# INVESTIGATING THE ROLE OF CLK3 IN HPV ASSOCIATED CANCERS

**Maxmilan Jeyakumar**

**September 2021**

A thesis submitted to the University of Kent for the degree of PhD

Cell Biology in Division of Natural Sciences.

School of Biosciences

# Declaration

No part of this thesis has been submitted in support of an application for any degree or qualification of the University of Kent or any other University or institute of learning

Maxmilan Jeyakumar

October 2021

# Acknowledgements

Firstly, I would like to thank Dr Tim Fenton for his guidance and support throughout my PhD. His patience and encouragement were greatly appreciated.

Thank you to the members of Dr Tim Fenton's lab for their advice and support over the years. In particular, I would like to thank Dr Nicola Smith for her support and valuable friendship. I would like to thank Dr Luke Williams for all his advice and support over the years. Many thanks to Dr Ben Goult for his valued input as my secondary supervisor. I would also like to thank Dr Chris Toseland and his lab members for their assistance and support with implementing the Baculovirus expression system. Many thanks to Dr Chris Mulligan for his assistance with protein expression in bacteria. Many thanks to Prof Mark Wass for his assistance in identifying the gatekeeper domain of CLK3.

Thanks to the Debbie fund and Rosetree Trust for funding this project, and to all our collaborators including Prof Michelle Garrett, Dr Clive McCarthy (Kinsensus Ltd), and Prof John Doorbar (University of Cambridge).

Finally, a big thank you to my fianceé and my parents for all their support and encouragement over the course of my PhD.

Insanity is doing the same thing, over and over again, but expecting different results.

*Albert Einstein*

# I Contents

# II List of Figures

# IV Abbreviations

| | |
|---|---|
| ADP | Adenosine diphosphate |
| APS | Ammonium persulphate |
| Arg | Arginine |
| ASKA | Analogue-Sensitive Kinase Allele Technology |
| ATP | Adenosine triphosphate |
| BCA | Bicinchoninic acid assay |
| bp | Base pair |
| BPV | Bovine papillomavirus |
| BSA | Bovine serum albumin |
| CDK | Cyclin dependent kinase |
| CIN | Cervical intraepithelial neoplasia |
| CLK | CDC like kinase |
| $CoCl_2$ | Cobalt chloride |
| $CO_2$ | Carbon dioxide |
| DAPI | 4',6-Diamidino-2-phenylindole |
| DFO | Deferoxamine mesylate |
| $dH_2O$ | Distilled water |
| DMEM | Dulbecco's modified Eagle media |
| DMSO | Dimethyl sulphoxide |
| DNA | Deoxyribonucleic acid |
| dsDNA | double strand DNA |
| dNTPs | deoxy-nucleotide-triphosphates |
| DTT | Dithiothreitol |
| E6AP | E6 associated protein |
| ECL | Enhanced chemiluminescence |
| EDTA | Ethlenediaminetetra-acetic acid |
| EGF | Epidermal growth factor |
| FACS | Fluorescence activated cell sorting |
| FBS | Fetal bovine serum |
| GAPDH | Glyceraldehyde 3-phosphate dehydrogenase |
| Gly | Glycine |
| GFP | Green fluorescent protein |
| HBES | Hepes-buffered Earles' Salts |
| HCl | Hydrochloric acid |
| His | Histidine |
| HNSCC | Head and neck squamous cell carcinoma |
| HRP | Horseradish peroxidase |
| HPV | human papillomavirus |
| KCl | Potassium chloride |
| KO | Knockout |

| | |
|---|---|
| LB | Luria-Bertani |
| $MgCl_2$ | Magnesium Chloride |
| M-PER | Mammalian Protein extraction reagent |
| mRNA | Messenger ribonucleic acid |
| MAGUK | Membrane-associated guanylate kinase |
| min | Minute |
| ml | Mililiter |
| mM | Milimolar |
| M | Molar |
| NaCl | Sodium Chloride |
| $NaN_3$ | Sodium azide |
| NEB | New England Biolabs |
| NIKS | Normal immortalised keratinocytes |
| ng | Nanogram |
| nm | Nanometer |
| nM | Nanomolar |
| NP-40 | Nonident P-40 |
| ORF | Open reading frame |
| PDM | PDZ binding motif |
| PBS | Phosphate buffered saline |
| PRC | Polycomb repressor complex |
| PCR | Polymerase chain reaction |
| PVDF | Polyvinylidene fluoride |
| PHD | prolyl-4-hydroxylase |
| pVHL | Von Hippel-Lindau protein |
| RNA | Ribonucleic acid |
| RNAi | Ribonucleic acid interference |
| Rpm | Revelutions per minute |
| Rb | Retinoblastoma protein |
| qRT-PCR | Reverse transcriptatse quantitative PCR |
| SCC | squamous cell carcinoma |
| SDS | Sodium dodecyl sulphate |
| Ser | Serine |
| shRNA | Short hairpin RNA |
| siRNA | Small interfering RNA |
| SRSF | Serine-arginine rich splicing factor |
| TBP | TATA box binding protein |
| TBS | Tris-buffered saline |
| TBST | Tris-buffered saline + Tween 20 |
| TEMED | N,N,N′,N′-tetramethylenediamine |
| Thr | Threonine |
| Tris | Tris[hydroxymethyl]amino-methane |
| VEGF | Vascular endothelial growth factor |
| VLP | Virus like particle |

| | |
|---|---|
| URR | Upstream regulatory region |
| WT | wild-type |
| µ | micro |
| µg | microgram |
| µl | microliter |
| µM | micromolar |

# V ABSTRACT

Human papillomaviruses (HPV) are responsible for over 95% of cervical cancers and over 20% of oropharyngeal cancers. The continued expression of high-risk HPV oncogenes E6 and E7 are considered essential for transformation and maintenance of tumours. Targeting kinases essential for the expression of E6/E7 presents an opportunity for novel therapies in HPV-associated cancers. Previously, various HPV+ cell lines showed survival dependency on CLK3, with CLK3 siRNA and shRNA knockdown resulting in poor cell viability, decreased transcription of E6/E7, and subsequent loss of protein expression. Cells were also found to arrest at G1, and enter senescence. Here, further validation was done in immortalised keratinocytes (NIKS) stably transduced with HPV16 URR E6E7, with stronger loss of E6/E7 mRNA and E7 protein being observed in CRISPR/Cas9 edited $CLK3^{-/-}$ NIKS. Wild-type but not CLK3 mutant lacking kinase activity (K186R) was able to restore E7 expression in $CLK3^{-/-}$ NIKS. Loss of CLK3 also resulted in reduced c-MYC protein, which was demonstrated to regulate HPV16 E7, with siRNA knockdown of c-MYC resulting in loss of HPV16 E7. In addition, c-MYC overexpression was shown to upregulate the oncoprotein. CLK3 inhibition with a new compound (KIN101982-001) was successful in reducing both E6/E7 and c-MYC levels. These findings show that CLK3 is a potential novel therapeutic target for HPV-driven cancers, as well as HPV infections, and the kinase has the potential to be targeted in c-MYC driven cancers to disturb the stability of the infamous oncoprotein.

# Chapter 1

# Introduction

# 1 Introduction

## 1.1 Human papillomavirus

Papillomaviridae (PV) are small non-enveloped double-stranded DNA viruses with a virion size of ~55nm in diameter. The virus is known to infect a wide range of species, including birds, reptiles, marsupials, and mammals (Bravo, et al., 2010). PV types are identified based on the L1 ORF, which is the most conserved gene within the genome. A new PV is only recognised if the complete genome has been cloned and the L1 ORF sequence differs by more than 10% from its closest known PV type. Differences between 2-10% homology define a new subtype, and any difference less than 2% define a new variant. Sixteen groups of the virus have been classified based on their genotype, with all human papillomaviruses falling into five groups (Alpha-, Beta-, Gamma-, Mu-, and Nu-) (De Villiers, et al., 2004). Over 400 subtypes of the virus have been identified, including over 200 that can infect humans (PaVE online database).

## 1.2 Cervical cancer and HPV

The first descriptions of cervical cancer dates back as far as 400 BC. During the mid-19$^{th}$ century, pioneer work by an Italian surgeon named Rigoni-Stern led to the initial association of an infectious agent being linked to cervical cancer. He observed a low incidence of cervical cancer among nuns, which led to the demonstration that cervical cancer was common among sex workers and among women whose husbands had frequent

interactions with prostitutes (Rigoni-Stern, 1842, Mak, et al., 2004, Bayo, et al., 2002). Following this, cervical cancer was considered highly transmissible through sexual intercourse. It was not until 1976 when Zur Hausen, a German virologist, first reported HPV DNA in cervical cancer and warts. Further work by Zur Hausen and his team confirmed the structure of the virus and laid the foundations to our current understanding of the role HPV plays in cancer development and most importantly, the development of the HPV vaccine.

HPV infections have a broad range of symptoms and vary greatly in severity. Most types of the virus cause cutaneous infections, typically resulting in benign lesions with low severity. The most prevalent benign lesions include common warts, plantar warts, flat warts, and pigmented warts (Sterling, et al., 2001). Diseases associated with cutaneous HPV infections are usually a low risk to the patient and are rarely fatal. In addition to causing cutaneous infections, some of the HPVs are associated with mucosal epithelial infections. In fact, over forty types of HPVs are known to infect various regions of the mucosal membranes, leading to a range of diseases from genital warts to malignant carcinomas. These 'mucosal' HPVs are usually classed into either a 'low-risk' or 'high-risk' group, based on their ability to cause malignancies. HPV 6 and 11 are the most common low-risk subtypes and are associated with 90% of genital warts (Giuliano, et al., 2015, Gross, et al., 2004). Most of these infections are cleared quickly by the immune system and do not persist in individuals.

The International Agency for Research on Cancer has further classified HPV types into risk groups: Group 1 defined as “carcinogenic to humans” (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59); Group 2A as “probably carcinogenic to humans (68); and Group 2B as “possibly carcinogenic to humans” (26, 53, 66, 67, 70, 73, and 82) (IARC, 2012). They are known to be associated with cancer of the cervix, anus, penis, vagina, vulva, and oropharynx. High-risk types HPV-16 and HPV-18 are the most prevalent in cancer, being responsible for over 70% of cervical cancers (Doorbar, et al., 2012). The association between HPV and cervical cancer has been extensively investigated. It is now apparent that HPV can be detected in 11-12% of all women showing no symptoms, and this number is much higher in younger women (50-80%) (Stanley, 2010). The detection of HPV increases with the severity of the disease, with 50-70% detection in CIN1 (cervical intraepithelial neoplasia grade 1), rising to between 90-100% for CIN3 and invasive cervical cancer (Bruni, et al., 2010, Guan, et al., 2012).

## 1.3 Head and Neck Cancer and HPV

The involvement of HPV in a subgroup of head and neck cancers was first hypothesised in 1983 (Syrjänen, et al., 1983). Since then, evidence has emerged identifying specific sites within the oral cavity that are frequently associated with HPV-linked cancers. The most common site by some margin is the oropharynx, which includes the base of the tongue and palatine tonsils and in which HPV accounts for as many as 70% of squamous cell carcinomas in parts of the USA and Western Europe, while HPV is linked to much smaller

proportion (<4%) of tumours in the oral cavity, larynx, and sinonasal mucosa (Gillison, et al., 2000, Syrjänen, et al., 2016). Head and neck squamous cell carcinoma (HNSCC) can be classed into two morphological types, either keratinised or non-keratinised. Majority of the HNSCC are keratinised, and mostly occur in elderly men with significant association to smoking and alcohol consumption. Whereas nonkeratinised HNSCC are less common and usually occur in men aged 40-55, with little association to smoking or alcohol consumption. Here, HPV association is considered to be a significant risk factor. In fact, HPV-16 is associated with 90% of HPV-positive oropharyngeal cancers. A decline in HPV-negative HNSCCs have been observed in high-income countries in recent years, with the most epidemiological data from the USA, which can be attributed to the declining trend in smoking and alcohol consumption (Chaturvedi, et al., 2008, Chaturvedi, et al., 2011,). However, there has been an increase in HPV-positive HNSCCs (Curado, et al., 2013, Ferlay, et al., 2015), which is thought to be due to a greater number of sexual partners and an increase in the frequency of oral sex (Gillison, et al., 2008, D'Souza, et al., 2009).

## 1.4 HPV Burden

HPV is responsible for 570,000 cases of cancer in women and 60,000 cases in men globally per year, and at least 250,000 deaths (De Martel, et al., 2017). Cervical cancer still remains the second most common cause of death from cancer among women in developing countries, and fourth globally despite an overall reduction in the incidence. The reduction in incidence is largely due to an increase in HPV screening and the effective vaccination

programmes that have been introduced in high-income countries and the vast majority of the cervical burden and deaths now occur in low/middle-income countries. Cervarix (GSK) was the first vaccine to be introduced in the UK in 2008, with protection against HPV16 and HPV18. In 2012, Cervarix was discontinued in the UK, and Gardasil (Merck), which provides additional immunity against HPV6 and HPV11, was introduced. Gradasil-9 (Merck), which protects against five additional HPV high-risk strains (31, 33, 45, 52 & 58) is currently available, however is not yet used by the NHS.

Clinical trials have shown that HPV vaccines are highly effective in preventing cervical infections when given prior to first exposure to the virus, and they have also been shown to reduce infections in other tissues including the anus and oral regions (Kreimer, et al., 2011, Herrero, et al., 2013). Gradasil-9 was found to be nearly 100% effective in preventing cervical, vulvar, and vaginal infections and precancers caused by all HPV types it targets (reviewed in Chatterjee, 2014). Despite the success of HPV vaccines, there still remains a number of barriers against implementing the vaccines for the vulnerable population, especially those in developing and third world countries. Such obstacles include high vaccine cost, inadequate delivery infrastructure, and lack of awareness about cervical cancer and early screening tools (reviewed in Graham, et al., 2011). Furthermore, there has also been an increase in HPV-associated head and neck cancers (Curado, et al., 2013, Ferlay, et al., 2015).

Most women with precancerous cervical lesions are treated with the loop electrosurgical excision procedure (LEEP). Treatment for precancerous lesions and genital warts includes topical medicines, surgical excision, cryosurgery, and laser therapy. Treatment of HPV-related cancers depend on the stage of the disease and the available options are surgery, radiotherapy, chemotherapy, and some targeted therapy. Bevacizumab has been approved for treatment against advanced cervical cancer (Tewari, et al., 2017). More recently, Pembrolizumab, an anti-PD-L1 antibody, has been approved for positive metastatic or unresectable cervical cancer (Colombo et al., 2021). In addition, currently there are over 60 clinical trials that involve HPV vaccines as monotherapies or in combination with other immunotherapeutics (Reviewed in Rumfield, et al., 2020). There is an urgent need to identify treatments to improve outcomes for patients with recurrent or metastatic HPV-associated cancer.

.

## 1.5 HPV Genome

The HPV genome consists of approximately 8000bp, with the coding region being divided into early and late regions (figure 1.5.1). The early region contains six genes (E1, E2, E4, E5, E6, E7), which are mostly expressed in the lower layers of the epithelium (Danos, et al., 1982). These genes encode proteins responsible for viral replication, transcription, immune system evasion, cell cycle control, and apoptosis inhibition. Two other Early genes, E3 and E8, were initially included, however, only E8 has been demonstrated to encode a protein in BPV-1 (Bovine papillomavirus), and E3 ORF was later identified as a sequencing

error in the BPV1 genome and was shown to be part of the E2 ORF (Lambert, et al., 1987, Choe, et al., 1989). The late region consists of two genes (L1 & L2), and expression of these genes are only seen in the granular, differentiated, layers. The L1 and L2 proteins form the viral capsid. Upstream of the coding region is the upstream regulatory region (URR), made up of the replication origin, as well as transcription promotor and enhancer sequences. There are variations in the genome between the HPV types and differences have been observed in the length of the sequence. E1, E2, L1, and L2 are highly conserved across all HPVs, with a reasonable degree of variation being observed in the remaining genes.

HPV16 is the most common cause of cervical cancer and therefore has been studied most extensively. The genome is made of 7960bp and consists of two main viral promotors, $P_{97}$ and $P_{670}$. $P_{97}$ is activated early in the viral cycle and is responsible for the expression of majority of the early genes. $P_{670}$ is activated after cellular differentiation and leads to the expression of late genes. The HPV16 genome also has three polyadenylation sites, one located just after the E5 gene and the other two found just after the L1 gene.

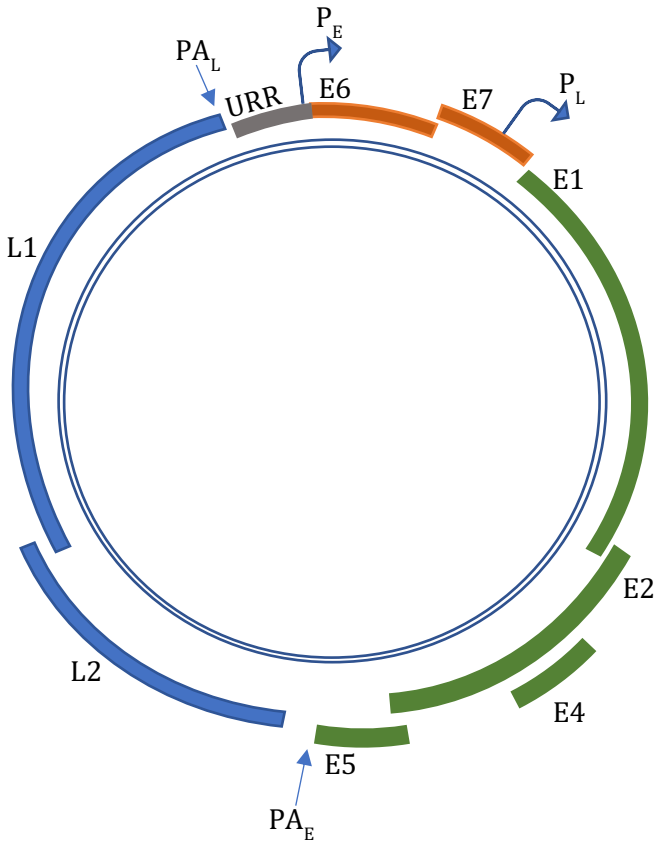



**Figure 1.5.1 Schematic diagram of the HPV16 genome**
The E6 and E7 oncogenes are in orange. All other early genes are shown in green. The late genes are shown in blue. $PA_E$ and $PA_L$ are the early and late poly adenylation sites, and $P_E$ and $P_L$ are early and late promoters. URR, the upstream regulatory region, is shown in grey.

## 1.6 HPV Replication cycle

The virus typically enters the host via a wound or abrasion in the epithelium, with the replication cycle initiating when the virus enters the basal epithelial layer (Figure 1.6.1). Due to the difficult nature of generating HPVs in vitro, the understanding of viral entry and host protein interaction still remains controversial. Studies have repeatedly found cell surface heparin sulfate to be the primary site of attachment of the L1 protein of HPV (Joyce et al., 1999, Giroglou et al., 2001). In addition, the following proteins have been

demonstrated to be involved in the viral binding and uptake: EGFR/KGFR, tetraspanins CD63 and CD151, α6 and ß4 integrins, and annexin A2 heterotetramer (Culp, et al., 2006, Woodham, et al., 2012, Aksoy, et al., 2017). Following initial binding, the virus is understood to be internalized into the host cell via a clatherin independent mechanism (Schelhaas, et al., 2012). The study found that viral internalisation involved various kinases including PAK-1, PKC, and PI3K, and utilised an endocytic pathway related to micropinocytosis, in an actin-dependent manner. The process of intracellular trafficking of the virus to the nucleus is poorly understood. The virus has been shown to travel through the Golgi and the ER before entering the nucleus, and the viral L2 protein was found to facilitate its movement via multiple cellular protein interactions including SNX17, SNX27, and γ-secretase (Marušič, 2012, Pim, et al., 2015, Zhang, et al., 2014).

The life cycle of the virus is highly dependent on the host cell's differentiation. Viral replication can be divided into four stages that utilise different mechanisms; amplification, establishment, maintenance and vegetative amplification. Initially, post-infection, following a few rounds of DNA replication cycle, the viral genomic copy number is maintained in low quantities (50-100 copies) (Maglennon, et al., 2011). The cell continues to proliferate to create a pool of infected basal cells, where the viral genome undergoes synchronous replication along with the host genome (McBride, 2008). This pool of infected basal cells is important in maintaining a persistent long-term viral infection in the host, and the low viral load helps evade the immune system. The replication cycle of the virus can only proceed further once host cell differentiation takes place. In uninfected keratinocytes, cell cycle

termination takes place once mitosis has occurred and the resulting daughter cell has migrated up the suprabasal layers. In HPV infected cells, however, the migrated cells are forced to continue differentiating.

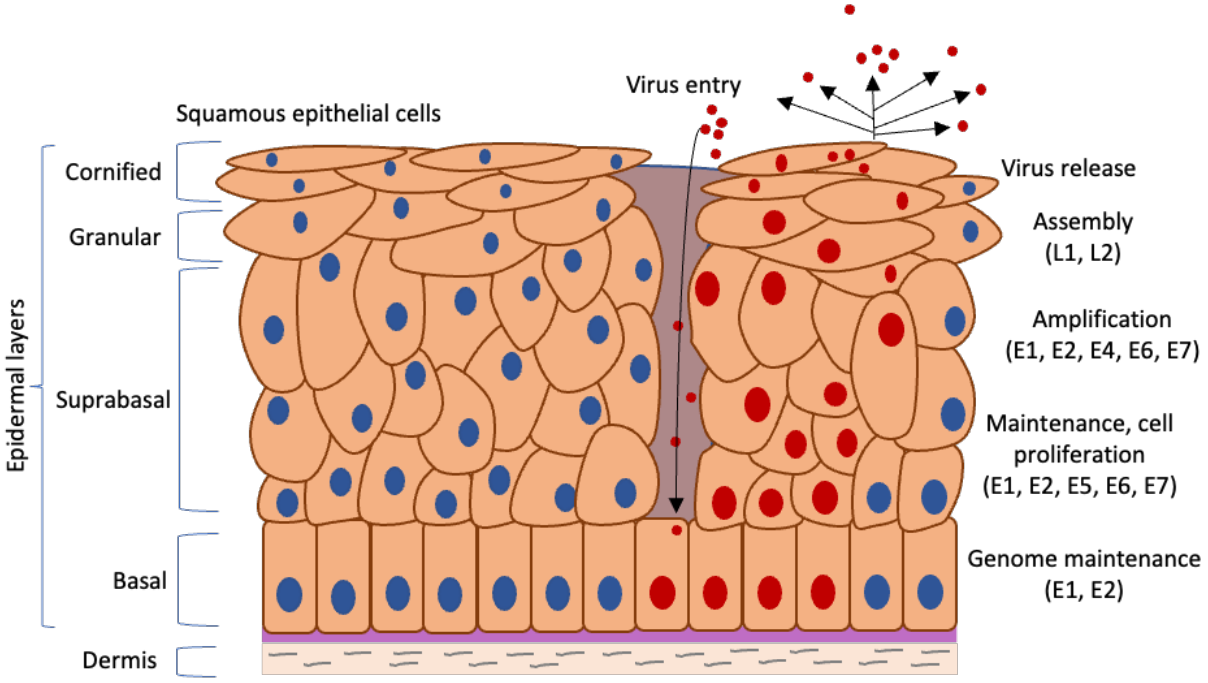



**Figure 1.6.1 HPV life cycle**
The epidermal layers are shown separated into basal, suprabasal, granular and cornified layers. The HPV gene expression at each layer is stated. Red cells represent infected cells, with the small red circles representing viral particles. Blue cells represent uninfected cells.

The last stage is vegetative amplification, which occurs exclusively in differentiated cells and generates a high copy number of viral genomes. Here, cells are maintained in a pseudo S phase by E6 and E7. For successful viral genome amplification, the virus requires the host replication machinery. However, DNA replication is significantly suppressed in differentiated cells, and HPV viral replication is thought to take place at G2 phase of the cell cycle in differentiated cells, and therefore lack access to the S-phase host replication machinery. HPVs are thought to mimic DNA damage and initiate viral DNA replication by

utilising the DNA damage response (DDR) pathway. It was shown that HPV E2 protein interacts with BRD4 proteins at fragile regions of the host genome susceptible to replication stress (Jang, et al., 2014). This interaction is thought to initiate a DDR signalling response and recruit factors that aid in DDR-dependent replication foci to complete late amplification of the viral DNA. As a result, when the infected cell enters the upper epithelium, the viral genome replication is considerably accelerated to thousands of copies per cell (Gadducci, et al., 2012). Finally, viral capsid proteins L1 and L2 are synthesised in the granular layer, which package the viral genome into virions. The virions are then released to further infect other keratinocytes.

## 1.7 HPV proteins

The genes to be transcribed from the HPV episomal genome is determined by which stage of differentiation the host cell is at. Initially following infection, E1 and E2 proteins, which are known as replication proteins, are expressed at high levels. The E1 protein is a ~70 kDa ATP-dependent helicase consisting of four domains. It binds to the origin of replication to initiate the replication of the viral genome (Yang, et al., 1993). E1 alone shows minimal DNA binding affinity and requires the presence E2 to initiate replication. E2, in addition to being a loading protein, also acts as an inhibitor and prevents the unwinding of DNA. Displacement of E2 with cellular chaperone proteins results in increased DNA unwinding by the E1-replication complex. E1 also possesses ATPase activity, which provides energy to drive helicase function. Unwinding of the DNA allows access and interaction with

other cellular proteins such as DNA polymerase α primase, topoisomerase 1, and polymerase δ (Castro-Muñoz, et al., 2019).

The E2 is a ~42 kDa protein made up of three domains and functions as a dimer in the regulation of viral DNA replication and transcription. In addition to its interaction with E1, E2 also interacts with cellular transcription factors such as p300 and NFK/1 and regulates viral transcription (review, McBride, A., 2013). Depending on the promoter it interacts with, E2 can either serve to upregulate or suppress gene transcription. This is true for p97, the promotor of E6/E7, which is either activated at low levels of E2, or repressed when E2 is expressed at higher levels (Steger., et al, 1997). Furthermore, E2 also interacts with multiple cellular promoters, which regulates the expression of cellular proteins required during the virus life cycle. This includes the downregulation of telomerase reverse transcriptase by HPV18 E2 (Lee, et al., 2002) and the interaction of HPV16 E2 with Sp1 to transactivate the SRSF1 promotor (Mole, et al., 2009).

E5 is a ~10 kDa hydrophobic protein that plays a significant role in evading the immune system and viral persistence. It is a membrane-bound protein that is expressed from an early stage of infection in the basal layer. Its major contribution is the downregulation of MHC class I in the infected cells, thereby avoiding recognition by CD8+ T cells (Ashrafi, et al., 2006, Zhang, et al., 2003). Detection of E5 protein in precancerous and malignant cervical lesions suggests that the protein may have oncogenic activities (Chang et al., 2001, Sahab, et al., 2012). This is supported further by evidence of E5 causing an

increase in EGF receptors by 2-5 fold in keratinocytes expressing HPV16 E5, by interfering with EGFR internalisation and degradation of EGFRs, which results in increased DNA synthesis and cell proliferation (Straight, et al., 1993). More recently, E5 from high-risk HPVs was found to possess viroporin functions (Wetherill, et al., 2012). Furthermore, HPV16 E5 has been shown to enhance cell proliferation by downregulating tumour suppressor proteins p21 and p27 by transcriptional and post-translational mechanisms, which further highlights its oncogenic role (Tsao, et al., 1996, Pedroza-Saavedra, et al., 2010).

E4 is a ~10 kDa protein expressed in the middle and upper epithelial layers (Maglennon, et al, 2011). The protein is translated from a spliced E1^E4 mRNA, where the first five amino acids are derived from the E1 ORF. Due to the high abundance of the protein in cells that support viral genome amplification and the lack of E4 in high-grade lesions and carcinoma, it has been suggested as a biomarker for productive infection (Griffin, et al., 2012, Yajid, et al, 2017). There has yet been no evidence to suggest a functional role for E4 during the early stages of the virus life cycle, supported by no detection of E1^E4 levels during the early stages (Doorbar., 2013). E4 is known to associate with keratin and disrupt the cytoskeleton network of the cell, which is considered important in the viral release. In addition, high-risk HPV E4 is thought to play a part in cell growth arrest at G2 by inhibiting the nuclear accumulation of cyclin/cdk1 (Davy, et al., 2002).

The virus capsid is formed of two structural proteins, L1 and L2. The major capsid protein, L1, is a ~55 kDa protein that can self-assemble into virus-like particles (VLPs)

(Kirnbauer, et al., 1992). These VLPs are highly immunogenic and are recognised by innate B-cells (Bachmann, et al., 1993), and these early discoveries marked a pivotal point in the development of current VLP-based vaccines against HPV. The viral capsid consists of 360 copies of L1, arranged as 72 pentamers. L1 is essential for the initial interaction between the virus and the host cell, interacting with heparin sulfate carbohydrates on proteoglycans (Joyce, et al., 1999). This results in a conformational change, exposing the L2 protein.

L2 is the minor capsid protein, with a molecular weight of ~55 kDa. The protein co-assembles with L1 into VLPs, outnumbered by a ratio of 5:1. Following initial interaction of the viral capsid with the host cell, the L2 N terminal is cleaved by furin, which is an essential step for completion of HPV infection (Day, et al., 2008). This leads to further conformational changes and additional interactions between L2 and an unknown receptor, facilitating the release of the L2/genome complex into the host via endocytosis.

