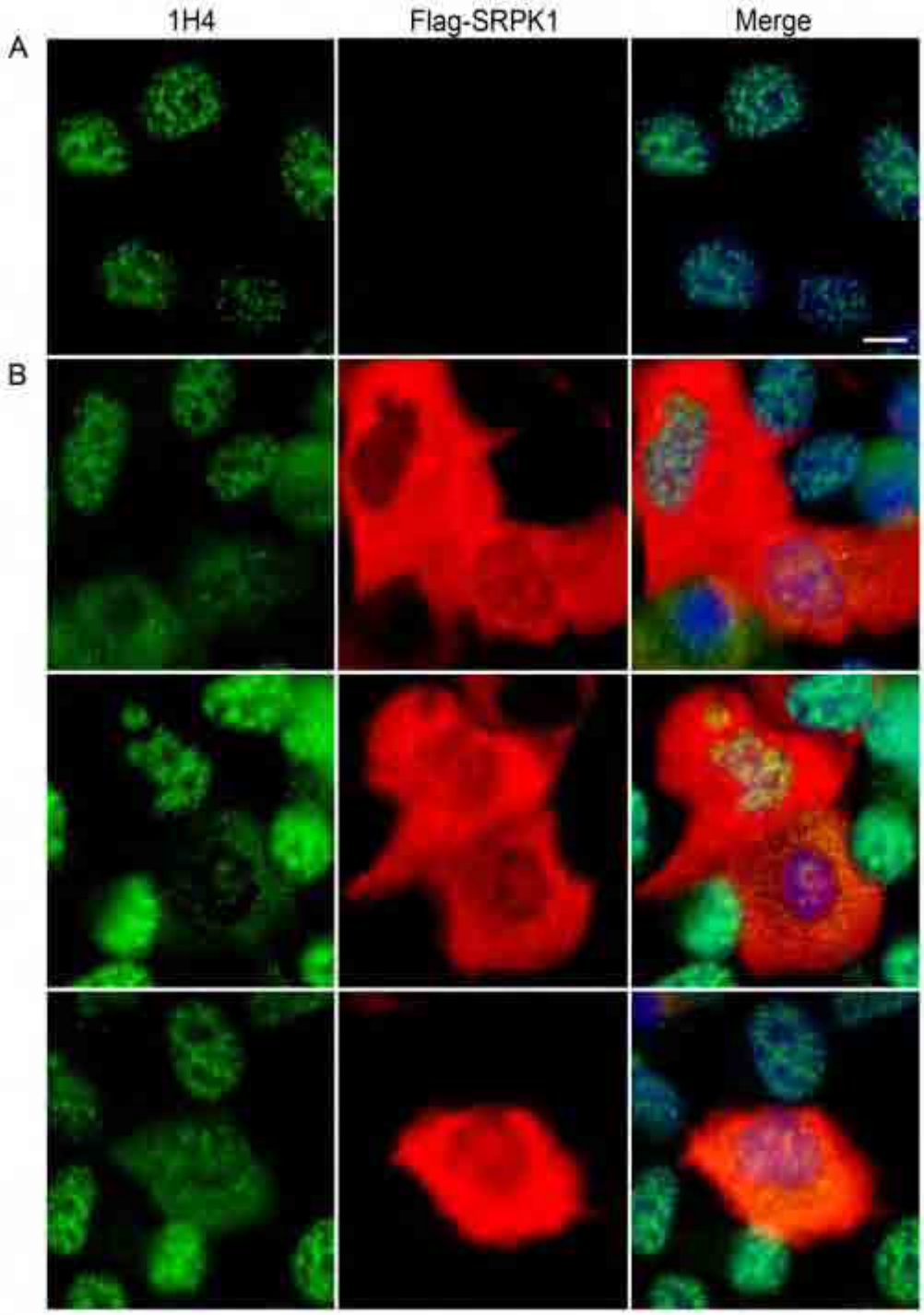

Figure 5.8 Localisation of phosphorylated SR proteins upon Flag-SRPK1 expression. (**A**) H1299 cells were transfected with Flag empty vector. (**B**) H1299 cells were transfected with Flag-SRPK1, three different staining patterns for the SR proteins were observed under the same conditions. Cells were stained with mouse anti-phospho SR protein (1H4) (green) and rabbit anti-Flag (red), nuclei were stained with DAPI (blue). Scale bar indicates 10 μm.

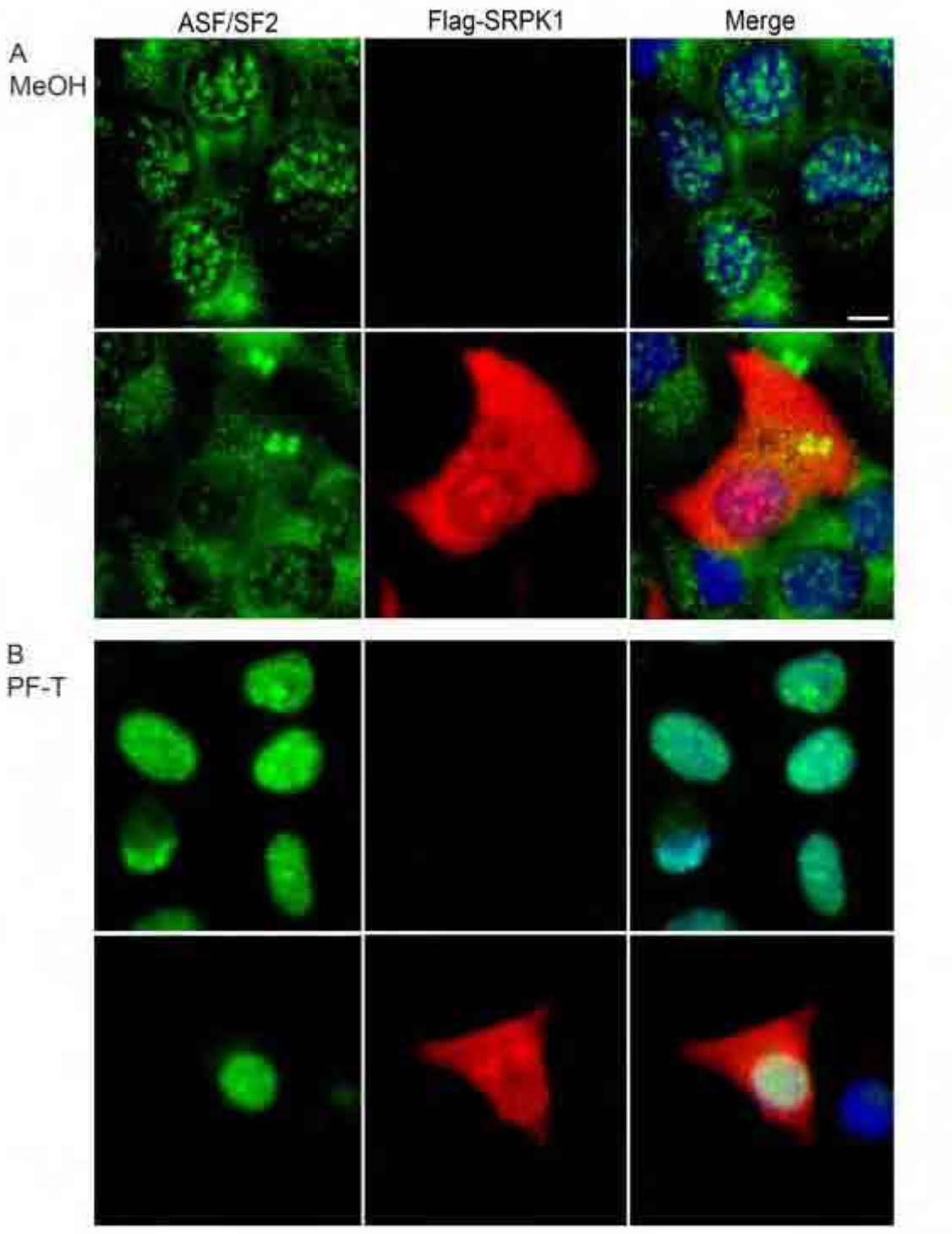

Figure 5.9 Localisation of ASF/SF2 upon expression of Flag-SRPK1. (**A**) H1299 cells were transfected in top image with Flag empty vector or in bottom image with Flag-SRPK1 and fixed using methanol. (**B**) H1299 cells were transfected in top image with Flag empty vector or in bottom image with Flag-SRPK1 and fixed using paraformaldehyde followed by 0.1% Triton®. Cells were stained with mouse anti-ASF/SF2 (green) and rabbit anti-Flag (red), nuclei were stained with DAPI (blue). Scale bar indicates 10 µm.

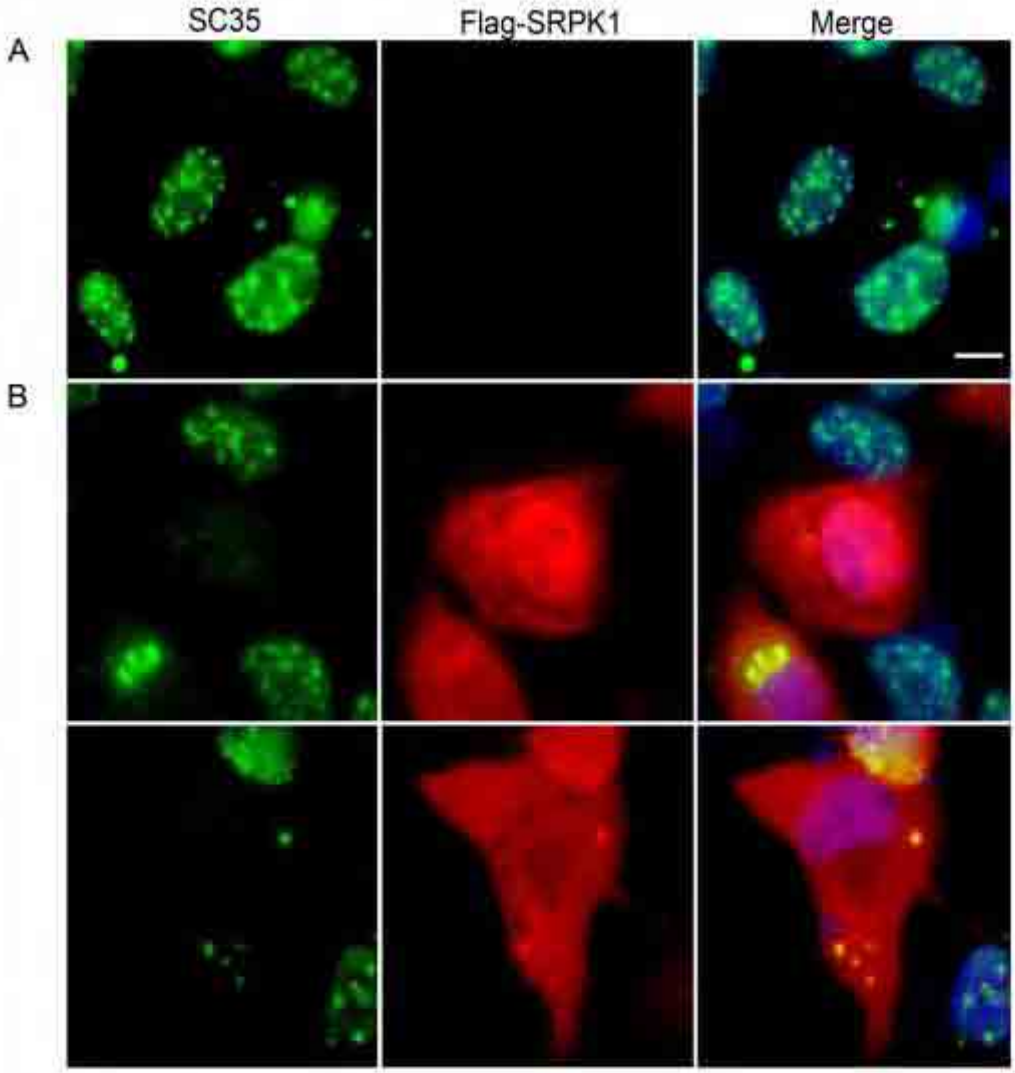

Figure 5.10 Localisation of SC35 upon expression of Flag-SRPK1. (**A**) H1299 cells transfected with Flag empty vector. (**B**) H1299 cells transfected with Flag-SRPK1, two alternative images under the same conditions. Cells were stained with mouse anti-SC35 (green) and rabbit anti-Flag (red), nuclei were stained with DAPI (blue). Scale bar indicates 10 μm.

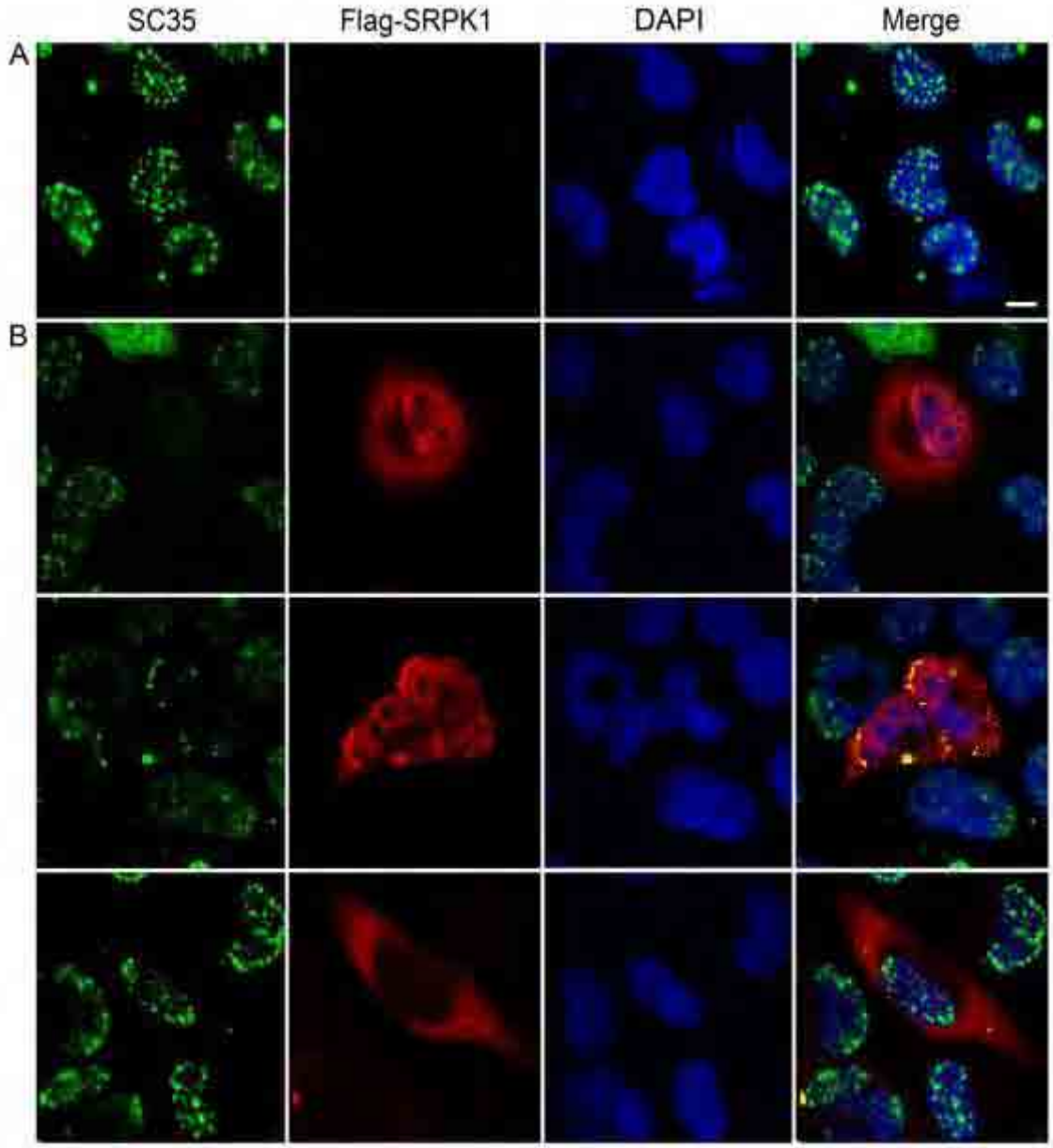

Figure 5.11 Localisation of SC35 upon expression of Flag-SRPK1. (**A**) H1299 cells transfected with Flag empty vector. (**B**) H1299 cells transfected with Flag-SRPK1, three alternative images under the same conditions. Cells were stained with mouse anti-SC35 (green) and rabbit anti-Flag (red), nuclei were stained with DAPI (blue). Cells were analysed by confocal microscopy, scale bar indicates 10 µm.

5.5.1 Phosphorylated SR proteins

The antibody against the phosphorylated SR proteins (1H4) (Figure 5.8) showed a mainly nuclear distribution of the SR proteins, with distinct speckles in the nucleus, corresponding to splicing speckles, when fixed using cold methanol for 10 min. Upon expression of Flag-tagged SRPK1, when the SRPK1 is expressed in the nucleus, the SR protein speckles become more diffuse and the staining becomes more cytoplasmic. The change in localisation occurs because the phosphorylation state of the SR proteins changes and the proteins are released from the splicing speckles (Koizumi et al., 1999). When the SRPK1 is cytoplasmic, no visible effect is observed upon 1H4 staining.

5.5.2 ASF/SF2

The prototypical SR protein ASF/SF2 had different localisations under different fixation conditions. With 10 min of cold methanol, ASF/SF2 was present in both the nucleus and the cytoplasm, with distinct speckles visible in the nucleus (Figure 5.9A). Upon expression of the Flag-SRPK1 vector, nuclear ASF/SF2 staining is markedly decreased and staining increases in the cytoplasm. Under different fixation conditions of 20 min of paraformaldehyde followed by 15 min of 0.1% Triton® X-100, the ASF/SF2 has diffuse nuclear staining with speckles of more intense staining (Figure 5.9B). The effect of SRPK1 overexpression on ASF/SF2 localisation was less pronounced under these staining conditions, with the speckles becoming less intense against the diffuse nuclear staining, some cytoplasmic staining was observed, however this was difficult to capture in an image.

5.5.3 SC35

The SR protein SC35 is considered to be a non-shuttling SR protein and is confined to the nucleus, SC35 is generally used as a marker of splicing speckles (Fu and Maniatis, 1990). As the overexpression of SRPK1 leads to the loss of nuclear speckles and cytoplasmic

accumulation of the phosphorylated SR proteins and ASF/SF2, cells transfected with Flag-SRPK1 were stained for endogenous phosphorylated SC35. The cells were fixed with 4% paraformaldehyde followed by acetone permeabilisation, giving the characteristic nuclear speckled pattern of SC35 (Figure 5.10). The presence of nuclear SRPK1 led to loss to the nuclear speckles in around 70% of cells and in around 25% of cells a complete loss of nuclear SC35 staining. When the SRPK1 was not located in the nucleus, the distribution of SC35 was not altered. Confocal analysis showed that there was a small amount of colocalisation between the SRPK1 and SC35 in the cytoplasm and around the nuclear membrane (Figure 5.11).

The loss of SC35 speckles which occurred when Flag-SRPK1 was expressed was specific to SRPK, as the expression of other plasmids including GFP and Flag-MCM7, which can be present in the nucleus, did not cause any change in SC35 localisation (Figure 5.12).

5.5.4 HA-tagged SR proteins

The HA-tagged SR proteins did not alter in their cellular localisation upon transfection with Flag-SRPK1 (Figure 5.13), confocal microscopy showed colocalisation between the HA-tagged SR proteins and Flag-SRPK1, especially with HA-ASF/SF2 (Figure 5.13B).

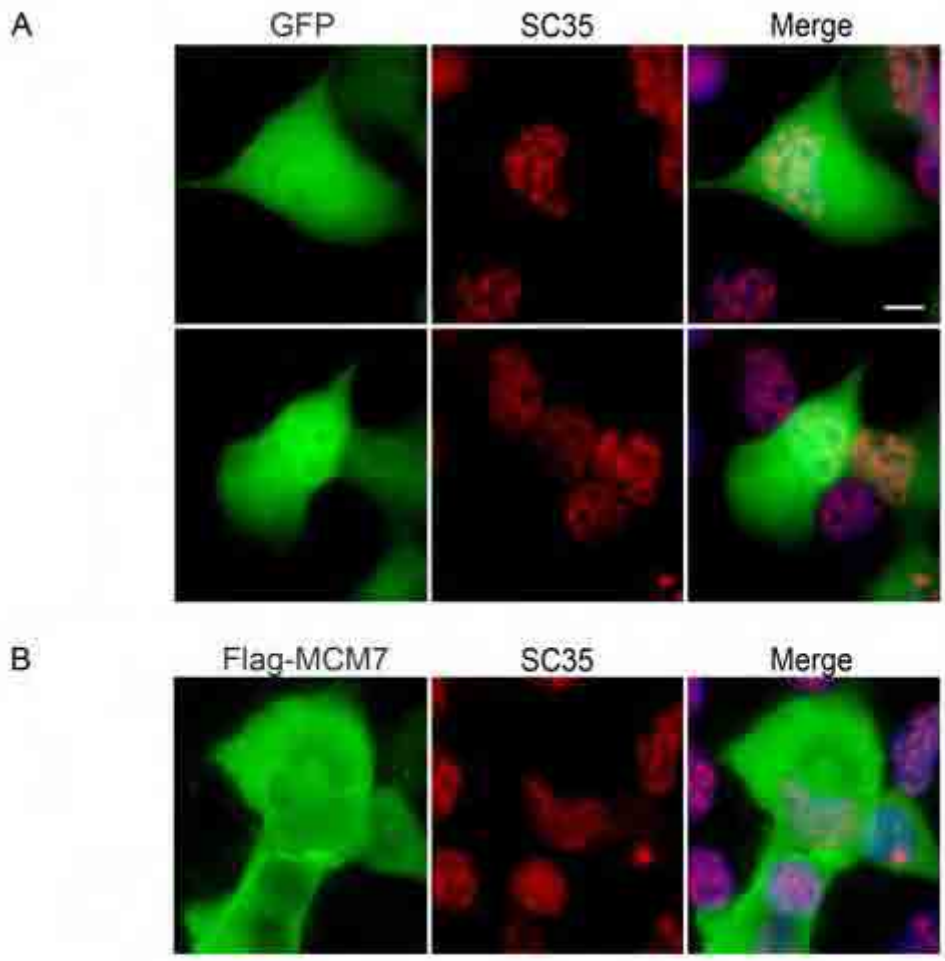

Figure 5.12 Localisation of SC35 upon expression of GFP and Flag-MCM7. (**A**) H1299 cells were transfected with GFP-C1 and cells were stained with mouse anti-SC35 (red), two alternative images under the same conditions. (**B**) H1299 cells were transfected with Flag-MCM7. Cells were stained with mouse anti-SC35 (red) and rabbit anti-Flag (green), nuclei were stained with DAPI (blue). Scale bar indicates 10 μm.