## 1.8 HPV Oncoproteins

With the exception of the oncogenic activities of E5, E6 and E7 are considered the two major oncoproteins responsible for the initial establishment and progression of high-risk HPV-associated cancers. The oncogenic activities of high-risk HPVs are attributed to the functional importance of E6 and E7. It has been suggested that both E6 and E7 evolved from a common ancestor (Cole and Danos, 1997, Van Dooslaer, et al., 2009), and consistent

with this, some PVs encode E6 and no E7 (Stevens, et al., 2008, Gottschling, et al., 2011) whereas others encode E7 and not E6 (Chen, et al., 2007, Nobre, et al., 2009).

The two oncogenes contribute to many of the aspects of the hallmark model (Weinberg, et al., 2011) including angiogenesis, uncontrolled proliferation, activating invasion and metastasis, and resisting cell death. Several studies have shown that high-risk E6 and E7 are together sufficient to induce the immortalisation of primary human keratinocytes, which are the natural host cells of the virus *in vivo* (Hawley-Nelson, et al., 1989, Münger et al., 1989). Studies have also demonstrated that E6 and E7 together can induce various types of tumours, depending on the tissue in which they are expressed (Griep, et al., 1993, Comerford, et al., 1995). The continued proliferation and survival of cervical cancer cell lines requires the ongoing expression of E6 and E7. When HPV or BPV E2 gene or siRNAs against HPV E6 and E7 are introduced into cervical carcinoma cells, they undergo cell growth arrest and/or cell death (McBride, et al., 1991, Hwang, et al., 1993., Dowhanick, et al., 1995, Jiang, et al., 2002). Furthermore, repression of HPV18 E6 and E7 in HeLa cervical cancer cells by BPV-E2 results in growth arrest and cellular senescence due to reactivation of the p53 and Rb tumour suppressor pathways (Goodwin, et al., 2000, Horner, et al., 2004, Johung, et al., 2007). Repression of E7 in transgenic mice was sufficient enough to cause regression of high-grade cervical dysplasia and tumours (Jabbar, et al., 2009).

E6 is a ~18 kDa nuclear protein and comprised of two zinc finger metal-binding domains linked by a central helix. The oncoprotein has a wide range of functions and targets multiple cellular factors. The most well-known function of E6 is the regulation of p53 levels

and activity. p53 is a transcription factor that controls a wide array of cellular processes including cellular senescence, apoptosis, and cell cycle regulation (reviewed in Kastenhuber, et al., 2017). p21, a key target of p53, is an important inhibitor of CDKs and is responsible for strictly regulating the cell cycle checkpoints (reviewed in Shamloo, et al., 2019). High-risk HPV E6 binds to E3 ubiquitin ligase UBE3A/E6AP to form the stable E6/E6AP complex, which then flags p53 for rapid proteasomal degradation (Scheffner, et al., 1993, Huibregtse, et al., 1993). This interference prevents apoptosis, as well as DNA repair and G2/M arrest. Low-risk HPV E6s were also shown to bind p53, but they are unable to target the tumour suppressor for degradation (Crook, et al., 1991). In addition, E6 also binds to histone acetyltransferase p300, which prevents the acetylation of p53. Acetylation of p53 results in conformational changes leading to altered transcriptional activity and stability of p53, therefore preventing acetylation results in p53 being targeted for ubiquitination and subsequent degradation (Xie, et al., 2014).

E6-induced loss of p53 is considered the most important function of E6, however, studies have highlighted numerous other functions of E6. In fact, E6 has been shown to interact with 153 cellular proteins (White, et al., 2012). Some of these functions include binding with FADD to inhibit Fas-mediated apoptosis, promotion of cell migration and invasion by binding to fibulin-1, binding with DNA repair protein XRCC1 to interfere with DNA repair efficiency, and binding to the DNA methyltransferase MGMT to cause increased genomic instability and accelerated progression towards carcinogenesis (Yim, et al., 2005, Howie, et al., 2009). High-risk HPV E6 proteins also contain a PDZ binding motif (PBM),

which is absent in low-risk HPV E6s, that interacts with multiple PDZ containing targets resulting in their degradation (Ganti, et al., 2015). The PBM is usually found at the extreme C-terminus of the protein, but can also be found at internal sites (Songyang, et al., 1997). Such E6 targets include several of the membrane-associated guanylate kinase (MAGUK) superfamily proteins known to have multiple PDZ domains (Thomas, et al., 2002). Examples include hDlg and hScrib, which are part of tight junctions that function in epithelial cell adhesion and control apicobasal polarity (Grm & Banks, et al., 2004, Thomas, et al., 2005, Banks, et al., 2012). E6 targets hDlg and hScrib for degradation, which results in loss of cellular polarity, defective cell-cell adhesion and promotion of tumourigenesis.

E7 is a ~16 kDa phosphoprotein with 3 conserved regions (CR). Sections of CR1 and majority of CR2 have sequence similarities with the adenovirus E1A protein and the large T antigen of SV40 (Phelps, et al., 1988, Barbosa, et al., 1990). The most studied role of high-risk HPV E7s is their interaction with Rb, where they bind to its 'pocket domains' (Imai, et al., 1991), which is essential for its tumour suppressing activities. E7 also interacts with and inactivates other members of the Rb family, p107 and p130 (Dyson, et al., 1992, Davis, et al., 1993). Normally, hypophosphorylated pRb binds E2F1-3a/b/DP heterodimers, p107 and p130 associate with E2F4 and E2F5, thereby suppressing the transcription function of these pocket proteins, which prevents the transition from G1 to S phase in the cell cycle. Phosphorylation of pRb by cyclin D/CDK4/6 complexes releases E2Fs, allowing the cell cycle to further progress. The binding of E7 to Rb disrupts the Rb-E2F interaction, which releases the E2Fs in their active forms to promote the expression of several genes involved in the

cell cycle, such as cyclin A and cyclin E. This results in premature progression towards S phase and DNA replication, and Rb-E2F disruption maintains the cell in this pseudo-S phase for optimal viral amplification. In contrast, low-risk HPV E7 binds Rb with a 10-fold lower affinity (Heck, et al., 1992), which may contribute to their lack of oncogenic activity.

HPV16 E7 has been reported to directly associate with E2F6-containing polycomb transcriptional repressor complexes (PRCs). PRCs associate with the histone H3 lysine 27 trimethyl (H3K27me3) modification to transcriptionally silence chromatins. PRC-regulated genes include INK4A-ARF tumour suppressor locus, which encodes $p16^{INK4A}$ (an inhibitor of CDK4 and 6), and $p14^{ARF}$ (an inhibitor of mdm2-mediated p53 degradation), both of which are key regulators of cellular senescence (Rayess, et al., 2011). The H3K27me3 modification can be removed by one of two histone demethylases, KDM6A and KDM6B, both of which are dramatically induced by HPV16 E7 (McLaughlin-Drubin, et al., 2011). KDM6B has been demonstrated to remove H3K27me3 from the $p16^{INK4A}$ promoter, upregulating p16 expression, leading to RAS/RAF mediated oncogene-induced senescence via Rb activation (Agger, et al., 2009, Barradas, et al., 2009). p16 expression is also regulated by Rb. Phosphorylation of Rb results in increased levels of p16, which in turn inhibits CDK4/6 from phosphorylating Rb, thereby increasing the levels of hypophosphorylated Rb, which then drives down p16 expression, resulting in a feedback loop between p16 and Rb. E7 overexpression of p16 through KDM6B induction should result in cellular senescence, however, the protein has evolved to additionally target Rb for degradation through ubiquitination (Boyer, et al., 1996). This means, despite the overexpression of p16, cells do

not enter senescence and high levels of p16 act as an excellent biomarker for high-risk HPV associated cancers (Munger, et al., 2013).

In addition to Rb degradation, high-risk HPV E7s can interfere with the cell cycle through other means. This includes interaction with cyclin-dependent kinase inhibitors p21 and p27, suppressing their activity on cyclin-dependent kinases, which drive the cell cycle (Funk, et al., 1997, Zefrass-Thome, et al., 1996). HPV16 E7 has also been shown to directly associate with cdk2/cyclin A and cyclin E complexes, promoting HPV replication through increased cdk2 activity (He, et al., 2003, Nguyen, et al., 2008). In addition, E7 was found to upregulate the expression of SIX1, a transcription factor necessary during early embryogenesis, but is usually lost in differentiated cells (Liu, et al., 2014). Upregulation of SIX1 by E7 is shown to promote tumour growth and metastasis. Furthermore, E7 is known to bind with various other cellular proteins, however, interactions seem to vary depending on the HPV type (White, et al., 2012).

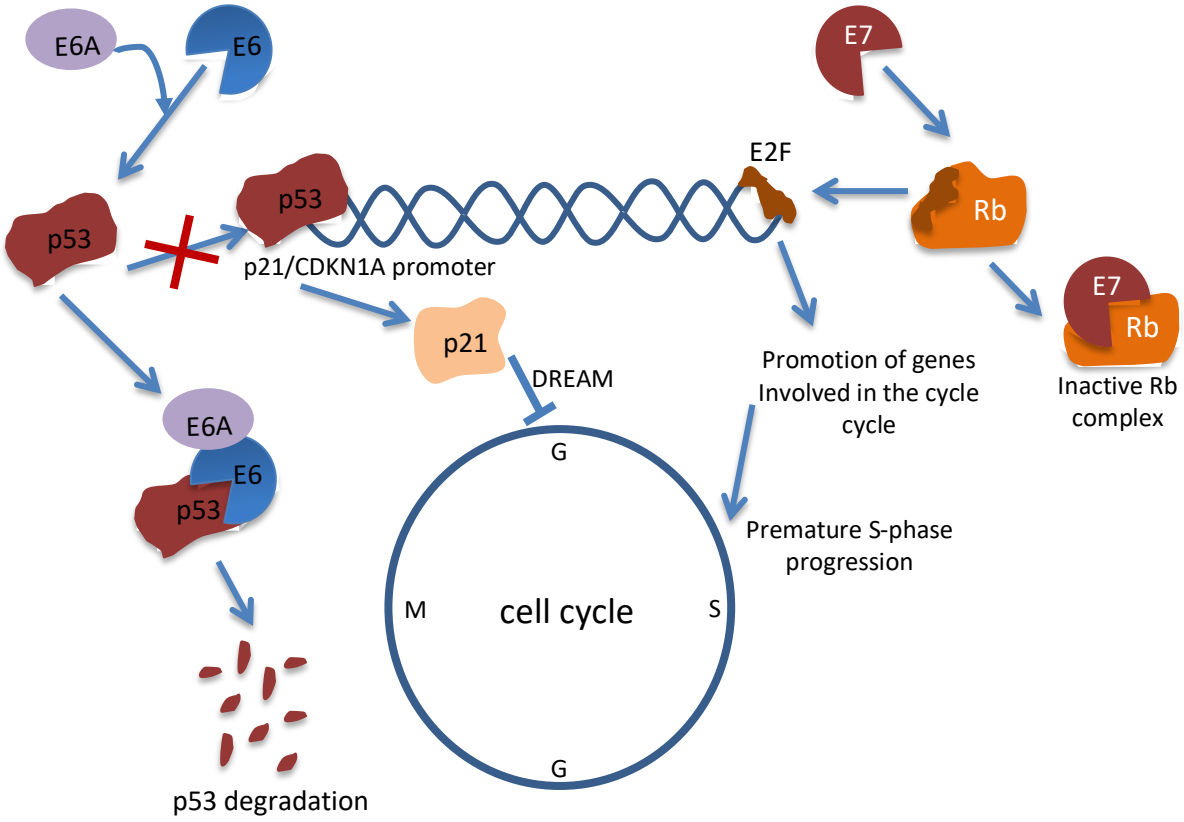



**Figure 1.8.1 Main functions of HPV oncogenes E6 and E7**
E6 inhibits p53-dependent growth arrest and apoptosis. E7 binds to Rb and disrupts the Rb-E2F interactions, resulting in premature progression towards S phase and DNA replication.

## 1.9 High-risk HPV genome integration

High-risk HPVs are often found integrated in premalignant tumours and HPV-associated cancers. The integration event is not part of the normal viral life cycle, as the integration is accompanied by the loss of one or more of the viral genes needed for the production of infectious virions. About 90% of High-risk HPV infections regress spontaneously within several months. In the remaining 10% of the cases, the infection may progress to preneoplastic lesions or invasive cancers. Normally, the latency period between

initial infection and low-grade cervical dysplasia is 5 years, and further development of invasive cancer is an additional 15 years. During this latency period, the virus continues to remain in episomes, and the low titre enables the virus to evade the host immune system.

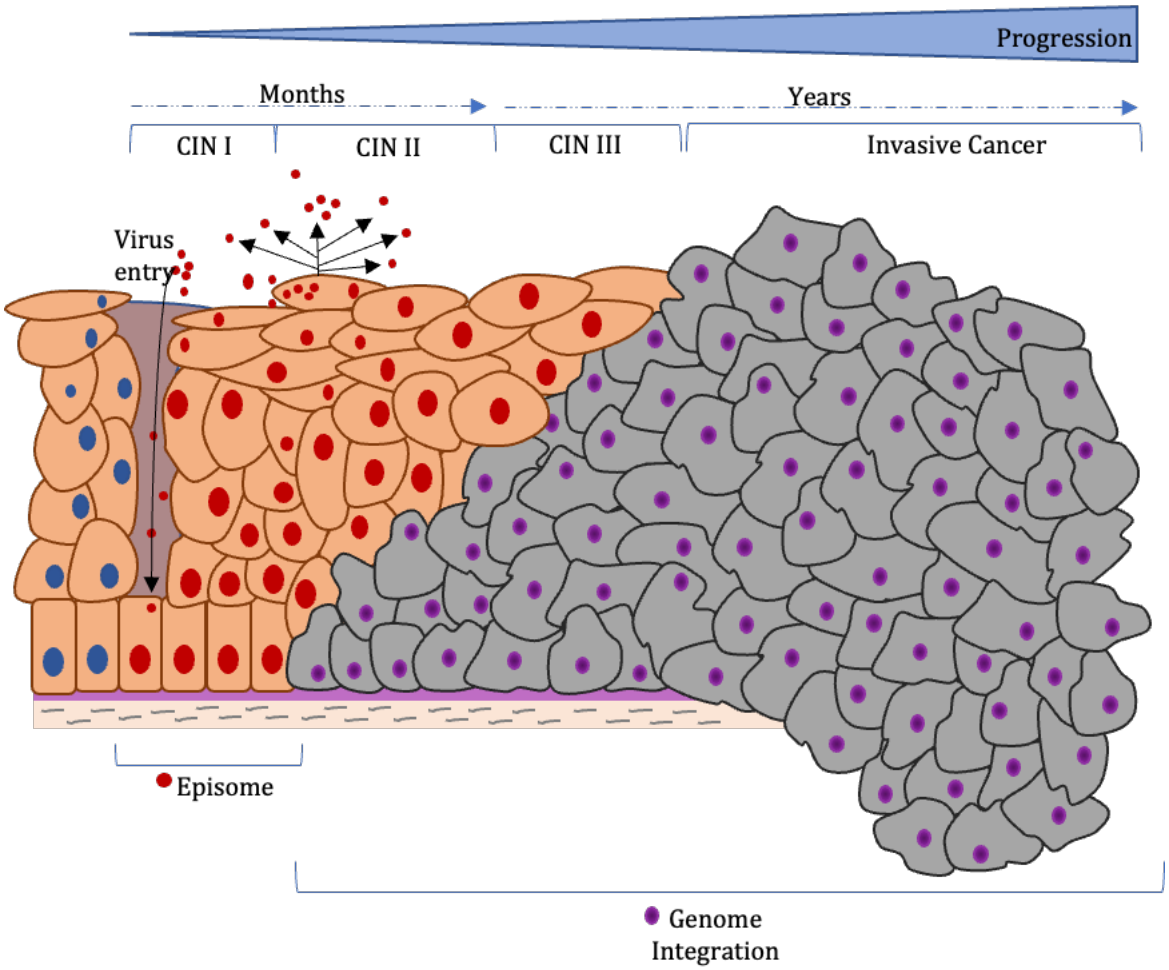



**Figure 1.9.1 Tumour progression following HPV infection**
Adapted from Woodman, et al., 2007. Blue represents normal uninfected cells, red represents cells with episomal viral DNA, and purple represents cells with integrated nuclei. Following initial infection, the virus is maintained in episomic form. High-risk HPV infections progress to cervical intraepithelial neoplasia grade II and III. Untreated lesions progress further to microinvasive and invasive cancer following the integration of the HPV genome into the host.

Integration of the viral genome can be detected in premalignant lesions; however, the detection of integration increases significantly as cells progress to tumourigenesis (Shukla, et al., 2014). Integration is not part of the normal virus life cycle, as it disrupts the production of new complete viral genomes, therefore meaning an end to any further transmission of the virus. The integration event can either be of a single viral genome (Type I), or multiple tandem head-to-tail repeats of the genome (Type 2) (fig 1.9.2 C, D) (Jeon, et al., 1995). The head-to-tail repeats of the genome is a common feature of cervical cancer cell line CaSki (Baker, et al., 1987). Various mechanisms have been proposed as to how the virus uses host DNA and replication machinery to integrate into the host's genome. The most widely accepted mechanism is the "looping" model, where integration of the HPV genome results in viral-host DNA concatemers (Akagi, et al., 2014). This frequently disrupts genes involved in oncogenesis, amplification of oncogenes, and chromosomal rearrangement and/or genome instability. Another study found micro-homology-rich zones between viral and host genome at the integration sites and suggested a microhomology-mediated break-induced replication (MMBIR) as a possible integration mechanism (Hu, et al., 2015).

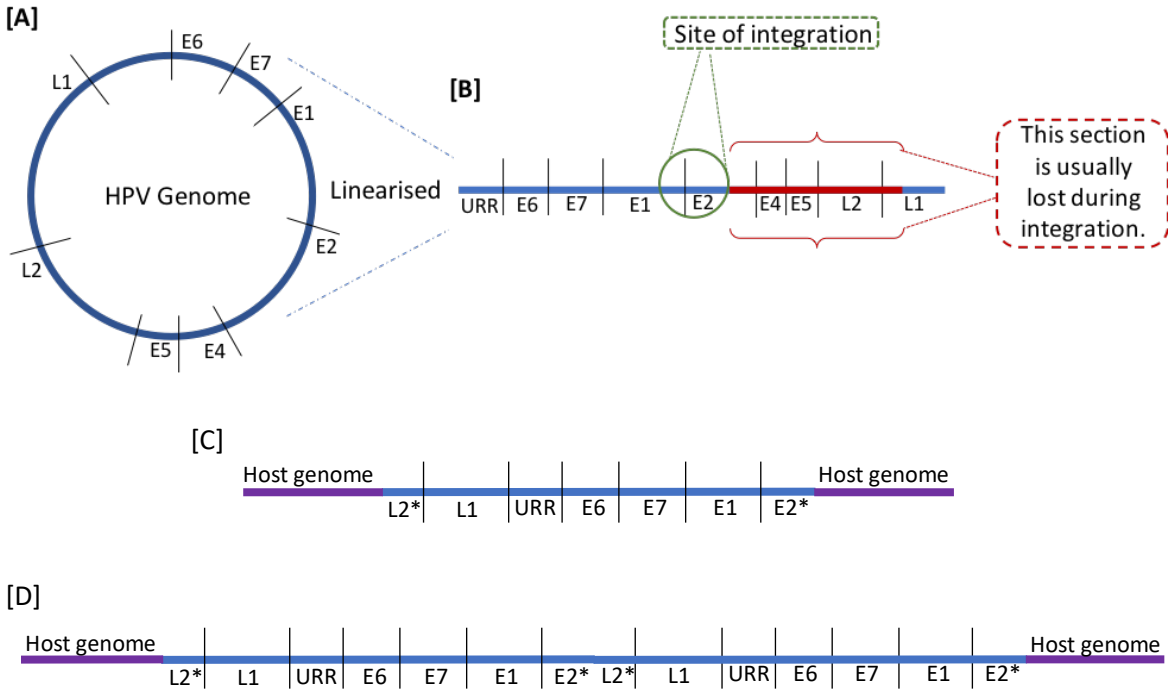



**Figure 1.9.2 Integration of HPV genome**
A) Circular HPV genome. B) Linear HPV genome with the most common site of integration annotated. The red section is usually lost during integration. C) Single viral genome is integrated into the host genome. D) multiple tandem head-to-tail repeats of the viral genome are integrated into the hose genome.

Host genome integration usually results in some level of disruption in the viral DNA. This is random, however, when the integration results in the disruption of the E2 ORF, the cells gain a selective advantage. Loss of the expression of the regulatory protein E2 is a vital step during the neoplastic transformation of HPV-infected cells. E2 functions to control the expression of E6 and E7, therefore disruption of E2 results in the overexpression of E6 and E7. Therefore, when analysing the integration of cancerous lesions, the ORFs of E6 and E7

typically remain unaffected, and disruptions are observed across the E1/E2 region, as well as adjacent ORFs, interrupting the expression of E2, E4, E5, and part of L2 (fig 1.9.2B).

Although the integration event is key, it is not seen across all HPV-associated cancers, as some cancers have been shown to possess extrachromosomal viral DNA instead (Vinokurova, et al., 2008, Parfenov, et al., 2014). In these circumstances, however, the viral genome has been shown to acquire additional mutations or epigenetic alterations to dysregulate E2 function, such as methylation of E2 binding sites, and allow E6/E7 overexpression (Reuschenbach, et al., 2015). The end result of E6/E7 overexpression is increased cellular proliferation and genomic instability, and the acquisition of permanent host DNA damage. Accumulation of sufficient mutations over the years results in the progression towards the formation of tumours.

## 1.10 The role of mRNA Splicing in HPV gene expression

mRNA splicing plays an important role in the life cycle of the HPV virus. At least 20 different mRNAs are known to be expressed from the HPV genome, and this is made possible through extensive alternative splicing (Graham et al., 2017). The E6/E7 polycistronic transcript, in particular, undergoes extensive splicing. The transcript is composed of two introns and three exons, with numerous splice sites that allow the transcript to have the potential to produce at least four different mRNAs (E6, E6*I, E6*II and E6*X) (Chen, et al., 2014). All E6* proteins share the first 44 amino acids of the full-

length E6, with truncations within E6 or frame shits into the E7 ORF making up the remaining regions. Splicing plays a key role in driving HPV-associated cancers as it determines the equilibrium of E6/E7 within cells, thereby controlling their activity levels. Without any splicing, the E6 and E7 coding regions are separated by only two nucleotides, which is not sufficient for successful assembly of the ribosome to perform translation of E7. E6*I is known to be the most abundant isoform in cervical cancer cells and is thought to allow for the translation of E7 to occur (Tang, et al, 2006). Splicing leading to the translation of E6*I results in premature termination of the E6 coding region and provides a gap of 144bp between E6*I and E7 ORFs. This is sufficient for successful ribosomal assembly and translation of E7.

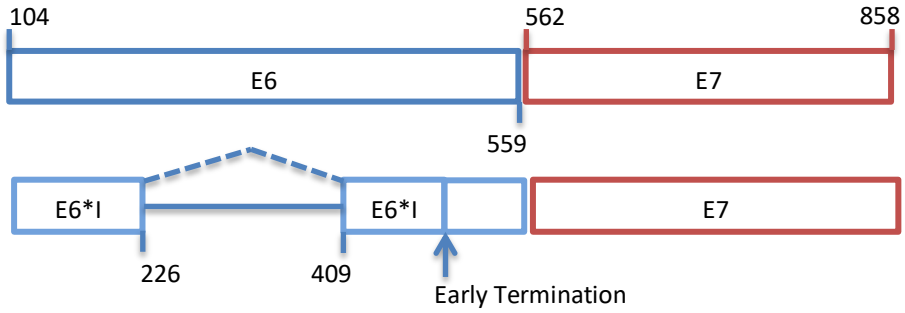



**Figure 1.10.1 Alternative splicing of HPV16 E6/E7 pre-mRNA**
Splice donor site at position 226 and the splice acceptor site at position 409 results in a frameshift and premature stop codon to create E6*I. This allows translation of E7 to follow, with a gap of 153 nucleotides between E6*I and E7.

## 1.10.1 Splicing

Two types of splicing can occur: constitutive and alternative splicing (Figure 1.10.2). Constitutive splicing involves the removal of all introns from pre-mRNAs and the joining of

exons to create mRNA. Alternative splicing, on the other hand, rises from the existence of multiple 5'- and 3'- splice sites within the pre-mRNA. This allows different isoforms of the mature mRNA to be created as a result of intron and exon retention, exon skipping or the use of pseudo splice sites to alter exon sizes. These isoforms can encode proteins with varying characteristics and different functions. The mechanism of alternative splicing is used by viral and mammalian genomes to maximise the production of different proteins from single RNA strands. This explains how a total of 20,000-25,000 human genes can produce well over 100,000 proteins.

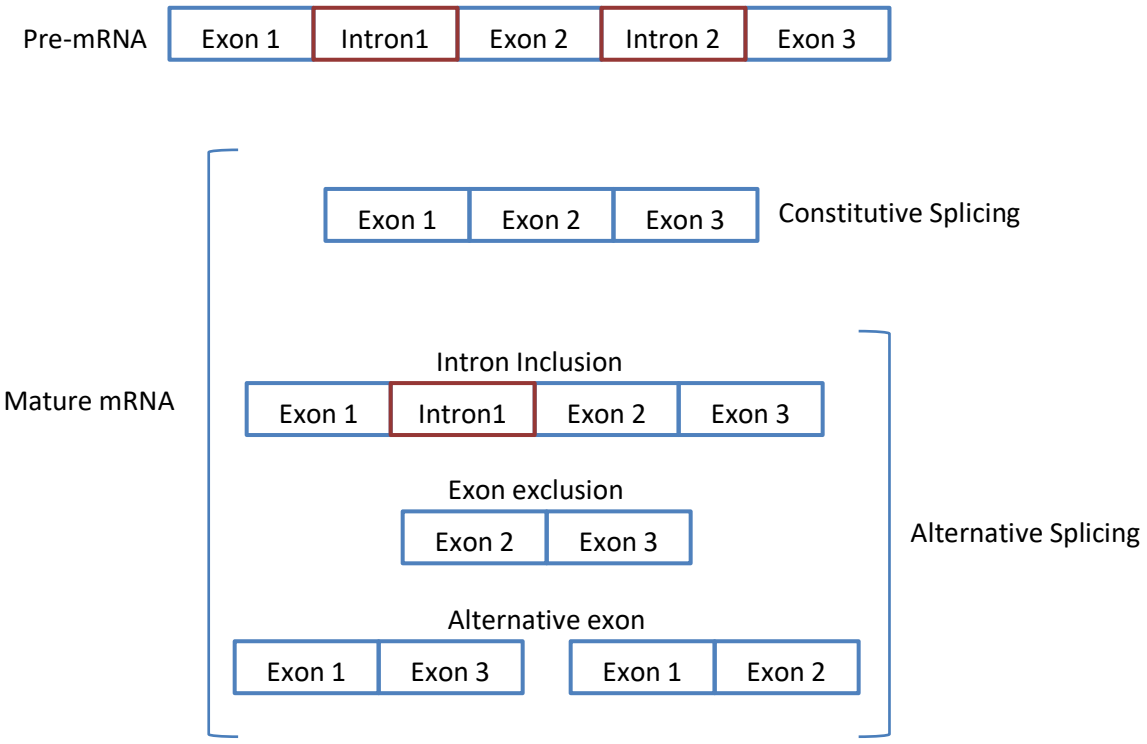



**Figure 1.10.2 Splicing**
Schematic diagram of constitutive splicing and alternative splicing. Introns are in red, and exons are in blue.

A ribonucleoprotein complex known as the spliceosome carries out the splicing. The complex is made of small nuclear riboprotein particles (snRNPs U1, U2, U4, U5, and U6). The initial step of splicing involves the recognition and complementary binding of the 5'-splice site by the U1 snRNP. This is followed by the binding of splicing factor 1 to the branch point, the binding of U2AF65 to the polypyrimidine tract upstream of the 3' splice site, and the binding of U2AF35 to the 5'splice site. This is termed the early complex. Following this, the splicing factor 1 is replaced at the branch point by U2, which is stabilised by U2AF and U4.U5.U6 tri-snRNP. Finally, U1 and U4 are released, allowing U2, U5 and U6 to form the active site for splicing and is known as the C complex. The 5' splice site forms a loop by binding to the branch point, following which the 3'end of the exon cleaves the 5'end of the adjacent exon. This allows the loop structure containing the intron separating the exons to be cleaved out and degraded, following which the two exons join together to form the mature mRNA (Graveley., 2000, Papasaikas, et al., 2016).

### 1.10.2 Control of Splicing

Splicing is controlled by two families of proteins: serine-arginine-rich splicing factors (SR proteins) and heterogeneous nuclear ribonucleoproteins (hnRNP proteins). Early work with SRSF1 and SRSF2 demonstrated the role of SR proteins in the promotion of U1 and U2 binding to splice sites, and they were also considered important for the shift from the early complex to the mature spliceosome (Fu et al., 1992). It is also understood that SR proteins binding to ESEs can influence splice site selection. SR proteins were found to have opposing

activities depending on how they interact with the splice sites, as their binding position was found to determine whether they promoted exon inclusion or exon skipping (Erkelenz, et al., 2013). In addition, early splice complex formation and stabilisation are controlled by SR proteins binding to the ESEs. Furthermore, SR proteins are also implicated to play a role in the recruitment of various splicing factors to the spliceosome and during the catalytic steps of splicing (Zhou, et al., 2013). hnRNP proteins have inhibitory effects on splicing by binding to exonic and intronic splicing suppressors.