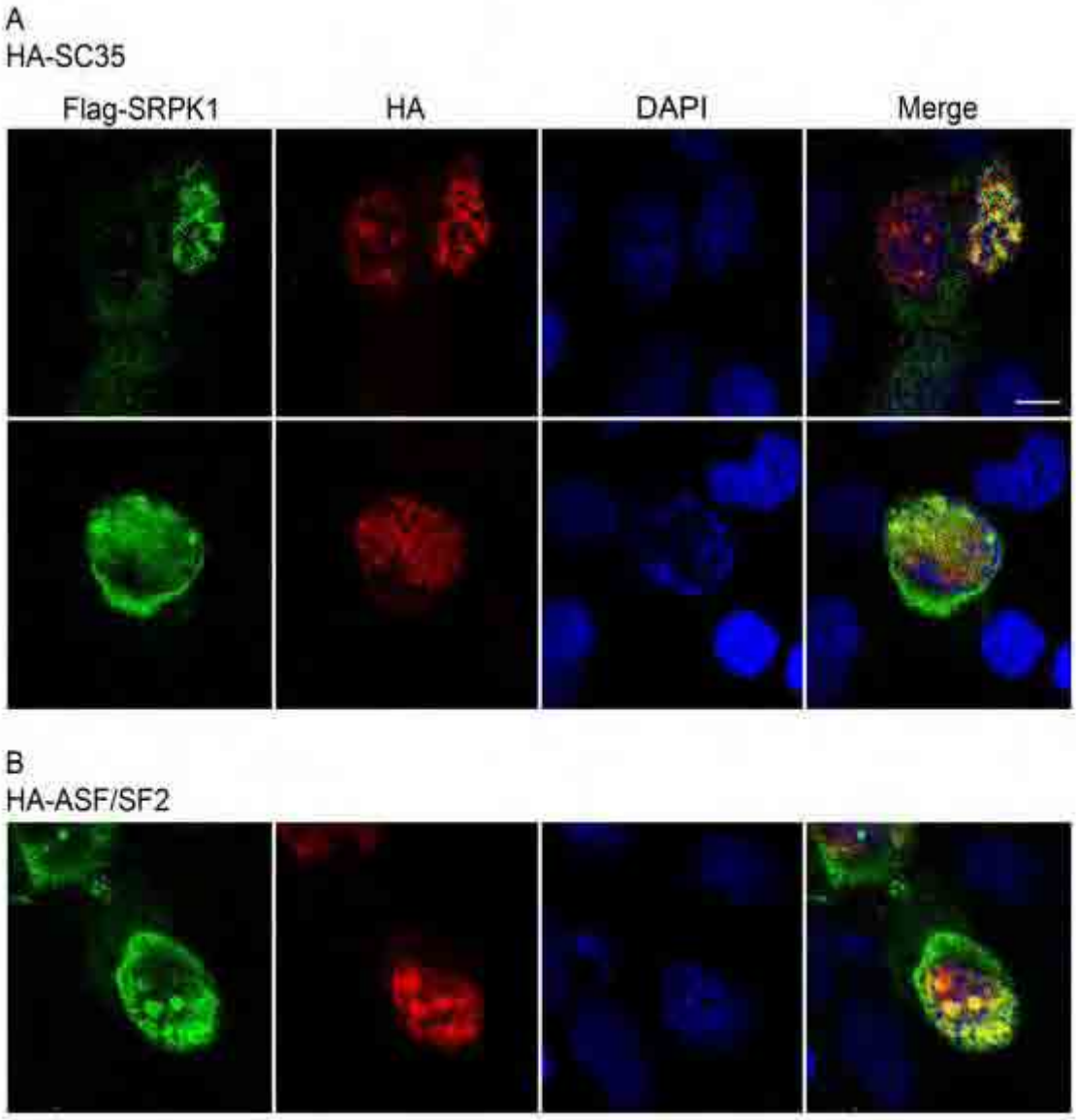

Figure 5.13 Localisation of HA-tagged SR proteins with expression of Flag-SRPK1. H1299 cells were transfected with Flag-SRPK1 and (**A**) HA-SC35 (two alternative images under the same conditions) or (**B**) HA-ASF/SF2. Cells were stained with mouse anti-Flag (M2) (green) and rabbit anti-HA (red), nuclei were stained with DAPI (blue). Cells were analysed by confocal microscopy, scale bar indicates 10 μm.

5.6 Inhibition of SRPK1 kinase activity does not inhibit SRPK1-mediated relocalisation

5.6.1 HPV1 E1^E4

As HPV1 E1^E4 has been shown to inhibit the kinase activity of SRPK1 *in vitro* and *in vivo*, and SRPK1 has been implicated in defining SR protein localisation, the effect of HPV1 E1^E4 on SRPK1-mediated distribution of SR proteins was analysed. Initially the effect of HPV1 E1^E4 on the endogenous SR proteins was examined. HPV1 E1^E4 was transfected into H1299 and HeLa cells and 24 h post transfection the cells were washed with 0.1% Triton® X-100 and then fixed with 4% paraformaldehyde followed by acetone permeabilisation. The cells were stained for endogenous SC35 (Figure 5.14A), ASF/SF2 (Figure 5.14B) and phosphorylated SR proteins (1H4) (Figure 5.14C). When HPV1 E1^E4 formed nuclear inclusions, a small proportion of the SR proteins costained with these HPV1 E4 inclusion bodies. *In vitro* HPV1 E1^E4 does not interact with any SR proteins (data not shown) so the recruitment of SR proteins to E4 inclusions may occur via the formation of a complex with other proteins. Apart from the recruitment to E4 inclusions, the distribution of the SR proteins is not altered in the presence of HPV1 E1^E4, with the inhibition of endogenous SRPK1 kinase activity by HPV1 E1^E4 having no effect on SR protein localisation.

The overexpression of SRPK1 caused a relocalisation of SR proteins, therefore HPV1 E1^E4 was cotransfected into H1299 and HeLa cells with Flag-SRPK1 to investigate whether the inhibition of SRPK1 activity could disrupt this relocalisation (Figure 5.15). Twenty-four hours post transfection the cells were fixed with 4% paraformaldehyde, permeabilised with 0.1% Triton® X-100 and then the cells were stained with antibodies against Flag and endogenous SC35. As HPV1 E1^E4 relocalised SRPK1 to inclusions and a triplicate stain was difficult to perform due to the antibody subtypes, a cell was considered to be HPV1

E1^E4 positive if SRPK1 was observed to be sequestered to inclusion bodies. In the presence of HPV1 E1^E4 the expression of nuclear Flag-SRPK1 still caused the loss of SC35 nuclear speckles in 60% of cells and caused a complete loss of nuclear SC35 staining in 30% of the cellular population. The sequestration of SRPK1 to E4 inclusion bodies in the nucleus caused a reduction in the loss of SC35 speckles (Figure 5.15C), whereas sequestration of SRPK1 to E4 inclusion bodies in the cytoplasm had no effect on SC35 speckle loss (Figure 5.15A and B). When SRPK1 was not located in the nucleus, no effect was observed on SC35 staining (Figure 5.15D). HPV1 E1^E4 does not inhibit the redistribution of SC35 upon SRPK1 overexpression in the nucleus, except when the nuclear SRPK1 is sequestered to E4 inclusion bodies.

5.6.2 SRPIN340

The small molecule inhibitor SRPIN340 has been shown to specifically inhibit SRPK1 and SRPK2 (Fukuhara et al., 2006). To determine whether the kinase activity of SRPK1 was required to cause the relocalisation of SC35 upon SRPK1 overexpression, 10 μM SRPIN340 was added to cells for 20 h following transfection of Flag-SRPK1. The cells were fixed with 4% paraformaldehyde, permeabilised with 0.1% Triton® X-100 and then were stained with antibodies against Flag and endogenous SC35 (Figure 5.16). Cells had an equivalent amount of DMSO (0.1% v/v) added, to check any effects observed were not due to solvent effects. Addition of SRPIN340 had no negative effect on the redistribution of SC35 from speckles upon SRPK1 overexpression. With SRPIN340, cells with nuclear SRPK1 had a loss of SC35 speckles in 27% of the cells and a complete loss of nuclear SC35 in 67% of the cells. A higher proportion of the cell population had a loss of nuclear SC35 staining when compared to a Flag-SRPK1 transfection in the absence of DMSO. However between the SRPIN340 and DMSO treated cells there was no change in the loss of nuclear SC35 staining, suggesting that the DMSO not SRPIN340, promoted the loss of SC35 speckles.

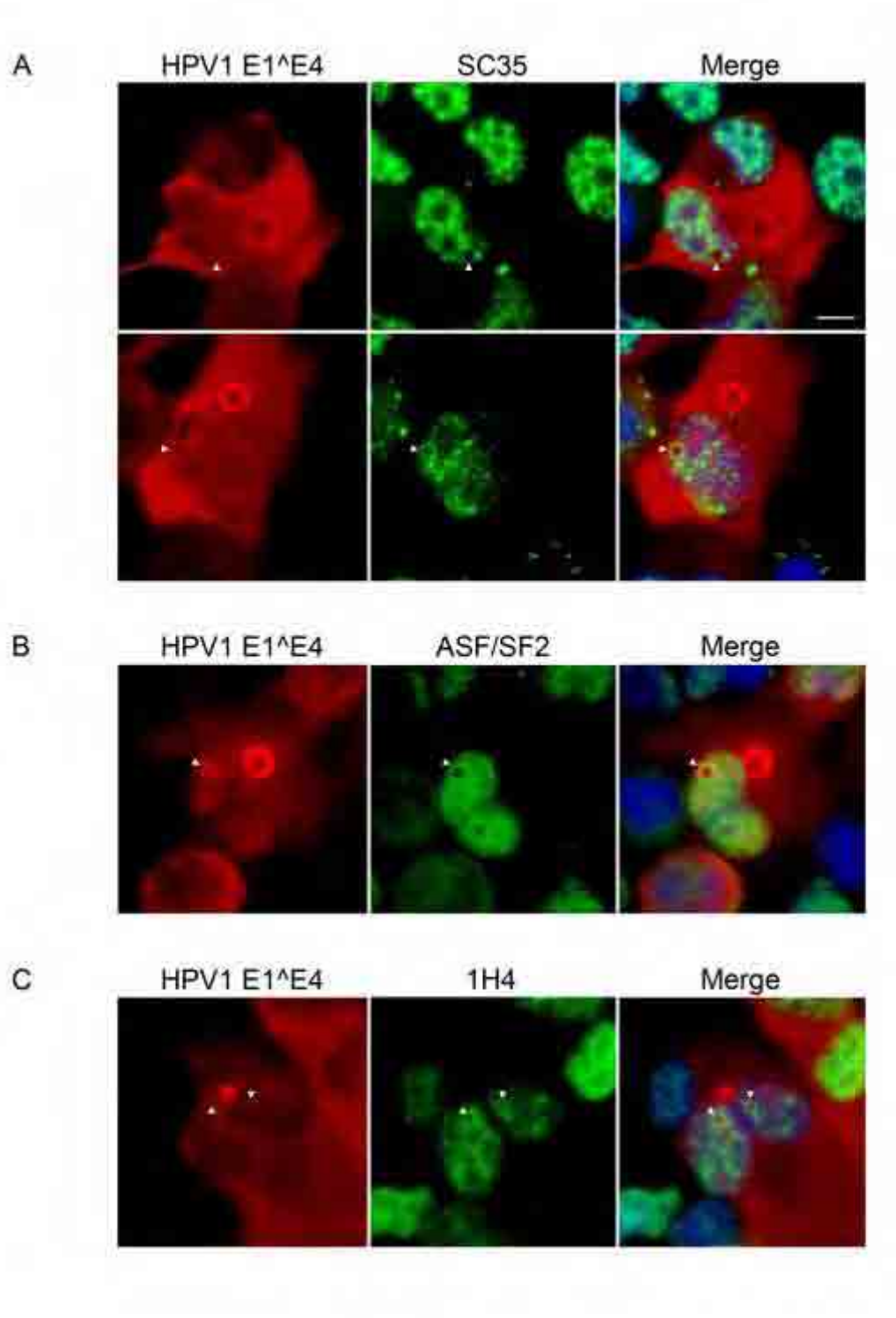

Figure 5.14 Localisation of SR proteins with HPV1 E1^E4 in H1299 cells. H1299 cells were transfected with HPV1 E1^E4. Cells were stained with mouse 4.37 (red) and (**A**) mouse anti-SC35 (green) two alternative inmages under the same conditions, (**B**) mouse anti-ASF/SF2 (green) or (**C**) mouse anti-phospho SR protein (1H4) (green); nuclei were stained with DAPI (blue). Arrowheads indicate nuclear E4 inclusions and scale bar indicates 10 µm.

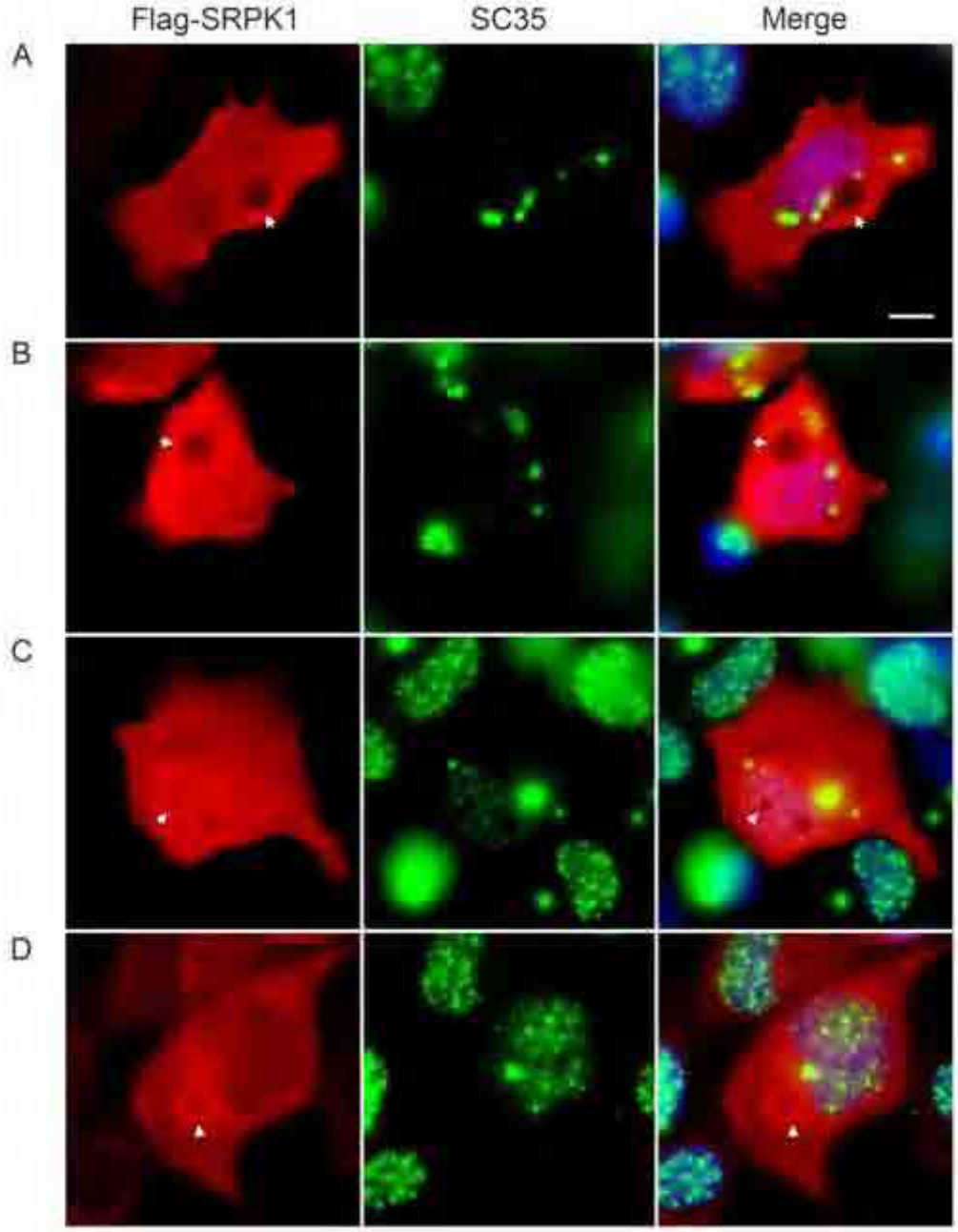

Figure 5.15 Localisation of SC35 with Flag-SRPK1 and HPV1 E1^E4. H1299 cells were transfected with Flag-SRPK1 and HPV1 E1^E4. Cells were stained with rabbit anti-Flag (red) and mouse anti SC35 (green), nuclei were stained with DAPI (blue). (**A** and **B**) Nuclear SRPK1 staining with cytoplasmic HPV1 E4 inclusions. (**C**) Nuclear SRPK1 with nuclear HPV1 E4 inclusions. (**D**) Cytoplasmic SRPK1 and cytoplasmic HPV1 E4 inclusions. Arrowheads indicate HPV1 E4 inclusions. Scale bar indicates 10 μm.

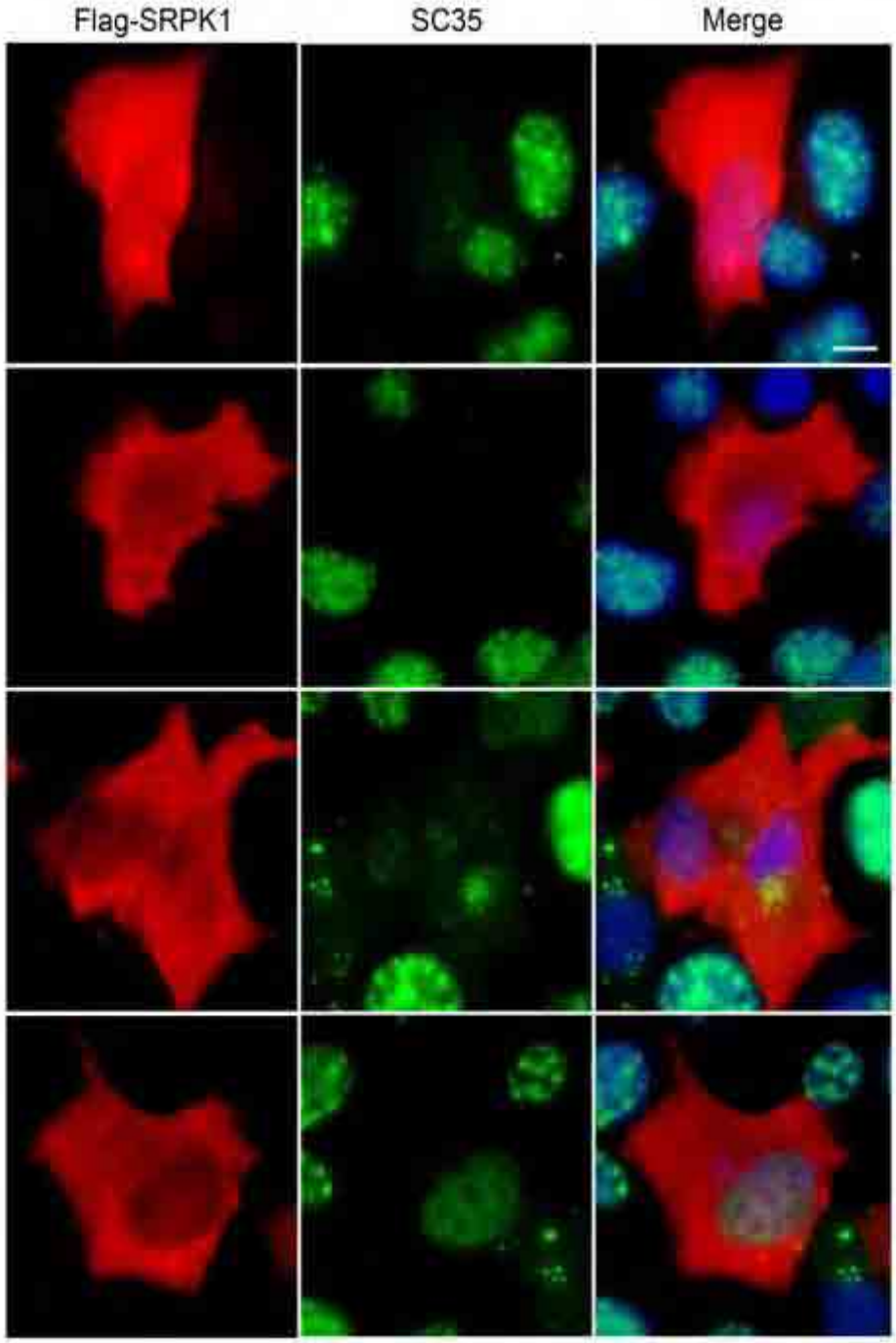

Figure 5.16 Localisation of SC35 with Flag-SRPK1 expression and SRPIN340 treatment. H1299 cells were transfected with Flag-SRPK1 and treated with 10 μM SRPIN340, four alternative images under the same conditions. Cells were stained with rabbit anti-Flag (red) and mouse anti SC35 (green), nuclei were stained with DAPI (blue). Scale bar indicates 10 μm.

Inhibition of SRPK1 kinase activity, either by cotransfection with HPV1 E1^E4 or by treatment with SRPIN340 had no effect on the relocalisation of SR proteins caused by the overexpression of SRPK, except when the SRPK1 was recruited to nuclear E4 inclusions.

5.7 The effect of HPV1 E1^E4 on SRPK1 function in alternative splicing

The phosphorylation state of the RS domain of SR proteins is altered during pre-mRNA splicing, with the initial recruitment of SR proteins to splice sites requiring the RS domain to be hyperphosphorylated (Gui et al., 1994a). The precise role of SRPK1 in splicing is yet to be determined, therefore the interaction between SRPK1 and HPV1 E1^E4 may act to modulate splice site selection and allow for preferential viral genome splicing.

5.7.1 The minigene system

The minigene system allows the identification and *in vivo* analysis of *cis*-acting regulatory elements and *trans*-acting factors which establish splicing efficiency and regulate alternative splicing. Minigenes contain an exon and flanking introns in an expression plasmid and allow the easy analysis of factors which affect the alternative splicing of that splice junction. Minigene constructs were obtained for X16 (mouse SRp20) (Figure 5.17A), which allowed a mammalian splice junction to be investigated and the adenovirus E1A gene (Figure 5.17B), which allowed a viral splice site to be investigated. The X16 minigene contained exon 4 of X16 flanked by part of exons 3 and 5 and the corresponding introns. Two alternatively spliced variants are produced; these are termed either ‘exon 4 included’ or ‘exon 4 skipped’ and can easily be distinguished between when resolved by native PAGE (Figure 5.17A). The E1A minigene produced five alternatively spliced variants, termed 9S, 10S, 11S, 12S and 13S, from the three 5’ splice sites and two 3’ acceptor sites as shown in figure 5.17B.