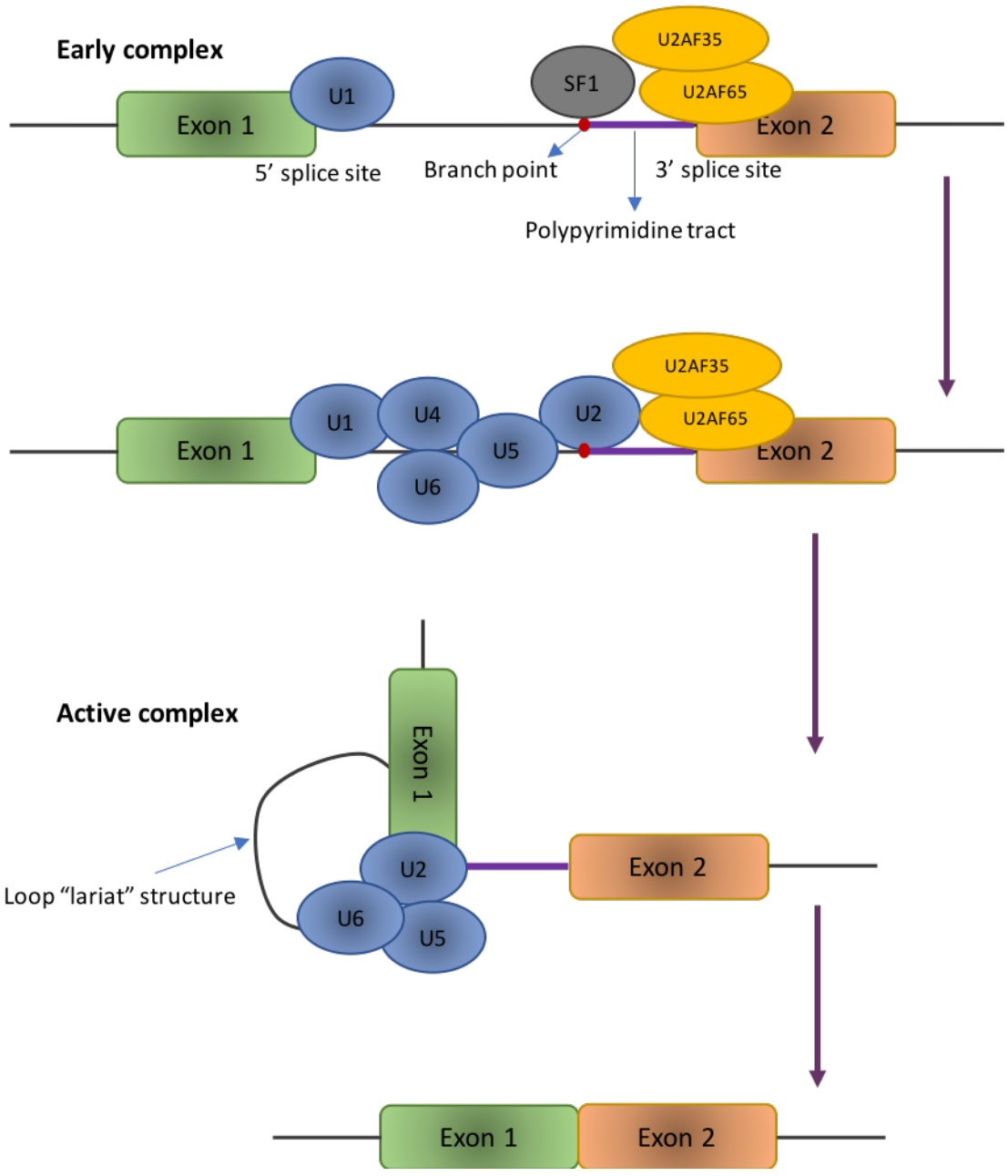



**Figure 1.10.3 Schematic diagram of the splicing mechanism**

### 1.10.3 SR Proteins

In addition to their involvement in constitutive and alternative splicing, SR proteins are also essential for transcription, translation, mRNA stability, and RNA processing (Long, et al., 2009). In total, there are twelve well characterised SR proteins, named SRSF1-12 (Manley, et al., 2010). Inactive forms of these proteins, along with other splice factors are stored in nuclear speckles within the nucleoplasm (Misteli, et al., 1997). They are so-called due to the long stretch of repeated serine/arginine dipeptide repeats (50-300 residue) found in the c-terminal tail. The proteins also contain one or two RNA recognition motifs (RRM) in their N-terminal domain, which are involved in recognising and binding to splicing enhancer sequences on RNA. The Arg-Ser domains undergo extensive phosphorylation, which is known to affect their interaction with other proteins, RNAs, and pre-mRNAs, as well as intracellular localisation and trafficking (Xiao, et al., 1997, Caceres, et al., 1998, Misteli, et al., 1998).

SR proteins are known to undergo three types of posttranslational modifications, including phosphorylation, methylation and acetylation. Phosphorylation of the first RS domain (RS1) triggers their initial import into the nucleus and accumulation into speckles. This event is controlled by SR-protein kinases (SRPKs) (Kuroyanagi, et al., 1998). The RS2 domain is phosphorylated by a family of cdc2-like kinases (CLKs), which dissolves the SR nuclear speckles and redistributes the SR proteins to the site of transcription (Colwill, et al., 1996). It is also known that SR proteins can be phosphorylated by other kinases such as PKC, Akt, DYRKs, and topoisomerase I, however, only SRPKs and CLKs have been directly shown

to phosphorylate them *in vivo* (Zhou, et al., 2013). Unlike a typical phosphorylation reaction, where a kinase dissociates from its substrate post-transfer of a phosphate in a 'hit and run' like fashion, SRPK and CLKs appear to bind strongly to SR proteins, and behave like 'polymerising kinases'. This results in hyperphosphorylated SR proteins, as several phosphates are transferred without dissociation of the kinases (Ghosh, et al., 2011). The importance of SR protein phosphorylation has been extensively studied. Extensive phosphorylation results in a highly negative charge within the RS domains, which prevents any non-specific interactions (Yeakley et al., 1999). Evidence also suggests that SR proteins exhibit distinct functions in relation to their phosphorylation status, as partially phosphorylated SR proteins have been shown to bind target RNA with the strongest affinity (compared to unphosphorylated and fully phosphorylated states) (Cho et al., 2011, Shin et al., 2004). In addition, Cho et al demonstrated that SRSF1 undergoes further phosphorylation after binding to target RNA, where a change in interaction from intermolecular to intramolecular was seen, and the additional phosphorylation step was necessary for promoting early complex formation during splicing. This suggests that most, if not all, SR proteins undergo alterations in their phosphorylation state to function differently during splicing.

A methylation site of SR proteins was identified near the c-terminus, adjacent to the phosphorylation site. It is understood that this stable methylation blocks the phosphorylation of SR proteins, preventing their nuclear import (Yun et al., 2000). In addition, SR proteins are also subject to extensive lysine acetylation. For example,

acetyltransferase Tip60 was found to promote degradation of SRSF2, whereas deacetylase HDAC6 was demonstrated to act as a positive regulator of the protein (Edmond, et al., 2011).

SR proteins have various other non-splicing functions including chromatin remodelling, transcriptional regulation, genome stability maintenance, nucleolar stress, cell cycle progression, and apoptosis control (Li, et al., 2005, Das, et al., 2007, Das, et al, 2014, Loomis, et al., 2009). It is therefore not surprising that SR proteins are overexpressed in a range of cancers (Stickeler, et al., 1999, Jia, et al., 2010, Cohen-Eliav, et al., 2013). A study in 2007 found that SRSF1 exhibited oncogenic properties, as overexpression of the gene in nude mice resulted in large tumours (Karni, et al., 2007). The study also found that overexpression of SRSF1 increased the proliferation rate of cells and allowed them to escape apoptosis. Since then, SR proteins have been repeatedly shown to be overexpressed in a wide range of cancers (Das, et al., 2014, Cohen-Eliav, et al., 2013). In addition, increased levels of SRSF1-3 were reported in patients with cervical cancers (Fregoso, et al., 2013). HP16 E2 was also demonstrated to control the expression of SRSF2-4 (Mole, et al., 2009).

More recently, McFarlane et al identified SRSF2 to be the key splicing factor directly involved in the expression of E6/E7 (McFarlane, et al., 2015). They demonstrated that SRSF2 depletion led to a significant reduction in the levels of E6/E7 mRNAs, resulting in reduced E7 protein expression. In addition to increased p53 levels, which would indicate a loss in E6 levels, they found a significant reduction in cell proliferation and induction of apoptosis in

the tumour cells. SRSF2 was demonstrated to be associated with a large number of gene promoters via the 7SK complex where its switch from the 7SK non-coding RNA to nascent RNA induces the relocation of a key RNA Polymerase II (RNA POL II) C-terminal domain kinase P-TEFb from the 7Sk complex to the RNA Pol II complex (Lin, et al., 2008, Ji, et al., 2013). RNA Pol II is frequently paused at promoter sites, and the recruitment of the P-TEFb cyclin-dependent kinase via the 7SK complex, along with other transcription elongation factors, is essential for productive elongation. A study comparing this function across seven SR proteins found SRSF2 to be the only one to possess this unique involvement in elongation control (Mo, et al., 2013). It is likely that SRSF2 is involved in the control of E6/E7 via splicing control and/or transcriptional control via RNA Pol II pause-release switch.

## 1.11 CDC2-like kinases

The CLK family is a group of CDC2-like kinases that function with dual-specificity. They possess the ability to phosphorylate their substrates on serine/threonine residues, and also undergo autophosphorylation on tyrosine residues. The first member of the CLK family was discovered in 1991 (Ben-David, et al., 1991), and other members shortly after (Hanes, et al., 1994, Nayler, et al., 1997). Until the discovery of CLKs, only two major classes of kinases were known, serine/threonine kinases and tyrosine kinases. CLK1 was the first kinase to be categorised into a new third class of kinases, known as dual-specificity kinases. The CLK family all share a highly conserved unique motif (EHLAMMERILG) within the catalytic domain and therefore are also referred to as LAMMER kinases (Yun et al., 1994). All

members of the CLK family can be found in most tissues localised to the nucleus, with their expression levels varying across tissue types (Hanes, et al., 1994, Lew, et al., 1995). The kinase family is known to undergo alternative splicing to produce a full-length catalytically active form, and a truncated, kinase-deficient form (Duncan, et al., 1995).

CLK1 was first shown to interact with serine-arginine rich (SR) proteins in 1991 (Howell, et al., 1991), which pioneered the possible idea of the CLK family being involved in the regulation of alternative splicing. SR proteins, as mentioned previously, are essential splicing factors involved in the regulation of RNA splicing. The interaction between the CLK family and SR proteins was confirmed in studies that demonstrated the phosphorylation ability of CLKs on SRSF1-3 (Colwill, et al., 1996, Nayler, et al., 1997, Bullock, et al., 2009). CLKs are involved in the phosphorylation of the RS2 domain of SR proteins, which dissolves the SR nuclear speckles and redistributes the proteins to the site of transcription (Colwill, et al., 1996). Similar to SRPKs, CLKs phosphorylate Arg-Ser repeats of SR proteins, however, unlike SRPKs, CLKs also phosphorylate Ser-Pro dipeptides (Aubol, et al., 2013). It is also interesting that despite not possessing a docking grove within the kinase domain-like SRPKs, CLKs are still able to bind with high affinity to SR proteins, and have been demonstrated to efficiently phosphorylate up to 18 serines within the SRSF1 RS domain (Velazquez-Dones, et al., 2005). Additional work by Aubol et al found the N-terminus of CLK1 to be responsible for inducing the hyperphosphorylation of SR proteins, as well as inducing the binding of SRSF1 to RNA (Aubol et al., 2014). Additionally, CLK1 inhibition has been shown to directly inhibit phosphorylation and dissociation of SR nuclear speckles, as well as inhibition of SRSF-

1 dependent splicing (Muraki, et al., 2004, Araki, et al., 2015). All CLKs have a similar N-terminus that resembles the RS domain of SR proteins; therefore it is likely that other members of the CLK family function in a similar fashion to phosphorylate SR proteins.

This idea of the CLK family being involved in alternative splicing gathered further momentum in 1998 when a study identified a role for CLK2 and CLK3 in the alternative splicing of viral genes (Duncan, et al., 1998). An adenovirus reporter construct was found to produce a significantly greater quantity of 9S RNAs in the presence of either kinase, compared to the favoured 12S and 13S RNAs in their absence. The co-transfection of either one of the kinases with the construct seemed to result in a shift in the splice donor site to a more distal one, resulting in the changes observed with the RNA products. Furthermore, CLK1 was identified to regulate the alternative splicing of Influenza A M1 pre-mRNA into M2 mRNA (Zu, et al., 2015).

### 1.11.1 CLKs and cancer

The significance of abnormal splicing in the development of cancers has been demonstrated recently (El Marabti, et al., 2018, Escobar-Hoyos, et al., 2019). Kinases involved in splicing, in particular, have become an attractive target for targeted therapy (Corkery, et al., 2015). Dysregulation of splicing kinases has been demonstrated to play a key role in tumourigenesis. CLK1 inhibition and siRNA knockdown was found to affect cellular viability, invasion and migration in gastric cancer (Babu, et al., 2020). Suppression of CLK2 in some breast cancer cells have shown decreased tumour growth in murine

xenograft models (Yoshida, et al., 2015). Additionally, loss of CLK2 has also led to decreased tumour growth and prolonged survival in glioblastoma xenografts (Park, et al., 2016). More recently, a study reported the significance of CLK3 in cholangiocarcinoma, demonstrating the inhibition of cancer progression by targeting a CLK3-dependent c-MYC pathway to halt purine metabolism (Zhou, et al., 2020).

Various investigations confirming the involvement of CLKs in several cancers have led to the recent development of successful small-molecule inhibitors (Moyano, et al., 2020). SM08502, targeting CLKs, has recently entered Phase I clinical trials for patients with advanced solid tumours (NCT03355066). T-025, another orally administered inhibitor of CLKs, has been suggested as a novel therapy for MYC-driven breast cancer patients (Iwai, et al., 2018). CC-671, selectively targeting TTK and CLK2, is currently considered for clinical trials against triple-negative breast cancer patients following evidence of significant efficacy in xenograft models (Zhu, et al, 2018).

### 1.11.2 CLK3

CLK3, unlike CLK1, has been studied less extensively. Two transcript variants of CLK3 have been identified. Variant 1 (NM_001130028.2) consists of 2713bp and is predicted to encode a protein of 638 amino acids in length, with a molecular weight of 73.5kDa. There has been no evidence of variant 1 being successfully translated in cells. Variant 2 (NM_003992.5) consists of 1846bp and differs in the 5’ UTR and coding region compared to

variant 1. Variant 2 encodes a catalytically active protein of 490 amino acids in length with a molecular weight of 58.6kDa. Variant 2 also encodes a truncated, catalytically inactive form of CLK3, with a deletion of 97bp that results in a frameshift and the loss of 338 amino acids. Like other family members, CLK3 has been shown to phosphorylate SR proteins (Duncan, et al., 1998).

## 1.12 Protein Kinase Inhibitors as Targeted Cancer Therapeutics

Frequently, cancers become exquisitely dependent on certain pathways and therefore on the specific proteins for their survival; a phenomenon famously described by Bernard Weinstein as oncogene addiction (Weinstein and Joe, 2008). The development of cancer often involves somatic mutations that lead either directly or indirectly to dysregulation of protein kinases and these have proven amenable to targeting with selective inhibitors for cancer therapy. The c-ABL inhibitor Imatinib was the first small molecule to be approved by the FDA as a kinase inhibitor for the treatment of chronic myelogenous leukaemia (CML, which is driven by oncogenic BCR-ABL fusion protein formed from the Philadelphia chromosome translocation) in 2001 (Carroll, et al., 1997). The significant reduction in CML mortality led to the development of a flurry of kinase inhibitors against oncogenic kinases. 89 kinase inhibitors have received FDA approval for cancer treatment, with many more currently in clinical trials (Bhullar, et al., 2018, Center for Drug Evaluation and research, 2021, MRC PPU online) Further examples from the clinic include BRAF inhibitors (vermurafenib and dabrafenib) for BRAF-mutant melanoma (Hauschild, et

al., 2012, McArthur, et al., 2014) and EGFR inhibitors (erlotinib, gefitinib, afatinib) for *EGFR*-mutant non-small cell lung cancer (Maemondo, et al., 2010, Zhou, et al., 2011, Sequist, et al., 2013). Importantly, the kinase targets of inhibitors are not always mutationally activated themselves, as evidenced by the successful deployment of CDK4/6 inhibitors for the treatment of $ER^+$ breast cancer (Shah, et al., 2019). A previous PhD student in our group (Dr Luke Williams) hypothesised that HPV-transformed cells may become dependent upon one or more kinases and set out to identify such dependency through collaboration with the group of Prof Chris Lord (Institute of Cancer Research).

## 1.13 HPV-positive cancer cell lines display a dependence on CLK3 for viability and HPV oncogene expression

A small interfering RNA (siRNA) screening of the human kinome and selected kinase-associated for DNA repair genes in 135 cancer cell lines (Cambell et al., 2016) included 6 HPV-positive lines (three cervical and three HNSCC) provided by our group. Analysis of the screen z-scores to identify genes with a differential effect on the viability of HPV-positive lines when compared with all other lines in the screen (Figure 1.13.1) revealed a small set of genes that were taken forward for validation by Luke Wiliams, including CLK3. During initial validation, siRNA knockdown of CLK3 was carried out in seven HPV positive cancer cell lines (Williams & Fenton, unpublished). Six of the cell lines (SiHa, CaSki, 93VU147T, HeLa, SCC047, SCC154) showed a significant loss of cell viability with at least two siGENOME® siRNAs (Thermo Fisher) targeting different regions of the kinase and similar

data were obtained with ON-TARGET Plus® siRNAs (Thermo Fisher), which are modified to enhance specificity. The initial observations were further validated with the use of stable cell lines transduced with tetracycline-inducible CLK3 shRNAs.

Upon further investigation, knockdown of CLK3 led to a significant loss in HPV E7 protein and E6/E7 mRNA levels (figure 1.13.2), with nascent RNA labelling revealing an effect on the rate of E6/E7 mRNA transcription rather than on E6/E7 mRNA stability. It is well documented that loss of E7 leads to cell growth arrest and/or apoptosis (Steele, et al., 1992, Von Knebel Doeberitz, et al., 1992). As expected, the E6/E7 loss was accompanied by increased levels of pRb and p53 and induction of senescence following sustained cell cycle arrest.

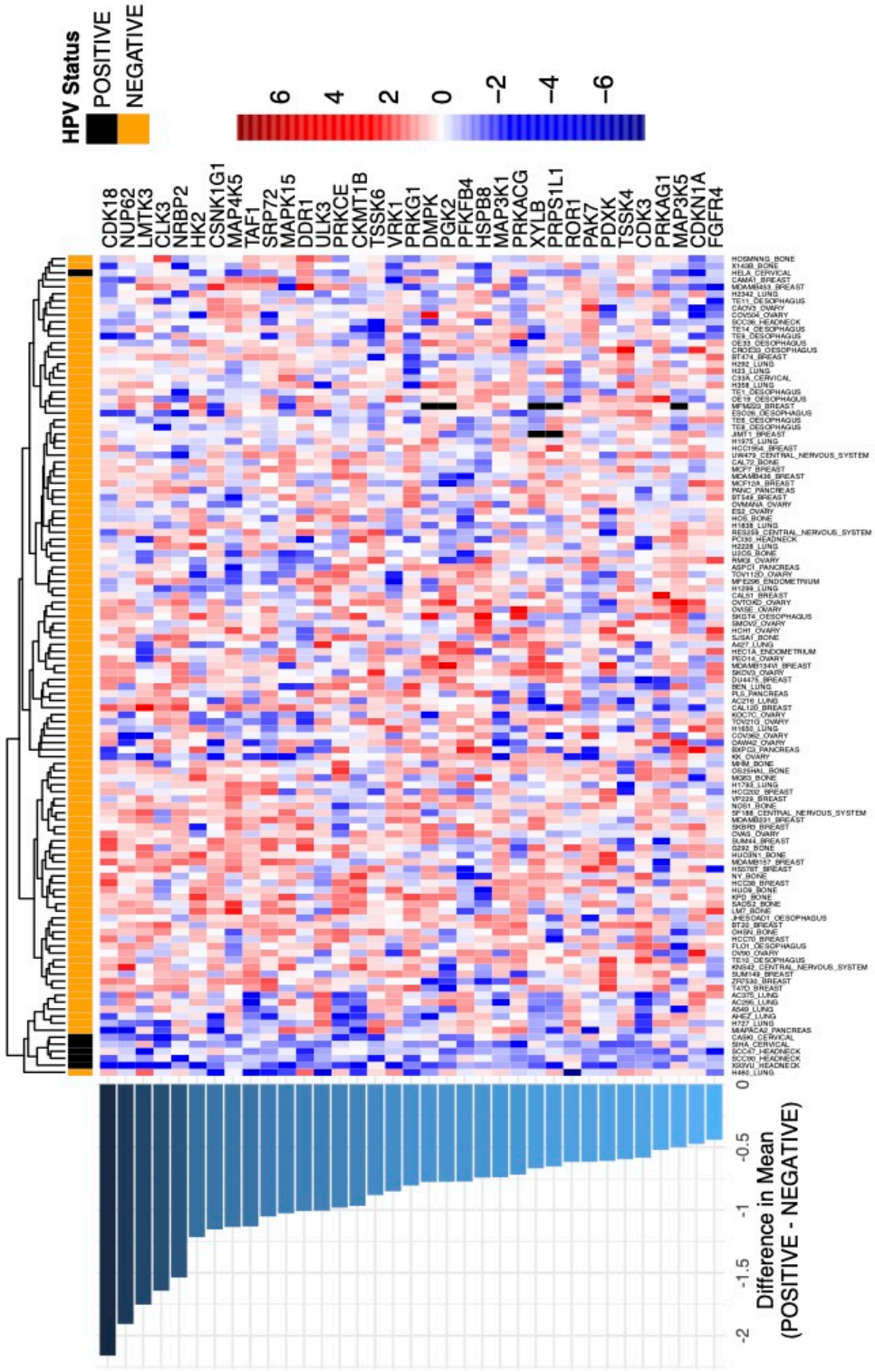



**Figure 1.13.1 Z-score from 136 cell line panel of the siRNA screen.**
HPV positive cells are represented in black. CLK3 was among the genes that had the largest difference in z-score in HPV positive cell lines compared to other HPV negative cell lines.

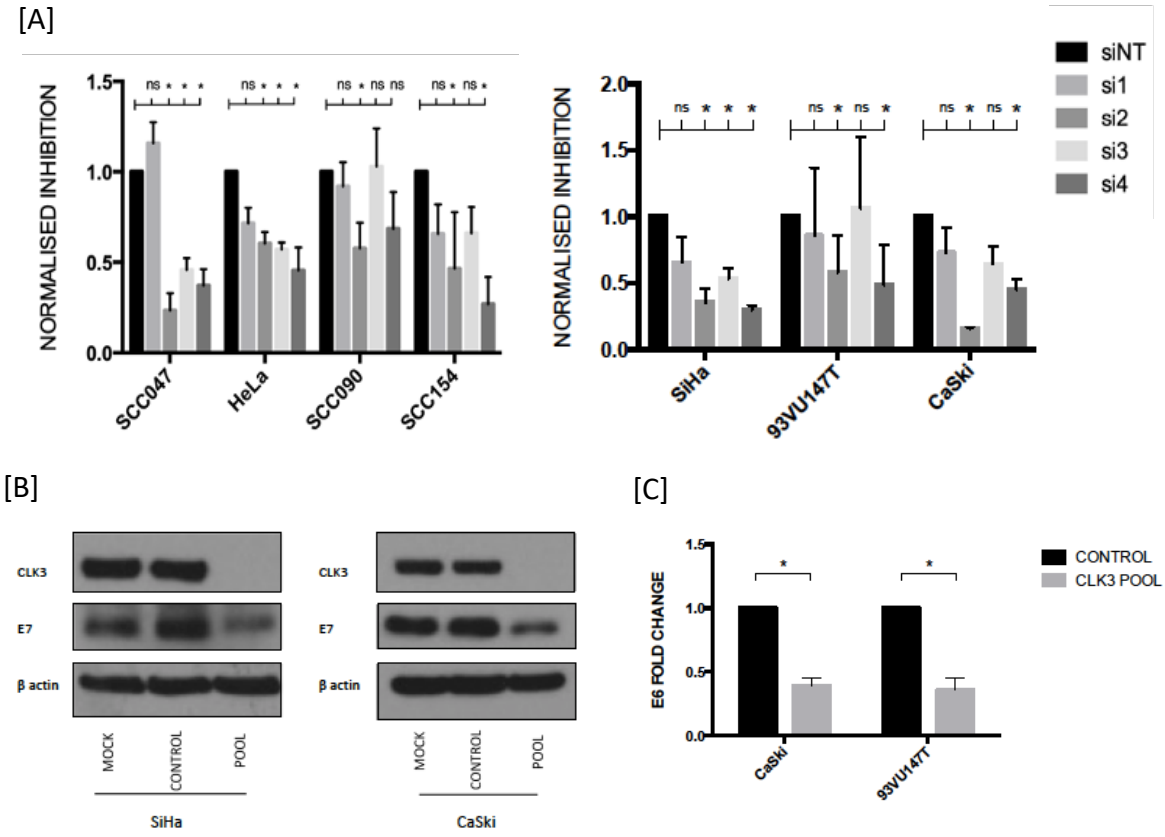



**Figure 1.13.2 CLK3 knockdown in HPV positive cancer cells.**
A) Viability of seven HPV positive cell lines following transfection of siGENOME siRNA to CLK3 five days post transfection. B) Western blot following transfection of SMARTpool siRNAs against CLK3 to determine E7 effect in SiHa and CaSki cell lines. MOCK is transfection of siRNA, CONTROL is transfection of non-targeting control siRNA, and POOL is transfection of SMARTpool siRNAs targeting CLK3. Molecular weight of CLK3 is approximately 59kDa. E7 is approximately 18kDa in molecular weight. β-Actin was used as a loading control. C) qRT PCR of E6/E7 mRNA levels from three independent experiments. CONTROL is non-targeting control siRNA, CLK3 POOL is SMARTpool of siRNA targeting CLK3 (Williams, 2021).

## 1.14 CLK3 and c-MYC

Since the discovery of CLKs, the main known function of the kinases has been to phosphorylate and redistribute SR proteins within the nucleus to the site of transcription. More recently, however, CLK3 was identified to play a significant role in controlling the stability of c-MYC in cholangiocarcinoma cells (Zhou, et al., 2020). The study identified USP13 as a new substrate of CLK3. WT-CLK3 was found to directly phosphorylate USP13 at position 708 (tyrosine). Furthermore, this phosphorylation event was found to promote the binding of USP13 to c-MYC, which prevented the Fbxl14-mediated c-MYC ubiquitination and thereby degradation.

c-MYC is an oncoprotein widely overexpressed in various cancers and contributes to an estimated 40% of tumours (Dang, et al., 2010). c-MYC is also known to be upregulated in many cervical cancers. It is a transcription factor responsible for the expression of various genes necessary for proliferation, and repression of antiproliferative genes. It is also known that c-MYC regulated genes are involved in almost every essential cellular function. A study by Peter and collegues demonstrated that HPV integration often takes place within the MYC locus (chromosome band 8q24), which is why the oncogene expression is often altered in cervical cancers (Peter, et al., 2006).

c-MYC forms a heterodimer complex with MAX to recognise enhancer box (E-box) DNA elements ‘CANNTG’. The complex is known to have the highest affinity for the E-box sequence CACGTG (Blackwood, et al., 1991, Blackwell, et al., 1990, Blackwell, et al., 1993).

Studies have also found that the affinity of MYC/MAX complex to the E-box can be influenced by nucleotides immediately flanking the core hexamer (Perna, et al., 2012). To date, thousands of c-MYC target genes have been identified by various screening studies including SAGE, DNA microarray, and subtractive hybridisation (Coller, et al., 2000, Lewis, et al., 1997, Menssen, et al., 2002). Recent studies have found that in addition to recognising specific E-box sequences, c-MYC also accumulates in the promoter regions of active genes, resulting in transcriptional amplification (Nie, et al., 2012, Lin, et al., 2012). Furthermore, c-MYC has also been implicated in playing a key role in RNA poly II pause release at its target genes, by the recruitment of the positive transcription elongation factor (P-TEFb (Rahl, et al., 2010). It was therefore hypothesised that CLK3 could control the levels of E6/E7 mRNA by stabilising c-MYC, which could then act on the promoter region of E6/E7 to increase transcription rate or be involved in the elongation step by controlling RNA Pol II pause release.

## 1.15 Objectives

The main focus of this PhD was to further understand the relationship between CLK3 and E6/E7, and strengthen the initial findings. This study will use various HPV-associated cancer cell lines, as well as near-normal immortalised human keratinocyte (NIKS) to further understand the importance of CLK3 in HPV16 E6/E7 expression, and to determine any other key proteins involved.

Key aims:

i. To utilise NIKS, which have undergone CRISPR/Cas9 editing to create CLK3 knockout (CLK3 KO) NIKS, to further strengthen the relationship between CLK3 and HPV16 E6/E7.

ii. To determine how CLK3 interacts with E6/E7, focusing on two known substrates of the kinase, SRSF2 and c-MYC. This will also involve further validating the functional role CLK3 plays with SRSF2 and c-MYC.

**iii.** To investigate other unknown substrates of CLK3 by utilising Shokat's Analogue-Sensitive Kinase Allele (ASKA) technology.

# Chapter 2

# Materials and Methods

# 2 Materials and methods

## 2.1 Cell Culture

### 2.1.1 Cell Lines

| | |
|---|---|
| HeLa | HPV-related endocervical adenocarcinoma from a 30yr old female. |
| CaSki | HPV-related cervical squamous cell carcinoma, derived from metastatic site: small intestine from a 40yr old female. |
| 93Vu-147T | HPV-related floor of mouth squamous cell carcinoma from a 58yr old male. |
| UMSCC104 | HPV-related floor of mouth squamous cell carcinoma from a 56yr old male. |
| UMSCC047 | HPV-related tongue squamous cell carcinoma from a 53yr old male. |
| C33A | HPV-negative cervical squamous cell carcinoma from a 66yr old female. |
| SiHa | HPV-related cervical cell carcinoma from a 55yr old female. |
| NIKS | A spontaneously immortalized keratinocyte cell line, isolated from the BC-1-Ep strain of normal human neonatal foreskin keratinocytes (Allen-Hoffmann et al, 2000). |
| 3T3-J2 | A subclone of the original 3T3-Swiss Albino fibroblast cell line. |
| HEK 293T | Highly transfectable derivative of human embryonic kidney 293 cells. |

### 2.1.2 Materials

| | |
|---|---|
| Ham's F12 Medium | PAN Biotech (#P04-14500) |
| DMEM (4.5g/l glucose) | PAN Biotech (#P04-04510) |
| DMEM, No Glucose | ThermoFisher (#11966025) |
| Trypsin-EDTA (0.05%) | ThermoFisher (#25300062) |
| PBS tablets | OXOD (#BR0014G) |
| FBS, Supreme | Pan Biotech (#P30-3031) |
| EGF | Sigma (#A8626-1G) |
| Hydrocortisone | Sigma (#H088801G) |
| Adenine | Sigma (#A8626-1G) |
| Earle's Balanced salt solution 10X | Sigma (#E7510) |
| Cholera toxin | Merck Millipore (#227036) |
| Insulin from bovine pancreas | Sigma (#16634-50MG) |
| Penicillin/Streptomycin | PAN Biotech (#P06-07100) |
| G418 Sulfate | Agilent (#200049) |
| FuGENE HD Transfection Reagent | Promega (#E2311) |
| Mitomycin C from Streptomyces caespitosus | Sigma (#M4287-2MG) |
| Cobalt (II) chloride hexahydrate | Sigma (#C8661-25G) |
| Deferoxamine mesylate salt | Sigma (#D9533-1G0 |
| Gibco™ Opti-MEM™ I Reduced Serum Medium | ThermoFisher (#31985062) |
| Doxycycline hyclate | Alfa Aesar (#J60579.14) |
| DMSO | Fisher Scientific (#10103483) |
| Cycloheximide, 95% | FisherSci (#15214714) |
| Pierce Firefly Luciferase Glow Assay Kit | ThermoFisher (#16176) |

| | |
|---|---|
| Lipofectamine 3000 | ThermoFisher (#L3000008) |

### 2.1.3 Stock solutions and media

| | |
|---|---|
| 100X Adenine | Dissolved 121mg adenine in 50ml 0.05N HCl by stirring for 1hr. Stored at -20$^{o}$C in 5ml aliquots. |
| 100X Cholera toxin | Add 1.2ml sterile $H_2O$ to 1mg vial. Dilute 50µl in 50ml HBES containing 0.1% BSA. Stored at 4$^{o}$C. |
| 100X Insulin | Dissolved 10mg in 20ml 0.05N HCl. Stored at -20$^{o}$C. |
| 100X Hydrocortisone | Dissolved 25mg in 5ml of cold 100% Ethanol to make 5mg/ml stock solution. 0.8ml of stock solution was added to 100ml HBES with 5% FBS. Stored at -20$^{o}$C in 5ml aliquots. |
| 100X EGF | Dissolved 100µg vial in 10ml of sterile $H_2O$. Added to 90ml HBES containing 0.1% BSA. Stored at -20$^{o}$C in 5ml aliquots. |
| HBES | Added 50ml Earle's 10X to less than 400ml of $dH_2O$, and 11g of $NaHCO_3$ was dissolved in the solution. Added 12.5ml of 1M Hepes buffer and $dH_2O$ was used to bring the total volume up to 500ml. Stored at 4$^{o}$C. |
| FC Medium | 67.6% Ham's F12 medium, 22.4% DMEM, 5% FBS Supreme, 1% 100X adenine, 1% 100X cholera toxin, 1% 100X insulin, 1% 100X hydrocortisone, 1% Pen-Strep. Filtered using a 500ml 0.22µm vacuum bottle filter. Stored at 4$^{o}$C. |
| Culture Medium | 10% FBS supreme and 1% pen-Strep were added to 500ml DMEM. Stored at 4$^{o}$C. |
| Freezing media | 10% DMSO added to Culture medium. Made fresh every time prior to cell freezing. |
| 1X PBS | Dissolved 1 tablet of PBS in 100ml of sterile $H_2O$. Autoclaved and stored at 4$^{o}$C. |

All cancer cells were grown in DMEM with 10% bovine calf serum. NIKS require a layer of mitomycin C treated 3T3-J2 cells to act as feeder cells to aid in their growth. NIKS were plated 3-4hrs after seeding the feeder cells, and cultured in FC medium. All cells were incubated in a humidified incubator at 37$^{o}$C, with 5-10% $CO_2$, and media was replaced every 48-72hrs.

### 2.1.4 Mitomycin C treatment

3T3-J2 cells were grown to 80% confluency, and fresh medium containing 2% Mitomycin C was added and incubated for 3.5-4hrs. The medium was removed and cells were washed with prewarmed 1X PBS. They were then trypsinised with prewarmed 0.05% trypsin-EDTA and pelleted. They were resuspended in freezing media at $2x10^6$ cells per ml and stored at -80$^{o}$C for 48-72hrs, prior to being transferred into nitrogen storage.

### 2.1.5 Passaging Cells

Cells were normally passaged before they reached a fully confluent state to avoid senescence or apoptosis. All cells were washed with prewarmed 1X PBS and then typsinised with prewarmed 0.05% trypsin-EDTA. They require varying degree of incubation depending on the cell type. When passaging the NIKS, the feeder cells were removed by two separate events of trypsinisation. Feeder cells adhere weakly to the surface, and therefore were separated from the NIKS with a short exposure to trypsin. Cells were pelleted at 300G in a centrifuge for 5 minutes. The medium was then aspirated and the pellet was re-suspended

in fresh media. A haemocytometer was used to calculate the cells per ml, and re-seeded accordingly to the required density.

### 2.1.6 Long term storage of Cells

After pelleting cells, they were resuspended in freezing media and stored in cryovials, suitable for liquid nitrogen dewars, at $1x10^6$ cells per ml. They cells were placed in polystyrene boxes at -80$^oC$ for a few days, before being transferred into liquid nitrogen storage. When resuscitating the cells, the frozen cryovials were thawed out quickly in a water bath at 37$^oC$ and the cell suspension was added to 5ml of prewarmed medium. Cells were then pelleted and the medium was removed, following which cells were resuspended in fresh media and plated into 25$cm^2$ flasks.

### 2.1.7 Preparing cell lysates for RNA and Protein extraction

Medium from the flask was removed and cells were washed with cold 1X PBS. Prewarmed 0.05% trypsin-EDTA was used to trypsinise the cells and they were pelleted at 300G. Excess medium was aspirated out and the cell pellet was snap frozen in dry ice and stored at -80$^oC$ until RNA or protein extraction.

### 2.1.8 Transfection

In preparation of transfection, cells were plated at $4x10^5$ in 6 well plates. On day two, transfection was performed using the FuGENE HD transfection reagent. 2µg of the plasmid was used per transfection, with 8µl of FuGENE HD, made up to 200µl with Optimem. The DNA:FuGENE:Optimem mixture was incubated at room temperature for 15

minutes. The cells were washed with PBS and replaced with 1.8ml of media without any penicillin/streptomycin. The DNA:FuGENE:Optimem mixture was added drop by drop to the cells and were incubated at 37$^{o}$C in a humidified incubator with 5-10% $CO_2$. 48hrs post-transfection, the media was removed, and the cells were washed with PBS and the media was replaced. If the plasmid being used has a GFP tag, efficiency of transfection was monitored 24hrs and 48hrs post transfection using a Etaluma Lumascope 720.

### 2.1.9 Antibiotic Selection

G418 Sulfate was used for positive selection of cells expressing the neomycin resistance gene following transfection. Cells were treated with 500ng/ml and the media was changed every 48hrs. Treatment continued until all the cells in the negative control were dead, typically 7-10 days.

Doxycycline hyclate was used to induce shRNA knockdown of CLK3 in SiHa cells, which had been stably integrated with TtRMVIR retroviral vector. The cells were provided by Luke Williams. These cells express shRNA targeting CLK3 in the presence of doxycycline. Cells were treated with 500ng/ml and the media was changed every 48hrs.

### 2.1.10 Luciferase Assay

Initially, the URR region from UM-SCC047 cells was cloned into pGL3 firefly luciferase reporter plasmid. Cells were plated out in 96-well plates 24hrs prior to transfection. For NIKS, cells were plated at a concentration of 5,000 cells per well, along

with 2,000 feeders. For HEK 293T cells, 10,000 cells were plated per well. The following day, the cells were co-transfected with URR-pGL3 firefly and pcDNA-RLuc8 Renilla luciferase control plasmid, at a ratio of 5:1. Empty pGL3 firefly plasmid was used as control.

48hrs post transfection, cells were washed with PBS, and lysed with 25µl of lysis buffer on a shaker at room temperature for 10-15 minutes. 100µl of Luciferase assay reagent II (LAR II) is added into each well of an opaque 96-well plate. Following lysis, using a multichannel pipette, 20µl of cell lysate was added to each well of LAR II and mixed well. The luminescence was recorded on a plate reader to measure firefly luciferase activity, and immediately afterwards, 100µl of prepared Stop & Glo Reagent was added to each well and mixed by pipetting up and down repeatedly using a multichannel pipette. Renilla luciferase activity is measured using the plate reader to account for transfection efficiency.

### 2.1.11 Hypoxic chamber & hypoxic mimetics

The hypoxic chamber was set at 0.1% oxygen and maintained at 37$^{o}$C, and prior to experiment setup, all the media were transferred into the chamber. Cells were spit within the chamber, and kept in a hypoxic chamber for up to 24hrs.

Cobalt chloride was freshly made in media on the day of use, and DFO (Deferoxamine mesylate) stock was made at 10mM stocks and frozen at -20$^{o}$C. Following initial optimisation, 200µM $CoCl_2$ and 150µM DFO were used to treat cells for up to a period of 24hrs. Cells were split 24hrs prior to treatment.

### 2.1.12 siRNA Cell viability assay

The siRNA was purchased from Dharmacon (ON-TARGETplus Non-targeting pool (D-001810-10-05), ON-TARGETplus Human MYC siRNA – SMARTpool (L-003282-02-0005), siGENOME Human PLK1 siRNA – SMARTpool (M-003290-01-0005)). Cell viability assay was carried out via reverse transfection. Each condition was carried out in triplicates in 96-well plates. 0.2µl of Lipofectamine 3000 was added to 19.8µl of Optimem and incubated at room temperature for 5 minutes. 20µl of additional optimum was added to each well. 0.2µM siRNA stocks diluted in 19.8µl of Optimem was added to each well and the mixture was incubated at room temperature for 10-15minutes. Mock controls had siRNA volume replaced with Optimem. Cells were added to each well at 2000 cells were well in 140µl of DMEM with 10% FBS without any antibiotics.

At day 5 post transfection, the transfection media in each well was replaced with 100µl of DMEM with 10% FBS, along with 20µl of CellTiter 96 Aqueous One reagent. Cells were then incubated at 37$^{o}$C for 1 hour and absorbance was read at 490nm using a plate reader.

When determining normalised inhibition values, blanked values from wells without any cells were subtracted from the raw absorbance values. Then, the average absorbance of PLK1 siRNA wells was substracted to account for transfection efficiency. Following this, values for the wells transfected with siRNA targeting c-MYC were divided by the average

values of wells transfected with the non-targeting siRNA control to obtain normalised inhibition relative to non-targeting siRNA.

### 2.1.13 siRNA knockdown

24hrs prior to siRNA transfection, 200,000 cells were plated in each well of a 6-well plate in 2ml of media. On the day of transfection, the media was removed, washed twice with warm PBS, and replaced with 1.5ml of antibiotics free media. 5µl of lipofectamine 3000 was added to 245µl of Optimem in an Eppendorf tube, and 3µl of 20µM stock siRNA was added to 247µl of Optimem in a separate tube. Both tubes were combined and incubated for 10-15minutes, following which the mixture was added drop-by-drop to each well, for a final concentration of 20nM siRNA per well. The cells were incubated for 48-72hrs and then harvested for either protein or RNA extraction for western blot and qRT-PCR analysis, respectively.

## 2.2 RNA

### 2.2.1 Materials

| | |
|---|---|
| RNase free DNase set | Qiagen (#79254) |
| RNeasy mini kit | Qiagen (#74014) |
| Monarch Total RNA Miniprep Kit | NEB (#T2010S) |
| GoScript ™ Reverse Transcription system | Promega (#A5000) |
| LunaScript RT SuperMix Kit | NEB (#E3010L) |
| X10 Powerup SYBR master mix | ThermoFisher (#15340939) |

| | |
|---|---|
| Adhesive Film for qPCR | ThermoFisher (#AB1170) |
| Microamp 96-well rxn plate | ThermoFisher (#N8010560) |

### 2.2.2 Reagents and Stock solutions

| | |
|---|---|
| qPCR master mix | 5µl SYBR green, 0.3µM Forward primer, 0.3µM Reverse primer, 10ng DNA (4µl), made up to 10µl for each qPCR reaction with nuclease-free water. |
| RT PCR Mix (Promega GoScript) | 4µl GoScript Reaction buffer, 2µl $MgCl_2$, 1µl PCR nucleotide mix, 20u recombinant RNasin Ribonuclease inhibitor, 1µl GoScript reverse transcriptase, made up to 15µl for each reaction with nuclease-free water. |
| RT PCR Mix (LunaScript) | 4µl LunaScript RT Supermix, up to 1µg RNA sample, made up to 20µl total volume per reaction with nuclease-free water. |

### 2.2.3 qPCR primers

All primers were purchased through IDT.

| | |
|---|---|
| CLK3 Forward | AGAAGCCAACAGAGCAGTAAG |
| CLK3 Reverse | CCCGGTAGTACTCAAAGAGAATG |
| GAPDH Forward | GTCATCCATGACAACTTTGGTA |
| GAPDH Reverse | GGATGATGTTCTGGAGAGC |
| TBP Forward | CCCATGACTCCCATGACC |
| TBP Reverse | TTTACAACCAAGATTCACTGTGG |
| c-MYC Forward | GCCTGCGATGATTTATACTCAC |
| c-MYC Reverse | AAACAGAGTAAGAGAGCCG |
| GFP Forward | GAACCGCATCGAGCTGAA |
| GFP Reverse | TGCTTGTCGGCCATGATATAG |
| HIF-1a Forward | TCTGGGTTGAAACTCAAGCAACTG |
| HIF-1a Reverse | CAACCGGTTTAAGGACACATTCTG |

| | |
|---|---|
| HPV16 E6/E7 Forward | ACTGCAATGTTTCAGGACCCA |
| HPV16 E6/E7 Reverse | TCAGGACACAGTGGCTTTT |
| p16 Forward | AACACCGCTTCTGCCTTT |
| p16 Reverse | TCATGAAGTCGACAGCTTCC |
| SRSF2 Forward | GTGCTTGGCTGTTTCCTGTTT |
| SRSF2 Reverse | CACTGTATGCTCCGTTATTTATATGC |

### 2.2.4 Extraction and quantification

The cell lysate samples were prepared according to chapter 2.1.7. The RNeasy Mini Kit was used to extract the RNA and the given protocol was strictly followed. An additional DNase digest step was completed to remove any genomic DNA contamination. Later, Monarach Total RNA Miniprep kit was used instead for RNA extraction, as this was considered more efficient and produced RNA of higher yield and quality.

Following extraction, RNA samples were quantified using the NanoDrop 2000. The system was blanked with RNase free water provided with the extraction kit and an average was taken from three readings of the RNA samples. The integrity of the sample was determined by evaluating the A260/280 ratio (compares the absorbance at 260 vs 280nm), which is a measure of contamination by proteins, and the A260/230 ratio (260 vs 230nm), which informs of the presence of any organic contaminants. A ratio of ~ 2.0 is generally accepted as pure for RNA. Samples were stored at -80$^{o}$C.

### 2.2.5 Reverse Transcriptase PCR

RNA samples were kept on ice throughout. The Promega GoScript reverse transcriptase kit was used to synthesise cDNA from the RNA samples. The protocol provided with the kit was followed MultiGene OptiMax thermal cycler from Corning, inc. was used to run the 2 step reactions during the synthesis. The temperatures and cycles required are stated below. Later, the LunaScript RT SuperMix kit was used for RT-PCR, as this was considered more efficient. When possible, the use of 1µg RNA was preferred during RT-PCR, and the newly prepared cDNA was stored at -20$^{o}$C.

### 2.2.6 Quantitative PCR

Mastermix for qPCR was prepared and plated in 96-well plates in a UV sterilising PCR hood. Prior to use, the hood was sterilised with UV for 30mins. Negative controls were included for individual set of primers. cDNA was diluted and 10ng of total cDNA was added to each reaction Plates were spun for 30secs on a bench plate centrifuge and were run on the ThermoFisher QuantStudio 3 RT PCR system. The following temperatures and cycles were used:

50$^{o}$C for 2mins

95$^{o}$C for 10mins

95$^{o}$C for 15secs
60$^{o}$C for 1min
95$^{o}$C for 15secs
60$^{o}$C for 1min
} 40 cycles

95$^{o}$C for 1sec

QuantStudio Design and Analysis software was used to analyse the results. The melt curve data was checked to identify any possible contaminants as they would appear with peaks at different temperatures. Negative control samples' ct values, along with melt curves were analysed to confirm sign of contamination. The data was exported as an excel file and dCt, ddCt and fold expression changes were calculated. Initially GAPDH was used as the housekeeping gene, however, it was found to have altered expression in hypoxia, therefore TBP was preferred.

## 2.3 Plasmid cloning

### 2.3.1 Materials

| | |
|---|---|
| 1kb DNA ladder | NEB (#N3232L) |
| 1Kb Plus DNA ladder | ThermoFisher (#10787018) |
| Ethidium bromide | Sigma (#E1510-10ML) |
| Anza ™ T4 DNA Ligase Master Mix | ThermoFisher (#IVGN2108) |
| 5-alpha Competent *E. coli* | NEB (#C29871) |
| Sodium Chloride | Fisher Scientific (#S/3160/60) |
| Bacto Yeast Extract | BD (#212750) |
| Bacto Tryptone | BD (#211705) |
| Agar Technical | OXOID (#LP0012) |
| Ampicillin | Melford (#A0104) |
| Kanamycin sulfate | Sigma (#60615-5G) |
| HotstarTaq Master Mix Kit | Qiagen (#203443) |
| QIAquick PCR & Gel Cleanup kit | Qiagen (#28704) |

| | |
|---|---|
| Monarch DNA Gel Extraction kit | NEB (#T1020S) |
| KAPA HiFi HotStart ReadyMix (2X) | Roche (#KK2601) |
| Plasmid Mini Kit | Qiagen (#12123) |
| Plasmid Midi Kit | Qiagen (#12143) |
| Plasmid Maxi kit | Qiagen (#12162). |
| Gateway™ LR Clonase™ II Enzyme mix | ThermoFIsher (#11791020) |
| Zero Blunt™ TOPO™ PCR Cloning kit | ThermoFisher (#450245) |
| One Shot™ TOP10 Chemically Competent *E. coli* | ThermoFisher (#C404010) |

### 2.3.2 Reagents and Stock solutions

| | |
|---|---|
| Agarose gel | 1% Agarose, made up in 1X TAE buffer. Microwave until agarose is completely dissolved. Add 0.0025% ethidium bromide (10mg/ml) once it cools down to 50°C. |
| 50X TAE buffer | 2M Tris, 1M glacial acetic acid, 50mM EDTA, made up in $dH_2O$. Stored at room temperature. |
| LB Medium | 1% Tryptone, 1% NaCl, 0.5% yeast extract, made up in $dH_2O$. Autoclaved and stored at 4°C. |
| SOB Medium | 2% Tryptone, 0.5% NaCl, 0.5% yeast extract, 2.5 mM KCl, 10mM NaCl, 10mM $MgCl_2$, 10mM $MgSO_4$ made up in $dH_2O$. Autoclaved and stored at 4°C. |
| SOC Medium | Autoclave SOB medium and then 10 mM $MgCl_2$ and 20mM glucose added once cooled. Stored at 4°C. |
| Agar plate | 1% Tryptone, 1% NaCl, 0.5% yeast extract, 1.2% agar, made up in $dH_2O$. Autoclaved. Antibiotic added after media cools to below 40°C. 20ml of the media is poured into 10cm agar plates under sterile conditions and left to set at room temperature, then stored at 4°C. |

| | |
|---|---|
| Inoue Solution | 55mM $MnCl_2$, 15mM $CaCl_2$, 250mM KCl, 10mM PIPES solution, made up in $dH_2O$. Filter-sterilised. Stored at $4^oC$. |

### 2.3.3 Plasmids

| | |
|---|---|
| pBABE-puro | Retroviral expression vector. Selection marker is puromycin. It was a gift from Hartmut Land & Jay Morgenstern & Bob Weinberg. CO-CLK3 was transfected in. (Addgene plasmid #1764). |
| pcDNA3.1 (+) | Vector for mammalian cell expression. Selection marker is Neomycin, and has ampicillin bacterial resistance. (Addgene plasmid #V790-20). |
| pEGFP-N1 | Vector for mammalian cell expression. Selection marker is Neomycin, and has kanamycin bacterial resistance. It was used to transfect CO-CLK3 in NIKS for overexpression (Addgene plasmid #6085-1). |
| pEN-Tmcs | Entry vector with Tetracycline response element. It has gentamycin bacterial resistance. It was used to clone CO-CLK3 into pSLIK-Hygro. It was a kind gift from Iain Fraser (Addgene plasmid #25751). |
| pSLIK-Hygro | Third generation lentiviral vector for Tetracycline inducible CO-CLK3 expression. Selection marker is hygromycin, and has ampicillin and chloramphenicol bacterial resistance. The plasmid also has the *ccdb* gene. It was a kind gift from Iain Fraser (Addgene plasmid #25737). |
| pET-28a (+) | Plasmid for bacterial expression in DH5alpha. It was used for CLK3 expression in DH5alpha cells. It has kanamycin bacterial resistance. It was modified to have 6xHis tag and FLAG tag. It was a kind gift from Chris Mulligan (Addgene plasmid #69864-3). |
| pFastBacHTB-NotI | Plasmid used as a donor plasmid for making bacmid for CLK3 expression in Sf9 cells. It has ampicillin and gentamycin bacterial resistance. It was a kind gift from Chris Toseland. |

| | |
|---|---|
| pGL3-basic | Luciferase reporter vector. It was used for transfecting the HPV16 URR sequence in to determine CLK3 effect on the promoter of E6/E7. It has ampicillin bacterial resistance It was purchased from Promega (Addgene plasmid #E1751). |
| pcDNA-RLuc8 | Mammalian expression vector expressing the Renilla Luciferase reporter gene. Co-transfected with pGL3-basic as a control to determine transfection efficiency. The plasmid was a gift from Sanjiv Sam Gambhir (Addgene plasmid #87121). |
| pMYC-GFP | Mammalian/bacterial expression plasmid with Ampicillin resistance. It was used to overexpress c-MYC. The plasmid was a gift from Olufunmilayo Olopade (Addgene plasmid #42142). |

### 2.3.4 Restriction Enzymes

| | | |
|---|---|---|
| AscI | 5'…GG*CGCGCC…3'<br>3'…CCGCGC*GG…5' | NEB (#R0558S) |
| SalI | 5'…G*TCGAC…3'<br>3'…CAGCT*G---5' | NEB (#R0138S) |
| NotI | 5'…GC*GGCCGC…3'<br>3'…CGCCGG*CG…5' | NEB (#R0189S) |
| XhoI | 5'…C*TCGAG…3'<br>3'…GAGCT*C…5' | NEB (#R0146S) |
| SpeI | 5'…A*CTAGT…3'<br>3'…TGATC*A…5' | Promega (#R6591) |
| BamHI | 5'…G*GATCC…3'<br>3'…CCTAG*G…5' | Promega (#R6021) |
| HindIII | 5'…A*AGCTT…3'<br>3'…TTCGA*A…5' | Promega (#R6041) |

### 2.3.5 List of Primers

All primers were purchased through IDT.

| | |
|---|---|
| ATATGGCGCGCCCACCACTGTAAAAGATACCGC | Adds AscI site to 5'-CO-CLK3 for cloning into pET-28 (+). |
| CCCTAACTGACACACATTCCAC | Reverse primer for amplifying CO-CLK3 out of pBABE-puro (includes SalI site on 3'). |
| GATCGGATCCATGCACCACTGTAAAAGATACC | Adds BamHI site to 5'-CO-CLK3 for cloning into pFastBacHTB-NotI. |
| GATCGAGTCGACTCACCTGCTAGGATTC | Adds SalI site to 3'-CO-CLK3 for cloning into pFASTBACHB-NotI. |
| CGCGGTCGACACACCATGCACCACTGTAAAAGATACC | Adds SalI site to 5'-CO-CLK3 for cloning into pEGFP-N1. |
| GATCGGATCCCCCCTGCTAGGATTCCTAGAAG | Reverse primer for cloning CO-CLK3 into pEGFP-N1, forms GFP fusion protein on the C-terminus. |
| CTGATGACTAGTACACCATGCACCACTGTAAAAGATACC | Adds SpeI site to 5'-CO-CLK3 for cloning insert into pEN-Tmcs. |
| GAGTCGCGGCCGCTTTACTTGTAC | Amplify CO-CLK3-GFP insert out of pEGFP-N1, including the NotI site. |

### 2.3.6 Restriction Enzyme Digest

5-10 units of the enzyme was used to digest a maximum of 1µg of DNA. The manufacturer's protocol was used to set up the digest, with enzymes from NEB requiring a final reaction volume of 50µl and enzymes from Promega requiring 20µl. Enzymes were

kept on ice and added to the reaction last. The final reaction mixture was mixed well by either pipetting up and down, or by flicking the tube. The mixture was then centrifuged briefly and placed in a 37$^{o}$C water bath. When using NEB enzymes, the reaction was incubated for up to 1hr. All enzymes purchased from Promega required a longer incubation period, up to 4hrs. Enzymes with compatible buffers were used simultaneously to perform double digest of the DNA. Otherwise, sequential digest was performed. Sequential digest was done using the enzyme that required the buffer with the lower NaCl concentration first. The different buffer recipes can be found on the enzyme datasheet. The salt concentration was increased after the initial digest to optimize the buffer for the second enzyme. This was done to increase the efficiency of the digest by avoiding column purification and loss of sample during the process. After the digest was complete, loading dye was added to the reaction mixture and the DNA was separated using electrophoresis using a 1% agarose gel. The gel was used to confirm successful digestion, as well as to separate the DNA according to its size. This step is essential when digesting a plasmid, so the insert and the backbone can be separated.

The cut DNA bands were purified using either the QIAquick PCR & Gel Cleanup kit or the Monarch DNA Gel extraction kit. After purification, the DNA sample was quantified using the NanoDrop 2000. The elution buffer from the purification kit was used as the blank.

### 2.3.7 Ligation

The Anza T4 DNA ligase master mix was used ligate DNA insert into a vector following restriction enzyme digestion. 5µl of the master mix was used, with the final reaction volume being 20µl. 50ng of linearised vector DNA was used with sufficient DNA insert. DNA insert amount was calculated with a 3:1 insert to vector molar ratio. The reaction mixture was incubated at room temperature for 15 minutes. A vector alone ligation was set up each time as a negative control, to show the efficiency of digestion, as well as the level of background colonies post transformation.

### 2.3.8 Transformation

All transformations were done using either NEB DH5-alpha competent *E. coli* or Chemically competent cells produced in the lab using those cells (chapter 2.3.15). Cells were stored at -80$^{o}$C and thawed on ice for 20-30 minutes before use. 5µl of the completed ligation reaction was mixed with 50µl of competent cells and incubated on ice for 30 minutes. The tube was mixed by gentle flicking rather than pipetting, Heat shock of the cells was performed by placing the cells in a 42$^{o}$C water bath for 30secs and then immediately being placed on ice for 2 minutes. 500µl of SOC media was added to the cells and placed in a bacterial shaker at 37$^{o}$C for 45 minutes, shaking at 200rpm. This outgrowth step is not essential if the plasmid contained the gene for ampicillin resistance, but is required for all other plasmids.

Following outgrowth, the cells were spun on a microcentrifuge at 500G for 2 minutes, and 450µl of the media was removed following the spin. The remaining mixture was mixed well and 50µl of the mixture was placed on a 10cm agar plate with the appropriate antibiotic. A sterile glass rod was used to spread the cells across the plate. This was performed under sterile conditions. The efficiency of digestion/ligation and transformation will determine the volume of mixture to be used on the plate. Lower efficiencies require higher volume of mixture. The plate was incubated overnight at 37$^{o}$C.