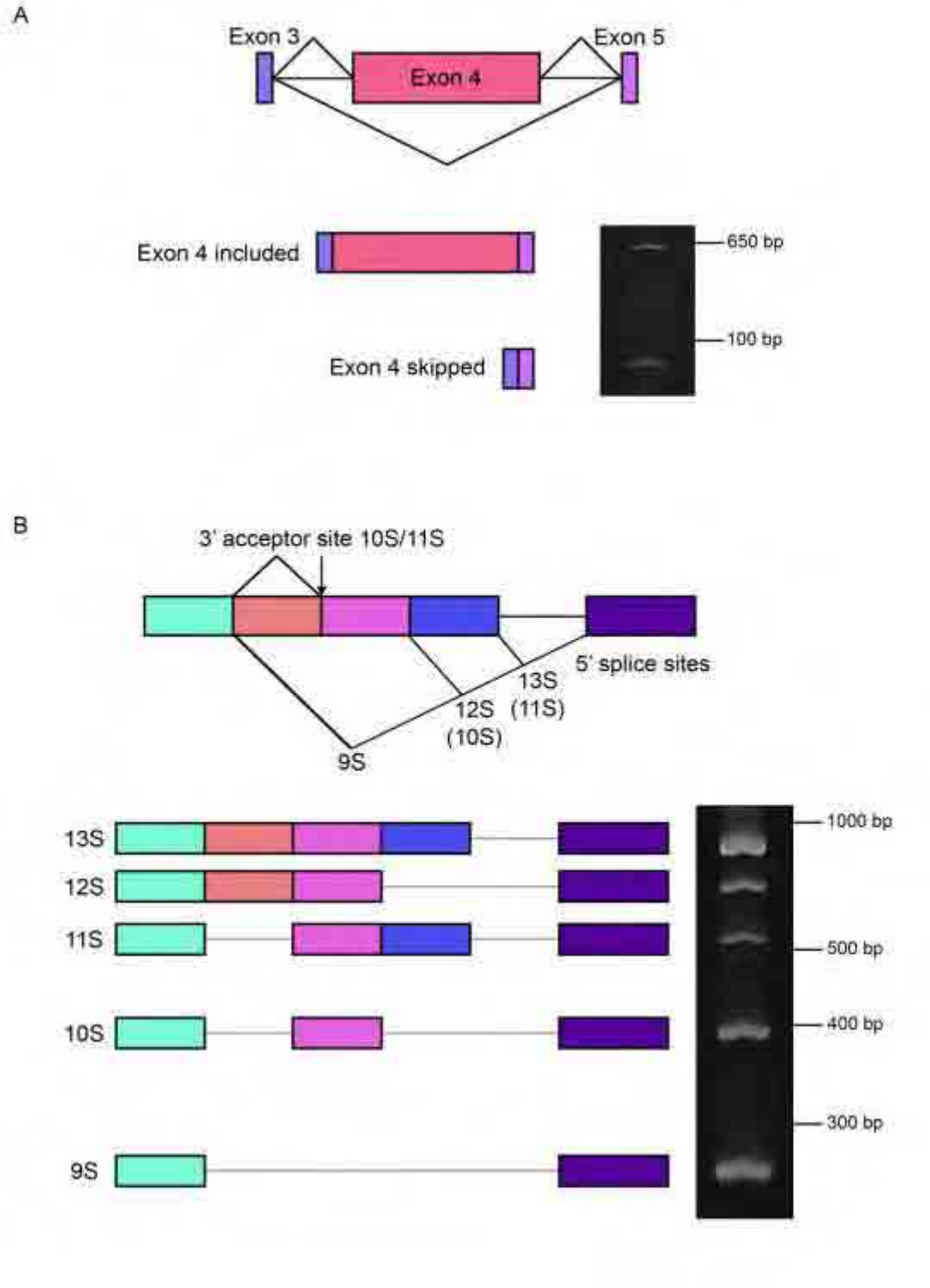

Figure 5.17 X16 and E1A minigene constructs. (**A**) The alternative splicing of the X16 minigene. Ethidium bromide stained native PAGE showing the two different splice variants. (**B**) The alternative splicing of the E1A minigene. Ethidium bromide stained native PAGE showing the five alternative splice variants.

5.7.2 The effect of SR proteins and SRPK1 on splice site selection

The SR proteins SRp20 and ASF/SF2 are known to differentially regulate the X16 minigene, with SRp20 promoting exon 4 inclusion and ASF/SF2 promoting exon 4 skipping (Jumaa and Nielsen, 1997). To aid with quantification α-^{32}P dCTP was added to the PCR reaction, this allowed the resolved native PAGE gel to be analysed using the phosphorimager. When the X16 minigene was transfected into 293 cells along with GFP, GFP-ASF/SF2, HA-SRp20 and Flag-SRPK1, the GFP-containing cells produced a mixture of the two splice variants, with 56% of the spliced RNA containing exon 4. The GFP-ASF/SF2 expressing cells, produced mainly the exon 4 skipped variant (93%), and those expressing HA-SRp20 expressed mainly the exon 4 inclusion variant (84%) (Figure 5.18). These results confirm that ASF/SF2 and SRp20 differently regulate the X16 minigene ($p = 0.00002$). The expression of Flag-SRPK1 had no significant effect on the splice site selection of the X16 minigene ($p = 0.67$-0.87), however the anti-Flag Western blot showed that the Flag-SRPK1 transfection was not very efficient.

The SR protein ASF/SF2 is known to activate proximal 5' splice sites, prevent abnormal exon skipping and promote exon inclusion with the E1A minigene (Caceres et al., 1994). When the E1A minigene was transfected into 293 cells along with GFP, GFP-ASF/SF2, HA-SRp20 and Flag-SRPK1, the cells expressing GFP produced a mixture of the five splice variants, the main forms produced are the 13S (33%) and 9S (28%) (Figure 5.19). Expression of GFP-ASF/SF2 increased the amount of 13S produced to 55% of the total RNA compared to the GFP transfected cells ($p = 0.003$); however all of the other splice variants were present. Expression of HA-SRp20 had no significant effect on the splice variants produced when compared to the GFP expressing cells ($p = 0.79$ for 13S). The expression of Flag-SRPK1 had no significant effect on the splice variants produced ($p = 0.77$-0.93 for 13S).

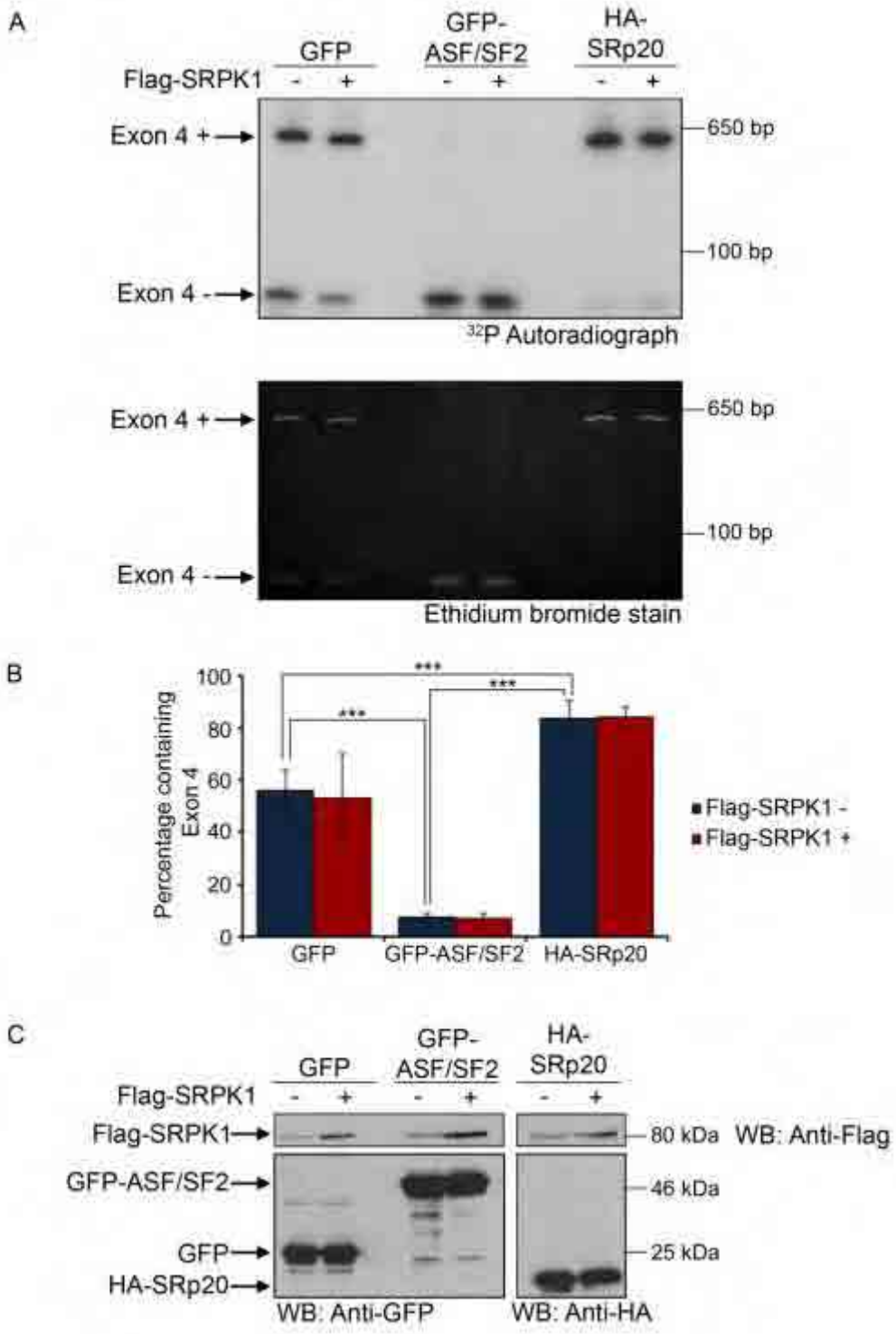

Figure 5.18 X16 minigene with ASF/SF2, SRp20 and SRPK1. GFP, GFP-ASF/SF2 and HA-SRp20 in the presence and absence of Flag-SRPK1 were co-transfected with the X16 minigene into 293 cells. (**A**) ^{32}P autoradiograph and ethidium bromide stained native PAGE gel showing the alternative splice variants. (**B**) Graph showing the percentage of exon 4 included RNA produced. The data show the mean ± standard deviation of five repeats. *** = $p < 0.0001$. (**C**) Western blot showing transfection levels of GFP, GFP-ASF/SF2, HA-SRp20 and Flag-SRPK1.

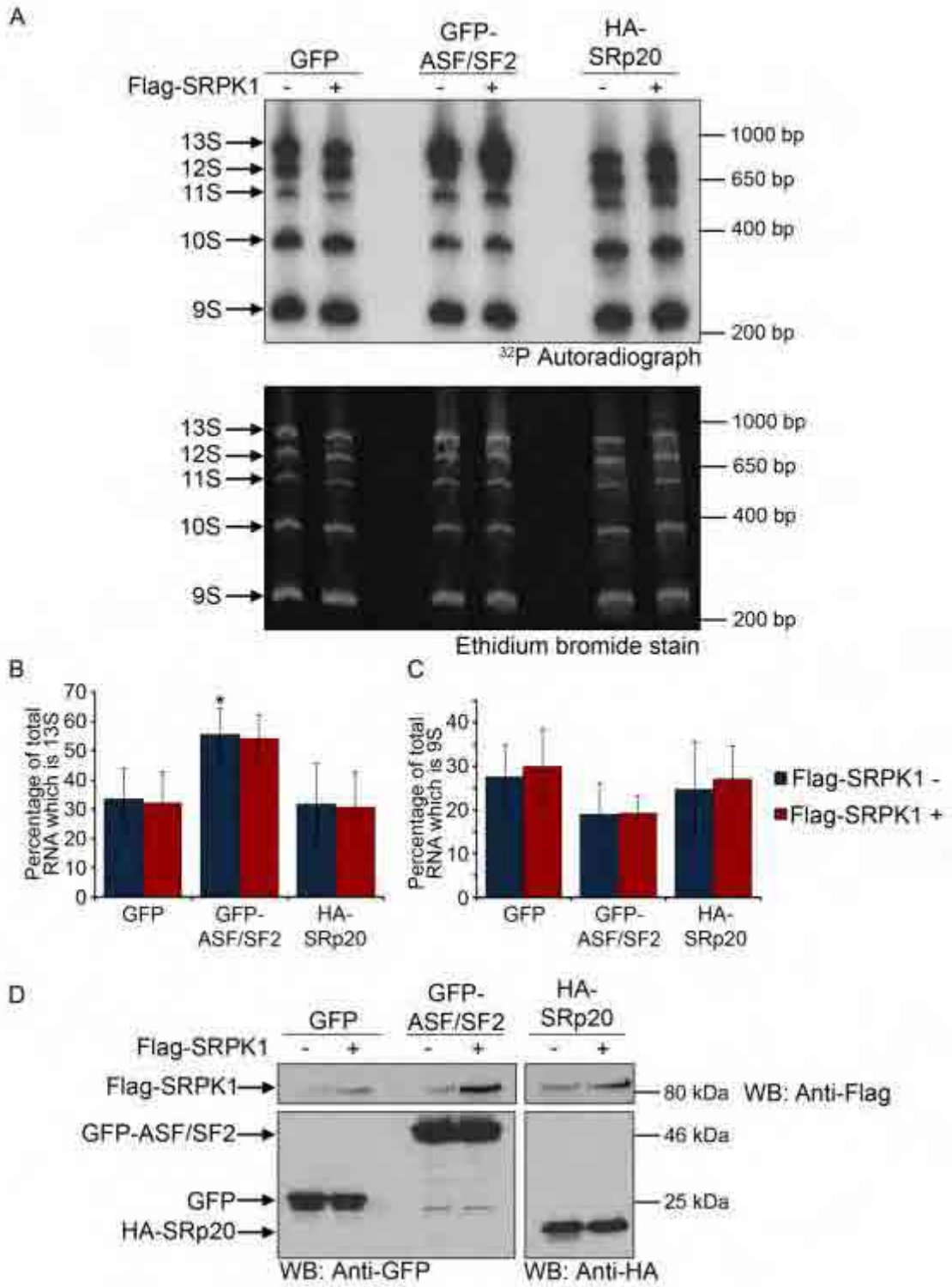

Figure 5.19 E1A minigene with ASF/SF2, SRp20 and SRPK1. GFP, GFP-ASF/SF2 and HA-SRp20 in the presence and absence of Flag-SRPK1 were co-transfected with the E1A minigene into 293 cells. (**A**) ^{32}P autoradiograph and ethidium bromide stained native PAGE gel showing the alternative splice variants. (**B**) Graph showing the percentage of spliced RNA on the gel which was 13S. * = $p < 0.005$. (**C**) Graph showing the percentage of spliced RNA on the gel which was 9S. Graphs show the mean ± standard deviation of five repeats. (**D**) Western blot showing transfection levels of GFP, GFP-ASF/SF2, HA-SRp20 and Flag-SRPK1.

5.7.3 The effect of HPV1 E1^E4 on splice site selection

The previous experiments showed that different SR proteins affect the choice of splice site with two different minigenes and that the overexpression of SRPK1 had no effect on the splice variants produced. However, interference with activity of endogenous SRPK1 might affect the splicing reactions within the minigene system. Therefore, to investigate whether HPV1 E1^E4 had any effect on splice site selection, HPV1 E1^E4 was transfected into 293 cells with Flag empty vector or Flag-SRPK1 and GFP, GFP-ASF/SF2 or HA-SRp20 and the X16 minigene (Figure 5.20). The expression of HPV1 E1^E4, with and without Flag-SRPK1 had no significant effect on the splice site selection with the X16 minigene, even in the presence of either of the two SR proteins ($p = 0.31$-91). In some experimental repeats the presence of HPV1 E1^E4 decreased the production of the exon 4 included variant in the presence of HA-SRp20 and Flag-SRPK1; however this result was not consistent between repeats. This data suggests that if HPV1 E1^E4 is controlling host splicing, this is on a cell-by-cell basis and is not the major functional consequence of the interaction of HPV1 E1^E4 and SRPK1. However as the X16 minigene is a mouse gene, the regulation of X16 splicing may not be performed by exactly the same mechanisms as for a human gene, therefore for any conclusions about the effect of HPV1 E1^E4 on host splicing requires additional experiments using human genes to be performed. The expression of HPV1 E1^E4 had no effect on splice site selection with the Adenovirus E1A minigene either (data not shown), this suggests that HPV1 E1^E4 is not controlling viral gene splicing. However the splicing of HPV genes may be regulated by different mechanisms to the Adenovirus E1A gene, therefore further investigation is required to determine the effect of HPV1 E1^E4 on viral splice site selection.

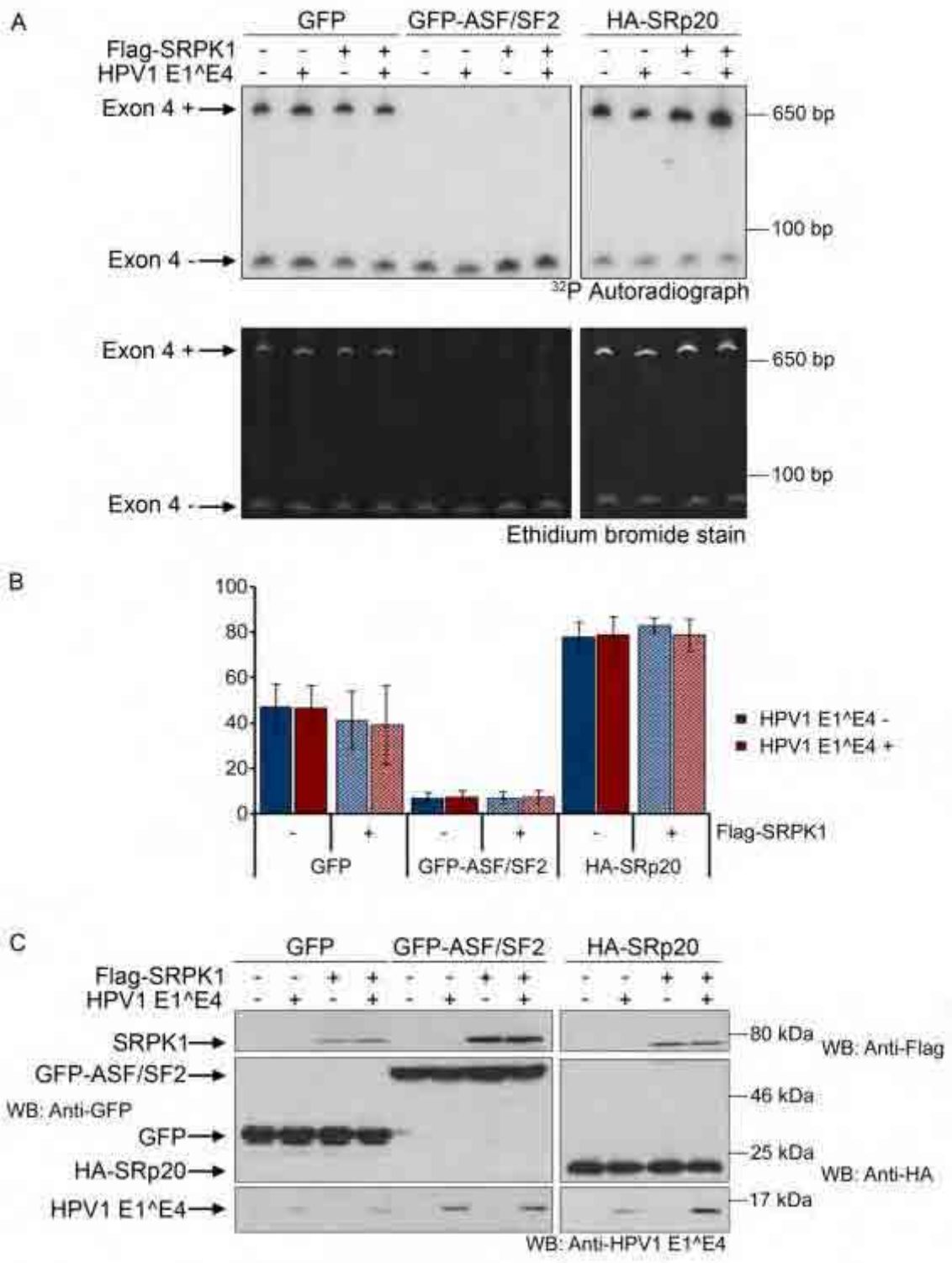

Figure 5.20 X16 minigene with HPV1 E1^E4. HPV1 E1^E4 with Flag empty vector or Flag-SRPK1 and GFP, GFP-ASF/SF2 or HA-SRp20 were co-transfected with the X16 minigene into 293 cells. (**A**) ^{32}P autoradiograph and ethidium bromide stained native PAGE gel showing the alternative splice variants. (**B**) Graph showing the percentage of the spliced RNA on the gel which was exon 4 included. Graph shows the mean ± standard deviation of four repeats. (**C**) Western blot showing transfection levels of GFP, GFP-ASF/SF2, HA-SRp20, Flag-SRPK1 and HPV1 E1^E4.

5.8 Discussion

The findings of co-expression of HPV1 E1^E4 and SRPK1 in cells has shown that HPV1 E1^E4 inhibits SRPK1 activity *in vivo*. The level of inhibition within the *in vivo* experiment however is much reduced than that observed *in vitro* (36% inhibition compared to 65%). This may be because not all of the SRPK1 is in complex with HPV1 E1^E4, as observed using confocal microscopy of immunostained cells (Figure 5.3). Interestingly, SRPK1 that was coprecipitated with the HPV1 E1^E4 antibody was able to phosphorylate ASF/SF2 to a small extent, suggesting that SRPK1 may be able to bind to HPV1 E1^E4 in a manner which does not always cause the loss of kinase activity.

It has been shown previously that HPV1 E1^E4 causes the relocalisation of exogenous SRPK1 to E4 inclusion bodies (Bell et al., 2007), and this relocalisation is observed with endogenous SRPK1 as well (Figure 5.2). The SRPK1 was also colocalised with HPV1 E1^E4 which was not in inclusions, demonstrating that the relocalisation of SRPK1 is not dependent upon the formation of E4 inclusions. In warts, although numerous E4 inclusions are present in cells in the lower suprabasal layers there is no evidence of SRPK1 sequestration; however the level of SRPK1 expression is low in these cells. Hence, sequestration may be dependent on specific E4 posttranslational modifications that occur only in the more superficial cells.

The distribution of HPV5 E1^E4 has not been reported previously and here it is shown that HPV5 E1^E4 was located in the cytoplasm and nucleus, with several characteristic staining patterns (Figure 5.4). The different staining patterns of HPV5 E1^E4 may be cell cycle regulated. HPV5 E1^E4 formed distinct cytoplasmic aggregates, which may be similar to the E4 inclusions formed in infections with other cutaneous HPVs including HPV1 and HPV2 (Doorbar and Gallimore, 1987). SRPK1 was observed to be colocalised to these aggregates, demonstrating that HPV5 E1^E4 can cause the relocalisation of SRPK1. Additionally, the

presence of HPV5 E1^E4 caused SRPK1 to form aggregates which were not colocalised with HPV5 E1^E4. The nature of these SRPK1 aggregates are unknown, however SRPK1 has been shown to localise to stress granules under chemical stress (Raithatha and Roberts, unpublished data), therefore the SRPK1 aggregates formed with HPV5 E1^E4 may be stress granules.

The overexpression of SRPK1 has been shown to lead to the disassembly of nuclear speckles and the cytoplasmic accumulation of the SR proteins (Gui et al., 1994a). Studies using a kinase dead SRPK1, have confirmed that the kinase activity of SRPK1 is required to cause this relocalisation (H. Y. Wang et al., 1998). The loss of nuclear speckles and nuclear staining of multiple SR proteins was observed with the overexpression of SRPK1 and this was most obvious when SRPK1 was in the nucleus and concurs with the findings of others that indicate SRPK1 must be present in the nucleus to cause the loss of nuclear speckles. The sequestration of SRPK1 to cytoplasmic E4 inclusions however, did not prevent the translocation of SRPK1 into the nucleus. In contrast, in cells where SRPK1 is sequestered to nuclear E4 inclusions the extent of disassembly of nuclear speckles in these cells was often reduced; therefore the sequestration of SRPK1 to E4 nuclear inclusions may inhibit this form of the kinase from phosphorylating SR proteins within speckles as occurs in cells expressing a kinase dead SRPK1. Surprisingly inhibition of SRPK1 with SRPIN340 also did not impede the disassembly of nuclear speckles with SRPK1 overexpression; this was unexpected but may be due to the conditions used in the experiment, as the SRPIN340 was added 4 h into the transfection. The relocalisation caused by SRPK1 overexpression may occur before this time and not be rescued by the inhibition of SRPK1 activity.

This study confirmed that SRp20 and ASF/SF2 differentially regulate the X16 minigene; however the overexpression of SRPK1 had no effect on the splice isoform produced. This is in contrast to the observations of Stoss et al., (1999), who showed that CLK2 can promote the

exclusion of the alternative exon 4; and demonstrates the specificity of the kinases *in vivo*. In accordance with Yomoda et al., (2008), SRPK1 overexpression had no effect on the splice isoforms produced with the E1A minigene, again demonstrating the functions of the different kinases in splicing regulation. The SR proteins ASF/SF2 and SRp20 produced different splice isoforms with the E1A minigene, however unlike Caceres et al., (1994); all isoforms were still present with the expression of ASF/SF2, whereas their study demonstrated a complete switch to the largest 13S isoform. The discrepancies between the studies may be due to differences in experimental conditions e.g. transfection efficiency. The data presented in this chapter confirms that different SR proteins perform different roles in splice site selection, with some splice sites only having ESEs which are recognised by specific SR proteins.

Somewhat surprisingly, the addition of HPV1 E1^E4 to the minigene systems had no overall effect on the splicing isoforms produced. However in the presence of SRp20, different results were observed in individual repeats with the X16 minigene, with some showing a decrease in the amount of exon 4 included isoform in the presence of HPV1 E1^E4. The reasons for this are unclear, but since overexpression of SRPK1 did not have any effect upon splicing in the X16 minigene system it could be that only a specific population of endogenous SRPK1 is involved and HPV1 E1^E4 may only target this form in a subpopulation of cells. It might also be possible that any effect of HPV1 E1^E4 upon SRPK is dependent on posttranslational modification of the viral protein e.g. N-terminal processing (Knight et al., 2004) and additional forms are required alongside the full-length protein. It is also possible that HPV1 E1^E4 is unable to target and inhibit the form of SRPK1 that is involved in splicing reactions but perhaps inhibits other functions of SRPK1.

CHAPTER 6 ROLE OF SRPK1 IN EPITHELIAL DIFFERENTIATION

6.1 Introduction

Multiple viruses have been shown to modulate SR proteins during infection, one such virus is HIV, which has been shown to induce the degradation, dephosphorylation and relocalisation of SR proteins. Interestingly overexpression of SRPK1 or SRPK2 in the presence of HIV, can avoid the virus mediated degradation of the SR proteins (Fukuhara et al., 2006), demonstrating that HIV is modulating SRPK activity. The phosphorylation of SR proteins by SRPKs has been shown to stabilise the proteins and accordingly inhibition of the SRPKs with the small molecule inhibitor SRPIN340, promotes their degradation (Fukuhara et al., 2006). Stabilisation of several SR proteins, including SRp75, by the overexpression of SRPK2, increases HIV production; however inhibition of the SRPKs with SRPIN340 has little effect on suppressing HIV replication (Fukuhara et al., 2006). In contrast the inhibition of SRPKs with SRPIN340 has been shown to inhibit the propagation of other viruses, including Sindbis virus, severe acute respiratory syndrome coronavirus (SARS), cytomegalovirus (CMV) and hepatitis C virus (HCV), which are all acutely replicating viruses (Fukuhara et al., 2006, Karakama et al., 2010). The production of HCV 1b and 2a replicons were suppressed with treatment with SRPIN340, but not with a control compound which does not inhibit SRPKs (Karakama et al., 2010). The effect of this inhibition upon SRPKs was directed at the intracellular core protein of HCV whose expression was downregulated by treatment with SRPIN340, and following the knockdown of SRPK1 and SRPK2 with siRNA (Karakama et al., 2010). These findings demonstrate that SRPKs are required for the efficient replication of HCV and that the targeting of SRPKs is a potential anti-viral treatment option for HCV.

The organotypic raft system allows cultured keratinocytes to stratify and differentiate in a manner similar to normal epithelium such that the structure accurately mimics the in vivo physiology of the epidermis. The organotypic raft system most commonly used today is based on the technique developed by Asselineau and Prunieras (1984). Briefly, primary keratinocytes are seeded onto a collagen plug in which mouse fibroblasts are embedded. Once confluence is attained the structure is then lifted to the air-liquid interface using a metal grid support and the cultures maintained for up to 2 weeks to allow full differentiation to occur. Studies of the HPV life cycle have been seriously hampered by the lack of tissue culture systems in which to grow the virus, but the adaption of the organotypic raft system to HPV containing cells has been an invaluable technique. Alongside the use of rafts has been the development of keratinocyte based models that support maintenance and replication of viral episomes from high-risk HPV types (Frattini et al., 1996). Transfection of viral episomes into primary keratinocytes isolated from cervical, foreskin or tonsillar sites enables the establishment of stable cells lines that maintain the episomes. Upon differentiation in organotypic rafts the complete virus life cycle is recapitulated. Thus primary HFKs containing viral genomes grown as organotypic raft cultures allow late, differentiation-dependent virus functions to be studied, for example see (Knight et al., 2011).

Little is known of the function of SRPK1 in epidermal differentiation and there is growing evidence that the family of SRPK kinases have functions other than those connected to splicing and RNA metabolism (Giannakouros et al., 2011). Likewise the role these kinases play in the virus life cycle is also unknown, although it is known that the virus modulates SR protein expression and there is evidence that some of these SR proteins are required for HPV RNA splicing (Jia et al., 2009, Somberg and Schwartz, 2010). Identifying the role of SRPK1 in epidermal differentiation and in the life cycle may be a successful route to understand the function of E1^E4 association with SRPK1 and also to understand why some HPVs such as

HPV1 inhibit SRPK1. As SRPK1 has been shown to be required for the replication of several viruses (Fukuhara et al., 2006) understanding SRPK1 function in the HPV life cycle might identify a new antiviral therapeutic target for viral intervention. To these ends, this chapter examines the distribution of SRPK1 in organotypic raft cultures of normal primary human keratinocytes and in rafts of HPV18 genomes containing keratinocytes (Wilson et al., 2007). Additionally the effect of inhibition of SRPK by using the small molecular inhibitor SRPIN340 on early and late virus functions will also be examined

6.2 Distribution of SRPK1 upon epithelial differentiation

In primary epidermal tissue, SRPK1 has been shown to be present in the basal and suprabasal cells, with expression levels increasing as the cells become more differentiated (Bell et al., 2007). SRPK1 staining is however mostly absent from the fully differentiated cornified layer. This pattern of SRPK1 distribution suggests that epithelial terminal differentiation might be dependent on the regulation of SRPK1 function. To investigate this further, the organotypic raft system of primary human keratinocytes was used (Chapter 2, 2.6.5). When primary human foreskin keratinocytes (HFKs) were allowed to differentiate in the raft system and were fixed and sectioned, the different epithelial layers from basal through parabasal, spinous and granular to the cornified layer were clearly present when stained with haematoxylin and eosin (H and E) (Figure 6.1A). The haematoxylin stains nuclei and acidic structures blue and the eosin stains the cytoplasm and basic structures shades of pink.

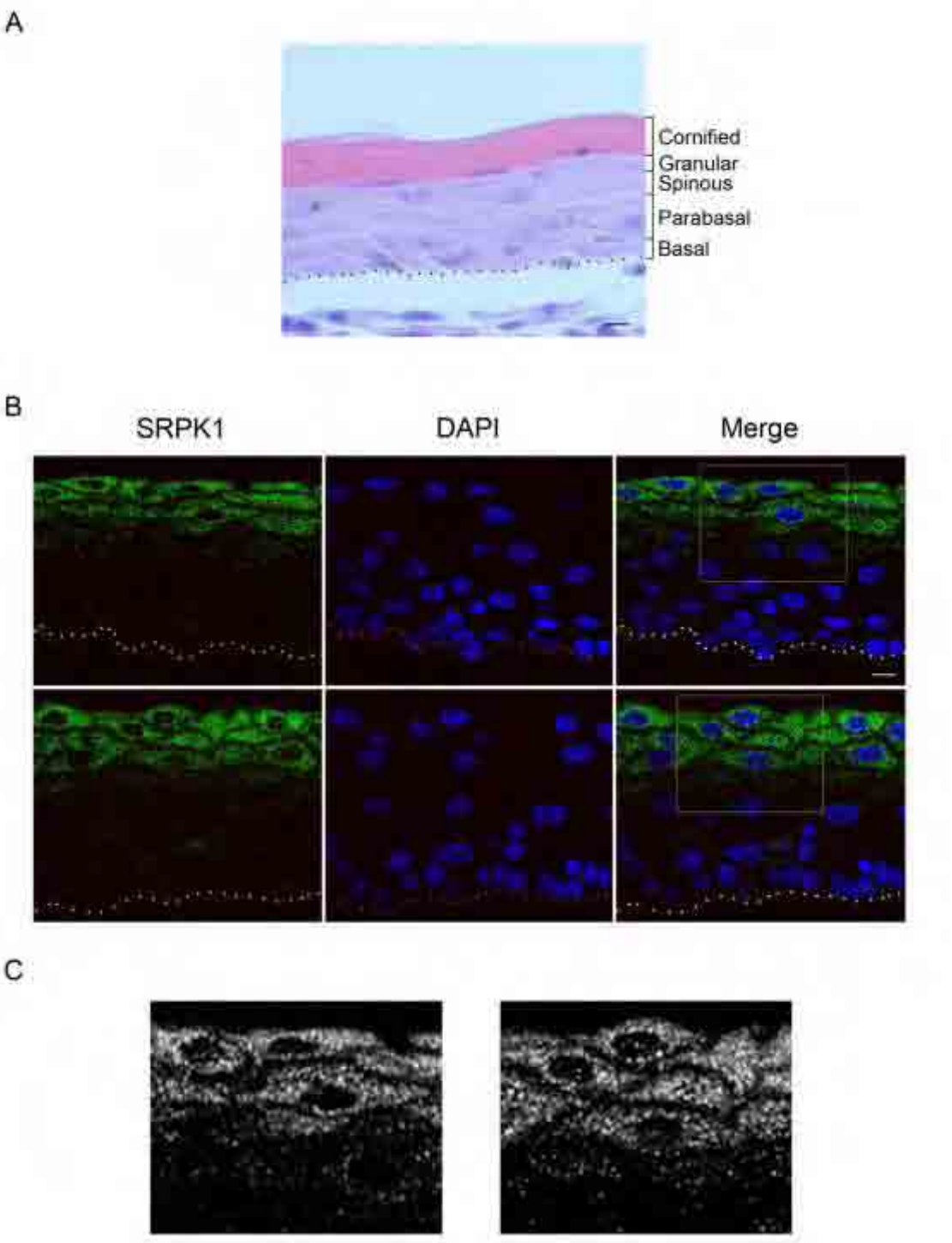

Figure 6.1 Expression of SRPK1 in organotypic rafts of primary human foreskin keratinocyte. (**A**) H and E stained section of untransfected HFKs; the layers of stratified epithelium are indicated. (**B**) HFKs stained with anti-SRPK1 monoclonal antibody (green) and nuclei stained with DAPI (blue). (**C**) Expanded boxed area of B, showing SRPK1 staining in upper stratified layers. Immunofluorescence staining was analysed by confocal microscopy, scale bar indicates 10 µm.

The raft sections were subjected to antigen retrieval using a low-temperature EDTA method (Watson et al., 2002) and then stained with an anti-SRPK1 monoclonal antibody (Figure 6.1B). Similar to the primary tissue, SRPK1 was present at low levels in the basal and parabasal layers. The expression of SRPK1 was upregulated in correlation with the morphological differentiation of the cells, with SRPK1 expression being highest in the granular layer (Figure 6.1C). SRPK1 staining had a grainy cytoplasmic appearance and unlike in monolayer culture, SRPK1 was not present in the nuclei of the differentiated cells.

6.2.1 Effect of HPV18 on SRPK1 distribution

When HPV18 containing HFKs are grown in the raft system, there a several differences in the morphology of the epithelia produced. These include the retention of nuclei in the upper stratified layers and thickening of the parabasal and spinous layers (Figure 6.2A). To determine if the distribution of SRPK1 is altered in cells containing HPV18 genomes, the HPV18 DNA-containing rafts were stained with the SRPK1 monoclonal antibody and SRPK1 distribution was compared to the pattern in non-HPV containing cells (Figure 6.2B). The distribution of SRPK1 expression was comparable to the differentiation dependent expression in the absence of HPV18. However, in cells in the upper spinous and granular layers, where SRPK1 expression is highest, SRPK1 formed prominent aggregates in the cytoplasm and these were often observed to be distributed around the nuclear membrane (Figure 6.2C). Akin to the untransfected cells, there was no accumulation of SRPK1 in the nuclei. Whilst SRPK1 aggregates are present in untransfected cells, in the HPV18 containing cells, the aggregates are more often larger and more numerous suggesting that the presence of HPV18 was modulating SRPK1 subcellular distribution.

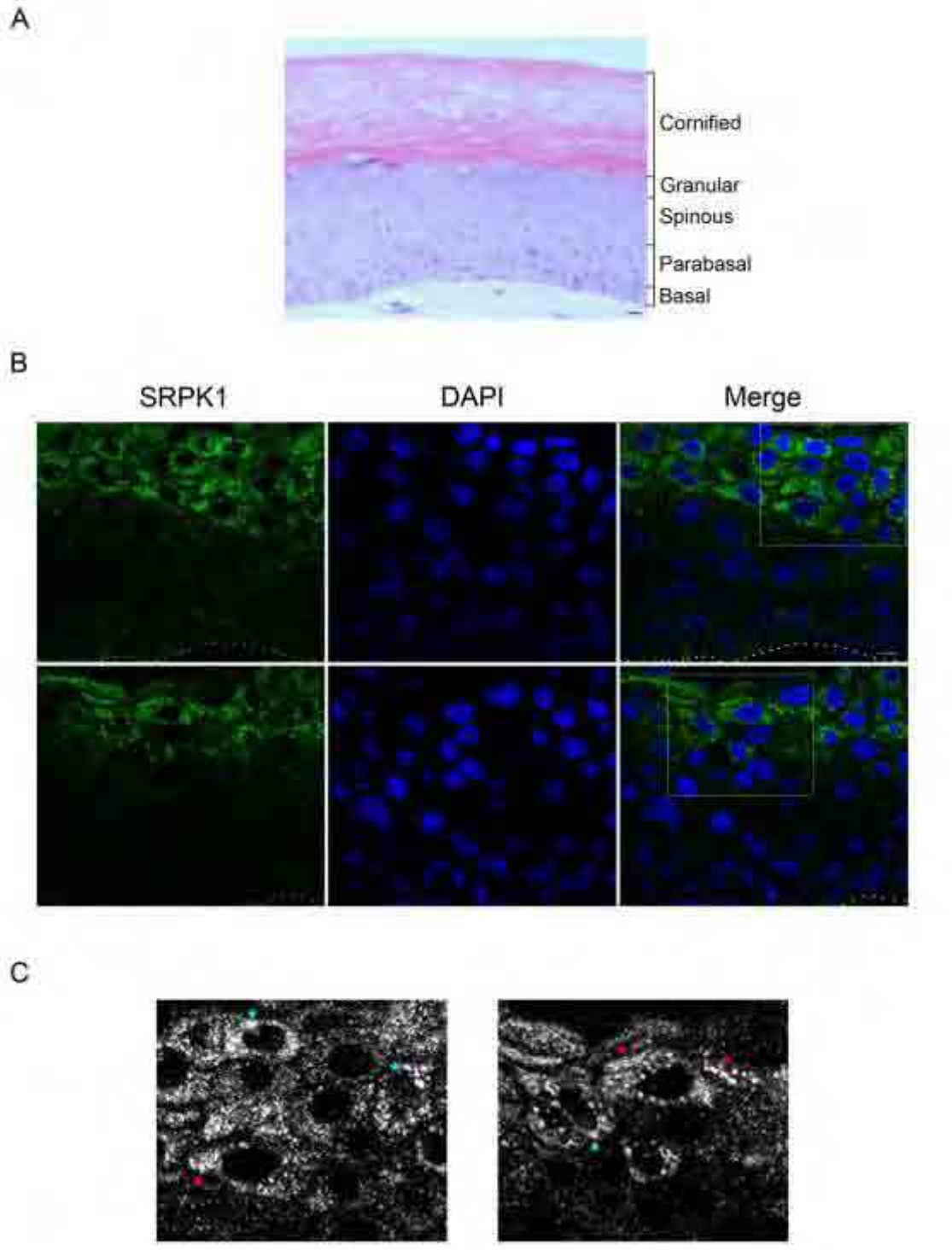

Figure 6.2 SRPK1 expression in HPV18 genome containing organotypic rafts. (**A**) H and E stained organotypic raft of HFKs containing the HPV18 genome, layers of stratified epithelium are indicated. (**B**) HPV18 containing HFKs stained with anti-SRPK1 monoclonal antibody (green) and nuclei stained with DAPI (blue). (**C**) Expanded boxed area of B, showing SRPK1 staining in upper stratified layers. Red arrowheads indicate SRPK1 aggregates, blue arrowheads indicate SRPK1 aggregates at nuclear membrane. Immunofluorescence staining was analysed by confocal microscopy, scale bar indicates 10 μm.

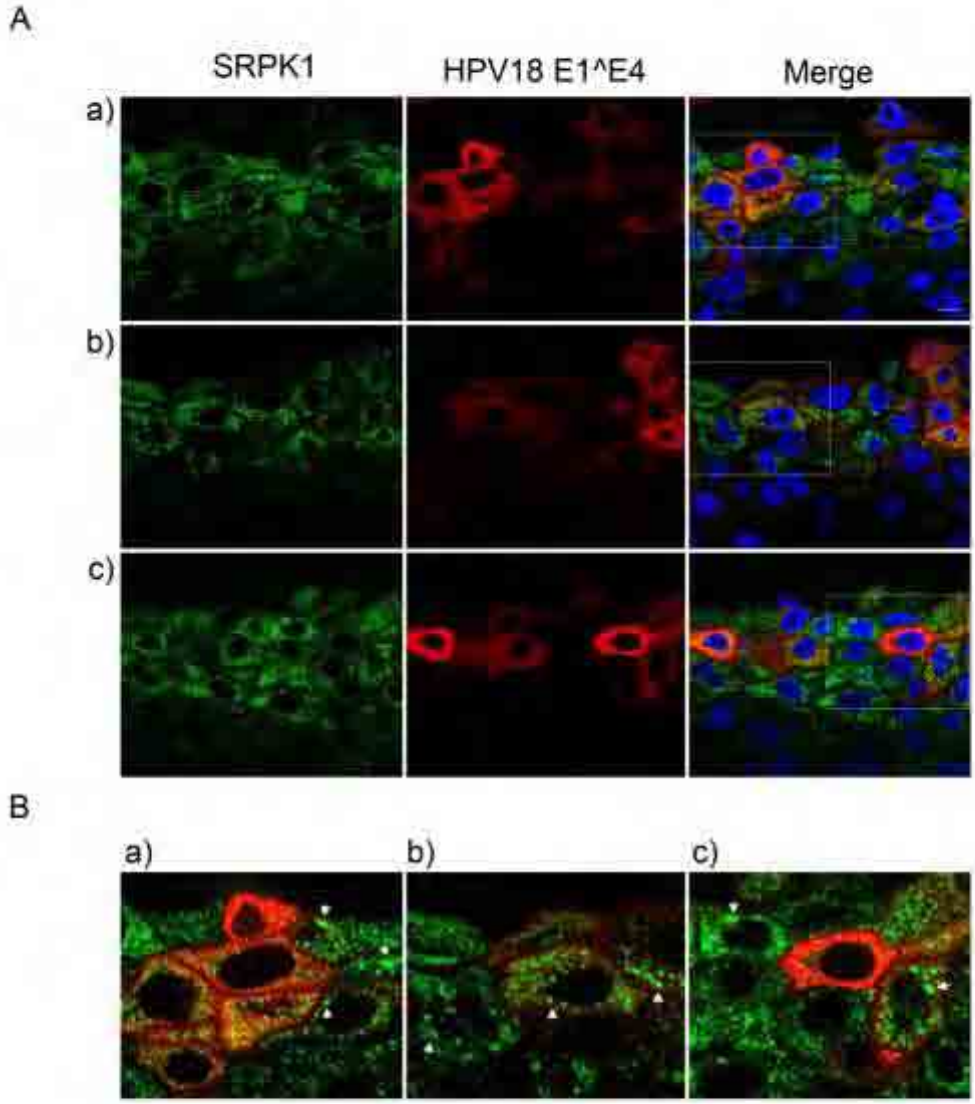

Figure 6.3 Localisation of SRPK1 and E1^E4 in HPV18 genome containing rafts. (**A**) HPV18 genome containing HFKs stained with mouse anti-SRPK1 (green) and rabbit anti-HPV18 E1^E4 (R424, red), nuclei are stained with DAPI (blue). (**B**) Expanded boxed areas of panels a) to c), showing SRPK1 and HPV18 E1^E4 staining. Arrowheads indicate SRPK1 aggregates. Immunofluorescence staining was analysed by confocal microscopy, scale bar indicates 10 μm.

Since SRPK1 accumulates to E4 inclusions within HPV1 warts and HPV18 E1^E4 interacts with SRPK1 *in vitro* (Chapter 3, 3.4) (Bell et al., 2007), it is possible that HPV18 E4 is modulating the distribution of SRPK1. Therefore to determine whether E4 is present in the SRPK1 aggregates HPV18 containing rafts were co-stained for SRPK1 and HPV18 E1^E4 (Figure 6.3A). The SRPK1 aggregates were present both in cells that were positive for E1^E4 and those which were negative (Figure 6.3B). This data suggests that HPV18 E1^E4 does not cause the aggregation of SRPK1 in the upper layers of the epithelium.

6.3 The effect of SRPK inhibition on HPV18 genome containing HFKs

6.3.1 In monolayer cell culture

The presence of HPV18 genomes has been shown to induce changes in the subcellular localisation of SRPK1, suggesting that SRPK1 function may be utilised by HPV18 for some part of the virus life cycle, or perhaps the function of this kinase is inhibited during the infectious cycle. To determine the importance of SRPKs in the HPV18 life cycle the small molecule inhibitor SRPIN340 can be used to inhibit SRPK kinase activity *in vivo* (Fukuhara et al., 2006). Initially, SRPIN340 was added to HPV18 containing HFKs growing in monolayer, in increasing concentrations to determine whether there was any toxic effect of SRPIN340 on cell growth and to determine if the inhibitor modulated viral episome replication (Figure 6.4). The SRPIN340 was dissolved in DMSO, so an equivalent amount of DMSO was added to controls for each SRPIN340 concentration to account for any solvent effects. To determine whether SRPIN340 had any toxic effects on the cells, a toxicity assay was performed (Chapter 2, 2.6.3) (Figure 6.4A). Cells were treated with different concentrations of SRPIN340 (1, 5, 10, 25 and 40 μM) and viable cells counted after 48 hours.