### 2.3.9 Colony PCR

Successful colonies are confirmed using colony PCR. The plate was initially compared with the vector only control to determine the level of background colonies present. The size of the colonies can also act as an indication towards false positive colonies. Small colonies usually represent false positives. Colony PCR was performed on a large number of colonies (10 per transformed plate) to increase the likelihood of identifying successful colonies, especially when there is increased level of background colonies present. Hotstar Taq DNA polymerase was used for the PCR reaction, with forward and reverse primers designed to amplify part of or the full insert. Reactions were set up according to the manufacturer's protocol with a reaction volume of 10µl. Sterile pipette tips were used to pick colonies, which were then mixed separately into their own PCR reaction mix. The pipette tips were then placed in 15cm falcon tubes containing 5ml of LB broth and appropriate antibiotic. The PCR reaction was completed in a thermocycler with the following temperatures:

| | | | |
|---|---|---|---|
| Initial activation step | 15mins | 95°C | |
| Denaturation | 30secs | 94°C | 30 cycles |
| Annealing | 30secs | 50-68°C | |
| Extension | 1min/kb | 72°C | |
| Final Extension | 10mins | 72°C | |

After PCR, samples were run on a 1% agarose gel to confirm the size of the amplified product. Overnight culture was set up at 37°C in a bacterial shaker using the successful colony in the 15ml falcon tube.

### 2.3.10 Plasmid purification

Plasmid purification was completed using the Qiagen plasmid Mini kit. When a greater yield of plasmid was required, Midi or Maxi kits were used. The manufacturer's protocol was strictly followed. 500µl of the overnight bacterial culture were added to 50% glycerol solution to make glycerol stocks of the bacteria. The glycerol stock was stored at -80°C. Following purification, the concentration of the purified plasmid determined using Nanodrop 2000. Samples were sent, with appropriate primers, to Eurofinsgenomics for sequencing. The data was compared with the known sequence of the insert/vector to identify any mutations and correct orientation and positioning of insert within the plasmid.

### 2.3.11 Site-directed mutagenesis

This PCR based technique was used to introduce specific nucleotide substitutions to a sequence of interest. Primers were designed to insert mutations within specific positions of the gene, and the KAPA HiFi HotStart ReadyMIx PCR kit was used to perform the PCR. The following conditions were used to run the PCR in a thermocycler:

| | | | |
|---|---|---|---|
| Initial activation step | 3mins | 95°C | |
| Denaturation | 20secs | 98°C | } 25 cycles |
| Annealing | 15secs | 60-75°C | |
| Extension | 45secs/kb | 72°C | |
| Final Extension | 1min/kb | 72°C | |

To maintain high fidelity, the number of cycles was kept at 25. Gradient PCR was used to determine the ideal annealing temperature for each set of primers. This was done by adjusting each block in the thermocycler to vary from 60 -72°C. PCR based site-directed mutagenesis involves two step PCR.

### 2.3.12 List of Primers

All primers were purchased through IDT.

| | |
|---|---|
| ATGTGCATCGCCGGCGAGCTGCT | Forward primer for inserting F236G mutation on CO-CLK3. |
| AGCAGCTCGCCGGCGATGCACAT | Reverse primer for inserting F236G mutation on CO-CLK3. |
| ATGTGCATCGCCGCCGAGCTGCT | Forward primer for inserting F236A mutation on CO-CLK3. |

| | |
|---|---|
| AGCAGCTCGGCGGCGATGCACAT | Reverse primer for inserting F236A mutation on CO-CLK3. |

## 2.3.13 Gateway Cloning

Gateway cloning was an alternative technique to traditional enzyme cloning used *In vitro* for the recombination between an entry clone and a destination vector. The method utilises the bacteriophage lambda site-specific recombination system, which facilitates the integration of lambda into the *E. coli.* The cloning involves the use of complementary *att* recombination sites and enzymes that mediate the recombination. pEN-Tmcs plasmid containing the CO-CLK3-GFP insert flanked by *att*L sites was created by traditional cloning. This was the entry plasmid used to transfer the CO-CLK3-GFP into pSLIK-Hygro, a destination vector containing *att*R sites. The cloning step involved mixing 150ng of the entry vector with 150ng of the destination vector in TE buffer, at pH 8.0, and reacted the mixture with 2µl of LR Clonase™ II enzyme mix at 25°C. The incubation period for this reaction is usually one hour, but this particular cloning required overnight incubation as the destination vector was over 10kb in size.

## 2.3.14 TOPO Cloning

Topo cloning was carried out using the PCR Blunt II-TOPO vector, which was part of the Zero Blunt® TOPO™ PCR Cloning kit. The vector contains the *ccdB* gene for positive selection, along with Kanamycin and Invitrogen Zeocin resistance genes for selection in *E. coli.* TOPO cloning was used to clone blunt end PCR products, specifically as part of making plasmids for performing standard curve assays with qPCR. The protocol involved setting up

a reaction with 1µl of the vector, along with fresh PCR product, and 1µl of salt solution, made up to 6µl total volume with sterile $H_2O$. The reaction is incubated at room temperature for 5 minutes, followed by transformation of 2µl of the mixture in One Shot *E. coli* cells. The following primers were used to amplify cDNA prior to TOPO cloning:

| | |
|---|---|
| CLK3 Forward | TAAGCGATACCGCTCCCCTG |
| CLK3 Reverse | GCTCATGGTCAAATGTGGCA |
| SRSF2 Forward | ACTAGGCGCAGTTGTGTAGC |
| SRSF2 Reverse | GTGCTCACTGTATGCTCCGTT |
| HPV16 E6/E7 Forward | ACTGCAATGTTTCAGGACCCA |
| HPV16 E6/E7 Reverse | TCAGGACACAGTGGCTTTT |
| TBP Forward | TTGAGGAAGTTGCTGAGAAGAG |
| TBP Reverse | CAGATAGCAGCACGGTATGAG |
| GAPDH Forward | GTCATCCATGACAACTTTGGTA |
| GAPDH Reverse | GGATGATGTTCTGGAGAGC |

## 2.3.15 Chemically Competent cells

Steaked out glycerol stock of 5-alpha competent cells from NEB onto an LB plate without any antibiotics and incubated overnight at 37$^{o}$C. A single colony was picked and inoculated in 500ml of sterile SOB medium at 18$^{0}$C, shaking at 220rpm until OD reached 0.6. This took approximately 35-40hrs. The culture was then placed on ice for 10 minutes, followed by a 10 minute spin in a pre-chilled centrifuge at 3220g. The medium was drained and the pellet was resuspended in 5ml of ice cold inoue buffer, then an additional 75ml. cells were then placed on ice for 10 minutes and spun down again at 3200g. Following this, the cells were resuspended in 20ml inoue buffer containing 7% DMSO, and incubated on ice for a further 10 minutes. At this stage, they were aliquoted at 100µl and immediately

placed on dry ice and then stored at -80$^{o}$C. Unless otherwise stated, all transformations were completed using these in house chemically competent cells.

## 2.4 Immunofluorescence Microscopy

### 2.4.1 Materials

| | |
|---|---|
| ProLong Gold antifade reagent with DAPI | Invitrogen (#P10144) |
| NucBlue Fixed Cell Stain ReadyProbes | Invitrogen (#R37606) |
| ProLong Gold antifade reagent | Invitrogen (#P36934) |
| Pierce 16% Formaldehyde (w/v) | ThermoFisher (#28906) |
| BSA | Sigma (#A2153-100G) |
| Triton X-100 | Sigma (#X100-100ML) |
| Tween-20 | Sigma (#P941650ML) |

### 2.4.2 Reagents and Stock solutions

| | |
|---|---|
| Fixing solution | Add one vial of Pierce 16% Formaldehyde to 3ml of 1X PBS. |
| Permeabilising buffer | Add 0.3% Triton X-100 to 1X TBS. |
| Blocking buffer* | Add 2% BSA (w/v) to 1X TBS. Filter sterilised. |
| Wash buffer* | Add 1% BSA (w/v) and 0.05% Tween-20 to 1X TBS. Filter sterilised. |

### 2.4.3 Antibodies

| | |
|---|---|
| HPV16 E7 (NM2)– anti-mouse | 1:50 Santa Cruz Biotechnology (#sc-65711) |
| p16 (F12) – anti-mouse | 1:50 Santa Cruz Biotechnology (#sc-1661) |

| | |
|---|---|
| SRSF2 – anti-mouse | 1:250 Abcam (#1b11826) |
| pSRSF2 – anti-mouse | 1:250 Santa Cruz Biotechnology (#sc-53518) |
| c-MYC (9E10) – anti-mouse | 1:200 Santa Cruz Biotechnology (#sc-40) |
| Alexa Fluor 594 donkey anti-rabbit | 1:1000 Invitrogen ((#A21207) |
| Alexa Fluor 488 donkey anti-mouse | 1:1000 Invitrogen (A21202) |
| Alexa Fluor 647 goat anti-mouse | 1:1000 ThermoFisher (#A28181) |

### 2.4.4 Fixing Cells

Prior to use, coverslips were autoclaved and placed in 6-well plates. The coverslips were washed with 70% Ethanol for 10 minutes and washed twice with 1X PBS. Cells were plated as stated previously. Prior to fixing, the media was removed and cells were washed twice with 1X PBS. Following this, 600µl of freshly prepared fixing solution was added to each well and incubated for 15 minutes at room temperature. After incubation, cells were washed three times with 1X TBS. Plates can be stored in 1X TBS at 4$^{o}$C for up to two weeks.

### 2.4.5 Permeabilisation

Cells were permeabilised for 15 minutes at room temperature using 1ml of freshly made permeabilising buffer per well. Following this, wells were washed twice with filtered 1X TBS.

### 2.4.6 Blocking and antibody staining

Non-specific binding sites were blocked by adding 1ml of blocking buffer for 1hr at room temperature. Following this, without additional washing, Cells were incubated with primary antibody made up in blocking buffer. 25µl of the primary antibody dilution was added on a glass slide and the coverslips were inverted on to the slide. Following overnight incubation at 4$^{o}$C, the coverslips were washed 3 times for 10 minutes each with washing buffer. Coverslips were then incubated with fluorophore-conjugated secondary antibody made up in blocking buffer for 1hr at room temperature in the dark. Following incubation, coverslips were washed twice for 10 minutes with washing buffer and then finally for 10 minutes with filtered 1X TBS. Two drops of NucBlue fixed cell readyprobe was added directly into each 6 well plate with 1ml 1X TBS and incubated at room temperature for 10 minutes. All wash steps following secondary antibody incubation were performed in the dark. Coverslips were then mounted on glass slides using Prolong Gold antifade reagent and kept overnight at room temperature in the dark.

### 2.4.7 Microscopes, Imaging and software

The following equipment and software were used for various aspects of microscopy.

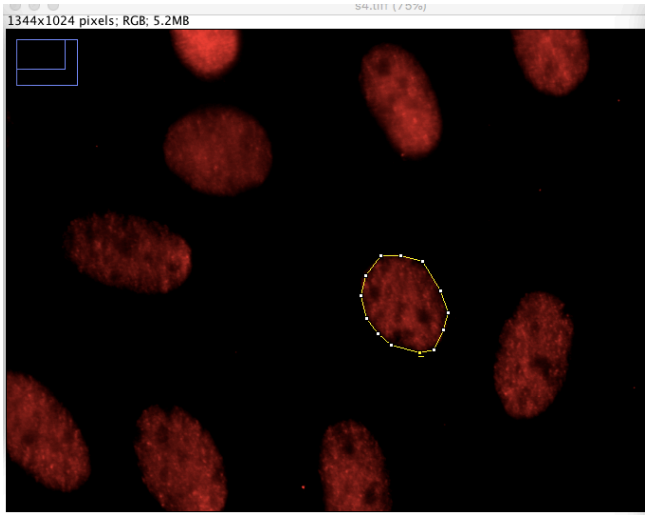
Etaluma Lumascope 720 – software: Lumaview72

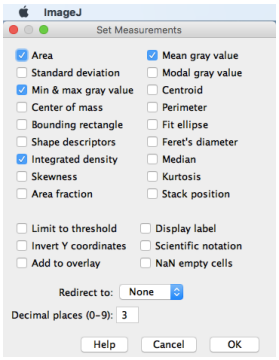
OLYMPUS CKX53 – software: GXCapture-T

OLYMPUS U-TV1X-2 – software: SmartCapture 3

Zeiss LSM 880 Elyra confocal microscope

### 2.4.8 Image J quantification




| | Area | Mean | Min | Max | IntDen | RawIntDe |
|---|---|---|---|---|---|---|
| 1 | 12268 | 30.473 | 0 | 68 | 373840 | 373840 |
| 2 | 13986 | 29.960 | 0 | 76 | 419019 | 419019 |
| 3 | 10765 | 40.072 | 8 | 79 | 431380 | 431380 |
| 4 | 13643 | 28.859 | 0 | 62 | 393725 | 393725 |
| 5 | 12212 | 32.620 | 0 | 72 | 398353 | 398353 |

Image J software was used to quantify the fluorescence intensity within individual cells. The measurement parameters Area, Min & max gray value, integrated density and mean gray values were measured for each selected cell. The cell was selected as shown in the above image, carefully selecting around the edge of the nucleus, and a measurement was recorded. In addition, a region next to the nucleus was selected and measured to record any background fluorescence. In total, 10 separate nuclei were measured for each value, and the mean fluorescence intensity was calculated by subtracting the intensity of the background region.

## 2.5 Protein

### 2.5.1 Materials

| | |
|---|---|
| Skim Milk Powder | OXOID (#LP0031) |
| Sodium Deoxycholate | Sigma (D6760-10G) |
| Luminata Forte Western HRP Substrate | Merck Millipore (#WBLUF0100) |

| | |
|---|---|
| Clarity Western ECL Substrate | BIO-RAD (#1705060) |
| BSA | Sigma (#A2153-100G) |
| Acrylamide/Bisacrylaminde (30%) | Alfa Aesar (#J61505) |
| Trans-blot Turbo RTA Midi PVDF Transfer kit | BIO-RAD (#1704273) |
| M-PER | Thermo Scientific (#78501) |
| Tween 20 | Sigma (#P941650ML) |
| Sodium azide | Sigma (#S2002-5G) |
| TEMED | Sigma (#T9281-25ML) |
| SDS | Melford (#B2008) |
| Ammonium Persulfate | Sigma (#A3678-25G) |
| PhosSTOP phosphatase inhibitors | Merck (#4906837001) |
| cOmplete Protease Inhibitor Cocktail | Merck (#11697498001) |
| Benzonase | Sigma (#E1014-5KU) |
| Pierce BCA Protein Assay Kit | ThermoFisher (#23227) |
| PageRuler Plus 10-250kDa Protein Ladder | ThermoFisher (#26619) |

### 2.5.2 Reagents and Stock solutions

| | |
|---|---|
| 10X TBS buffer (1L) | 24g Tris-HCl, 5.6g Tris Base, 88g NaCl, made up in $dH_2O$, adjust to pH 7.6. |
| TBST buffer | 0.1% Tween 20, 10% 10X TBS, made up in $dH_2O$. |
| Blocking buffer | 5% Milk powder, made up in 1X TBST |
| Running buffer | 25mM Tris-HCl, 200mM Glycine, 0.1% SDS, made up in $dH_2O$. |
| Transfer buffer | 20% Methanol, 20% transfer buffer, made up in $dH_2O$. |

| | |
|---|---|
| Coomassie destain | 10% methanol, 5% acetic acid, made up in $dH_2O$. |
| Coomassie stain | 0.25% Coomassie brilliant blue dye in Coomassie destain |
| Stripping buffer | 1.5% glycine, 1% SDS, 10% Tween 20, made up in $dH_2O$. Adjust to pH 2.2. |
| 12% separating gel | 40% Acrylamide (30%), 26% 1.5M Tris pH 8.8, 0.1% SDS, 0.1% APS, 0.01% TEMED, made up in $dH_2O$. |
| 6% stacking gel | 14% Acrylamide (30%), 26% 0.5% Tris pH 6.8, 0.1% SDS, 0.1% APS, 0.01% TEMED, made up in $dH_2O$. |
| 4X Sample buffer | 20% 1M Tris-HCl pH 6.8, 8% SDS, 40% glycerol, 4% Bromophenol blue, made up in $dH_2O$, add 400mM DTT before use. |
| Protein extraction buffer | 1X protease inhibitor cocktail, 1X phosphatase inhibitor cocktail, 0.1% benzonase, made up in M-PER. |
| RIPA buffer | 50mM Tris-HCL pH 8.0, 150mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS. |
| Primary antibody (WB) | 5% BSA, made up in 1X TBST. |
| Secondary antibody (WB) | 5% Milk powder, made up in 1X TBST. |

### 2.5.3 Antibodies

| | |
|---|---|
| CLK3 – anti-rabbit | 1:1000 Cell Signaling Technology (#3256S) |
| HPV16 E7 (ED17) – anti-mouse | 1:500 Santa Cruz Biotechnology (#sc-6981) |
| HPV16 E7 (NM2)– anti-mouse | 1:500 Santa Cruz Biotechnology (#sc-65711) |
| p16 (F12) – anti-mouse | 1:500 Santa Cruz Biotechnology (#sc-1661) |
| c-MYC (G-4) – anti-mouse | 1:500 Santa Cruz Biotechnology (#B1618) |
| c-MYC (9E10) – anti-mouse | 1:500 Santa Cruz Biotechnology (#sc-40) |
| SRSF2 – anti-rabbit | 1:2000 Abcam (#ab28428) |

| | |
|---|---|
| β-actin HRP (AC-15) | 1:100,000 Novus Biologicals (#NB600-501H) |
| GFP-anti-rabbit | Cell Signaling Technology (#2555S) |
| HIF-1a – anti-mouse | 1:2000 BD Biosciences (#610958) |
| FLAG tag – anti-mouse | CUSABIO (#MA000021M0m) |
| DKXRB HRP Secondary | 1:10,000 Invitrogen (#A16023) |
| DKXMO HRP Secondary | 1:10,000 Invitrogen (#A16011) |

### 2.5.4 Extraction

The cell lysate samples were prepared according to chapter 2.1.7. Initially M-PER buffer was used for the lysis, but was later switched to RIPA for higher yield in protein. The buffer was prepared fresh prior to use by the addition of protease and phosphatase inhibitor cocktails, as well as benzonase. The cell pellets were resuspended in the buffer and kept on ice for 30 minutes, before being spun at max speed (16.3G) at 4$^{o}$C on a temperature-controlled microcentrifuge for 30 minutes. The supernatant was transferred to a fresh Eppendorf tube and stored at -20$^{o}$C.

### 2.5.5 BCA Assay

The BCA assay was carried out using the ThermoFisher Pierce BCA protein assay kit. The assay was used to quantify protein samples. Albumin standards provided with the kit were made according to the protocol and stored at -20$^{o}$C. Three sets of each standard and sample were added in clear flat-bottom 96-well plates. 200µl of working reagent was added using multichannel pipette in each well. The plate was covered in foil, stored at 37$^{0}$C for 30

minutes, and then cooled for 10 minutes at room temperature. Absorbance was measured at 562nm on a microplate reader. The protein concentration was calculated using the standard curve.

### 2.5.6 Western Blot

Equal amounts of protein samples were heated to 50$^{o}$C for 10mins* in sample buffer containing 10% 1M DTT. Gels were run using the Bio-Rad mini trans-blot cell at 150V. 5µl of PageRuler Plus ladder was loaded in a well to act as a molecular weight marker.

PVDF membrane was soaked in 100% methanol for 1 minute and equilibrated in freshly made cold transfer buffer, along with a pair of transfer stacks. Prior to transfer, the gel was assembled in a Trans-Blot SD Semi-Dry cell along with the PVDF membrane and the transfer stacks. They were tightly packed without any air bubbles and loaded onto a Trans-Blot Turbo system and ran for 7 minutes. After the completion of the transfer, the membrane was removed and blocked in blocking buffer for 1hr at room temperature on a rocker.

After blocking, the blot was washed with TBST before being incubated overnight in the primary antibody solution against the target protein at 4$^{o}$C on a roller. The blot was washed with TBST at 10 minute interval for 3-5 times before being incubated in the HRP-conjugated secondary antibody for 1hr at room temperature. The blot was then washed with TBST again at 5 minute intervals for 3-5 times before imaging.

Imaging of the blot was done using the G:BOX Chemi gel imaging system from syngene. When the expression of the protein was low, the blots were developed in a darkroom by exposing the blots to films. Luminata Forte Western HRP Substrate was used as the primary ECL reagent. However, Clarity Western ECL Substrate was used as a substitute when detecting proteins with higher level of expression. As a loading control, β-actin conjugated with HRP was incubated with the blot for 1hr, followed by 10 minute 3-5 washes with TBST. Exposure to ECL substrate was reduced to 1 minute as overexposure results in bands with very high intensity.

*Initially all protein samples were boiled at 95$^{o}$C, but later following optimisation, 50$^{o}$C was identified as the ideal temperature for maximum protein yield.

## 2.6 Protein production

A codon optimised sequence of CLK3 was created in the lab and was used in all of the following expressions as the primary sequence for the gene. A sequence coding for a functionally inactive form of CLK3, with the substitution mutation at K186R, was also used along with the functional form. In addition, two mutant forms of CLK3 with a substitution at position 236 were also used to make recombinant proteins. All four forms of CLK3 were based on the codon-optimised sequence. Initially, protein expression was attempted using BL21 (DH3) *E. coli*, however, the expression was unsuccessful. Following this, the Bac-to-

Bac Baculovirus expression system was used, which proved to be successful. Both techniques will be described in detail.

The first step in creating the mutant forms was to identify the gatekeeper domain of CLK3. The software PyMOL was used to align the structure of CLK3 and similar kinases with known gatekeeper domains (Chapter 5.2). The gatekeeper domain was identified to be a phenylalanine at position 236. Following this, site directed mutagenesis was carried out to replace the phenylalanine with either a glycine (mutant 1) or alanine (mutant 2). Site directed mutagenesis is described in further detail in chapter 2.3.11

### 2.6.1 Materials

| | |
|---|---|
| BL21 DE3 cells | ThermoFisher (#10362100) |
| IPTG, Isopropyl B-D-1-thiogalactopyranoside | Sigma (#16758-1G) |
| Gibco Sf-900™ III SFM | ThermoFisher (#12658027) |
| Gibco Sf9 cells | ThermoFisher (#12659017) |
| Celfectin™ II Reagent | ThermoFisher (#10362100) |
| MAX Efficiency™ DH10B competent cells | ThermoFisher (#18297010) |
| Fisherbrand™ shaker flasks 125ml | FisherSci (#11735373) |
| Fisherbrand™ shaker flasks 250ml | FisherSci (#11765253) |
| X-gal | (Sigma #11680293001) |
| Kanamycin sulfate | Sigma (#60615-5G) |
| Tetracycline | Sigma (#87128-25G) |
| Gibco Gentamicin | Fisher scientific (#15750078) |

### 2.6.2 Reagents and Stock solutions

| | |
|---|---|
| Buffer A | 50mM Tris.HCl pH 7.5, 40mM Imidazole, 500mM NaCl, 1mM DTT, made up in $ddH_2O$. |
| Buffer B | 50mM Tris.HCl pH 7.5, 400mM Imidazole, 500mM NaCl, 1mM DTT, made up in $ddH_2O$. |

### 2.6.3 Protein expression in BL21 (DE3) *E. coli*

The gene of interest was inserted into the pET-28a (+) plasmid and transformed into competent BL21 (DE3) *E. coli.* The cloning and transformation techniques are described in detail in chapter 2.3.8. A colony from the overnight plates was resuspended in 10ml of SOC media with kanamycin. The culture was incubated at 37$^{o}$C overnight to produce a starter culture. 1L of SOC media was inoculated with the starter culture and incubated at 37$^{o}$C until $OD_{600}$ reached 0.6. 400µM IPTG was added to induce expression of protein for 3hrs at 37$^{o}$C. Post expression, cells were spun using a temperature controlled 4$^{o}$C centrifuge at 5000rpm. The supernatant was removed and the cell pellet was suspended in 20ml of lysis buffer. The resuspended cells were incubated on ice for 30 minutes and then sonicated for 5 minutes whilst being kept on ice, with 10 second pulses and 30 second intervals. After sonication, the cells were centrifuged at 10,000G in an ultracentrifuge for 20 minutes. Both the supernatant and the pellet were collected and stored at -20$^{o}$C.

An SDS-page gel was run with the supernatant and pellet samples to determine whether the protein was being expressed successfully. Lack of expression led to

optimisation by increasing the amount of IPTG used to 1M and decreasing the incubation temperature to 27$^{o}$C. Altering expression conditions did not resolve the issue as no evidence of protein expression was found.

### 2.6.4 Protein expression in Bac-to-Bac Baculovirus expression system

The initial step was to clone the gene of interest into the pFastBac donor plasmid. The donor plasmid contains the AcMNPV polyhedron promotor, along with 6x His tag and a TEV site. BamHI and XhoI restriction digest enzymes were used to clone the gene of interest from pBP plasmid. In total, four pFastBac donor plasmids were made: wild-type, kinase inactive, F236A mutant, and F236G mutant. The plasmids were transformed into competent *E. coli* and grown overnight on LB agar plates with ampicillin resistance. Successful colonies were confirmed by colony PCR and an overnight LB liquid culture was set up to extract recombinant pFastBac donor plasmid using the Qiagen plasmid Mini kit.

The recombinant pFastBac donor plasmid was then transformed into DH10Bac *E. coli*. The transformation protocol is described in detail in chapter 2.3.8. The cells contain the baculovirus shuttle vector (bacmid), bMON14272 (136kb), and the helper plasmid, pMON7124 (13.2kb). The bacmid contains a kanamycin resistance marker, as well as a segment of DNA encoding the LacZ$\alpha$ peptide, which forms blue colonies in the presence of X-gal and IPTG. The bacmid also contains an *att*Tn7 target site within the LacZ$\alpha$ coding region. The helper plasmid provides the transposition proteins during the transposition reaction that takes place between the Tn7 element on the pFastBac vector and the *att*Tn7

target site on the bacmid. This transposition reaction creates a recombinant bacmid with the gene of interest inserted within the LacZα region. Following transformation, the DH10Bac *E. coli* cells were plated on LB agar plates containing 50µg/ml kanamycin, 7µg/ml gentamicin, 10µg/ml tetracycline, 100µg/ml X-gal, and 40µg/ml IPTG to select for successful DH10Bac transformants. The plates were incubated for 48hrs at 37°C. Blue colonies were ignored as they still contained the intact LacZα gene, therefore the transposition wasn't successful in those colonies. White colonies were picked and re-streaked on LB agar plates with the same selection conditions and grown overnight at 37°C. The next day, successful colonies were picked and inoculated in a 3ml liquid LB culture containing 50µg/ml kanamycin, 7µg/ml gentamycin, and 10µg/ml tetracycline. The culture was set up early as the cells usually glow very slowly.

### 2.6.5 Recombinant bacmid purification

500µl of the overnight culture was used to make a glycerol stock of the bacmids. The remaining culture was spun at maximum speed on a benchtop microcentrifuge for 1min. The supernatant was removed and the cells were resuspended in 300µl of Qiagen Buffer P1 that doesn't contain the blue lysis indicator dye. This was followed by addition of 300µl of Qiagen Buffer P2, and the mixture was incubated at room temperature for 5mins. 300µl Qiagen buffer N3 was then added slowly and mixed gently, following which the mixture was placed on ice for 10mins. The mixture was then centrifuged at maximum speed for 10mins. The clear supernatant was transferred into a 1.5ml Eppendorf tube with 800µl isopropanol and left overnight at -20°C. This step can be reduced to 10mins on ice, but overnight storage

significantly produces a higher yield of the bacmid. The mixture was centrifuged at maximum speed for 15mins, following which all the supernatant was removed and the pellet was washed with 70% ethanol by inverting the tube several times. The mixture is centrifuged again at maximum speed for 10mins, and the wash step is repeated. The pellet is then air-dried for 10mins after careful removal of all the supernatant. The dried pellet is resuspended in 40µl of TE buffer and stored at 4°C until use for transfection. Long term storage is at -20°C. Freeze-thawing of bacmid was kept to a minimum as this significantly decreases transfection efficiency.

### 2.6.6 PCR analysis of recombinant bacmid

The following primers and PCR conditions were used to confirm successful cloning:

| | |
|---|---|
| pUC/M13 Forward | CCCAGTCACGACGTTGTAAAACG |
| pUC/M13 Reverse | AGCGGATAACAATTTCACACAGG |

| | | | |
|---|---|---|---|
| Initial denaturation | 3 mins | 93°C | |
| Denaturation | 45 secs | 94°C | 35 cycles |
| Annealing | 45 secs | 55°C | |
| Extension | 5 mins | 72°C | |
| Final extension | 7 mins | 72°C | |

Following PCR, the reaction was analysed by agrose gel electrophoresis. The expected size of the band was 3800bp. All tested samples had the correct expected size.

### 2.6.7 Sf9 cell culture

Transferred 25ml of SF-900™ III SFM into a 125ml sterile flask, and incubated it at 27°C. A vial of frozen Sf9 cells were thawed quickly in a 37°C water bath and transferred to the pre-

warmed 25ml media in the flask. The flask was placed in a 27°C shaking incubator without $CO_2$ at 100RPM. On day three, cells were counted using a DeNovix CellDrop™ cell counter. If the Viability was >80% with a cell density of ≥ $2x10^6$ cells/ml, they were passaged at a cell density of $4x10^5$ viable cells/ml in 30-50ml media. If the cell density was lower, the cells were passaged the following day. If the cell viability was below 80%, the cells were pelleted and then resuspended at $4x10^5$ cells/ml in 30-50ml media. If the cells continued to show poor viability (<80%), a new vial of Sf9 cells were thawed and the process was repeated.

Before passaging the cells, it is important to allow the cells to achieve a minimum of $2x10^6$ viable cells/ml. Cells were usually passaged at a density of $4x10^5$ cells/ml in 30-50ml media every 3-4 days. It was important to allow the cells to undergo 3-5 passages prior to any transfection. Transfection was only done on cells that have undergone under 25 passages, and cell viability was regularly measured. If the viability of the cells fell below 90%, they were discarded.

### 2.6.8 Bacmid Transfection

Sf9 cells were checked to confirm cell viability to be greater than 95%, and passage number to be between 5-20. The cells were plated at $1x10^6$ cells in 6-well plates. After 30mins, the cells would have attached. Add the transfection mix, made with 8µl of Celfectin II in 100µl medium, and the bacmid mix, made with 2µg of DNA and 100µl of medium, together. The mixture is incubated at room temperature for 15mins and then added dropwise to the cells. The cells are incubated at 27°C for 48-72hrs, or until signs of viral

infection are observed under the microscope. This includes an increase in cellular diameter, increased size of the nuclei, granular appearance with signs of viral budding, detachment, and cell lysis. Following successful transfection, the medium containing the virus is collected and centrifuged to remove any large debris. This P1 viral stock is stored at $4^{O}$C, protected from light.

### 2.6.9 Baculoviral stock amplification

The P1 stock is a low-titer stock, therefore was used to generate higher-titer P2, and then P3 stock. 500µl of the P1 stock was added to 25ml of Sf9 cells at $1x10^6$ cells per ml, and grown in suspension on a shaker at 27°C for 72hrs, or until significant loss of cell viability is seen. The media with the virus is collected and spun at 500g for 5mins to remove any cell debris and the supernatant is stored at 4°C, protected from light. P2 stock was used to make P3 stock following the same steps, with 50ml of sf9 cells used ($1x10^6$ cells per ml) with 5ml of P2 stock. After P3 stock collection, a small amount of the Sf9 cells were collected and lysed with sample buffer and boiled at 95°C for 10mins and the protein expression was confirmed by SDS-PAGE. P3 stock is spun and the cell pellet is frozen at -80°C until protein purification.

### 2.6.10 Protein purification

Following successful expression, frozen Sf9 cells are thawed at room temperature, and 1µM PMSF inhibitor is added. Cells are then sonicated on ice, for a period of 5mins with bursts of 15 seconds. The sonicated sample is spun at 18,000 rpm at 4°C in a Beckman Avanti

J-26S using a J 25.50 rotor for 20mins. Following this, the supernatant was collected, which should be clear, and stored at 4$^{o}$C. If particles are still present, the sample is re-spun again. A 5ml His-Trap HP column (GE healthcare) was equilibrated in buffer A. The clear supernatant is loaded onto column at 1ml/min using the sample pump and washed with buffer A. The protein was eluted using the same buffer with increasing imidazole concentrations from 40mM to 400mM (buffer B at 10, 20, 40, 60, 100%). Fractions were analysed using SDS-PAGE, and pure factions were concentrated using a Vivaspin tube with 10 kDa MWCO spinning at 3500 rpm at 4$^{o}$C. The purified protein samples were stored at -80$^{o}$C.

# Chapter 3

# Investigating the effects of CLK3 deletion on E6/E7 expression in keratinocytes

## 3.1 Validation of CLK3 and E6/E7 regulation

### 3.1.1 Introduction

The importance of CLK3 in HPV associated cancer cell lines has been well established by Luke Williams when initial knockdown studies demonstrated the loss of cell viability. Further experiments revealed that CLK3 knockdown resulted in reduction of E6/E7 mRNA levels, leading to the reactivation of the p53 and pRb pathways supressed by these oncoproteins. Cells were found to undergo cell cycle arrest in G1 phase and eventually enter senescence. This indicated that CLK3 could potentially be a useful molecular target for treatment of HPV-associated cancers.

Strong evidence has been provided to demonstrate that the effects on cell viability and loss of E6/E7 expression observed were directly due to the knockdown of CLK3 and not as a result of RNAi off target effects. Multiple siRNA sequences and inducible shRNA targeting CLK3 in several cell lines have delivered the same conclusions. This chapter will focus on CLK3 rescue experiments to further validate the relationship between the kinase and E6/E7. In addition, immortalised keratinocytes in which CRISPR-Cas9 gene editing was used to generate a knockout of CLK3 will be used to further support the previous CLK3 knockdown observations.

The main limitations with the use of RNAi to knockdown targets is the possible risk of off target effects. CRISPR is a powerful tool to completely block the expression of the gene

of interest, and eliminate any confounding effects from low level protein expression that could remain following RNAi silencing. In addition, due to its minimal off-target effects, CRISPR is the ideal tool for creating a clean system to study the effects of total loss of the target protein. However, as the resulting effect is permanent, the technology is not ideal for manipulating genes that are crucial for cell survival.

CRISPR editing of CLK3 was successfully achieved in Normal Immortal Keratinocytes (NIKS) (Sarastry, 2016). NIKS are a spontaneously immortalised cell line originating from human neonatal foreskin keratinocytes (Allen-Hoffmann et al., 2000). NIKS have a stable near-diploid karyotype, with an extra isochromosome of the long arm of chromosome 8. In organotypic raft cultures, NIKS are able to differentiate and form a stratified epithelium that is histologically identical to that formed by primary keratinocytes. HPV associated cancers are known to originate from keratinocytes, and the NIKS cell line is therefore a relevant system in which to study the HPV replication and HPV-driven carcinogenesis. In previous studies, the HPV 16 genome was stably transfected in NIKS, and following infection, the cells were found to display a phenotype similar to that of high- and low-grade cervical lesions (Flores, et al., 1999, Wechsler, et al., 2012). CRISPR knockout of CLK3 in NIKS had no significant effect on the cells, with continued growth and proliferation being observed. CLK3 is therefore not considered an essential gene in normal keratinocytes, which is a desirable characteristic for molecular targets in targeted therapy.

All following experiments were repeated at least three times unless stated otherwise.

## 3.2 Knockout of CLK3 results in loss of HPV E7

Following successful CRISPR knockout of CLK3, the cells were infected with a promoterless lentiviral vector (pLenti-CMV-GFP-DEST) containing the HPV16 URR and E6/E7 coding sequence cloned from the 93VU147T head and neck squamous cancer cell line and underwent selection to create a stable cell line in which E6 and E7 expressed is under the control of the viral promotor (Figure 3.2.1). These cells were provided to study the effect of CLK3 knockout (KO) on HPV E6/E7. Initial experiments were conducted to validate the cell lines and determine the effect of CLK3 KO on E6/E7 expression. Protein expression and western blot was carried out as described in chapter 2. Knockout of CLK3 was confirmed with the total loss of CLK3 observed in the western blot (Figure 3.2.2). It was previously known that siRNA knockdown of CLK3 in HPV-associated cancer cells resulted in reduced levels of HPV E7 protein (Williams, 2021). Consistent with this, HPV E7 protein was undetectable in the $CLK3^{-/-}$ NIKS, which also showed a much lower expressed of $p16^{INK4A}$ than WT NIKS expressing E6/E7. (Figure 3.2.2).

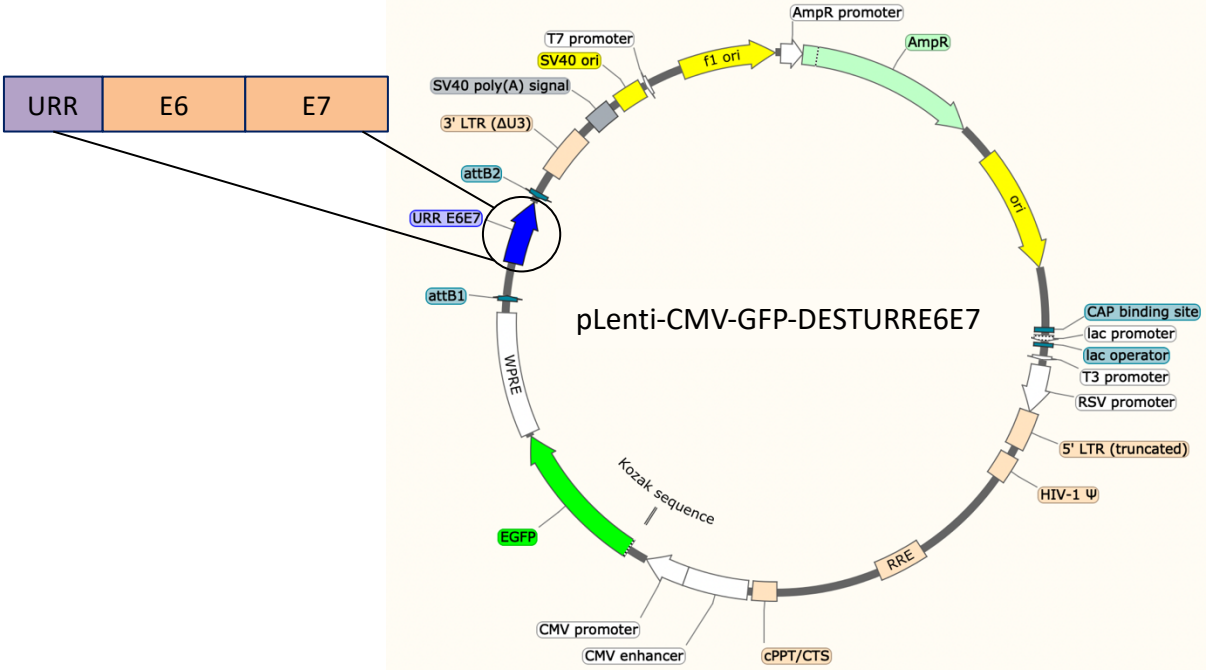



**Figure 3.2.1 Map of pLenti-CMV-GFP-DEST-URRE6/E7 plasmid**
Gateway cloning was performed to transfer the URR E6/E7 fragment of 93VU147T from a pENTR4 entry vector to pLenti-CMV-GFP destination vector (Sarastry, 2016).

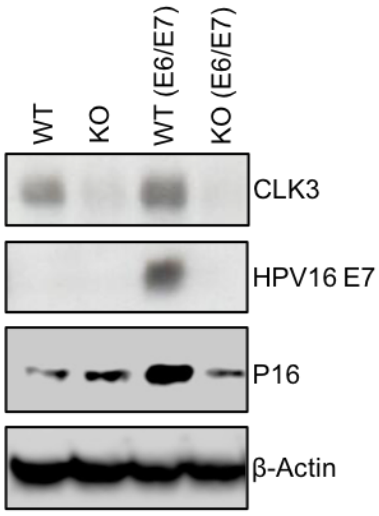



**Figure 3.2.2 CRISPR-Cas9 Knockout of CLK3 results in loss of HPV E7 protein**
Western blot completed to validate the loss of CLK3 following CRISPR/Cas9 editing of NIKS. Confirms the absence of detectable HPV16 E7 protein in CLK3$^{-/-}$ NIKS. 50µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of CLK3 is approximately 59kDa. E7 is approximately 18kDa in molecular weight. p16 expression further validates loss of HPV16 E7 in the KO cells. p16 is approximately 16kDa in molecular weight. β-Actin was used as a loading control. PageRuler plus prestained protein ladder was used as a molecular marker.

Furthermore, previous observations indicated that CLK3 was involved in the regulation of HPV E6/E7 mRNA, with the knockdown of CLK3 resulting in the reduction in mRNA levels by over 50% (Williams, 2021). qRT-PCR data from this experiment further strengthens these findings, with similar reductions of over 70% being observed in the E6/E7 mRNA levels (3.2.3, A). To eliminate the possibility that the lower levels of E6/E7 expressed in the CLK3$^{-/-}$ cells was due to a lower number of integrated copies of the URR-E6/E7 cassette in this cell line versus the WT NIKS, it was necessary to determine the DNA copy number of this cassette in both cell lines. DNA extracted from both cell lines was used to quantify the genome copy number of E6/E7 relative to the GAPDH copy number (Figure 3.2.3, B). In fact, CLK3$^{-/-}$ NIKS were found to contain on average one E6/E7 copy per cell, whereas the WT NIKS contained 0.5 copy per cell, indicating that the difference in mRNA expression observed could be even greater than initial observations.

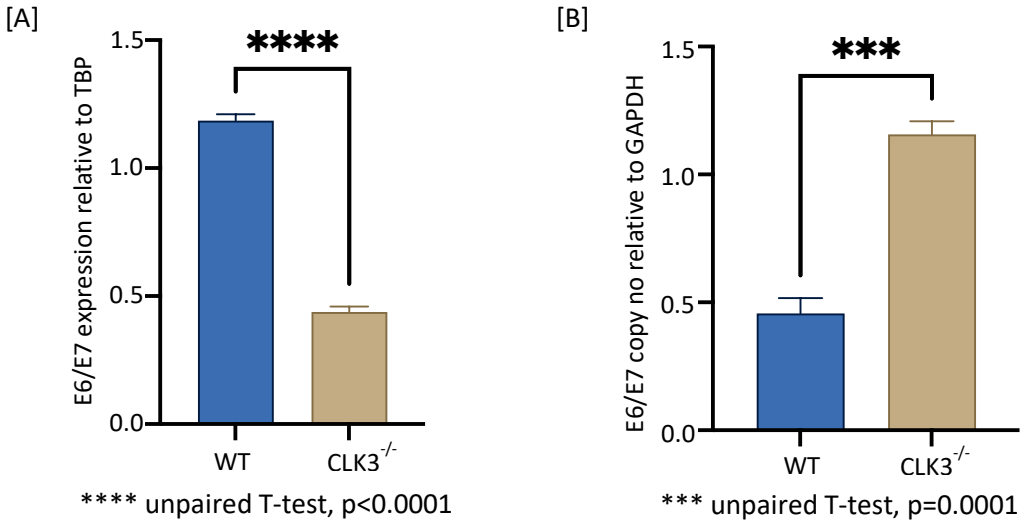



**Figure 3.2.3 CRISPR-Cas9 Knockout of CLK3 results in loss of HPV E6/E7 mRNA.**
A) qRT-PCR analysis of E6/E7 mRNA relative to TBP in wildtype and CLK3$^{-/-}$ NIKS. B) qPCR analysis of E6/E7 genome copy number relative to GAPDH in wildtype and CLK3$^{-/-}$ NIKS.

The plasmid used for the stable transduction of URR-E6/E7 also contains a separate expression cassette, in which GFP expression is driven by the constitutive CMV promoter (Figure 3.2.1), allowing cells in which the transgenes have integrated to be isolated using FACS. This provided an alternative method to normalise the expression of the plasmid in the wild-type and KO cells. Single parameter histograms allow cells expressing a particular marker to be identified (Fig 3.2.4, A). From the histograms, a noticeable difference in GFP expression between the two cell lines were observed. Gating allowed cells expressing GFP within a given range to be isolated. Cells were selected that expressed the same level of GFP, and following FACS, were grown and used in all future experiments. qRT-PCR analysis of the sorted cells confirmed a 8-16 fold difference in mRNA levels between the WT and CLK3$^{-/-}$ cells, further confirming the effect seen on E6/E7 levels was likely due to the absence of CLK3, as opposed to a reduced number of URR-E6/E7 copies per cell in the CLK3$^{-/-}$ NIKS (Fig 3.2.4, B).

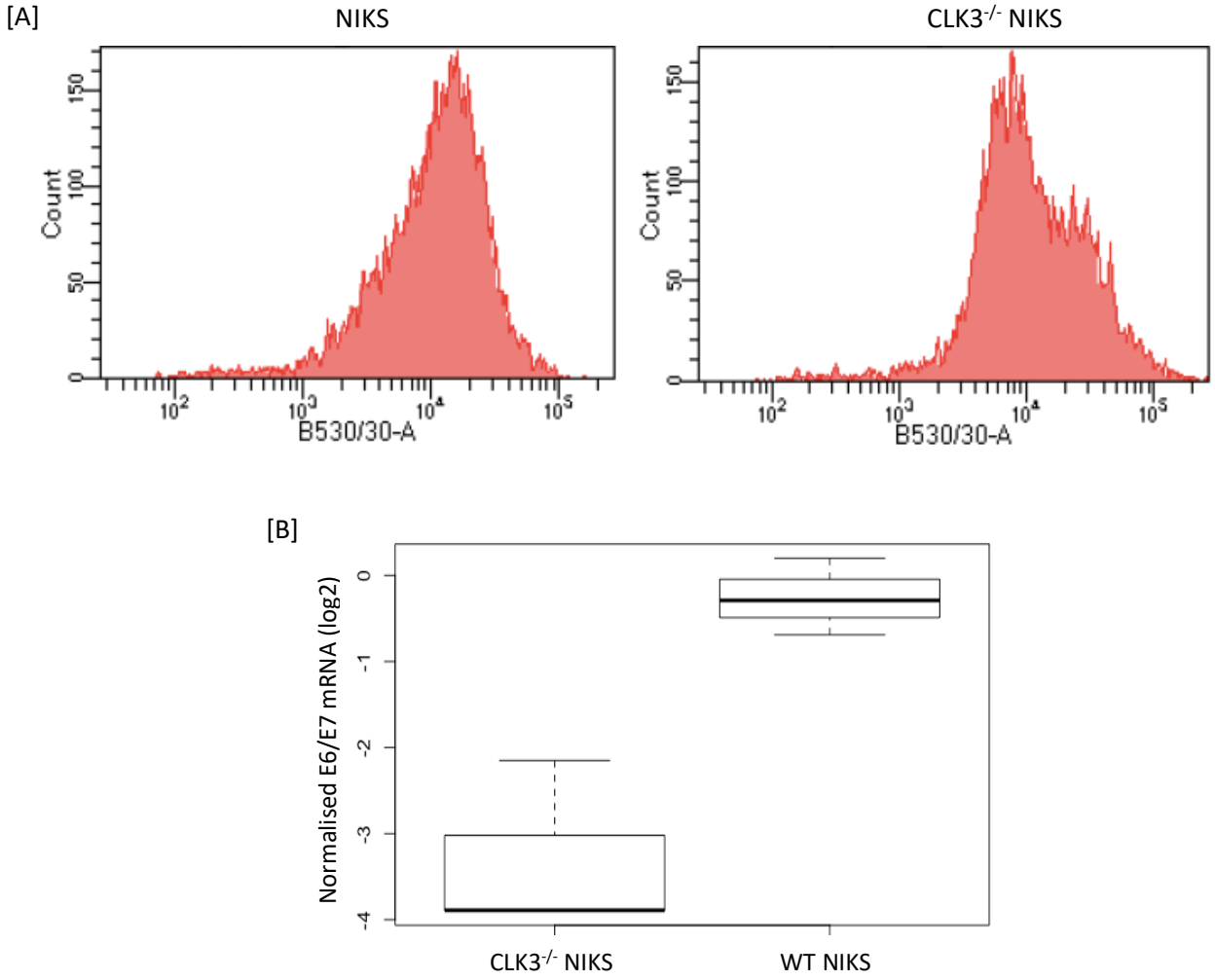



**Figure 3.2.4 FACS to normalise GFP in WT and CLK3$^{-/-}$ NIKS**
A) GFP expression in cells prior to FACS of wildtype and CLK3$^{-/-}$ NIKS, respectively. C) Box plot representing the fold change of E6/E7 mRNA in wildtype and CLK3$^{-/-}$ NIKS following FACS sorting, where GFP was normalised between the two cell lines.

Further validation was done with fluorescent immunocytochemistry (IF/ICC). Both WT and CLK3$^{-/-}$ NIKS with E6/E7 were stained with the HPV16 E7 antibody. The staining pattern further confirmed the importance of CLK3 in E7 expression, with minimal staining being observed in the KO NIKS. Image J software was used to quantify the intensity of the signals and confirm the significant difference between the cell lines. HPV positive cell lines

SiHa and CaSki, along with a HPV negative cell line, C33a, were used as controls to confirm the validity of the E7 antibody.

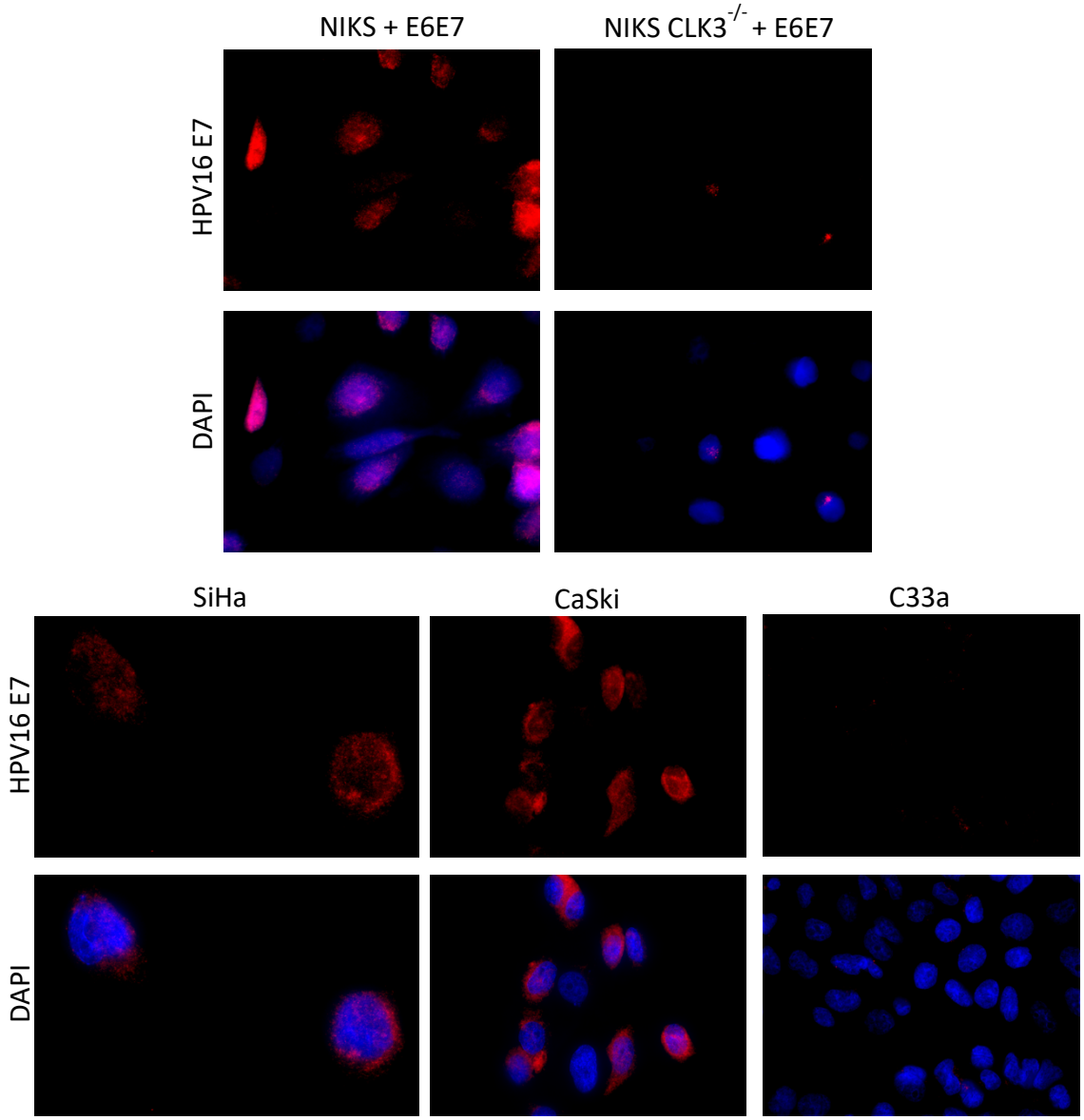



**Figure 3.2.5 Immunofluorescent staining confirms loss of E7 upon CLK3 knockout**
IF staining of two HPV positive cervical cancer cell lines (CaSki, SiHa) and one HPV negative cervical cancer cell line (C33A), along with WT and CLK3 KO NIKS with and without HPV16 URR E6/E7 following HPV16 E7 antibody staining (red). DAPI staining confirms presence of nuclei (blue). Microscopy images were at 60X magnification.

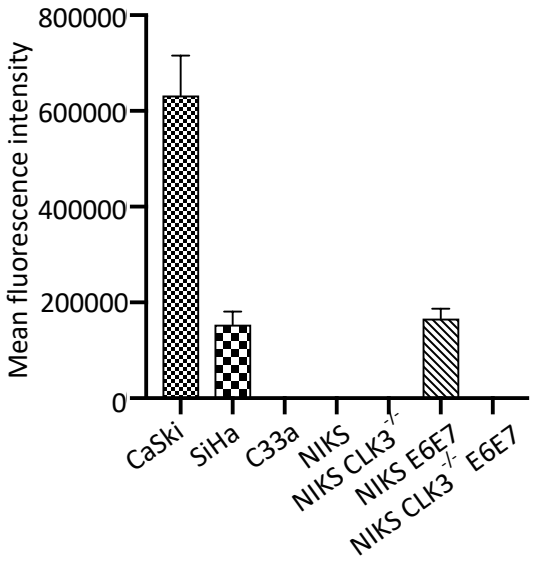



**Figure 3.2.6 Quantification of Immunofluorescent staining**
Image J analysis of the cells listed following HPV16 E7 antibody staining. The mean fluorescence intensity of 10 cells is represented by each bar.

p16 antibody was used alongside HPV16 E7 to further validate the results. IF/ICC was performed on various HPV positive and HPV negative cells to validate the p16 antibody. As expected, strong expression of p16 was detected in CaSki and HeLa cell lines, both of which are known HPV16 and HPV18 cervical carcinoma cell lines, respectively (Figure 3.2.7). SiHa, another HPV16 cervical carcinoma cell line, shows significantly lower p16 expression compared to CaSki and HeLa cells. SiHa is known to have only two copies of HPV16 integrated within each cell, therefore it is understandable that p16 expression is lower in this cell line. Minimal staining in C33A confirmed the use of p16 as a marker for HPV positive cells. Data from Image J quantification supports the visual observations. From the western blot in figure 3.2.2, p16 expression correlates with the expression of HPV16 E7 in the NIKS, with a significant loss of p16 observed in the CLK3$^{-/-}$ NIKS with the URR HPV16 E6/E7. In

addition, p16 mRNA levels were found to correlate with E6/E7 mRNA levels (figure 3.1.4a). Furthermore, the staining pattern observed in the CLK3$^{-/-}$ cells expressing the URR E6/E7 is consistent with the E7 staining patterns seen previously, with the mean difference in fluorescence intensity being up to 9-fold greater in the wild-type (Figure 3.2.8).

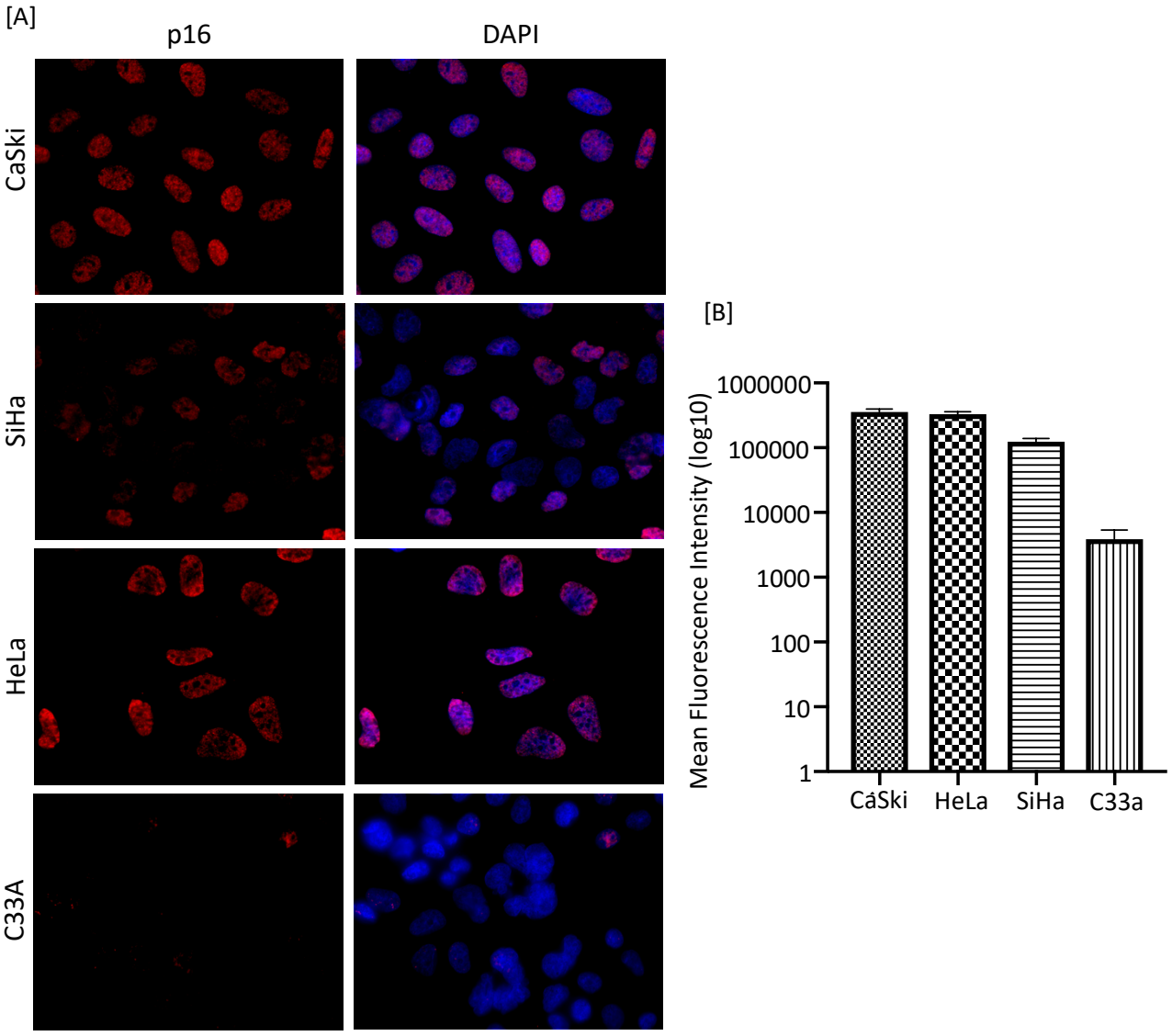



**Figure 3.2.7 Validation of p16 staining in cancer cell lines**
A) p16 antibody staining (red) of three HPV positive cervical cancer cell lines (CaSki, SiHa, HeLa) and one HPV negative cervical cancer cell line (C33A) DAPI nuclear staining is in blue. Microscopy images were at 60X magnification. B) Image J analysis. The mean fluorescence intensity of 10 cells being represented by each bar.