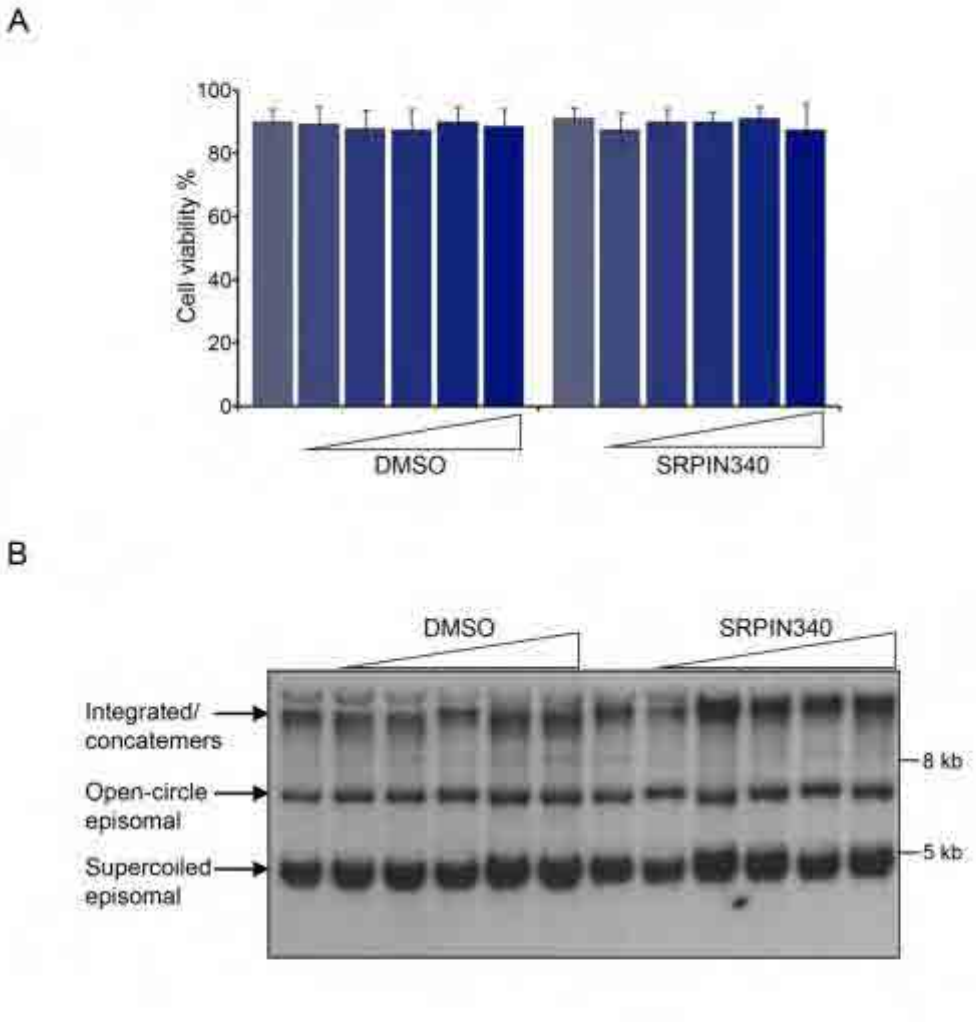

Figure 6.4 The effect of SRPIN340 on HPV18 containing HFKs. SRPIN340 of increasing concentrations was added to HFKs for 48 h, with equal amounts of DMSO added to control cells. (**A**) Cells were harvested and viable cells counted after trypan blue staining. Graph shows average ± standard deviation for 2 experimental repeats (10 counting chambers for each). (**B**) Autoradiograph of Southern blot of HPV DNA maintenance replication in SRPIN340 treated HPV18 containing HFKs. Blot probed with HPV18 DNA.

This showed that there was no significant change in cell viability with increasing SRPIN340 concentration, even at the highest concentration of 40 μM ($p = 0.63$) when compared to the DMSO control. Total DNA was extracted from the cells and analysed by Southern blot to monitor HPV18 genome maintenance (Chapter 2, 2.6.4) (Figure 6.4B). There was stable episome replication at all concentrations of SRPIN340 examined, suggesting that SRPK function was not required for the maintenance of HPV18 genomes in primary HFKs.

6.3.2 In organotypic rafts

The inhibition of SRPK activity had no effect on the maintenance of HPV18 genomes in monolayer culture, but since SRPK1 is most abundant in the upper layers of the epithelium where the virus is undergoing its vegetative cycle SRPK1 function may have more relevance to the later stages of the virus life cycle. To examine the effect of SRPIN340 on the stratification of the HPV genome containing cells SRPIN340 was added to cells growing in organotypic raft culture for the final seven days out of thirteen. The SRPIN340 was added in two concentrations to untransfected HFKs and cells containing HPV18 genomes. DMSO solvent controls were also included. After 13 days, rafts were formalin fixed and paraffin-embedded and sections analysed using H and E staining (Figure 6.5). Following SRPIN340 treatment, untransfected HFK stratified into multi-layered structure with a similar morphological appearance to untreated, or rafts treated with DMSO alone. However, rafts incubated with SRPIN340 showed slightly reduced thickness of the spinous and parabasal layers (Figure 6.5A). In the HPV18 genome containing HFKs, a dose dependent response was observed with SRPIN340 treatment. At 10 μM SRPIN340, the raft produced a stratified structure, however the different stratified layers were less defined than when compared to DMSO treated rafts. Additionally there was some disorder in the granular layer, as the cells were less granulated and more koilocytotic (Figure 6.5B).

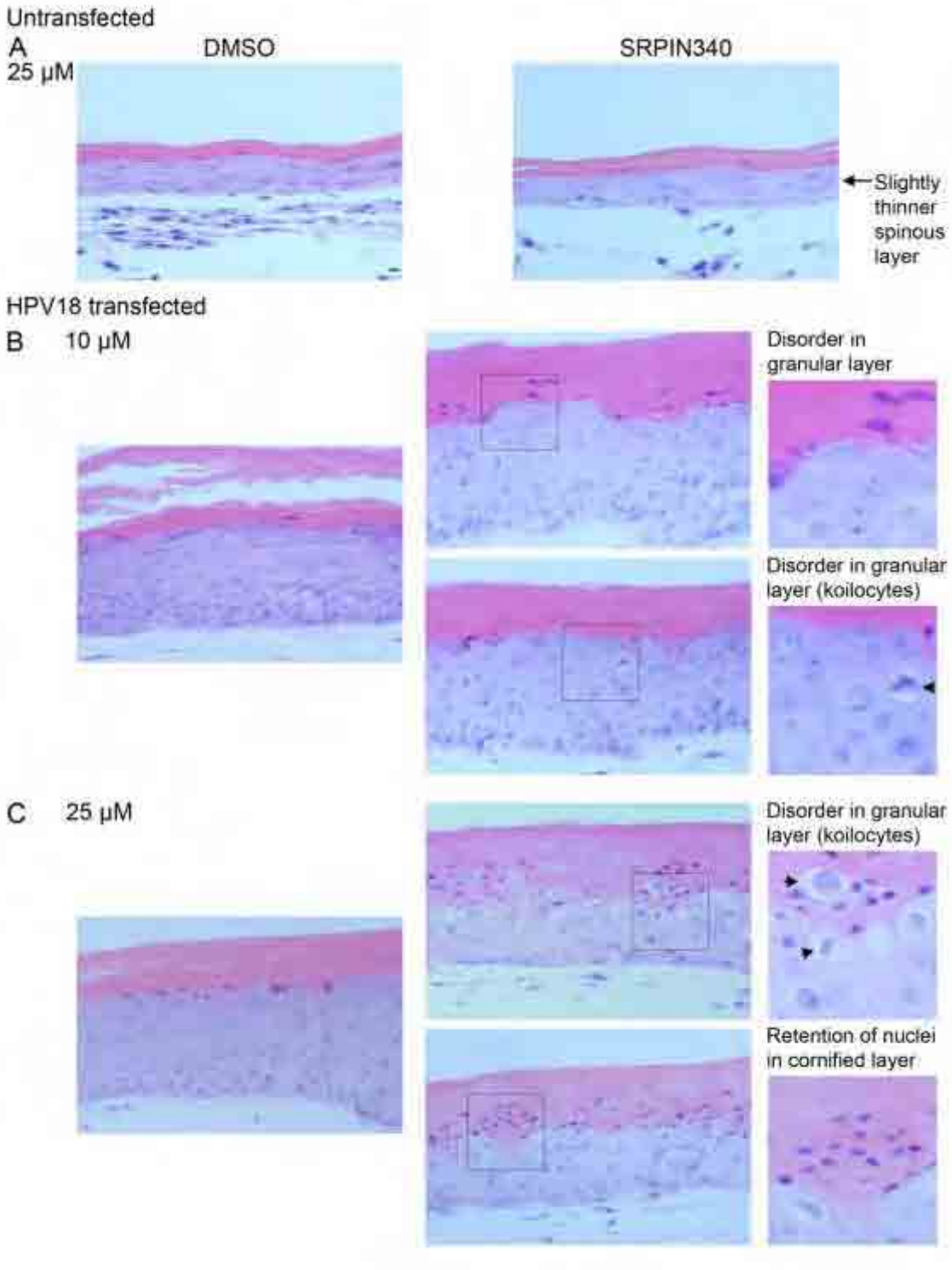

Figure 6.5 The effect of SRPIN340 on organotypic rafts. (**A**) Untransfected HFKs were treated with 25 µM SRPIN340 or an equal amount of DMSO. (**B**) HFKs transfected with HPV18 genomes were treated with 10 µM SRPIN340 or an equal amount of DMSO. (**C**) HFKs transfected with HPV18 genomes were treated with 25 µM SRPIN340 or an equal amount of DMSO. Arrowheads indicate koilocytes. Scale bar indicates 10 µm.

At the higher dose of SRPIN340 (25 μM), the thickness of the spinous and parabasal layers were reduced, and the disorder of the granular layer more marked at this concentration of SRPIN340. There was also a high degree of nuclei retention in the cornified layer (parakeratosis) and an increase in the amount of cells which are koilocytotic (Figure 6.5C). Koilocytic cells are epithelial cells which contain an acentric, hypercromatic, moderately enlarged nucleus which is displaced by a large perinuclear vacuole, koilocytic cells are a marker for the productive phase of HPV (Krawczyk et al., 2008).

The organotypic rafts were also treated with SRPIN340 for the complete thirteen days of growth of the rafts. Using these conditions of treatment gross changes in morphology were observed for rafts treated with SRPK inhibitor but also the DMSO vehicle control when compared to non-treated. There was a large amount of invasion of the underlying collagen layer, however as this was present in the DMSO treated cells as well as SRPIN340, this effect is likely to be due to the DMSO (data not shown). When SRPIN340 was added to the rafts for the last 2 days of growth, the rafts were thinner compared to the DMSO treated, with a small amount of disorder in the granular and spinous layers (data not shown).

The morphology of the organotypic rafts treated with SRPIN340 suggests that SRPKs are involved in some differentiation dependent functions of HPV18, as inhibition of SRPK function, resulted in thinner rafts with atypical stratification.

6.4 Discussion

The distribution and expression of SRPK1 in primary epidermal tissue and in primary HFKs grown in organotypic raft culture is similar. In the basal and parabasal layers the expression of SRPK1 is low; however as the cells differentiate the expression of SRPK1 increases up into the granular layer. In the fully differentiated cornified layer, SRPK1 staining is absent. The SRPK1 is only present in the cytoplasm of the differentiated cells, however in cultured

cells grown in monolayer, some cells have nuclear SRPK1 (Figure 5.2A). The nuclear transport mechanism for SRPK1 is unknown, however SRPK1 has been observed to be translocated to the nucleus at the end of the G_2 phase of cell cycle (J. H. Ding et al., 2006). The differentiated cells are no longer undergoing cell cycle; therefore there may be no longer be a need for nuclear SRPK1 in these cells. The SRPK1 staining has a grainy but fairly uniform appearance in the differentiated untransfected cells, but the presence of HPV18 genomes causes SRPK1 to form aggregates which are commonly located around the nuclear membrane. The formation of the SRPK1 aggregates is not dependent upon HPV18 E1^E4, as the aggregates are present in both cells positive and negative for HPV18 E1^E4 staining. The cause and the function of the SRPK1 aggregates is currently unknown, but they may be similar to the SRPK1-containing aggregates caused by transfection with HPV5 E1^E4 (Figure 5.5). HPV1 causes the relocalisation of SRPK1 to E4 inclusions; so the aggregation of SRPK1 into inclusions or aggregates may be a conserved strategy by which HPV modulates the functions of SRPK1. Staining of the rafts with an antibody against poly(A)-binding protein (PABP) a known constituent of stress granules, showed no colocalisation of stress granules with the SRPK1 aggregates (data not shown).

SRPIN340 has been shown to be non-toxic in heptato carcinoma cells (Huh7) and human embryonic kidney cells (Flp-In293) grown in cell culture up to 30 μM (Fukuhara et al., 2006, Karakama et al., 2010). This study shows that SRPIN340 is non-toxic to HPV18 genome containing primary HFKs up to 40 μM. The treatment of the HPV18 genome containing HFKs with SRPIN340, demonstrates that SRPKs are not required for the stable maintenance of episomal genomes in these cells. This suggests that SRPK regulation of SR proteins is not required for the efficient viral genome splicing, the other kinases able to phosphorylate SR proteins, such as the CLKs may perform this regulation. The treatment of organotypic rafts with SRPIN340 produces morphological changes in the stratification of the rafts. There is an

increased amount of koilocytosis and parakeratosis when SRPIN340 is added to the cells during differentiation. Since these morphological changes were not strongly evident in the DMSO controls then this suggests that SRPKs are modulating HPV18 replication in some as yet undefined manner. SRPIN340 caused the predominance of koilocytes, which are epithelial cells containing an acentric, hypercromatic, moderately enlarged nucleus that is displaced by a large perinuclear vacuole. The mechanism of koilocyte formation is largely unknown; however studies have linked the proteins E5 and E6 with their formation (Krawczyk et al., 2008). These cells are associated with virion production in the granular layers of the epithelium and are used as a marker of the productive phase of HPV. Additionally, treatment with SRPIN340 caused a high degree of parakeratosis, with nuclei retention in the cornified layer of the raft. Parakeratosis is beneficial to the virus, as the viral capsid is produced in the nucleus. The morphological changes observed with SRPIN340 treatment suggest that the inhibition of SRPK activity promotes the late vegative stages of the virus life cycle.

Time constraints did not allow for further analysis of the effect of SRPIN340 on the different viral functions that occur in suprabasal cells including vegetative events of viral genome amplification, capsid expression and virion production and infectivity.

CHAPTER 7 DISCUSSION

7.1 Overview of findings

The work in this thesis has sought to address the hypothesis that the interaction between E1^E4 and SRPK1 alters SRPK1 activity and that this association influences the functions of SRPK1 in the HPV life cycle. This study has uncovered the novel finding that the E1^E4 protein of HPV1 is a potent inhibitor of SRPK1 kinase activity *in vitro* and *in vivo* and that this inhibition is dependent upon E1^E4 binding to SRPK1. Other E1^E4 proteins also bind SRPK1 in vitro, including those of HPV16 and HPV18 (Bell et al., 2007), and as shown in this thesis the beta virus HPV5, however none of these influence SRPK1 activity *in vitro*. HPV1 E1^E4 inhibits phosphorylation of a range of SR protein substrates including the prototypical SR protein ASF/SF2, SRp20, SC35, SRp75 and 9G8 but with some evidence of differences in potency of inhibition by the E1^E4 protein. The serine-arginine rich beta HPV E2 proteins (5 and 8) are known to behave as SRPK1 substrates and in this study the E2 protein of HPV1 was also found to be phosphorylated by SRPK1 *in vitro* and HPV1 E1^E4 inhibited SRPK phosphorylation of this protein. Both SRPK1 binding partners HPV1 E1^E4 and HPV5 E1^E4 are substrates for the cellular kinase (Bell et al., 2007)(and this thesis) and overexpression of these E1^E4 proteins in cells sequesters SRPK1 to inclusions and aggregates, suggesting that altered SRPK1 localisation may modulate SRPK1 functions. However, upon overexpression of HPV1 E1^E4 only moderate effects on SR protein cellular localisation and on cellular and viral RNA splicing in minigene systems were observed. By making use of an HPV18-keratinocyte cell model (Wilson et al., 2007) grown as organotypic rafts it was found that SRPK1 localisation in the more differentiated cell layers is modified in comparison to cells negative for the virus. Moreover, addition of the small molecule inhibitor

of SRPK, SRPIN340 to these raft cultures enhanced the morphological features of HPV replication in a dose-dependent manner and suggests that HPV modulates SRPK activity to facilitate the virus life cycle.

7.2 Inhibition of SRPK1 by HPV1 E1^E4

The direct interaction between HPV1 E1^E4 and SRPK1 inhibited the kinase activity of SRPK1 both *in vitro* and *in vivo*. The biological significance of the inhibition of SRPK1 function is still unknown, however examination of the splicing regulation of mammalian and viral genes using the minigene system showed that the splice site selection was not altered significantly in the presence of HPV1 E1^E4. The splicing of HPV genes was not examined in this study and may be regulated by different proteins to the adenovirus gene analysed.

The role of SRPK1 overexpression in splice site selection has not been demonstrated by this study and others (Yomoda et al., 2008). Phosphorylation of the SR proteins by CLK overexpression has been shown to affect splice site selection (Stoss et al., 1999, Yomoda et al., 2008) however studies have not demonstrated a role for SRPK1. To determine whether SRPK1 does play a role in splice site selection, SRPIN340 or siRNA knockdown of SRPK1 could be used with the minigene system. If inhibition of SRPK by SRPIN340 does cause a change in splice site selection this could reveal whether E1^E4 targets specific pools of SRPK1 in the cell.

SR proteins have also been shown to have roles in mRNA export and protein translation, none of which have been examined in this study and perhaps the inhibition of SRPK1 activity by E1^E4 may be affecting one or more of these other functions of SR proteins. Therefore the investigation of the effect of HPV1 E1^E4 on these processes is required to determine the biological effects of this inhibition. This study has focussed on the effect of SRPK1 on SR proteins involved in RNA metabolism, however evidence is accumulating that SRPKs have

effects on other cellular activities including chromatin reorganisation, regulation of cyclins and metabolic signalling (Giannakouros et al., 2011).

In comparison to the *in vitro* findings, the potency of HPV1 E1^E4 as an inhibitor of SRPK1 is reduced *in vivo*. This may be due to posttranslational modifications of E4 such as phosphorylation which do not occur *in vitro*, impeding the ability of E4 to interact with and inhibit SRPK1. In natural HPV1 infections SRPK1 sequestration to cytoplasmic E4 inclusions is restricted to the more differentiated cells (Bell et al., 2007), cells in which the late viral genes are expressed. It is well known that the HPV1 E4 protein is extensively modified in these cells by proteolytic cleavage (Breitburd et al., 1987); in these cells the processed forms may contribute the inhibitory function of E4 and these smaller forms may be lacking in the established cells lines in which the experiments were performed.

As the sequestration of SRPK1 to E4 inclusions in a natural HPV1 infection coincides with late HPV functions, the inhibition of SRPK1 function may aid in the switch from early to late gene expression, as the SR proteins ASF/SF2 and SRp20 have been implicated in this regulation (Jia et al., 2009, Somberg and Schwartz, 2010).

There are other human members of the SRPK family which are found in epithelial cells (SRPK1a and SRPK2), and have been shown to have specific functions and interactions. SRPK2 has been shown to phosphorylate acinus and control the transcription of cyclin A1 (Jang et al., 2008), while SRPK1a has been shown to specifically interact with scaffold attachment factor B (Nikolakaki et al., 2001). The interaction of E1^E4 with these proteins should be examined to determine whether the inhibition caused by HPV1 E1^E4 is specific to SRPK1 and also whether any of the other E1^E4 proteins can inhibit these kinases. This would allow for more insight into the biological functions of the SRPK-E1^E4 interaction.

7.3 Relocalisation of SRPK1

Like the HPV1 E1^E4 protein, overexpression of the E1^E4 protein of HPV5 induces a relocalisation of SRPK1 to E4 containing cytoplasmic aggregates. In addition HPV18 causes aggregation of SRPK1 in the upper layers of differentiating epithelial cells demonstrating that the modulation of SRPK1 localisation may be a conserved feature among HPV types, although not always E4 dependent. The modulation of SRPK localisation may inhibit SRPK1 functions or may provide the virus with a source of the kinase which can be preferentially used for viral replication, either by phosphorylating viral proteins or cellular targets. It is therefore interesting that the HPV1 E2 protein is phosphorylated within the hinge region by SRPK1, as are E2 proteins of the beta viruses (Lai et al., 1999, Sekhar et al., 2009) (see section below).

The SRPK1 aggregates caused by the expression of HPV5 E1^E4 were localised close to periphery of the nucleus, suggesting that a complex with a membrane protein may be formed. One potential membrane bound protein which is a known substrate of SRPK1 is the lamin B receptor (LBR), which is involved in the disassembly and reassembly of the nuclear envelope (Takano et al., 2002). Disturbances of nuclear integrity might be relevant to easing the escape of newly made virions from the cells of the upper layers of warts. Abrogation of the interaction between LBR and heterochromatin by the human polyomavirus JC virus leads to a loss in nuclear envelope integrity and a promotion of virus escape (Okada et al., 2005).

The overexpression of SRPK1 has been shown to cause the disassembly of nuclear speckles (Colwill et al., 1996, Gui et al., 1994a, H. Y. Wang et al., 1998); the sequestration of SRPK1 to cytoplasmic HPV1 E4 inclusions did not inhibit the nuclear translocation of SRPK1 or the disassembly of the speckles. Sequestration to nuclear E4 inclusions however did reduce the level of disassembly observed, suggesting that HPV1 E1^E4 may only be targeting specific

pools of SRPK1 to modulate its activity. Interestingly these nuclear inclusions are also associated with ND10 body associated proteins (Roberts).

7.4 Role of SRPK1 in the late virus life cycle

Splicing of the early transcripts has been shown to require ASF/SF2 and SRp20 (Jia et al., 2009, Somberg and Schwartz, 2010). Within organotypic raft cultures of HPV18 genome containing keratinocytes however, the level of SRPK1 is low and only increases in the upper levels concomitant with late HPV functions. Perhaps therefore SRPK1 does not act upon the SR proteins in these early phases of the HPV18 life cycle. The SR proteins can be phosphorylated by the CLK family of kinases (Colwill et al., 1996); therefore it is likely that these kinases control these SR protein functions. The distribution of these kinases in the rafts should be examined. Further evidence that SRPK1 is not an important regulator of early HPV events came from the inhibitor studies which showed that SRPK functions are not required for the efficient maintenance of HPV18 episomes.

When the role of SRPKs was examined in the late virus life cycle, the inhibition of SRPK1 activity changed the growth morphology of organotypic rafts containing HPV18 genomes. The rafts had an increase in the amount of koilocytotic cells that are indicative of virus production and suggests that SRPK1 inhibition might be associated with enhanced virion production. The increased appearance of koilocytes correlated with an increase of parakeratosis, indicating that the nuclei were not being degraded. Preservation of nuclear integrity is a significant phase within the HPV lifecycle promoting the vegetative phase. As already mentioned the SRPK1 substrate LBR is involved in nuclear envelope disassembly. The phosphorylation of the LBR by CDK1 has been shown to induce the disassembly of the nuclear envelope and prevents the premature reassembly of the envelope during normal cell

cycle (Tseng and Chen, 2011). SRPK1 may therefore control the disassembly of the nuclear envelope in the upper layers of the epithelium, after the cells have exited cell cycle.

Further study is required to extend these interesting findings using SRPIN340. Whether there is an increase in virion production can be verified by immunostaining the raft with antibodies to the capsid proteins. Virus can be isolated from the rafts and infectivity assays performed to compare virus production between untreated and SRPIN340 treated structures. The effect of the inhibitor on viral genome amplification could be examined by using fluorescence insitu hybridisation. It would also be important to verify SRPK1 inhibition in treated rafts by isolating the protein and testing activity in *in vitro* kinase assays. If SRPKs are modulating the switch from early to late gene expression then they present a viable therapeutic target as upregulated production of the capsid proteins and downregulation of the E6 immunomodulatory protein could cause immune system activation.

7.5 Substrates of SRPK1

The E2 protein of HPV5 and HPV8 contains multiple RS repeats in the hinge region, but the HPV1 E2 is also a substrate for SRPK1 *in vitro*. The RS repeats of the beta papillomaviruses have been shown to modulate their subcellular localisation (Lai et al., 1999, Sekhar et al., 2009), however the effect of RS repeat phosphorylation on E2 functions are unknown. The HPV5 E2 protein has been shown to recruit the splicing machinery to pre-mRNAs transactivated by E2 itself (Lai et al., 1999) and the phosphorylation of the E2 hinge by SRPK1 may act to regulate this function. The phosphorylation of HPV1 E2 by SRPK1 *in vitro* is novel and the regulation of this phosphorylation by HPV1 E1^E4 may act to stabilise the E2 protein and also may modulate some of the E2 proteins functions in the later stages of the virus life cycle. It has been reported that the HPV16 E1^E4 protein can stabilise and modulate the subcellular localisation of the E2 protein (Davy 2009). Additionally the E1^E4

proteins of HPV1 and HPV5 are substrates for SRPK1 *in vitro* (Bell et al., 2007); the phosphorylation of these proteins may act to stabilise them and/or change their functions.

APPENDIX

Appendix 1A pGex-2T-HPV5 E1^E4 against HPV5b complete genome

(Accession D90252.1)

Base changes shown in bold and underlined

```
HPV5 E1^E4   AGCTCCACGCCTCCAGGGTCGCCAGGAAGACAAGCAGACACAGACACCACCGCCAAGACC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
HPV5b  3328  AGCTCCACGCCTCCAGGGTCGCCAGGAAGACAAGCAGACACAGACACCACCGCCAAGACC  3387

HPV5 E1^E4   CCCACCACCTCCACAACCGCCGTTGACTCCACGTCCAGACAGCTCACCACATCAAAACAG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
HPV5b  3388  CCCACCACCTCCACAACCGCCGTTGACTCCACGTCCAGACAGCTCACCACATCAAAACAG  3447

HPV5 E1^E4   CCACAACAAACCGAAACCAGAGGAAGAAGGTACGGACGGAGGCCCTCCAGCAAGTCAAGG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
HPV5b  3448  CCACAACAAACCGAAACCAGAGGAAGAAGGTACGGACGGAGGCCCTCCAGCAAGTCAAGG  3507

HPV5 E1^E4   AGATCGCAAACGCAGCAAAGGCGATCAAGGTCCCGACACCGGTCCCGGTCTCGGTCCCGG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
HPV5b  3508  AGATCGCAAACGCAGCAAAGGCGATCAAGGTCCCGACACCGGTCCCGGTCTCGGTCCCGG  3567

HPV5 E1^E4   TCGCGCTCCAAGTCCCAAACCCACACCACT**C**GGTCCACCACCAGGTCCCGGTCCACGTCG
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
HPV5b  3568  TCGCGCTCCAAGTCCCAAACCCACACCACT**T**GGTCCACCACCAGGTCCCGGTCCACGTCG  3627

HPV5 E1^E4   GTCGGCAAGACTCGGGCCCTTACAAGCAGATCGCGATCCAGGGGAAGGTCCCCAAGTACC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
HPV5b  3628  GTCGGCAAGACTCGGGCCCTTACAAGCAGATCGCGATCCAGGGGAAGGTCCCCAAGTACC  3687

HPV5 E1^E4   TGCAGAAGGGGAGGTGGAAGGTCACCCAGGCGGCGATCAAGGTCACCCTCCACCTACTCC
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
HPV5b  3688  TGCAGAAGGGGAGGTGGAAGGTCACCCAGGCGGCGATCAAGGTCACCCTCCACCTACTCC  3747

HPV5 E1^E4   TCCTGCACCACACAACGGTCACAGCGGGCACGGGCC**A**AAAGTCCAACAACCAGAGGGGCC
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
HPV5b  3748  TCCTGCACCACACAACGGTCACAGCGGGCACGGGCC**G**AAAGTCCAACAACCAGAGGGGCC  3807

HPV5 E1^E4   CGAGGGTCGAGAGGGTCACGAGGAGGGAGCCGTGGGGGGAGATTGCGGCGACGAGGAAGG
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
HPV5b  3808  CGAGGGTCGAGAGGGTCACGAGGAGGGAGCCGTGGGGGGAGATTGCGGCGACGAGGAAGG  3867

HPV5 E1^E4   TCA**TCC**TCCTCCTCCTCCTCCCCC**T**CCCACAAACGGTCACGAGGGGGGTCTGCTAAGCTC
             |||   |||||||||||||||||| |||||||||||||||||||||||||||||||||||
HPV5b  3868  TCA---TCCTCCTCCTCCTCCCCC**G**CCCACAAACGGTCACGAGGGGGGTCTGCTAAGCTC  3924

HPV5 E1^E4   CGTGGCGTCTCTCCTGGTGAAGTGGGAGGGTCACTTCGATCAGTTAGTTCAAAGCATACA
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
HPV5b  3925  CGTGGCGTCTCTCCTGGTGAAGTGGGAGGGTCACTTCGATCAGTTAGTTCAAAGCATACA  3984

HPV5 E1^E4   GGA**A**GACTTGGAAGATTACTGGAAGAAGCTCGCGACCCCCCAGTAA
             ||| ||||||||||||||||||||||||||||||||||||||||||
HPV5b  3985  GGA**C**GACTTGGAAGATTACTGGAAGAAGCTCGCGACCCCCCAGTAA  4030
```

Appendix 1B GST-HPV5 E4 against HPV5b E4 protein sequence

Amino acid changes shown in bold and underlined, RS dipeptides shown in grey

```
GST-HPV5 E4   APRLQGRQEDKQTQTPPPRPPPPPQPPLTPRPDSSPHQNSHNKPKPEEEGTDGGPPASQG
              APRLQGRQEDKQTQTPPPRPPPPPQPPLTPRPDSSPHQNSHNKPKPEEEGTDGGPPASQG
HPV5b E4      APRLQGRQEDKQTQTPPPRPPPPPQPPLTPRPDSSPHQNSHNKPKPEEEGTDGGPPASQG

GST-HPV5 E4   DRKRSKGDQGPDTGPGLGPGRAPSPKPTPLGPPPGPGPRRSARLGPLQADRDPGEGPQVP
              DRKRSKGDQGPDTGPGLGPGRAPSPKPTPLGPPPGPGPRRSARLGPLQADRDPGEGPQVP
HPV5b E4      DRKRSKGDQGPDTGPGLGPGRAPSPKPTPLGPPPGPGPRRSARLGPLQADRDPGEGPQVP

GST-HPV5 E4   AEGEVEGHPGGDQGHPPPTPPAPHNGHSGHGPKVQQPEGPEGREGHEEGAVGGDCGDEEG
              AEGEVEGHPGGDQGHPPPTPPAPHNGHSGHGPKVQQPEGPEGREGHEEGAVGGDCGDEEG
HPV5b E4      AEGEVEGHPGGDQGHPPPTPPAPHNGHSGHGPKVQQPEGPEGREGHEEGAVGGDCGDEEG

GST-HPV5 E4   HPPPPPPPPTNGHEGGLLSSVASLLVKWEGHFDQLVQSIQEDLEDYWKKLATPQ
              H PPPPPPPTNGHEGGLLSSVASLLVKWEGHFDQLVQSIQ+DLEDYWKKLATPQ
HPV5b E4      H-PPPPPPPTNGHEGGLLSSVASLLVKWEGHFDQLVQSIQDDLEDYWKKLATPQ
```

LIST OF REFERENCES

Akgul, B., Bauer, B., Zigrino, P., et al. (2011) Upregulation of lipocalin-2 in human papillomavirus-positive keratinocytes and cutaneous squamous cell carcinomas. **J Gen Virol**, 92: (Pt 2): 395-401.

Akgul, B., Cooke, J.C. and Storey, A. (2006) HPV-associated skin disease. **J Pathol**, 208: (2): 165-175.

Alam, M. and Ratner, D. (2001) Cutaneous squamous-cell carcinoma. **N Engl J Med**, 344: (13): 975-983.

Antson, A.A., Burns, J.E., Moroz, O.V., et al. (2000) Structure of the intact transactivation domain of the human papillomavirus E2 protein. **Nature**, 403: (6771): 805-809.

Ashmole, I., Gallimore, P.H. and Roberts, S. (1998) Identification of conserved hydrophobic C-terminal residues of the human papillomavirus type 1 E1E4 protein necessary for E4 oligomerisation in vivo. **Virology**, 240: (2): 221-231.

Asselineau, D. and Prunieras, M. (1984) Reconstruction of 'simplified' skin: control of fabrication. **Br J Dermatol**, 111 Suppl 27: 219-222.

Aubol, B.E. and Adams, J.A. (2011) Applying the brakes to multisite SR protein phosphorylation: substrate-induced effects on the splicing kinase SRPK1. **Biochemistry**, 50: (32): 6888-6900.

Aubol, B.E., Chakrabarti, S., Ngo, J., et al. (2003) Processive phosphorylation of alternative splicing factor/splicing factor 2. **Proc Natl Acad Sci U S A**, 100: (22): 12601-12606.

Baker, C. (1997) "Post-transcriptional regulation of papillomavirus gene expression". **Human Papillomaviruses: A compilation and analysis of nucleic acid and amino acid sequences**.

Baker, C. and Calef, C. (1997) "Maps of papillomavirus mRNA transcripts". **Human papillomaviruses: A compilation and analysis of nucleic acid and amino acid sequences**.

Band, V., De Caprio, J.A., Delmolino, L., et al. (1991) Loss of p53 protein in human papillomavirus type 16 E6-immortalized human mammary epithelial cells. **J Virol**, 65: (12): 6671-6676.

Banerjee, N.S., Wang, H.K., Broker, T.R., et al. (2011) Human papillomavirus (HPV) E7 induces prolonged G2 following S phase reentry in differentiated human keratinocytes. **J Biol Chem**, 286: (17): 15473-15482.

Barbosa, M.S., Lowy, D.R. and Schiller, J.T. (1989) Papillomavirus polypeptides E6 and E7 are zinc-binding proteins. **J Virol**, 63: (3): 1404-1407.

Barksdale, S. and Baker, C.C. (1995) Differentiation-specific alternative splicing of bovine papillomavirus late mRNAs. **J Virol**, 69: (10): 6553-6556.

Bastien, N. and McBride, A.A. (2000) Interaction of the papillomavirus E2 protein with mitotic chromosomes. **Virology**, 270: (1): 124-134.

Bedard, K.M., Daijogo, S. and Semler, B.L. (2007) A nucleo-cytoplasmic SR protein functions in viral IRES-mediated translation initiation. **EMBO J**, 26: (2): 459-467.

Bekkers, R.L., Massuger, L.F., Bulten, J., et al. (2004) Epidemiological and clinical aspects of human papillomavirus detection in the prevention of cervical cancer. **Rev Med Virol**, 14: (2): 95-105.

Bell, I., Martin, A. and Roberts, S. (2007) The E1^E4 protein of human papillomavirus interacts with the serine-arginine-specific protein kinase SRPK1. **J Virol**, 81: (11): 5437-5448.

Bernard, H.U. (2002) Gene expression of genital human papillomaviruses and considerations on potential antiviral approaches. **Antivir Ther**, 7: (4): 219-237.

Bernard, H.U., Burk, R.D., Chen, Z., et al. (2010) Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. **Virology**, 401: (1): 70-79.

Blaustein, M., Pelisch, F., Tanos, T., et al. (2005) Concerted regulation of nuclear and cytoplasmic activities of SR proteins by AKT. **Nat Struct Mol Biol**, 12: (12): 1037-1044.

Bodily, J. and Laimins, L.A. (2011) Persistence of human papillomavirus infection: keys to malignant progression. **Trends Microbiol**, 19: (1): 33-39.

Bonanni, P., Boccalini, S. and Bechini, A. (2009) Efficacy, duration of immunity and cross protection after HPV vaccination: a review of the evidence. **Vaccine**, 27 Suppl 1: A46-53.

Boucher, L., Ouzounis, C.A., Enright, A.J., et al. (2001) A genome-wide survey of RS domain proteins. **RNA**, 7: (12): 1693-1701.

Boukis, L.A., Liu, N., Furuyama, S., et al. (2004) Ser/Arg-rich protein-mediated communication between U1 and U2 small nuclear ribonucleoprotein particles. **J Biol Chem**, 279: (28): 29647-29653.

Boyer, S.N., Wazer, D.E. and Band, V. (1996) E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. **Cancer Res**, 56: (20): 4620-4624.

Breitburd, F., Croissant, O. and Orth, G. (1987) "Expression of human papillomavirus type 1 E4 gene products in warts". In Steinberg, B.;Brandsma, J. & Taichman, L.B. (Eds.) **Cancer cells: Papillomaviruses**. 115 - 122.

Breitburd, F., Kirnbauer, R., Hubbert, N.L., et al. (1995) Immunization with viruslike particles from cottontail rabbit papillomavirus (CRPV) can protect against experimental CRPV infection. **J Virol**, 69: (6): 3959-3963.

Brown, R.E., Breugelmans, J.G., Theodoratou, D., et al. (2006) Costs of detection and treatment of cervical cancer, cervical dysplasia and genital warts in the UK. **Curr Med Res Opin**, 22: (4): 663-670.

Bruzik, J.P. and Maniatis, T. (1995) Enhancer-dependent interaction between 5' and 3' splice sites in trans. **Proc Natl Acad Sci U S A**, 92: (15): 7056-7059.

Bryan, J.T. and Brown, D.R. (2000) Association of the human papillomavirus type 11 E1^E4 protein with cornified cell envelopes derived from infected genital epithelium. **Virology**, 277: (2): 262-269.

Bryan, J.T., Han, A., Fife, K.H., et al. (2000) The human papillomavirus type 11 E1E4 protein is phosphorylated in genital epithelium. **Virology**, 268: (2): 430-439.

Bryant, H.E., Wadd, S.E., Lamond, A.I., et al. (2001) Herpes simplex virus IE63 (ICP27) protein interacts with spliceosome-associated protein 145 and inhibits splicing prior to the first catalytic step. **J Virol**, 75: (9): 4376-4385.

Buck, C.B., Cheng, N., Thompson, C.D., et al. (2008) Arrangement of L2 within the papillomavirus capsid. **J Virol**, 82: (11): 5190-5197.

Caceres, J.F., Misteli, T., Screaton, G.R., et al. (1997) Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. **J Cell Biol**, 138: (2): 225-238.

Caceres, J.F., Screaton, G.R. and Krainer, A.R. (1998) A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. **Genes Dev**, 12: (1): 55-66.

Caceres, J.F., Stamm, S., Helfman, D.M., et al. (1994) Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. **Science**, 265: (5179): 1706-1709.

Campo, M.S., Graham, S.V., Cortese, M.S., et al. (2010) HPV-16 E5 down-regulates expression of surface HLA class I and reduces recognition by CD8 T cells. **Virology**, 407: (1): 137-142.

Carrillo-Infante, C., Abbadessa, G., Bagella, L., et al. (2007) Viral infections as a cause of cancer (review). **Int J Oncol**, 30: (6): 1521-1528.

Cavaloc, Y., Popielarz, M., Fuchs, J.P., et al. (1994) Characterization and cloning of the human splicing factor 9G8: a novel 35 kDa factor of the serine/arginine protein family. **EMBO J**, 13: (11): 2639-2649.

Cazalla, D., Zhu, J., Manche, L., et al. (2002) Nuclear export and retention signals in the RS domain of SR proteins. **Mol Cell Biol**, 22: (19): 6871-6882.

Chellappan, S., Kraus, V.B., Kroger, B., et al. (1992) Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. **Proc Natl Acad Sci U S A**, 89: (10): 4549-4553.

Chen, G. and Stenlund, A. (2000) Two patches of amino acids on the E2 DNA binding domain define the surface for interaction with E1. **J Virol**, 74: (3): 1506-1512.

Cheng, S., Schmidt-Grimminger, D.C., Murant, T., et al. (1995) Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes. **Genes Dev**, 9: (19): 2335-2349.

Chow, L.T., Reilly, S.S., Broker, T.R., et al. (1987) Identification and mapping of human papillomavirus type 1 RNA transcripts recovered from plantar warts and infected epithelial cell cultures. **J Virol**, 61: (6): 1913-1918.

Collins, S., Rollason, T.P., Young, L.S., et al. (2010) Cigarette smoking is an independent risk factor for cervical intraepithelial neoplasia in young women: a longitudinal study. **Eur J Cancer**, 46: (2): 405-411.

Colwill, K., Pawson, T., Andrews, B., et al. (1996) The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. **EMBO J**, 15: (2): 265-275.

Conrad, M., Bubb, V.J. and Schlegel, R. (1993) The human papillomavirus type 6 and 16 E5 proteins are membrane-associated proteins which associate with the 16-kilodalton pore-forming protein. **J Virol**, 67: (10): 6170-6178.

Cooper, T.A. (2005) Use of minigene systems to dissect alternative splicing elements. **Methods**, 37: (4): 331-340.

Crusius, K., Auvinen, E. and Alonso, A. (1997) Enhancement of EGF- and PMA-mediated MAP kinase activation in cells expressing the human papillomavirus type 16 E5 protein. **Oncogene**, 15: (12): 1437-1444.

Cumming, S.A., McPhillips, M.G., Veerapraditsin, T., et al. (2003) Activity of the human papillomavirus type 16 late negative regulatory element is partly due to four weak consensus 5' splice sites that bind a U1 snRNP-like complex. **J Virol**, 77: (9): 5167-5177.

Daub, H., Blencke, S., Habenberger, P., et al. (2002) Identification of SRPK1 and SRPK2 as the major cellular protein kinases phosphorylating hepatitis B virus core protein. **J Virol**, 76: (16): 8124-8137.

Davy, C., McIntosh, P., Jackson, D.J., et al. (2009) A novel interaction between the human papillomavirus type 16 E2 and E1^E4 proteins leads to stabilization of E2. **Virology**, 394: (2): 266-275.

Davy, C.E., Jackson, D.J., Wang, Q., et al. (2002) Identification of a G(2) arrest domain in the E1 wedge E4 protein of human papillomavirus type 16. **J Virol**, 76: (19): 9806-9818.

de Sanjose, S., Quint, W.G., Alemany, L., et al. (2010) Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. **Lancet Oncol**, 11: (11): 1048-1056.

de Villiers, E.M., Fauquet, C., Broker, T.R., et al. (2004) Classification of papillomaviruses. **Virology**, 324: (1): 17-27.

Demeret, C., Desaintes, C., Yaniv, M., et al. (1997) Different mechanisms contribute to the E2-mediated transcriptional repression of human papillomavirus type 18 viral oncogenes. **J Virol**, 71: (12): 9343-9349.

Deng, W., Jin, G., Lin, B.Y., et al. (2003) mRNA splicing regulates human papillomavirus type 11 E1 protein production and DNA replication. **J Virol**, 77: (19): 10213-10226.

Deng, W., Lin, B.Y., Jin, G., et al. (2004) Cyclin/CDK regulates the nucleocytoplasmic localization of the human papillomavirus E1 DNA helicase. **J Virol**, 78: (24): 13954-13965.

Ding, J., Doorbar, J., Li, B., et al. (2010) Expression of papillomavirus L1 proteins regulated by authentic gene codon usage is favoured in G2/M-like cells in differentiating keratinocytes. **Virology**, 399: (1): 46-58.

Ding, J.H., Zhong, X.Y., Hagopian, J.C., et al. (2006) Regulated cellular partitioning of SR protein-specific kinases in mammalian cells. **Mol Biol Cell**, 17: (2): 876-885.

DoHealth (2011) "HPV vaccination programme: Provisional data as submitted by PCTs, for first, second and third dose uptake, for month ending 30 June 2011, Available at www.dh.gov.uk.

Doorbar, J. (2006) Molecular biology of human papillomavirus infection and cervical cancer. **Clin Sci (Lond)**, 110: (5): 525-541.

Doorbar, J., Campbell, D., Grand, R.J., et al. (1986) Identification of the human papilloma virus-1a E4 gene products. **EMBO J**, 5: (2): 355-362.

Doorbar, J., Elston, R.C., Napthine, S., et al. (2000) The E1^E4 protein of human papillomavirus type 16 associates with a putative RNA helicase through sequences in its C terminus. **J Virol**, 74: (21): 10081-10095.

Doorbar, J., Ely, S., Sterling, J., et al. (1991) Specific interaction between HPV-16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. **Nature**, 352: (6338): 824-827.

Doorbar, J., Evans, H.S., Coneron, I., et al. (1988) Analysis of HPV-1 E4 gene expression using epitope-defined antibodies. **EMBO J**, 7: (3): 825-833.

Doorbar, J., Foo, C., Coleman, N., et al. (1997) Characterization of events during the late stages of HPV16 infection in vivo using high-affinity synthetic Fabs to E4. **Virology**, 238: (1): 40-52.

Doorbar, J. and Gallimore, P.H. (1987) Identification of proteins encoded by the L1 and L2 open reading frames of human papillomavirus 1a. **J Virol**, 61: (9): 2793-2799.

Dyson, N., Howley, P.M., Munger, K., et al. (1989) The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. **Science**, 243: (4893): 934-937.

Estmer Nilsson, C., Petersen-Mahrt, S., Durot, C., et al. (2001) The adenovirus E4-ORF4 splicing enhancer protein interacts with a subset of phosphorylated SR proteins. **EMBO J**, 20: (4): 864-871.

Fang, L., Budgeon, L.R., Doorbar, J., et al. (2006) The human papillomavirus type 11 E1/\E4 protein is not essential for viral genome amplification. **Virology**, 351: (2): 271-279.

Fay, J., Kelehan, P., Lambkin, H., et al. (2009) Increased expression of cellular RNA-binding proteins in HPV-induced neoplasia and cervical cancer. **J Med Virol**, 81: (5): 897-907.

Fehrmann, F., Klumpp, D.J. and Laimins, L.A. (2003) Human papillomavirus type 31 E5 protein supports cell cycle progression and activates late viral functions upon epithelial differentiation. **J Virol**, 77: (5): 2819-2831.

Finnen, R.L., Erickson, K.D., Chen, X.S., et al. (2003) Interactions between papillomavirus L1 and L2 capsid proteins. **J Virol**, 77: (8): 4818-4826.

Florin, L., Sapp, C., Streeck, R.E., et al. (2002) Assembly and translocation of papillomavirus capsid proteins. **J Virol**, 76: (19): 10009-10014.

Forouzanfar, M.H., Foreman, K.J., Delossantos, A.M., et al. (2011) Breast and cervical cancer in 187 countries between 1980 and 2010: a systematic analysis. **Lancet**.

Fradet-Turcotte, A., Moody, C., Laimins, L.A., et al. (2010) Nuclear export of human papillomavirus type 31 E1 is regulated by Cdk2 phosphorylation and required for viral genome maintenance. **J Virol**, 84: (22): 11747-11760.

Francis, D.A., Schmid, S.I. and Howley, P.M. (2000) Repression of the integrated papillomavirus E6/E7 promoter is required for growth suppression of cervical cancer cells. **J Virol**, 74: (6): 2679-2686.

Frattini, M.G. and Laimins, L.A. (1994) Binding of the human papillomavirus E1 origin-recognition protein is regulated through complex formation with the E2 enhancer-binding protein. **Proc Natl Acad Sci U S A**, 91: (26): 12398-12402.

Frattini, M.G., Lim, H.B. and Laimins, L.A. (1996) In vitro synthesis of oncogenic human papillomaviruses requires episomal genomes for differentiation-dependent late expression. **Proc Natl Acad Sci U S A**, 93: (7): 3062-3067.

Frazer, I.H. (2004) Prevention of cervical cancer through papillomavirus vaccination. **Nat Rev Immunol**, 4: (1): 46-54.

Fu, X.D. and Maniatis, T. (1990) Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. **Nature**, 343: (6257): 437-441.