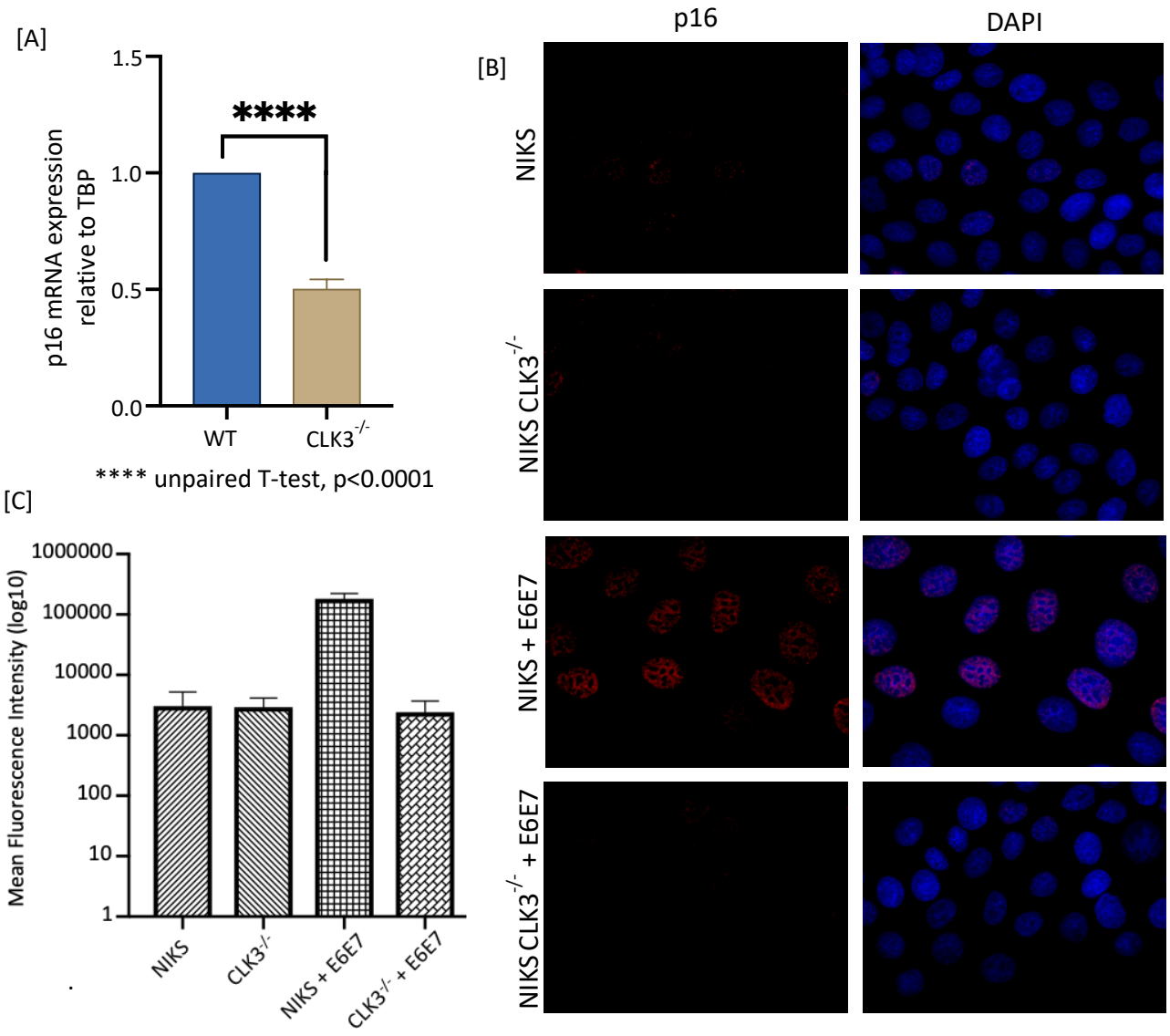



**Figure 3.2.8 Loss of p16 observed with loss of CLK3**

A) qRT-PCR analysis of p16 mRNA relative to TBP in WT and CLK3$^{-/-}$ NIKS. B) p16 staining of WT and CLK3$^{-/-}$ NIKS with and without HPV16 URR E6/E7. DAPI nuclear staining is in blue. Microscopy images were at 60X magnification. C) Image J analysis. The mean fluorescence intensity of 10 cells being represented by each bar.

Furthermore, using SiHa cells stably transduced with tertracycline-inducible CLK3 shRNA (prepared by Luke Williams) HPV16 E7 loss was observed from western blot analysis over 72hrs (Figure 3.2.9). In addition, IF staining with p16 antibody also revealed a total loss in p16 levels over a 72hr period of doxycycline induction to knockdown CLK3 (Figure 3.2.10). This added further evidence towards the relationship between CLK3 and HPV16 E6/E7.

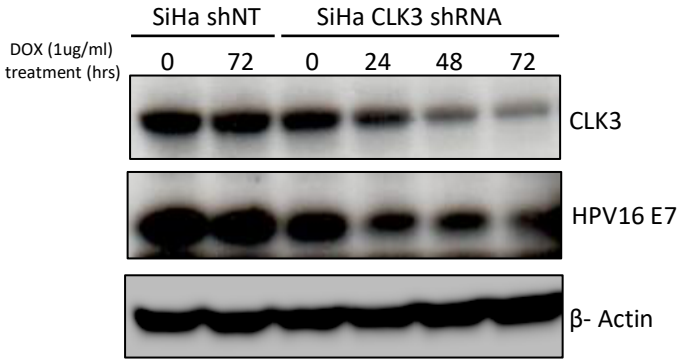



**Figure 3.2.9 ShRNA knockdown of CLK3 in SiHa results in loss of E7**
Western blot analysis of SiHa cells following siRNA knockdown of CLK3 over 72hrs. Cells were transfected with TtRMPVIR retroviral vector expressing shRNA targeting CLK3 and non-targeting shRNA (NT) vector. Cells stably expressing the plasmids were selected following single cell cloning (completed by Luke Williams). Cells were treated with 1µg/µl of doxycycline over 72hrs. 50µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of CLK3 is approximately 59kDa. E7 is approximately 18kDa in molecular weight. β-Actin was used as a loading control. PageRuler plus prestained protein ladder was used as a molecular marker. Loss of E7 is observed following shRNA knockdown of CLK3.

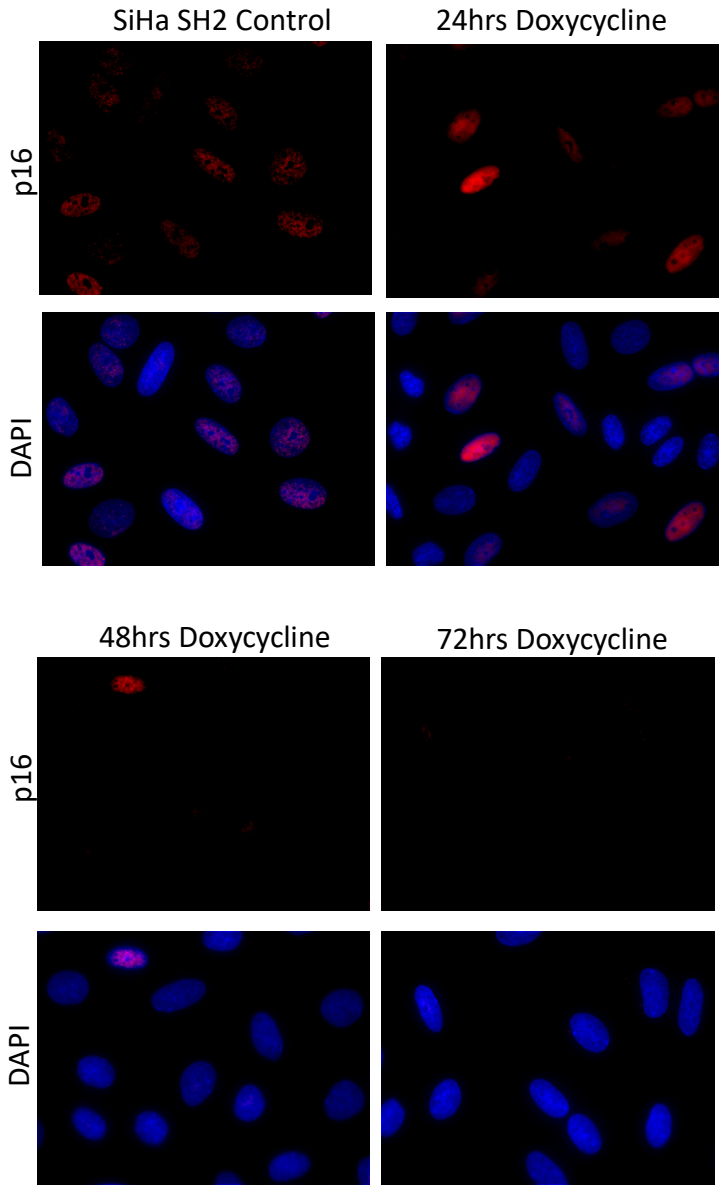



**Figure 3.2.10 Immunofluorescence following shRNA knockdown of CLK3 in SiHa**
SiHa cells were stained with p16 antibody over 72hrs of treatment with Doxycycline. Loss of p16 (red) is observed over 72hrs. DAPI staining indicates the nuclei (blue). Microscopy images were at 60X magnification.

## 3.3 Stable overexpression of CLK3

Thus far, robust evidence is presented detailing the effect of CLK3 KO on HPV E6/E7 mRNA, as well as E7 protein. The next aim was to overexpress CLK3 and determine the effect on HPV E6/E7, with the hypothesis that E6/E7 will be overexpressed under these conditions to reflect previous observations. Rescue experiments are typically carried out to conclusively verify the effects of RNAi knockdown on the targets. Initial rescue experiments were attempted using retro- or lentiviral constructs designed to allow either tetracycline-inducible or constitutive expression of a codon-optimised (CO-CLK3) CDS, in which changes to the nucleotide sequence were made to avoid RNAi-mediated silencing of the transgene. These attempts proved unsuccessful due to an inability to generate CLK3-containing retro- or lentiviruses, even using the inducible constructs (Williams, 2021).

Here, CO-CLK3, along with a kinase inactive mutant sequence (K186R), were directly transfected into various HPV positive cancer cell lines in an attempt to overexpress CLK3, using the pEGFP-N1 plasmid (Fig 3.3.1), from which a C-terminal GFP fusion protein of the gene of interest is expressed. Initial sub-cloning of CO-CLK3 and CO-CLK3 K186R from the pBABE-puro retroviral plasmid into this plasmid was completed following the methods detailed in chapter 2.3. One advantage with the use of the pEGFP-N1 plasmid was the opportunity to visualise and determine the level of transfection efficiency immediately. Confocal microscopy was used to visualise the expression of WT-CLK3-GFP and K186R-CLK3-GFP, which displayed nuclear localisation as expected (Figure 3.3.2, A). Transfection efficiency was poor, with the best results being seen in SiHa (10%) and 93Vu147T (15%).

There was no opportunity to carry out protein/RNA analysis following transient transfection as the proportion of cells overexpressing CLK3 were far too low to see any effect in the total population. pEGFP-N1 contains neomycin as a selective marker, thus following transfection, cells stably expressing the plasmid can be selected with G418 Sulfate. Complete selection was achieved by day 8, however, GFP expression was almost undetectable. Cells transfected with the empty pEGFP-N1 vector still showed strong GFP expression following selection. Protein analysis confirmed this, with no evidence of CO-CLK3-GFP being expressed in the cells following selection, however expression of GFP can still be observed in the cells with the empty vector (Fig 3.3.2, B). Given our previous unsuccessful attempts at the generation of stable or inducible CLK3-expressing cell lines, we reasoned that prolonged CLK3 overexpression is toxic and instead pursued an alternative strategy for the rescue experiments.

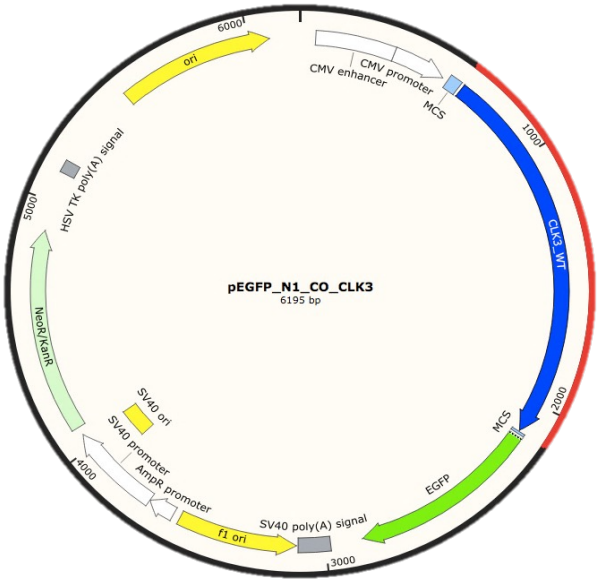



**Figure 3.3.1 Plasmid map of pEGFP-N1-CLK3**
Schematic diagram of the pEGFP-N1 plasmid with the CO-CLK3 insert (blue). The total length of CLK3-GFP is 2211bp, with the protein consisting of 737 amino acids, with a molecular weight of 86.2 KDa.

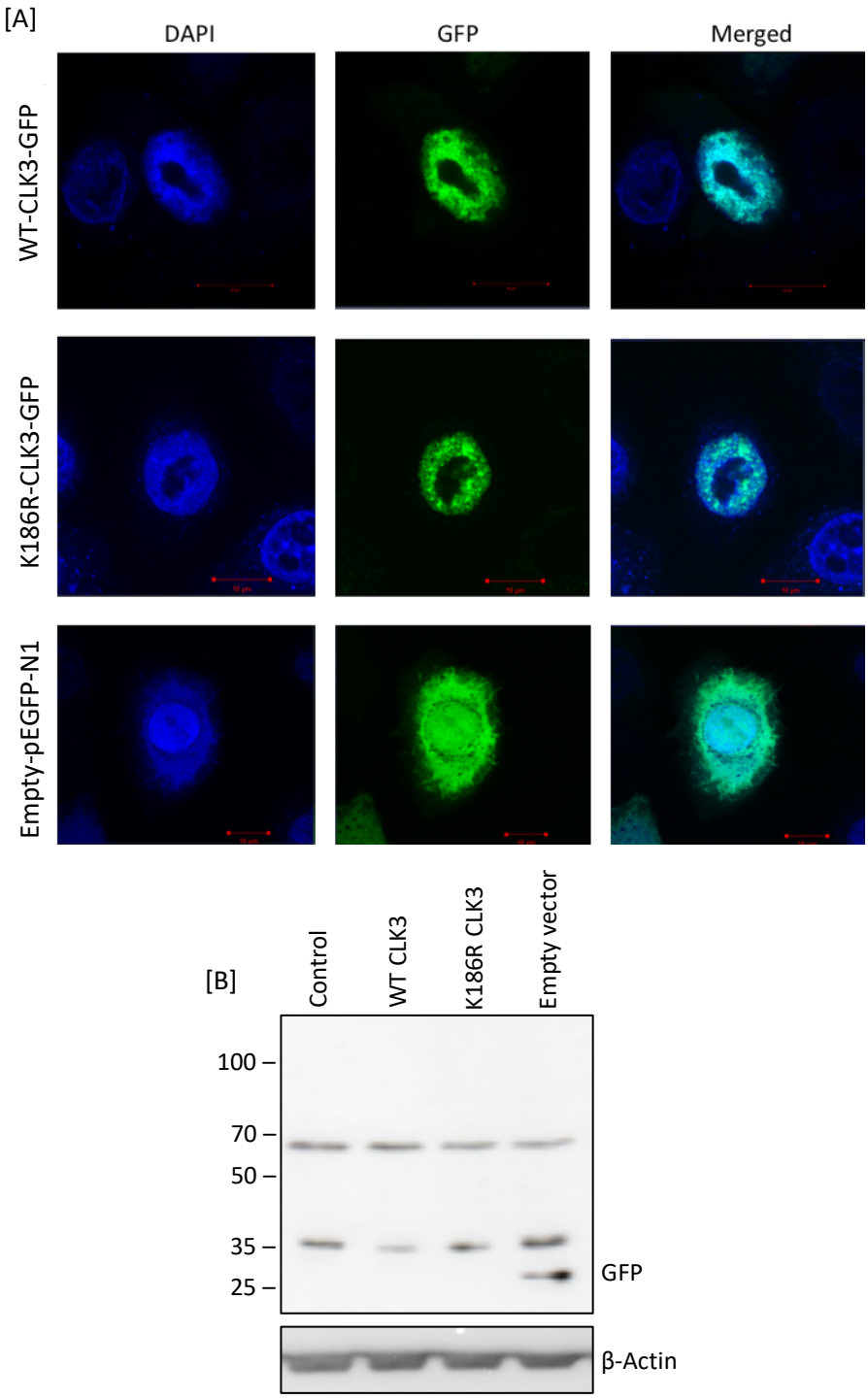



**Figure 3.3.2 Expression of CLK3 using pEGFP-N1**

A) Confocal microscopy showing expression of WT-CLK3-GFP, K186R-CLK3-GFP, and Empty-pEGFP-N1 in SiHa cell line. Microscopy images were at 100X magnification. B) Western blots confirming no evidence of CLK3-GFP being expressed in SiHa following G418 selection. 50µg of lysate was loaded in each well and separated by SDS page and probed with GFP anti-rabbit primary antibody overnight. Empty vector still has expression of GFP, with a molecular weight of 26.7 kDa. β-Actin was used as a loading control. PageRuler plus prestained protein ladder was used as a molecular marker.

## 3.4 Induction of endogenous CLK3 expression to observe effects on E7

Following various failed attempts to stably or inducibly overexpress CLK3 in several HPV-associated cancer cell lines and NIKS, an experiment was designed to examine the consequences on E6/E7 expression upon induction of endogenous CLK3 expression. Previously, studies have demonstrated CLK3 to be one of the genes overexpressed under hypoxic conditions (Bowler, et al., 2018). Here, hypoxic conditions were used to upregulate endogenous CLK3 expression in HPV-positive cancer cell lines and NIKS. It was hypothesised hypoxia-induced upregulation of CLK3 would result in increased levels of HPV E6/E7.

Hypoxia refers to a state when the demand of oxygen by cells is not met by the supply. Such hypoxic conditions activate various transcription factors, most importantly the hypoxia inducible factor (HIF), which plays a central role in the cellular hypoxic response. HIF consists of α and β subunits, with three α subunits (HIF-1α, HIF-2α, HIF-3α) and two β subunits (HiF-B and ARNT2) being identified thus far. The HIF-α subunit is regulated in an oxygen dependent manner, and is hydroxylated by dioxygenases called prolyl-4-hydroxylase (PHD) under normoxic conditions (Epstein, et al., 2001, Hirota, et al., 2005). This hydroxylation triggers an ubiquitination reaction by the von Hippel-Lindau protein (pVHL), resulting in proteolysis of HIF-α, which has a half-life of 6-8 minutes (Ivan, et al., 2001, Jaakola, et al., 2001). However, under hypoxic conditions, as a result of reduced hydroxylation by the oxygen sensitive PHD, HIF-1α becomes more stable, and translocates into the nucleus where it interacts with HIF-1β, to form the HIF-1 heterodimer. This transcription factor induces the transcription of hundreds of genes, including CLK3 (Dominguez Sanchez, 2017).

Hypoxia is a common feature in tumours, with the average oxygen levels being 1-2% and below in majority of the tumours. Cervical cancer has a mean oxygen level of 1.2% (Höckel, et al., 1991, Höckel, et al., 2001, Vaupel, et al., 2007). Tumours with increased hypoxic regions are more difficult to treat. Typically, radiotherapy involves using high-energy photons to directly target tumour sites and create reactive oxygen species that induce apoptosis by forming DNA-damaging molecules. However, in regions of very low oxygen (anoxia/severe hypoxia), radiotherapy is often found to be ineffective due to induction of cellular DNA damage responses, and usually results in tumour persistence and the development more aggressive tumours (Wouters, et a., 2007, Graham, et al., 2018). Severe hypoxia induces stress-mediated DNA damage repair (DDR), which also involves the stabilisation and activation of the p53 tumour suppressor gene. Hypoxia induced p53 has been shown to upregulate various genes (Lesczynska, et al., 2015). Evidence from this study suggests that CLK3 is a potential downstream target of this hypoxia induced p53-dependent targets. Expression of CLK3 was found to correlate significantly with these targets, and as a result, we opted to focus on inducing CLK3 in severe hypoxia.

In addition, hypoxia has been known to aid in the development of resistance against a number of chemotherapy agents (Teicher, et al., 1990). This resistance arises due to a number of factors (Brown, 1999). As a result of poor vasculature, drugs reaching hypoxic regions is low. In addition, these agents target rapidly dividing cells, however hypoxia slows down the cell cycle process, which allows tumour cells to escape treatment. Hypoxia also upregulates the expression of proteins that are involved in drug resistance, such as

metallothionein-IIA (MT-IIA) in cisplatin resistance, and periostin involved in developing resistance against arsenic trioxide (Murphy, et al., 1994, Liu, et al., 2017). Understanding whether CLK3 overexpression takes place in HPV-associated cancers under hypoxic conditions, and how this effects E6/E7 levels could prove to be significant when considering the kinase as a therapeutic target.

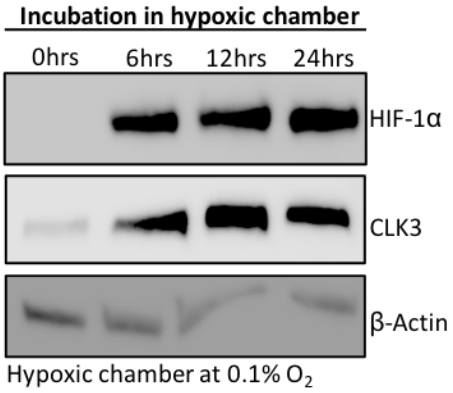



**Figure 3.4.1 Hypoxia induced in SiHa in a hypoxic chamber**
Western blot of SiHa cells maintained in a hypoxic chamber (0.1% $O_2$) for 24hrs. 50µg of protein was loaded for each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of CLK3 is approximately 59kDa. HIF-1α has a molecular weight of 120kDa. β-Actin was used as a loading control. PageRuler plus prestained protein ladder was used as a molecular marker.

Initially cells were exposed to hypoxia using a hypoxic chamber. Most of the optimisation was completed using SiHa cells. The oxygen level within the chamber was maintained at 0.1%, creating an anoxic environment. Cells were maintained in these conditions for 24hrs, with samples being obtained at various intervals. HIF-1α antibody staining confirmed the presence of stable HIF-1α at the first time point of 6hrs. It was also possible to observe a significant increase in CLK3 expression at 6hrs, with the levels remaining high throughout the 24hr period (Figure 3.4.1). Due to technical issues, the

hypoxic chamber became unavailable for the remainder of this project, therefore other means of inducing hypoxia had to be investigated.

Chemicals have been widely used to mimic the effects of hypoxia by stabilising HIF-1α. Cobalt (II) chloride and deferoxamine are considered two of the most effective hypoxic mimetics (Zhang, et al., 2011, Huang, et al., 2014). $CoCl_2$ is thought to inhibit the activity of PHDs by displacing $Fe^{2+}$ with $Co^{2+}$ in the PHD active site, thereby allowing HIF-1α to stabilise. This is the most widely accepted mechanism of $CoCl_2$, referred to as replacement hypothesis. Typically, HIF-1a stabilisation is observed within 2 hours of $CoCl_2$ treatment, and continues to last for 48hrs, with dosage ranging from 100 to 300μM (reviewed in Muñoz-Sánchez and Chánez-Cárdenas, 2018). Doses above 200-300μM have been observed to exert toxic cellular effects that result in cell death due to ROS DNA damage (Lison, et al., 2001, Mahey, et al, 2016). DFO, on the other hand, is an $Fe^{2+}$ chelator, and causes the inhibition of PHDs, thereby increasing the accumulation of HIF-1α. Studies have established the apoptotic effects of DFO at higher concentrations in various cell lines (Simonart, et al., 2000, Fan, et al., 2001, Guo, et al., 2006).

Following initial optimisation, 150μM of DFO, and 100-200μM of $CoCl_2$ were considered the optimum concentrations for this study to minimise any toxic effects. Further optimisation was carried out in SiHa cells, and overexpression of CLK3 was observed with both $CoCl_2$ and DFO (Figure 3.4.2). However, from the protein analysis, DFO treatment was found to resemble the results from the hypoxic chamber, with robust overexpression of

CLK3 being observed from 12hrs of treatment. qRT-PCR analysis of various HPV positive cancer cell lines found an overall increase in the mRNA levels of CLK3 following a 24hr treatment with DFO (Figure 3.4.2, C). Therefore, it was possible to conclude that CLK3 is upregulated under hypoxic conditions/mimetic treatment.

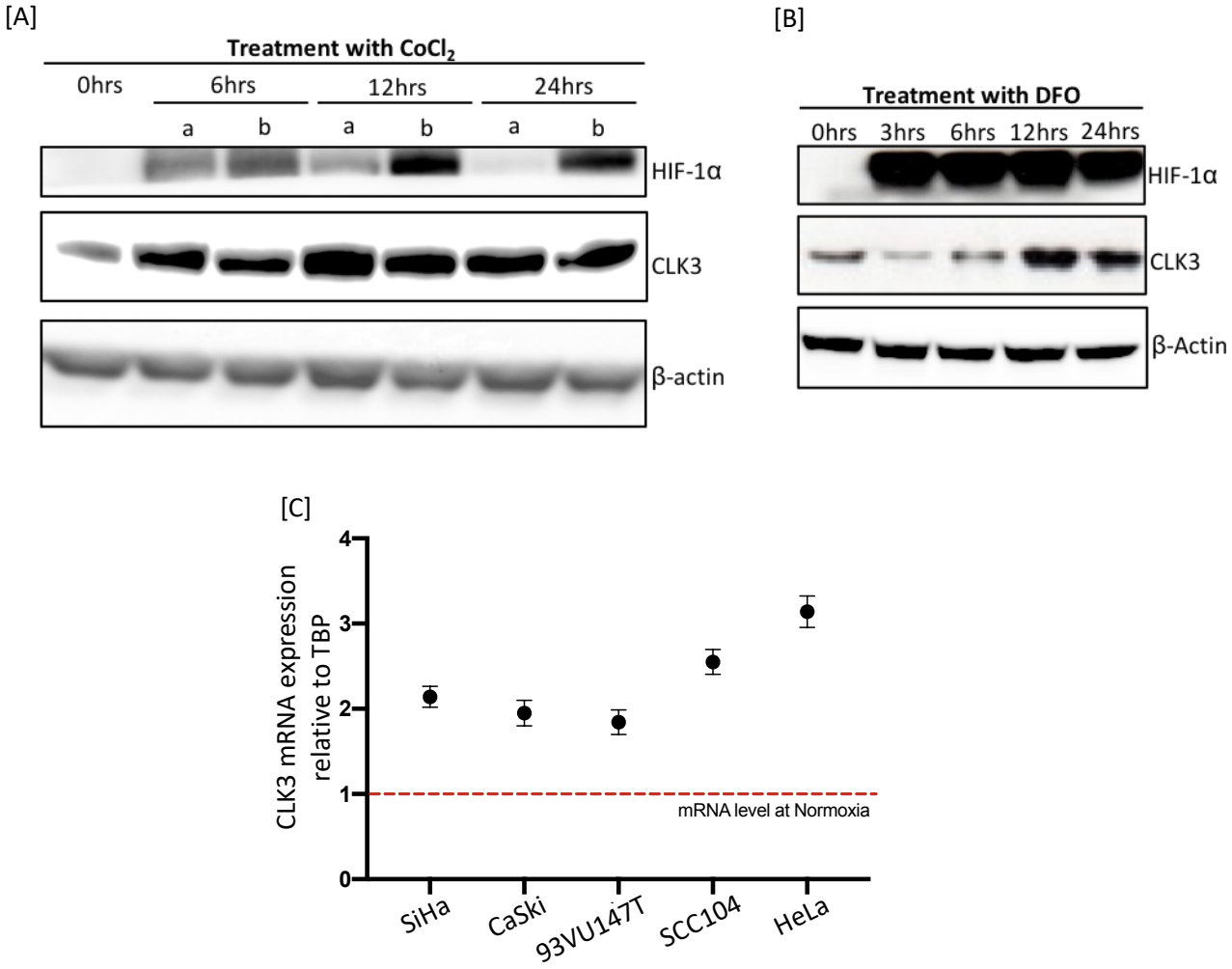



**Figure 3.4.2 Chemically induced hypoxia in SiHa**
Western blot of SiHa cells treated with either 100µM(a) or 200µM(b) $CoCl_2$ (A), or 150µM of DFO (B). 50µg of protein was loaded for each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of CLK3 is approximately 59kDa. HIF-1α has a molecular weight of 120kDa. β-Actin was used as a loading control. PageRuler plus prestained protein ladder was used as a molecular marker. C) qRT-PCR analysis of CLK3 mRNA following 24hrs of 150µM DFO treatment. Red line marks the control, cells grown in normoxic conditions with 25mM glucose. Three

HPV associated cervical cancer cells (SiHa, CaSki, HeLa), and two HPV associated head and neck cancer cells (UMSCC104, 93VU147T) were analysed.