Fukuhara, T., Hosoya, T., Shimizu, S., et al. (2006) Utilization of host SR protein kinases and RNA-splicing machinery during viral replication. **Proc Natl Acad Sci U S A**, 103: (30): 11329-11333.

Furth, P.A., Choe, W.T., Rex, J.H., et al. (1994) Sequences homologous to 5' splice sites are required for the inhibitory activity of papillomavirus late 3' untranslated regions. **Mol Cell Biol**, 14: (8): 5278-5289.

Gammoh, N., Grm, H.S., Massimi, P., et al. (2006) Regulation of human papillomavirus type 16 E7 activity through direct protein interaction with the E2 transcriptional activator. **J Virol**, 80: (4): 1787-1797.

Gammoh, N., Isaacson, E., Tomaic, V., et al. (2009) Inhibition of HPV-16 E7 oncogenic activity by HPV-16 E2. **Oncogene**, 28: (23): 2299-2304.

Garrido-Guerrero, E., Carrillo, E., Guido, M., et al. (1996) Different arrangement of human papillomavirus E2 binding sites distinguishes cutaneous types from those associated with mucosal lesions. **Arch Med Res**, 27: (3): 389-394.

Gauthier, J.M., Dillner, J. and Yaniv, M. (1991) Structural analysis of the human papillomavirus type 16-E2 transactivator with antipeptide antibodies reveals a high mobility region linking the transactivation and the DNA-binding domains. **Nucleic Acids Res**, 19: (25): 7073-7079.

Ge, H. and Manley, J.L. (1990) A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. **Cell**, 62: (1): 25-34.

Genther, S.M., Sterling, S., Duensing, S., et al. (2003) Quantitative role of the human papillomavirus type 16 E5 gene during the productive stage of the viral life cycle. **J Virol**, 77: (5): 2832-2842.

Ghosh, G. and Adams, J.A. (2011) Phosphorylation mechanism and structure of serine-arginine protein kinases. **FEBS J**, 278: (4): 587-597.

Giannakouros, T., Nikolakaki, E., Mylonis, I., et al. (2011) Serine-arginine protein kinases: a small protein kinase family with a large cellular presence. **FEBS J**, 278: (4): 570-586.

Grand, R.J., Doorbar, J., Smith, K.J., et al. (1989) Phosphorylation of the human papillomavirus type 1 E4 proteins in vivo and in vitro. **Virology**, 170: (1): 201-213.

Grassmann, K., Rapp, B., Maschek, H., et al. (1996) Identification of a differentiation-inducible promoter in the E7 open reading frame of human papillomavirus type 16 (HPV-16) in raft cultures of a new cell line containing high copy numbers of episomal HPV-16 DNA. **J Virol**, 70: (4): 2339-2349.

Graveley, B.R. (2000) Sorting out the complexity of SR protein functions. **RNA**, 6: (9): 1197-1211.

Green, J., Berrington de Gonzalez, A., Sweetland, S., et al. (2003) Risk factors for adenocarcinoma and squamous cell carcinoma of the cervix in women aged 20-44 years: the UK National Case-Control Study of Cervical Cancer. **Br J Cancer**, 89: (11): 2078-2086.

Gui, J.F., Lane, W.S. and Fu, X.D. (1994a) A serine kinase regulates intracellular localization of splicing factors in the cell cycle. **Nature**, 369: (6482): 678-682.

Gui, J.F., Tronchere, H., Chandler, S.D., et al. (1994b) Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. **Proc Natl Acad Sci U S A**, 91: (23): 10824-10828.

Hagopian, J.C., Ma, C.T., Meade, B.R., et al. (2008) Adaptable molecular interactions guide phosphorylation of the SR protein ASF/SF2 by SRPK1. **J Mol Biol**, 382: (4): 894-909.

Harper, D.M., Franco, E.L., Wheeler, C., et al. (2004) Efficacy of a bivalent L1 virus-like particle vaccine in prevention of infection with human papillomavirus types 16 and 18 in young women: a randomised controlled trial. **Lancet**, 364: (9447): 1757-1765.

Harwood, C.A. and Proby, C.M. (2002) Human papillomaviruses and non-melanoma skin cancer. **Curr Opin Infect Dis**, 15: (2): 101-114.

Hautbergue, G.M., Hung, M.L., Golovanov, A.P., et al. (2008) Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP. **Proc Natl Acad Sci U S A**, 105: (13): 5154-5159.

Hawley-Nelson, P., Vousden, K.H., Hubbert, N.L., et al. (1989) HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. **EMBO J**, 8: (12): 3905-3910.

Hayes, G.M., Carrigan, P.E. and Miller, L.J. (2007) Serine-arginine protein kinase 1 overexpression is associated with tumorigenic imbalance in mitogen-activated protein kinase pathways in breast, colonic, and pancreatic carcinomas. **Cancer Res**, 67: (5): 2072-2080.

Hegde, R.S., Grossman, S.R., Laimins, L.A., et al. (1992) Crystal structure at 1.7 A of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. **Nature**, 359: (6395): 505-512.

Hertel, K.J. and Maniatis, T. (1999) Serine-arginine (SR)-rich splicing factors have an exon-independent function in pre-mRNA splicing. **Proc Natl Acad Sci U S A**, 96: (6): 2651-2655.

Hickman, A.B. and Dyda, F. (2005) Binding and unwinding: SF3 viral helicases. **Curr Opin Struct Biol**, 15: (1): 77-85.

Hirochika, H., Broker, T.R. and Chow, L.T. (1987) Enhancers and trans-acting E2 transcriptional factors of papillomaviruses. **J Virol**, 61: (8): 2599-2606.

Howie, H.L., Katzenellenbogen, R.A. and Galloway, D.A. (2009) Papillomavirus E6 proteins. **Virology**, 384: (2): 324-334.

Huang, Y. and Steitz, J.A. (2001) Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. **Mol Cell**, 7: (4): 899-905.

Huang, Y., Yario, T.A. and Steitz, J.A. (2004) A molecular link between SR protein dephosphorylation and mRNA export. **Proc Natl Acad Sci U S A**, 101: (26): 9666-9670.

Hughes, F.J. and Romanos, M.A. (1993) E1 protein of human papillomavirus is a DNA helicase/ATPase. **Nucleic Acids Res**, 21: (25): 5817-5823.

Huh, K., Zhou, X., Hayakawa, H., et al. (2007) Human papillomavirus type 16 E7 oncoprotein associates with the cullin 2 ubiquitin ligase complex, which contributes to degradation of the retinoblastoma tumor suppressor. **J Virol**, 81: (18): 9737-9747.

Huibregtse, J.M., Scheffner, M. and Howley, P.M. (1991) A cellular protein mediates association of p53 with the E6 oncoprotein of human papillomavirus types 16 or 18. **EMBO J**, 10: (13): 4129-4135.

Huynh, N., Ma, C.T., Giang, N., et al. (2009) Allosteric interactions direct binding and phosphorylation of ASF/SF2 by SRPK1. **Biochemistry**, 48: (48): 11432-11440.

Jackson, S., Harwood, C., Thomas, M., et al. (2000) Role of Bak in UV-induced apoptosis in skin cancer and abrogation by HPV E6 proteins. **Genes Dev**, 14: (23): 3065-3073.

Jang, S.W., Yang, S.J., Ehlen, A., et al. (2008) Serine/arginine protein-specific kinase 2 promotes leukemia cell proliferation by phosphorylating acinus and regulating cyclin A1. **Cancer Res**, 68: (12): 4559-4570.

Jeon, S., Allen-Hoffmann, B.L. and Lambert, P.F. (1995) Integration of human papillomavirus type 16 into the human genome correlates with a selective growth advantage of cells. **J Virol**, 69: (5): 2989-2997.

Jia, R., Liu, X., Tao, M., et al. (2009) Control of the papillomavirus early-to-late switch by differentially expressed SRp20. **J Virol**, 83: (1): 167-180.

Joyce, J.G., Tung, J.S., Przysiecki, C.T., et al. (1999) The L1 major capsid protein of human papillomavirus type 11 recombinant virus-like particles interacts with heparin and cell-surface glycosaminoglycans on human keratinocytes. **J Biol Chem**, 274: (9): 5810-5822.

Jumaa, H. and Nielsen, P.J. (1997) The splicing factor SRp20 modifies splicing of its own mRNA and ASF/SF2 antagonizes this regulation. **EMBO J**, 16: (16): 5077-5085.

Jumaa, H. and Nielsen, P.J. (2000) Regulation of SRp20 exon 4 splicing. **Biochim Biophys Acta**, 1494: (1-2): 137-143.

Kamachi, M., Le, T.M., Kim, S.J., et al. (2002) Human autoimmune sera as molecular probes for the identification of an autoantigen kinase signaling pathway. **J Exp Med**, 196: (9): 1213-1225.

Karakama, Y., Sakamoto, N., Itsui, Y., et al. (2010) Inhibition of hepatitis C virus replication by a specific inhibitor of serine-arginine-rich protein kinase. **Antimicrob Agents Chemother**, 54: (8): 3179-3186.

Karni, R., de Stanchina, E., Lowe, S.W., et al. (2007) The gene encoding the splicing factor SF2/ASF is a proto-oncogene. **Nat Struct Mol Biol**, 14: (3): 185-193.

Kataoka, N., Bachorik, J.L. and Dreyfuss, G. (1999) Transportin-SR, a nuclear import receptor for SR proteins. **J Cell Biol**, 145: (6): 1145-1152.

Khan, J., Davy, C.E., McIntosh, P.B., et al. (2011) Role of calpain in the formation of human papillomavirus type 16 e1^e4 amyloid fibers and reorganization of the keratin network. **J Virol**, 85: (19): 9984-9997.

Kim, M.K., Kim, H.S., Kim, S.H., et al. (2010) Human papillomavirus type 16 E5 oncoprotein as a new target for cervical cancer treatment. **Biochem Pharmacol**, 80: (12): 1930-1935.

Kines, R.C., Thompson, C.D., Lowy, D.R., et al. (2009) The initial steps leading to papillomavirus infection occur on the basement membrane prior to cell surface binding. **Proc Natl Acad Sci U S A**, 106: (48): 20458-20463.

Kiviat, N.B. (1999) Papillomaviruses in non-melanoma skin cancer: epidemiological aspects. **Semin Cancer Biol**, 9: (6): 397-403.

Kjellberg, L., Hallmans, G., Ahren, A.M., et al. (2000) Smoking, diet, pregnancy and oral contraceptive use as risk factors for cervical intra-epithelial neoplasia in relation to human papillomavirus infection. **Br J Cancer**, 82: (7): 1332-1338.

Klumpp, D.J. and Laimins, L.A. (1999) Differentiation-induced changes in promoter usage for transcripts encoding the human papillomavirus type 31 replication protein E1. **Virology**, 257: (1): 239-246.

Knapp, A.A., McManus, P.M., Bockstall, K., et al. (2009) Identification of the nuclear localization and export signals of high risk HPV16 E7 oncoprotein. **Virology**, 383: (1): 60-68.

Knight, G.L., Grainger, J.R., Gallimore, P.H., et al. (2004) Cooperation between different forms of the human papillomavirus type 1 E4 protein to block cell cycle progression and cellular DNA synthesis. **J Virol**, 78: (24): 13920-13933.

Knight, G.L., Pugh, A.G., Yates, E., et al. (2011) A cyclin-binding motif in human papillomavirus type 18 (HPV18) E1^E4 is necessary for association with CDK-cyclin complexes and G2/M cell cycle arrest of keratinocytes, but is not required for differentiation-dependent viral genome amplification or L1 capsid protein expression. **Virology**, 412: (1): 196-210.

Knight, G.L., Turnell, A.S. and Roberts, S. (2006) Role for Wee1 in inhibition of G2-to-M transition through the cooperation of distinct human papillomavirus type 1 E4 proteins. **J Virol**, 80: (15): 7416-7426.

Koizumi, J., Okamoto, Y., Onogi, H., et al. (1999) The subcellular localization of SF2/ASF is regulated by direct interaction with SR protein kinases (SRPKs). **J Biol Chem**, 274: (16): 11125-11131.

Kojima, T., Zama, T., Wada, K., et al. (2001) Cloning of human PRP4 reveals interaction with Clk1. **J Biol Chem**, 276: (34): 32247-32256.

Koulova, A., Tsui, J., Irwin, K., et al. (2008) Country recommendations on the inclusion of HPV vaccines in national immunization programmes among high-income countries, June 2006-January 2008. **Vaccine**, 26: (51): 6529-6541.

Krainer, A.R., Conway, G.C. and Kozak, D. (1990) Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. **Genes Dev**, 4: (7): 1158-1171.

Krainer, A.R. and Maniatis, T. (1985) Multiple factors including the small nuclear ribonucleoproteins U1 and U2 are necessary for pre-mRNA splicing in vitro. **Cell**, 42: (3): 725-736.

Krawczyk, E., Suprynowicz, F.A., Liu, X., et al. (2008) Koilocytosis: a cooperative interaction between the human papillomavirus E5 and E6 oncoproteins. **Am J Pathol**, 173: (3): 682-688.

Kreimer, A.R., Clifford, G.M., Boyle, P., et al. (2005) Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. **Cancer Epidemiol Biomarkers Prev**, 14: (2): 467-475.

Kukimoto, I. and Kanda, T. (2001) Displacement of YY1 by differentiation-specific transcription factor hSkn-1a activates the P(670) promoter of human papillomavirus type 16. **J Virol**, 75: (19): 9302-9311.

Kuroyanagi, N., Onogi, H., Wakabayashi, T., et al. (1998) Novel SR-protein-specific kinase, SRPK2, disassembles nuclear speckles. **Biochem Biophys Res Commun**, 242: (2): 357-364.

Labourier, E., Rossi, F., Gallouzi, I.E., et al. (1998) Interaction between the N-terminal domain of human DNA topoisomerase I and the arginine-serine domain of its substrate determines phosphorylation of SF2/ASF splicing factor. **Nucleic Acids Res**, 26: (12): 2955-2962.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, 227: (5259): 680-685.

Lai, M.C., Lin, R.I., Huang, S.Y., et al. (2000) A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. **J Biol Chem**, 275: (11): 7950-7957.

Lai, M.C. and Tarn, W.Y. (2004) Hypophosphorylated ASF/SF2 binds TAP and is present in messenger ribonucleoproteins. **J Biol Chem**, 279: (30): 31745-31749.

Lai, M.C., Teh, B.H. and Tarn, W.Y. (1999) A human papillomavirus E2 transcriptional activator. The interactions with cellular splicing factors and potential function in pre-mRNA processing. **J Biol Chem**, 274: (17): 11832-11841.

Lazarczyk, M., Cassonnet, P., Pons, C., et al. (2009) The EVER proteins as a natural barrier against papillomaviruses: a new insight into the pathogenesis of human papillomavirus infections. **Microbiol Mol Biol Rev**, 73: (2): 348-370.

Lechner, M.S. and Laimins, L.A. (1994) Inhibition of p53 DNA binding by human papillomavirus E6 proteins. **J Virol**, 68: (7): 4262-4273.

Lehman, C.W. and Botchan, M.R. (1998) Segregation of viral plasmids depends on tethering to chromosomes and is regulated by phosphorylation. **Proc Natl Acad Sci U S A**, 95: (8): 4338-4343.

Lemaire, R., Prasad, J., Kashima, T., et al. (2002) Stability of a PKCI-1-related mRNA is controlled by the splicing factor ASF/SF2: a novel function for SR proteins. **Genes Dev**, 16: (5): 594-607.

Lewellyn, E.B. and Loeb, D.D. (2011) Serine phosphoacceptor sites within the core protein of hepatitis B virus contribute to genome replication pleiotropically. **PLoS One**, 6: (2): e17202.

Liao, J.B. (2006) Viruses and human cancer. **Yale J Biol Med**, 79: (3-4): 115-122.

Lin, S., Coutinho-Mansfield, G., Wang, D., et al. (2008) The splicing factor SC35 has an active role in transcriptional elongation. **Nat Struct Mol Biol**, 15: (8): 819-826.

Longworth, M.S. and Laimins, L.A. (2004a) The binding of histone deacetylases and the integrity of zinc finger-like motifs of the E7 protein are essential for the life cycle of human papillomavirus type 31. **J Virol**, 78: (7): 3533-3541.

Longworth, M.S. and Laimins, L.A. (2004b) Pathogenesis of human papillomaviruses in differentiating epithelia. **Microbiol Mol Biol Rev**, 68: (2): 362-372.

Loomis, R.J., Naoe, Y., Parker, J.B., et al. (2009) Chromatin binding of SRp20 and ASF/SF2 and dissociation from mitotic chromosomes is modulated by histone H3 serine 10 phosphorylation. **Mol Cell**, 33: (4): 450-461.

Lukasiewicz, R., Velazquez-Dones, A., Huynh, N., et al. (2007) Structurally unique yeast and mammalian serine-arginine protein kinases catalyze evolutionarily conserved phosphorylation reactions. **J Biol Chem**, 282: (32): 23036-23043.

Ma, C.T., Ghosh, G., Fu, X.D., et al. (2010) Mechanism of dephosphorylation of the SR protein ASF/SF2 by protein phosphatase 1. **J Mol Biol**, 403: (3): 386-404.

Ma, C.T., Velazquez-Dones, A., Hagopian, J.C., et al. (2008) Ordered multi-site phosphorylation of the splicing factor ASF/SF2 by SRPK1. **J Mol Biol**, 376: (1): 55-68.

Manley, J.L. and Krainer, A.R. (2010) A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). **Genes Dev**, 24: (11): 1073-1074.

Marur, S., D'Souza, G., Westra, W.H., et al. (2010) HPV-associated head and neck cancer: a virus-related cancer epidemic. **Lancet Oncol**, 11: (8): 781-789.

Marusic, M.B., Mencin, N., Licen, M., et al. (2010) Modification of human papillomavirus minor capsid protein L2 by sumoylation. **J Virol**, 84: (21): 11585-11589.

Masterson, P.J., Stanley, M.A., Lewis, A.P., et al. (1998) A C-terminal helicase domain of the human papillomavirus E1 protein binds E2 and the DNA polymerase alpha-primase p68 subunit. **J Virol**, 72: (9): 7407-7419.

Mathew, R., Hartmuth, K., Mohlmann, S., et al. (2008) Phosphorylation of human PRP28 by SRPK2 is required for integration of the U4/U6-U5 tri-snRNP into the spliceosome. **Nat Struct Mol Biol**, 15: (5): 435-443.

Mathews, C.K., van Holde, K.E. and Ahern, K.G. (2000) **Biochemistry**. Third Edition. San Francisco: Benjamin/Cummings.

McFarlane, M. and Graham, S.V. (2010) Human papillomavirus regulation of SR proteins. **Biochem Soc Trans**, 38: (4): 1116-1121.

McIntosh, P.B., Laskey, P., Sullivan, K., et al. (2010) E1^E4-mediated keratin phosphorylation and ubiquitylation: a mechanism for keratin depletion in HPV16-infected epithelium. **J Cell Sci**, 123: (Pt 16): 2810-2822.

McIntosh, P.B., Martin, S.R., Jackson, D.J., et al. (2008) Structural analysis reveals an amyloid form of the human papillomavirus type 16 E1^E4 protein and provides a molecular basis for its accumulation. **J Virol**, 82: (16): 8196-8203.

McPhillips, M.G., Veerapraditsin, T., Cumming, S.A., et al. (2004) SF2/ASF binds the human papillomavirus type 16 late RNA control element and is regulated during differentiation of virus-infected epithelial cells. **J Virol**, 78: (19): 10598-10605.

Michlewski, G., Sanford, J.R. and Caceres, J.F. (2008) The splicing factor SF2/ASF regulates translation initiation by enhancing phosphorylation of 4E-BP1. **Mol Cell**, 30: (2): 179-189.

Milligan, S.G., Veerapraditsin, T., Ahamet, B., et al. (2007) Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. **Virology**, 360: (1): 172-181.

Misteli, T., Caceres, J.F., Clement, J.Q., et al. (1998) Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription in vivo. **J Cell Biol**, 143: (2): 297-307.

Misteli, T. and Spector, D.L. (1999) RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. **Mol Cell**, 3: (6): 697-705.

Modis, Y., Trus, B.L. and Harrison, S.C. (2002) Atomic model of the papillomavirus capsid. **EMBO J**, 21: (18): 4754-4762.

Mole, S., McFarlane, M., Chuen-Im, T., et al. (2009a) RNA splicing factors regulated by HPV16 during cervical tumour progression. **J Pathol**.

Mole, S., Milligan, S.G. and Graham, S.V. (2009b) Human papillomavirus type 16 E2 protein transcriptionally activates the promoter of a key cellular splicing factor, SF2/ASF. **J Virol**, 83: (1): 357-367.

Moody, C.A. and Laimins, L.A. (2009) Human papillomaviruses activate the ATM DNA damage pathway for viral genome amplification upon differentiation. **PLoS Pathog**, 5: (10): e1000605.

Moody, C.A. and Laimins, L.A. (2010) Human papillomavirus oncoproteins: pathways to transformation. **Nat Rev Cancer**, 10: (8): 550-560.

Muench, P., Probst, S., Schuetz, J., et al. (2010) Cutaneous papillomavirus E6 proteins must interact with p300 and block p53-mediated apoptosis for cellular immortalization and tumorigenesis. **Cancer Res**, 70: (17): 6913-6924.

Munger, K., Werness, B.A., Dyson, N., et al. (1989) Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. **EMBO J**, 8: (13): 4099-4105.

Nakahara, T., Nishimura, A., Tanaka, M., et al. (2002) Modulation of the cell division cycle by human papillomavirus type 18 E4. **J Virol**, 76: (21): 10914-10920.

Nakahara, T., Peh, W.L., Doorbar, J., et al. (2005) Human papillomavirus type 16 E1^E4 contributes to multiple facets of the papillomavirus life cycle. **J Virol**, 79: (20): 13150-13165.

Ngo, J.C., Chakrabarti, S., Ding, J.H., et al. (2005) Interplay between SRPK and Clk/Sty kinases in phosphorylation of the splicing factor ASF/SF2 is regulated by a docking motif in ASF/SF2. **Mol Cell**, 20: (1): 77-89.

Ngo, J.C., Giang, K., Chakrabarti, S., et al. (2008) A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1. **Mol Cell**, 29: (5): 563-576.

Nguyen, C.L., Eichwald, C., Nibert, M.L., et al. (2007) Human papillomavirus type 16 E7 oncoprotein associates with the centrosomal component gamma-tubulin. **J Virol**, 81: (24): 13533-13543.

Nikolakaki, E., Kohen, R., Hartmann, A.M., et al. (2001) Cloning and characterization of an alternatively spliced form of SR protein kinase 1 that interacts specifically with scaffold attachment factor-B. **J Biol Chem**, 276: (43): 40175-40182.

Nomine, Y., Masson, M., Charbonnier, S., et al. (2006) Structural and functional analysis of E6 oncoprotein: insights in the molecular pathways of human papillomavirus-mediated pathogenesis. **Mol Cell**, 21: (5): 665-678.

Nordin, H., Jungnelius, M., Karlsson, R., et al. (2005) Kinetic studies of small molecule interactions with protein kinases using biosensor technology. **Anal Biochem**, 340: (2): 359-368.

Nowak, D.G., Amin, E.M., Rennel, E.S., et al. (2010) Regulation of vascular endothelial growth factor (VEGF) splicing from pro-angiogenic to anti-angiogenic isoforms: a novel therapeutic strategy for angiogenesis. **J Biol Chem**, 285: (8): 5532-5540.

Nowak, D.G., Woolard, J., Amin, E.M., et al. (2008) Expression of pro- and anti-angiogenic isoforms of VEGF is differentially regulated by splicing and growth factors. **J Cell Sci**, 121: (Pt 20): 3487-3495.

O'Keefe, R.T., Mayeda, A., Sadowski, C.L., et al. (1994) Disruption of pre-mRNA splicing in vivo results in reorganization of splicing factors. **J Cell Biol**, 124: (3): 249-260.

Okada, Y., Suzuki, T., Sunden, Y., et al. (2005) Dissociation of heterochromatin protein 1 from lamin B receptor induced by human polyomavirus agnoprotein: role in nuclear egress of viral particles. **EMBO Rep**, 6: (5): 452-457.

Okamoto, Y., Onogi, H., Honda, R., et al. (1998) cdc2 kinase-mediated phosphorylation of splicing factor SF2/ASF. **Biochem Biophys Res Commun**, 249: (3): 872-878.

Oliveira, J.G., Colf, L.A. and McBride, A.A. (2006) Variations in the association of papillomavirus E2 proteins with mitotic chromosomes. **Proc Natl Acad Sci U S A**, 103: (4): 1047-1052.

Ote, I., Lebrun, M., Vandevenne, P., et al. (2009) Varicella-zoster virus IE4 protein interacts with SR proteins and exports mRNAs through the TAP/NXF1 pathway. **PLoS One**, 4: (11): e7882.

Pagano, J.S., Blaser, M., Buendia, M.A., et al. (2004) Infectious agents and cancer: criteria for a causal relation. **Semin Cancer Biol**, 14: (6): 453-471.

Palermo-Dilts, D.A., Broker, T.R. and Chow, L.T. (1990) Human papillomavirus type 1 produces redundant as well as polycistronic mRNAs in plantar warts. **J Virol**, 64: (6): 3144-3149.

Papoutsopoulou, S., Nikolakaki, E. and Giannakouros, T. (1999) SRPK1 and LBR protein kinases show identical substrate specificities. **Biochem Biophys Res Commun**, 255: (3): 602-607.

Parkin, D.M. (2006) The global health burden of infection-associated cancers in the year 2002. **Int J Cancer**, 118: (12): 3030-3044.

Parkin, D.M. and Bray, F. (2006) Chapter 2: The burden of HPV-related cancers. **Vaccine**, 24 Suppl 3: S3/11-25.

Pedroza-Saavedra, A., Lam, E.W., Esquivel-Guadarrama, F., et al. (2010) The human papillomavirus type 16 E5 oncoprotein synergizes with EGF-receptor signaling to enhance cell cycle progression and the down-regulation of p27(Kip1). **Virology**, 400: (1): 44-52.

Peh, W.L., Brandsma, J.L., Christensen, N.D., et al. (2004) The viral E4 protein is required for the completion of the cottontail rabbit papillomavirus productive cycle in vivo. **J Virol**, 78: (4): 2142-2151.

Peh, W.L., Middleton, K., Christensen, N., et al. (2002) Life cycle heterogeneity in animal models of human papillomavirus-associated disease. **J Virol**, 76: (20): 10401-10416.

Pelisch, F., Gerez, J., Druker, J., et al. (2010) The serine/arginine-rich protein SF2/ASF regulates protein sumoylation. **Proc Natl Acad Sci U S A**, 107: (37): 16119-16124.

Peto, J., Gilham, C., Fletcher, O., et al. (2004) The cervical cancer epidemic that screening has prevented in the UK. **Lancet**, 364: (9430): 249-256.

Petti, L., Nilson, L.A. and DiMaio, D. (1991) Activation of the platelet-derived growth factor receptor by the bovine papillomavirus E5 transforming protein. **EMBO J**, 10: (4): 845-855.

Pfister, H. (2003) Chapter 8: Human papillomavirus and skin cancer. **J Natl Cancer Inst Monogr**, (31): 52-56.

Phelps, W.C., Yee, C.L., Munger, K., et al. (1988) The human papillomavirus type 16 E7 gene encodes transactivation and transformation functions similar to those of adenovirus E1A. **Cell**, 53: (4): 539-547.

Pim, D. and Banks, L. (2010) Interaction of viral oncoproteins with cellular target molecules: infection with high-risk vs low-risk human papillomaviruses. **APMIS**, 118: (6-7): 471-493.

Raj, K., Berguerand, S., Southern, S., et al. (2004) E1^E4 protein of human papillomavirus type 16 associates with mitochondria. **J Virol**, 78: (13): 7199-7207.

Remm, M., Remm, A. and Ustav, M. (1999) Human papillomavirus type 18 E1 protein is translated from polycistronic mRNA by a discontinuous scanning mechanism. **J Virol**, 73: (4): 3062-3070.

Roberts, S. (2006) "The E4 Protein - A Late Starter". In Saveria Campo, M. (Ed.) **Papillomavirus Research: From Natural History to Vaccines and Beyond**. Caister Academic Press 83 - 95.

Roberts, S., Ashmole, I., Gibson, L.J., et al. (1994) Mutational analysis of human papillomavirus E4 proteins: identification of structural features important in the formation of cytoplasmic E4/cytokeratin networks in epithelial cells. **J Virol**, 68: (10): 6432-6445.

Roberts, S., Ashmole, I., Johnson, G.D., et al. (1993) Cutaneous and mucosal human papillomavirus E4 proteins form intermediate filament-like structures in epithelial cells. **Virology**, 197: (1): 176-187.

Roberts, S., Ashmole, I., Rookes, S.M., et al. (1997) Mutational analysis of the human papillomavirus type 16 E1--E4 protein shows that the C terminus is dispensable for keratin cytoskeleton association but is involved in inducing disruption of the keratin filaments. **J Virol**, 71: (5): 3554-3562.

Roberts, S., Hillman, M.L., Knight, G.L., et al. (2003) The ND10 component promyelocytic leukemia protein relocates to human papillomavirus type 1 E4 intranuclear inclusion bodies in cultured keratinocytes and in warts. **J Virol**, 77: (1): 673-684.

Roberts, S., Kingsbury, S.R., Stoeber, K., et al. (2008) Identification of an arginine-rich motif in human papillomavirus type 1 E1;E4 protein necessary for E4-mediated inhibition of cellular DNA synthesis in vitro and in cells. **J Virol**, 82: (18): 9056-9064.

Roden, R.B., Day, P.M., Bronzo, B.K., et al. (2001) Positively charged termini of the L2 minor capsid protein are necessary for papillomavirus infection. **J Virol**, 75: (21): 10493-10497.

Rogel-Gaillard, C., Breitburd, F. and Orth, G. (1992) Human papillomavirus type 1 E4 proteins differing by their N-terminal ends have distinct cellular localizations when transiently expressed in vitro. **J Virol**, 66: (2): 816-823.

Roscigno, R.F. and Garcia-Blanco, M.A. (1995) SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. **RNA**, 1: (7): 692-706.

Rossi, F., Labourier, E., Forne, T., et al. (1996) Specific phosphorylation of SR proteins by mammalian DNA topoisomerase I. **Nature**, 381: (6577): 80-82.

Rous, P. (1911) A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells. **J Exp Med**, 13: (4): 397-411.

Rush, M., Zhao, X. and Schwartz, S. (2005) A splicing enhancer in the E4 coding region of human papillomavirus type 16 is required for early mRNA splicing and polyadenylation as well as inhibition of premature late gene expression. **J Virol**, 79: (18): 12002-12015.

Sanford, J.R., Ellis, J. and Caceres, J.F. (2005) Multiple roles of arginine/serine-rich splicing factors in RNA processing. **Biochem Soc Trans**, 33: (Pt 3): 443-446.

Sanford, J.R., Gray, N.K., Beckmann, K., et al. (2004) A novel role for shuttling SR proteins in mRNA translation. **Genes Dev**, 18: (7): 755-768.

Sapra, A.K., Anko, M.L., Grishina, I., et al. (2009) SR protein family members display diverse activities in the formation of nascent and mature mRNPs in vivo. **Mol Cell**, 34: (2): 179-190.

Scheffner, M., Huibregtse, J.M., Vierstra, R.D., et al. (1993) The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. **Cell**, 75: (3): 495-505.

Schiller, J.T., Day, P.M. and Kines, R.C. (2010) Current understanding of the mechanism of HPV infection. **Gynecol Oncol**, 118: (1 Suppl): S12-17.

Schmitt, A., Harry, J.B., Rapp, B., et al. (1994) Comparison of the properties of the E6 and E7 genes of low- and high-risk cutaneous papillomaviruses reveals strongly transforming and high Rb-binding activity for the E7 protein of the low-risk human papillomavirus type 1. **J Virol**, 68: (11): 7051-7059.

Sciabica, K.S., Dai, Q.J. and Sandri-Goldin, R.M. (2003) ICP27 interacts with SRPK1 to mediate HSV splicing inhibition by altering SR protein phosphorylation. **EMBO J**, 22: (7): 1608-1619.

Screaton, G.R., Caceres, J.F., Mayeda, A., et al. (1995) Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. **EMBO J**, 14: (17): 4336-4349.

Sedman, J. and Stenlund, A. (1998) The papillomavirus E1 protein forms a DNA-dependent hexameric complex with ATPase and DNA helicase activities. **J Virol**, 72: (8): 6893-6897.

Sekhar, V., Reed, S.C. and McBride, A.A. (2009) Interaction of the Beta Papillomavirus E2 Tethering Protein with Mitotic Chromosomes. **J Virol**.

Shepard, P.J. and Hertel, K.J. (2009) The SR protein family. **Genome Biol**, 10: (10): 242.

Siebel, C.W., Feng, L., Guthrie, C., et al. (1999) Conservation in budding yeast of a kinase specific for SR splicing factors. **Proc Natl Acad Sci U S A**, 96: (10): 5440-5445.

Sinha, R., Allemand, E., Zhang, Z., et al. (2010) Arginine methylation controls the subcellular localization and functions of the oncoprotein splicing factor SF2/ASF. **Mol Cell Biol**, 30: (11): 2762-2774.

Somberg, M., Rush, M., Fay, J., et al. (2009) Adenovirus E4orf4 induces HPV-16 late L1 mRNA production. **Virology**, 383: (2): 279-290.

Somberg, M. and Schwartz, S. (2010) Multiple ASF/SF2 sites in the human papillomavirus type 16 (HPV-16) E4-coding region promote splicing to the most commonly used 3'-splice site on the HPV-16 genome. **J Virol**, 84: (16): 8219-8230.

Spector, D.L. (1993) Macromolecular domains within the cell nucleus. **Annu Rev Cell Biol**, 9: 265-315.

Spector, D.L. (1996) Nuclear organization and gene expression. **Exp Cell Res**, 229: (2): 189-197.

Stacey, S.N., Jordan, D., Williamson, A.J., et al. (2000) Leaky scanning is the predominant mechanism for translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA. **J Virol**, 74: (16): 7284-7297.

Stanley, M. (2008) Immunobiology of HPV and HPV vaccines. **Gynecol Oncol**, 109: (2 Suppl): S15-21.

Stauffer, Y., Raj, K., Masternak, K., et al. (1998) Infectious human papillomavirus type 18 pseudovirions. **J Mol Biol**, 283: (3): 529-536.

Steger, G. and Corbach, S. (1997) Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. **J Virol**, 71: (1): 50-58.

Sterling, J.C., Handfield-Jones, S. and Hudson, P.M. (2001) Guidelines for the management of cutaneous warts. **Br J Dermatol**, 144: (1): 4-11.

Stevens, S.W., Ryan, D.E., Ge, H.Y., et al. (2002) Composition and functional characterization of the yeast spliceosomal penta-snRNP. **Mol Cell**, 9: (1): 31-44.

Stoss, O., Stoilov, P., Hartmann, A.M., et al. (1999) The in vivo minigene approach to analyze tissue-specific splicing. **Brain Res Brain Res Protoc**, 4: (3): 383-394.

Straight, S.W., Hinkle, P.M., Jewers, R.J., et al. (1993) The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. **J Virol**, 67: (8): 4521-4532.

Strauss, M.J., Shaw, E.W. and et al. (1949) Crystalline virus-like particles from skin papillomas characterized by intranuclear inclusion bodies. **Proc Soc Exp Biol Med**, 72: (1): 46-50.

Sun, L., Zhang, G., Lei, T., et al. (2008) Two different HPV-11E6 fusion proteins trap p53 in the cytoplasm and induce apoptosis. **Cancer Biol Ther**, 7: (12): 1909-1915.

Sun, S., Zhang, Z., Sinha, R., et al. (2010) SF2/ASF autoregulation involves multiple layers of post-transcriptional and translational control. **Nat Struct Mol Biol**, 17: (3): 306-312.

Sun, Y., Han, H. and McCance, D.J. (1998) Active domains of human papillomavirus type 11 E1 protein for origin replication. **J Gen Virol**, 79 (Pt 7): 1651-1658.

Sureau, A., Gattoni, R., Dooghe, Y., et al. (2001) SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. **EMBO J**, 20: (7): 1785-1796.

Suzich, J.A., Ghim, S.J., Palmer-Hill, F.J., et al. (1995) Systemic immunization with papillomavirus L1 protein completely prevents the development of viral mucosal papillomas. **Proc Natl Acad Sci U S A**, 92: (25): 11553-11557.

Swartz, J.E., Bor, Y.C., Misawa, Y., et al. (2007) The shuttling SR protein 9G8 plays a role in translation of unspliced mRNA containing a constitutive transport element. **J Biol Chem**, 282: (27): 19844-19853.

Takano, M., Takeuchi, M., Ito, H., et al. (2002) The binding of lamin B receptor to chromatin is regulated by phosphorylation in the RS region. **Eur J Biochem**, 269: (3): 943-953.

Takeuchi, M. and Yanagida, M. (1993) A mitotic role for a novel fission yeast protein kinase dsk1 with cell cycle stage dependent phosphorylation and localization. **Mol Biol Cell**, 4: (3): 247-260.

Tan, S.H., Leong, L.E., Walker, P.A., et al. (1994) The human papillomavirus type 16 E2 transcription factor binds with low cooperativity to two flanking sites and represses the E6 promoter through displacement of Sp1 and TFIID. **J Virol**, 68: (10): 6411-6420.

Tang, S., Tao, M., McCoy, J.P., Jr., et al. (2006) The E7 oncoprotein is translated from spliced E6*I transcripts in high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via translation reinitiation. **J Virol**, 80: (9): 4249-4263.

Tao, M., Kruhlak, M., Xia, S., et al. (2003) Signals that dictate nuclear localization of human papillomavirus type 16 oncoprotein E6 in living cells. **J Virol**, 77: (24): 13232-13247.

Tavalai, N. and Stamminger, T. (2008) New insights into the role of the subnuclear structure ND10 for viral infection. **Biochim Biophys Acta**, 1783: (11): 2207-2221.

Thomas, M., Glaunsinger, B., Pim, D., et al. (2001) HPV E6 and MAGUK protein interactions: determination of the molecular basis for specific protein recognition and degradation. **Oncogene**, 20: (39): 5431-5439.

Thomas, M.C. and Chiang, C.M. (2005) E6 oncoprotein represses p53-dependent gene activation via inhibition of protein acetylation independently of inducing p53 degradation. **Mol Cell**, 17: (2): 251-264.

Titolo, S., Brault, K., Majewski, J., et al. (2003) Characterization of the minimal DNA binding domain of the human papillomavirus e1 helicase: fluorescence anisotropy studies and characterization of a dimerization-defective mutant protein. **J Virol**, 77: (9): 5178-5191.

Titolo, S., Pelletier, A., Pulichino, A.M., et al. (2000) Identification of domains of the human papillomavirus type 11 E1 helicase involved in oligomerization and binding to the viral origin. **J Virol**, 74: (16): 7349-7361.

Tsao, Y.P., Li, L.Y., Tsai, T.C., et al. (1996) Human papillomavirus type 11 and 16 E5 represses p21(WafI/SdiI/CipI) gene expression in fibroblasts and keratinocytes. **J Virol**, 70: (11): 7535-7539.

Tseng, L.C. and Chen, R.H. (2011) Temporal control of nuclear envelope assembly by phosphorylation of lamin B receptor. **Mol Biol Cell**, 22: (18): 3306-3317.

Tumban, E., Peabody, J., Peabody, D.S., et al. (2011) A Pan-HPV Vaccine Based on Bacteriophage PP7 VLPs Displaying Broadly Cross-Neutralizing Epitopes from the HPV Minor Capsid Protein, L2. **PLoS One**, 6: (8): e23310.

Tungteakkhun, S.S. and Duerksen-Hughes, P.J. (2008) Cellular binding partners of the human papillomavirus E6 protein. **Arch Virol**, 153: (3): 397-408.

Velazquez-Dones, A., Hagopian, J.C., Ma, C.T., et al. (2005) Mass spectrometric and kinetic analysis of ASF/SF2 phosphorylation by SRPK1 and Clk/Sty. **J Biol Chem**, 280: (50): 41761-41768.

Villa, L.L. (2011) HPV prophylactic vaccination: The first years and what to expect from now. **Cancer Lett**, 305: (2): 106-112.

Villa, L.L., Costa, R.L., Petta, C.A., et al. (2006) High sustained efficacy of a prophylactic quadrivalent human papillomavirus types 6/11/16/18 L1 virus-like particle vaccine through 5 years of follow-up. **Br J Cancer**, 95: (11): 1459-1466.

Vousden, K.H. and Jat, P.S. (1989) Functional similarity between HPV16E7, SV40 large T and adenovirus E1a proteins. **Oncogene**, 4: (2): 153-158.

Walboomers, J.M., Jacobs, M.V., Manos, M.M., et al. (1999) Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. **J Pathol**, 189: (1): 12-19.

Wang, H.K., Duffy, A.A., Broker, T.R., et al. (2009a) Robust production and passaging of infectious HPV in squamous epithelium of primary human keratinocytes. **Genes Dev**, 23: (2): 181-194.

Wang, H.Y., Lin, W., Dyck, J.A., et al. (1998) SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells. **J Cell Biol**, 140: (4): 737-750.

Wang, Q., Griffin, H., Southern, S., et al. (2004) Functional analysis of the human papillomavirus type 16 E1^E4 protein provides a mechanism for in vivo and in vitro keratin filament reorganization. **J Virol**, 78: (2): 821-833.

Wang, Q., Kennedy, A., Das, P., et al. (2009b) Phosphorylation of the human papillomavirus type 16 E1--E4 protein at T57 by ERK triggers a structural change that enhances keratin binding and protein stability. **J Virol**, 83: (8): 3668-3683.

Wang, W.S., Lee, M.S., Tseng, C.E., et al. (2009c) Interaction between human papillomavirus type 5 E2 and polo-like kinase 1. **J Med Virol**, 81: (3): 536-544.

Wang, X., Meyers, C., Wang, H.K., et al. (2011) Construction of a full transcription map of human papillomavirus type 18 during productive viral infection. **J Virol**, 85: (16): 8080-8092.

Watanabe, S., Kanda, T. and Yoshiike, K. (1989) Human papillomavirus type 16 transformation of primary human embryonic fibroblasts requires expression of open reading frames E6 and E7. **J Virol**, 63: (2): 965-969.

Watson, R.A., Rollason, T.P., Reynolds, G.M., et al. (2002) Changes in expression of the human homologue of the Drosophila discs large tumour suppressor protein in high-grade premalignant cervical neoplasias. **Carcinogenesis**, 23: (11): 1791-1796.

WHO/ICO (2010) "Human papillomavirus and related cancers in world, Available at www.who.int/hpvcentre.

Williams, H.C., Pottier, A. and Strachan, D. (1993) The descriptive epidemiology of warts in British schoolchildren. **Br J Dermatol**, 128: (5): 504-511.

Wilson, R., Fehrmann, F. and Laimins, L.A. (2005) Role of the E1--E4 protein in the differentiation-dependent life cycle of human papillomavirus type 31. **J Virol**, 79: (11): 6732-6740.

Wilson, R. and Laimins, L.A. (2005) Differentiation of HPV-containing cells using organotypic "raft" culture or methylcellulose. **Methods Mol Med**, 119: 157-169.

Wilson, R., Ryan, G.B., Knight, G.L., et al. (2007) The full-length E1E4 protein of human papillomavirus type 18 modulates differentiation-dependent viral DNA amplification and late gene expression. **Virology**, 362: (2): 453-460.

Woodman, C.B., Collins, S.I. and Young, L.S. (2007) The natural history of cervical HPV infection: unresolved issues. **Nat Rev Cancer**, 7: (1): 11-22.

Wu, E.W., Clemens, K.E., Heck, D.V., et al. (1993) The human papillomavirus E7 oncoprotein and the cellular transcription factor E2F bind to separate sites on the retinoblastoma tumor suppressor protein. **J Virol**, 67: (4): 2402-2407.

Wu, H., Sun, S., Tu, K., et al. (2010) A splicing-independent function of SF2/ASF in microRNA processing. **Mol Cell**, 38: (1): 67-77.

Wu, J.Y. and Maniatis, T. (1993) Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. **Cell**, 75: (6): 1061-1070.

Xiao, S.H. and Manley, J.L. (1997) Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. **Genes Dev**, 11: (3): 334-344.

Yomoda, J., Muraki, M., Kataoka, N., et al. (2008) Combination of Clk family kinase and SRp75 modulates alternative splicing of Adenovirus E1A. **Genes Cells**, 13: (3): 233-244.

Zahler, A.M., Lane, W.S., Stolk, J.A., et al. (1992) SR proteins: a conserved family of pre-mRNA splicing factors. **Genes Dev**, 6: (5): 837-847.

Zhang, W.J. and Wu, J.Y. (1996) Functional properties of p54, a novel SR protein active in constitutive and alternative splicing. **Mol Cell Biol**, 16: (10): 5400-5408.

Zhang, Z. and Krainer, A.R. (2004) Involvement of SR proteins in mRNA surveillance. **Mol Cell**, 16: (4): 597-607.

Zhao, K.N., Hengst, K., Liu, W.J., et al. (2000) BPV1 E2 protein enhances packaging of full-length plasmid DNA in BPV1 pseudovirions. **Virology**, 272: (2): 382-393.

Zhao, X., Rush, M. and Schwartz, S. (2004) Identification of an hnRNP A1-dependent splicing silencer in the human papillomavirus type 16 L1 coding region that prevents premature expression of the late L1 gene. **J Virol**, 78: (20): 10888-10905.

Zheng, Y., Fu, X.D. and Ou, J.H. (2005) Suppression of hepatitis B virus replication by SRPK1 and SRPK2 via a pathway independent of the phosphorylation of the viral core protein. **Virology**, 342: (1): 150-158.

Zheng, Z.M. and Baker, C.C. (2006) Papillomavirus genome structure, expression, and post-transcriptional regulation. **Front Biosci**, 11: 2286-2302.

Zheng, Z.M., Tao, M., Yamanegi, K., et al. (2004) Splicing of a cap-proximal human Papillomavirus 16 E6E7 intron promotes E7 expression, but can be restrained by distance of the intron from its RNA 5' cap. **J Mol Biol**, 337: (5): 1091-1108.

Zhong, X.Y., Ding, J.H., Adams, J.A., et al. (2009a) Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones. **Genes Dev**, 23: (4): 482-495.

Zhong, X.Y., Wang, P., Han, J., et al. (2009b) SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. **Mol Cell**, 35: (1): 1-10.

Zhou, J., Liu, W.J., Peng, S.W., et al. (1999) Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. **J Virol**, 73: (6): 4972-4982.

zur Hausen, H. (1976) Condylomata acuminata and human genital cancer. **Cancer Res**, 36: (2 pt 2): 794.

zur Hausen, H. (2002) Papillomaviruses and cancer: from basic studies to clinical application. **Nat Rev Cancer**, 2: (5): 342-350.

zur Hausen, H. (2009) Papillomaviruses in the causation of human cancers - a brief historical account. **Virology**, 384: (2): 260-265.