SiHa and CaSki cells were initially treated with $CoCl_2$ to determine if induction of CLK3 results in upregulation of HPV16 E7. $CoCl_2$ treatment over 24hrs resulted in minimal changes in the overall level of CLK3 and E7 (Figure 3.4.2, A). Previous studies have determined that HPV E6/E7 expression was significantly reduced in hypoxic conditions only when the cells were starved of glucose (Hoppe-Seyler, et al., 2017, Bossler, et al., 2019). It was therefore hypothesised that a loss of CLK3 would be observed upon glucose starvation.

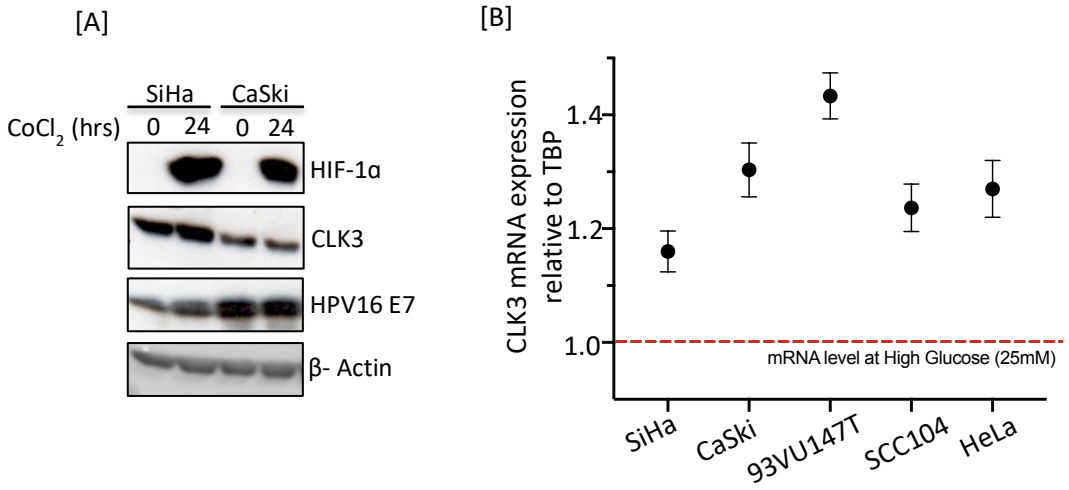



**Figure 3.4.3 $CoCl_2$ treatment/glucose starvation of HPV-associated cancer cell lines**
A) 200µM $CoCl_2$ treatment of SiHa and CaSki cells over 24hrs. 50µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of CLK3 is approximately 59kDa, HIF-1α has a molecular weight of 120kDa, E7 is approximately 18kDa in molecular weight β-Actin was used as a loading control. PageRuler plus prestained protein ladder was used as a molecular marker. B) qRT-PCR analysis of CLK3 mRNA following 24hrs in the absence of glucose. Red line marks the control, cells grown in normoxic conditions with 25mM glucose. Three HPV associated cervical cancer cells (SiHa, CaSki, HeLa), and two HPV associated head and neck cancer cells (UMSCC104, 93VU147T) were analysed.

To determine this, various HPV positive cells were starved of glucose over period of 24hrs, and qRT-PCR analysis was performed. CLK3 mRNA was however found to be increased under starved conditions (Figure 3.4.3, B). Following this, SiHa and CaSki cells were starved of glucose and treated with DFO for a period of 24hrs. The conditions were replicated from the previous experiments carried out by Hoppe-Seyler and Bossler. Protein analysis found CLK3 expression to be increased in glucose starved conditions, with similar levels being expressed under DFO treatment (figure 3.4.4, A). The levels of HPV16 E7 were significantly reduced in both cell lines under glucose starved conditions with DFO treatment, similar to the results of the previous studies (Hoppe-Seyler, et al., 2017, Bossler, et al., 2019). qRT-PCR analysis, however presented contradictory results, with minimal change in E6/E7 mRNA levels following $CoCl_2$ treatment, and an increase in E6/E7 mRNA levels following DFO treatment (figure 3.4.1, B). It is possible that upregulation of CLK3 observed with DFO treatment caused the increase in E6/E7 mRNA, and other confounding factors such as decreased translation or destabilisation of E7 could be responsible for the loss of HPV E7 at protein level, despite the increase in mRNA levels. Hoppe-Seyler and colleagues observed a loss of E6/E7 mRNA in both SiHa and CaSki at 1% oxygen levels, and the contradictory results observed here could be due to the difference in cellular response at 0.1% oxygen levels (DFO treatment) compared to 1% in the previous study. It is also likely that chemically induced hypoxia will vary on a molecular level to oxygen deprived hypoxic conditions, resulting in differences in protein expression. This has been observed in previous studies (Novak, et al, 2019). It was therefore not possible to conclusively state that hypoxia-mediated upregulation of CLK3 results in increased E6/E7 mRNA levels but we can

conclude that DFO treatment caused upregulation of both CLK3 protein expression and E6/E7 mRNA.

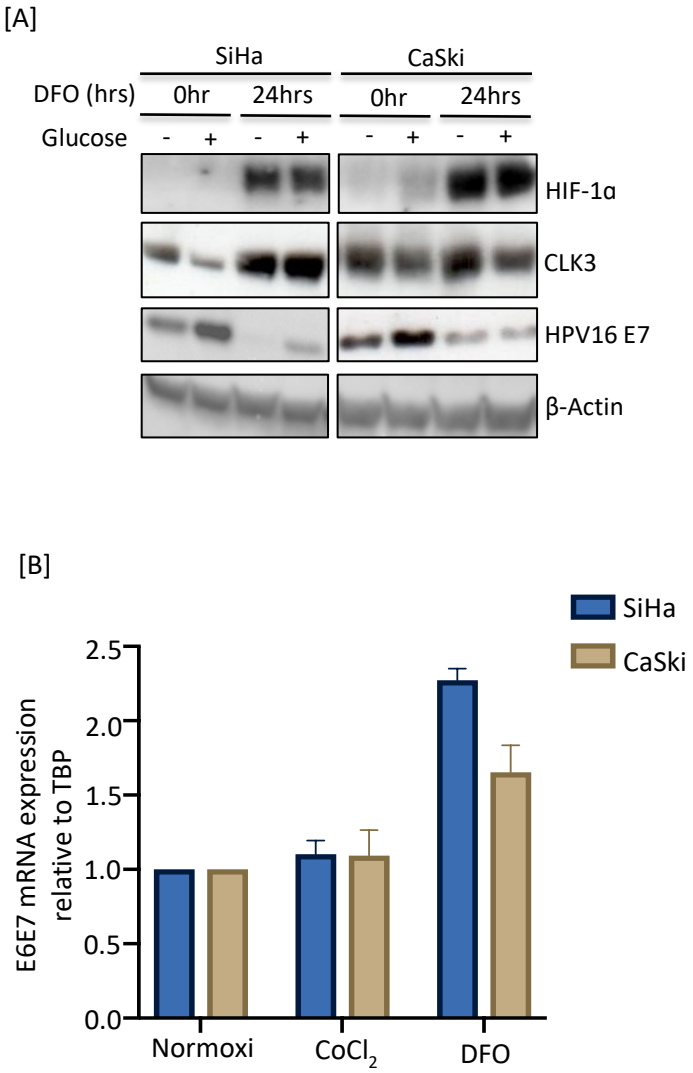



**Figure 3.4.4 Overexpression of CLK3 under chemically induced hypoxic conditions**
A) 150µM DFO treatment of SiHa and CaSki cells over 24hrs, with the presence and absence of 25mM glucose. 50µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of CLK3 is approximately 59kDa, HIF-1α has a molecular weight of 120kDa, E7 is approximately 18kDa in molecular weight β-Actin was used as a loading control. PageRuler plus prestained protein ladder was used as a molecular marker. C) qRT-PCR analysis of SiHa and CaSki cells following RNA extraction. Cells were either treated with 200µM $CoCl_2$ or 150µM DFO over 24hrs. Glucose concentration was maintained at 25µM throughout. E6/E7 levels were measured relative to TBP, and represented as fold change compared to normoxic conditions.

## 3.5 Transient overexpression of CLK3 upregulates E7

Overexpression of CLK3 faced various obstacles, with poor transfection efficiency and lack of stable expression resulting in no progress being made with rescue experiments. Induction of endogenous CLK3 by hypoxia mimetics was successful, however, it is impossible to directly relate changes in CLK3 to changes in E6/E7 due to the pleiotropic effects of hypoxia on cells. Here, immunostaining of cells with E7/p16 antibodies following transient transfection of CLK3 plasmid was considered as an alternate method to observe the rescue effects of CLK3 on E7 protein expression, *in situ*, thus overcoming the limitation of low transfection efficiency. The $CLK3^{-/-}$ NIKS provided a model without any background CLK3 expression, therefore any changes observed would be as a direct result of the transfected WT CLK3. The use of CLK3-GFP allowed successfully transfected cells overexpressing the kinase to be identified easily, and staining the cells with E7/p16 antibodies was used to assess the effect of CLK3 on E7 expression, on. A cell-by-cell basis. It was hypothesised that overexpression of WT-CLK3 will result in the upregulation of E7/p16, therefore validating the CLK3 knockdown/knockout effects previously seen.

Following transfection with pEGFP-N1-CLK3, cells were incubated for 48hrs, and then staining with HPV16 E7 antibody. Cells expressing GFP were analysed under a fluorescent microscope, and this confirmed the predicted increase in E7 protein levels, with significantly increased E7 staining intensity observed (Figure 3.5.1). Importantly, the use of the kinase-inactive CLK3 plasmid (pEGFP-N1-K186R-CLK3) demonstrated the importance of the CLK3 kinase function in E7 upregulation, as minimal differences were observed in these

cells compared to the control (pEGFP-N1). The observations were quantified using Image J analysis (Figure 3.5.2).

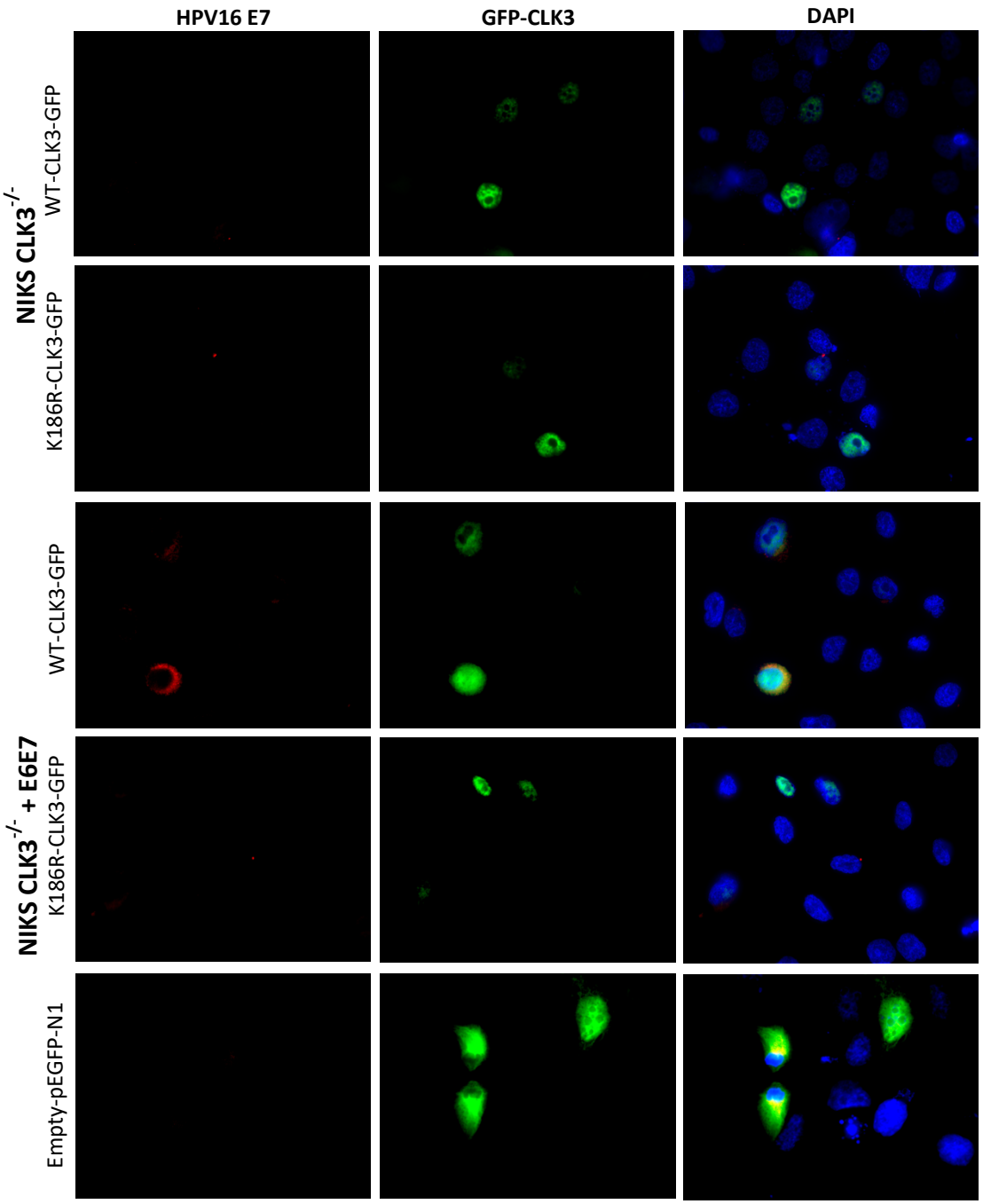



**Figure 3.5.1 E7 IF staining following transient overexpression of CLK3 in NIKS**
IF staining of NIKS with HPV16 E7 antibody following transient transfection of pEGFP-N1-CLK3-GFP, pEGFP-N1-KD-CLK3-GFP or Empty-pEGFP-N1 (Control) for 48hrs. HPV16 E7 is in red, cells expressing the plasmid are green (GFP), and DAPI staining indicates the nuclei (blue). Microscopy images were at 60X magnification.

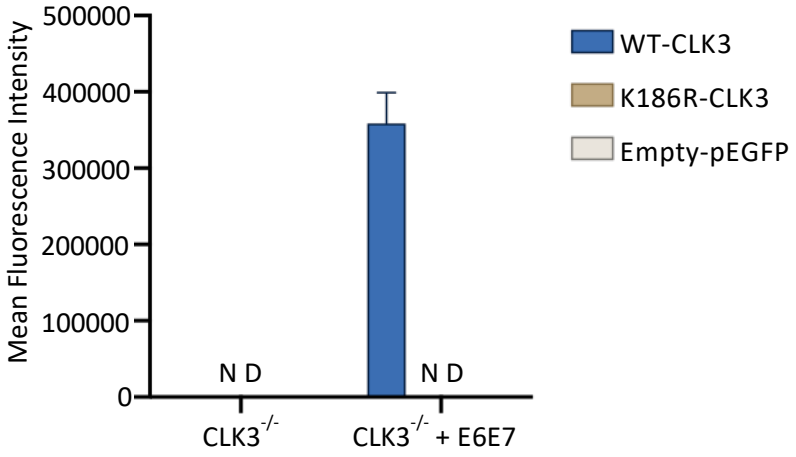



**Figure 3.5.2 Quantification of E7 immunostaining**
Image J quantification, with the mean fluorescence intensity of 10 cells being represented by each bar. E7 Antibody produces minimal background stains; therefore, signal was not detectable (N.D) for Empty-pEGFP control, K186R, and $CLK3^{-/-}$ cell line.

p16 levels showed a significant increase of up to 10-fold in mean fluorescence intensity compared to the K186R-CLK3 transfected cells (Fig 3.5.3, 3.5.4). The CLK3 KO NIKS with URR E6/E7 showed minimal E7 expression, and no expression of E7 was observed following transfection with the K186R-CLK3 plasmid. This provided a model with no background level staining to show the significant increase in E7 levels following transfection of WT-CLK3, which is emphasised in image J quantification (Fig 3.5.2). Supplementary figure 8.1.3 provides additional microscopy images from this experiment.

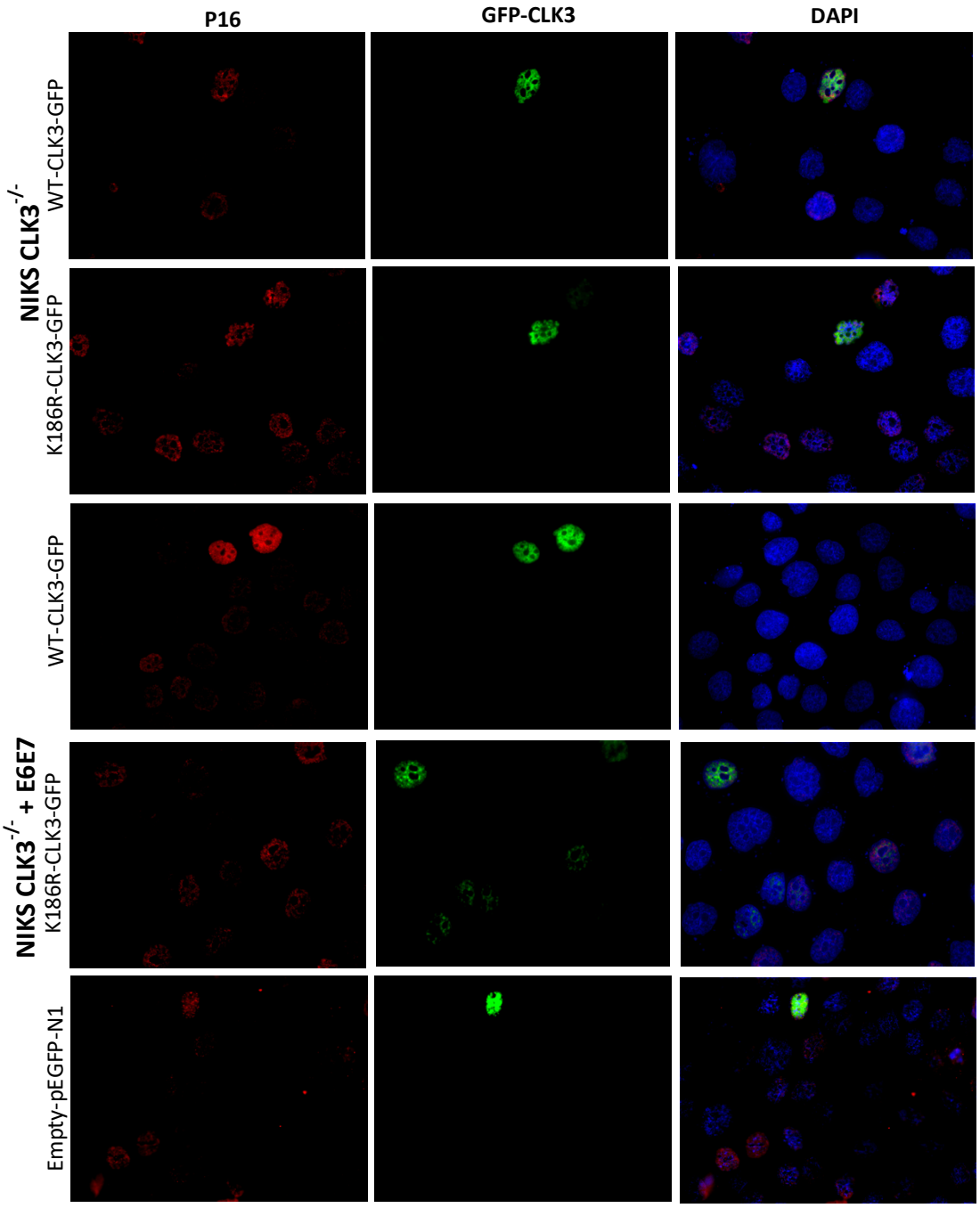



**Figure 3.5.3 p16 IF staining following transient overexpression of CLK3**
IF staining of NIKS with p16 antibody following transient transfection of pEGFP-N1-CLK3-GFP, pEGFP-N1-KD-CLK3-GFP or Empty-pEGFP-N1 (Control) for 48hrs. NIKS CLK3 KO and NIKS CLK3 KO E6/E7 were used. p16 is stained in red, cells expressing the plasmid are green (GFP), and DAPI staining indicates the nuclei (blue). Microscopy images were at 60X magnification.

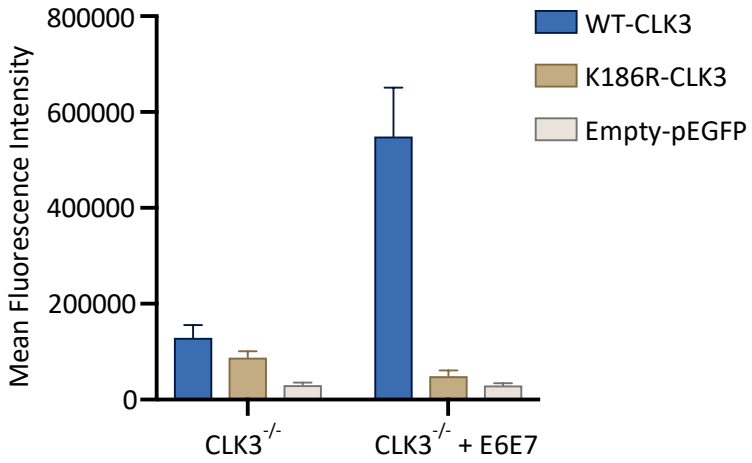



**Figure 3.5.4 Quantification of p16 immunostaining**
Image J quantification, with the mean fluorescence intensity of 10 cells being represented by each bar.

In addition to the stable transfection of HPV16 URR E6/E7, the WT and CLK3 KO NIKS were also stably transfected with the whole HPV16 genome. HPV16 genome Transfection and stable selection was completed by Prof John Doobar and colleagues (Department of Pathology, University of Cambridge) as described previously (Wechsler, et al., 2012) and the cells were returned to us for the CLK3 transfection experiments. Protein and mRNA analysis of these cells further reinforced the relationship between CLK3 and E6/E7, with significant reductions in E6/E7 mRNA and E7 protein being observed in CLK3$^{-/-}$ NIKS versus WT NIKS (Fig 3.5.5 A, B). In addition, significant loss of p16 mRNA levels were seen, further supporting the observations.

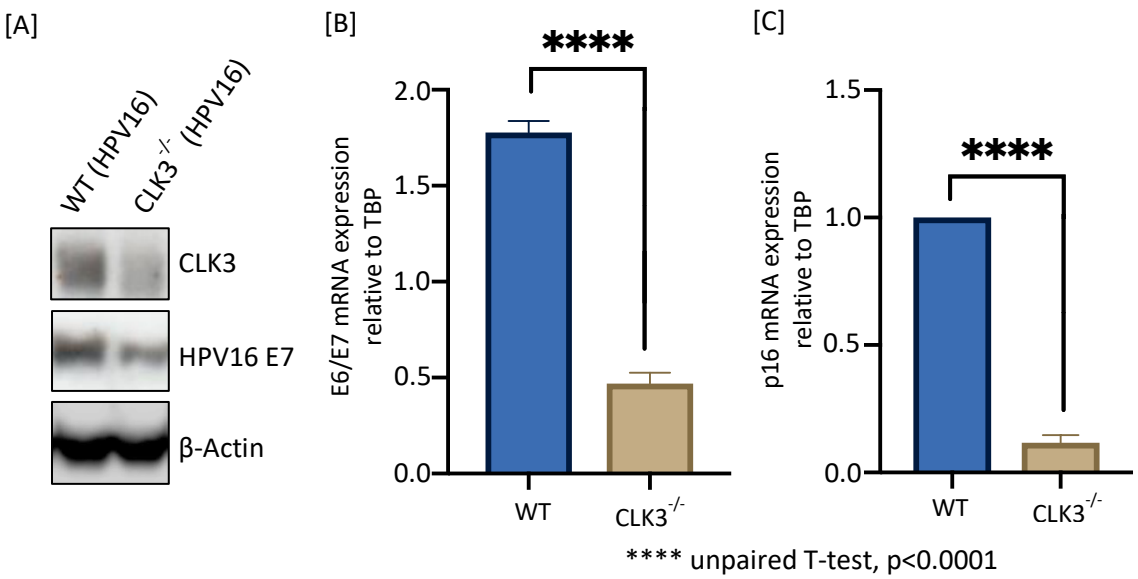



**Figure 3.5.5 WT and CLK3$^{-/-}$ NIKS with stably transfected HPV16 genome**
A) western blot. 50µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of CLK3 is approximately 59kDa. E7 is approximately 18kDa in molecular weight. β-Actin was used as a loading control. PageRuler plus prestained protein ladder was used as a molecular marker. B, C) qRT-PCR analysis of E6/E7 and p16 mRNA relative to TBP in wildtype and CLK3 KO NIKS with URR E6/E7.

Furthermore, transient transfection of WT and K186R CLK3 plasmids into CLK3$^{-/-}$ NIKS with the stably transfected HPV16 genome was carried out. IF staining with HPV16 E7 and p16 antibodies produced similar data to that of NIKS transfected with HPV16 URR E6/E7, with increased E7 and p16 staining being observed following transfection of WT-CLK3, which is not observed with the K186R-CLK3 (Figures 3.5.6, 3.5.7). This adds further support to the relationship between CLK3 and HPV16 E6/E7. Supplementary figure 8.1.3 provides additional microscopy images from this experiment.

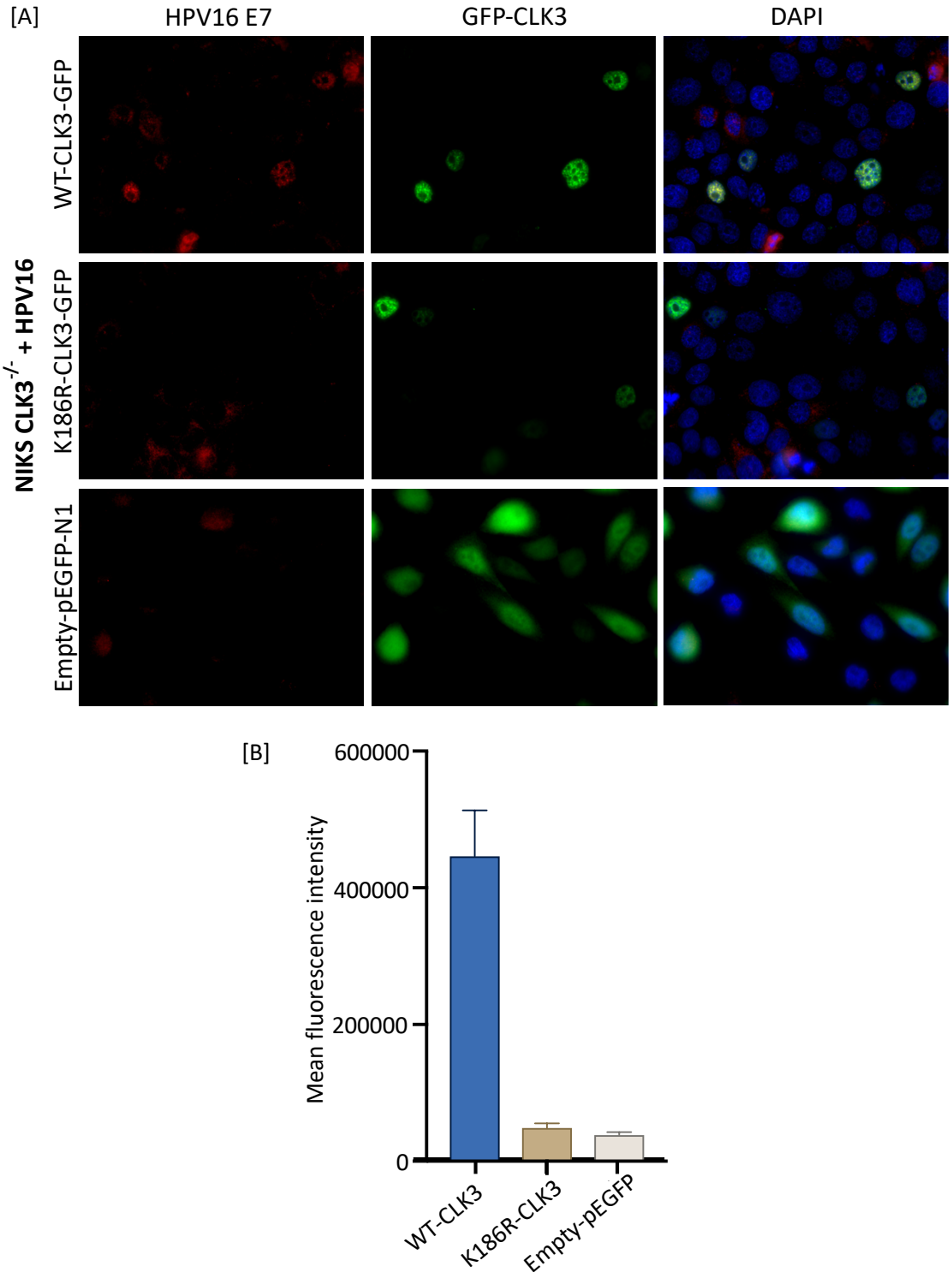



**Figure 3.5.6 HPV16 E7 IF staining of NIKS (+HPV16) transfected with CLK3**
(A) IF staining of WT and CLK3$^{-/-}$ NIKS with HPV16 E7 antibody (red) following 48hrs of transfection of WT-CLK3-GFP, K186R-CLK3-GFP, and Empty-pEGFP-N1 (control) plasmids. Cells expressing the plasmid are green (GFP), DAPI staining indicates the nuclei (blue). Microscopy images were at 60X magnification. (B) Image J analysis was used to quantify the fluorescence, with mean fluorescence intensity of 10 cells being represented by each bar.

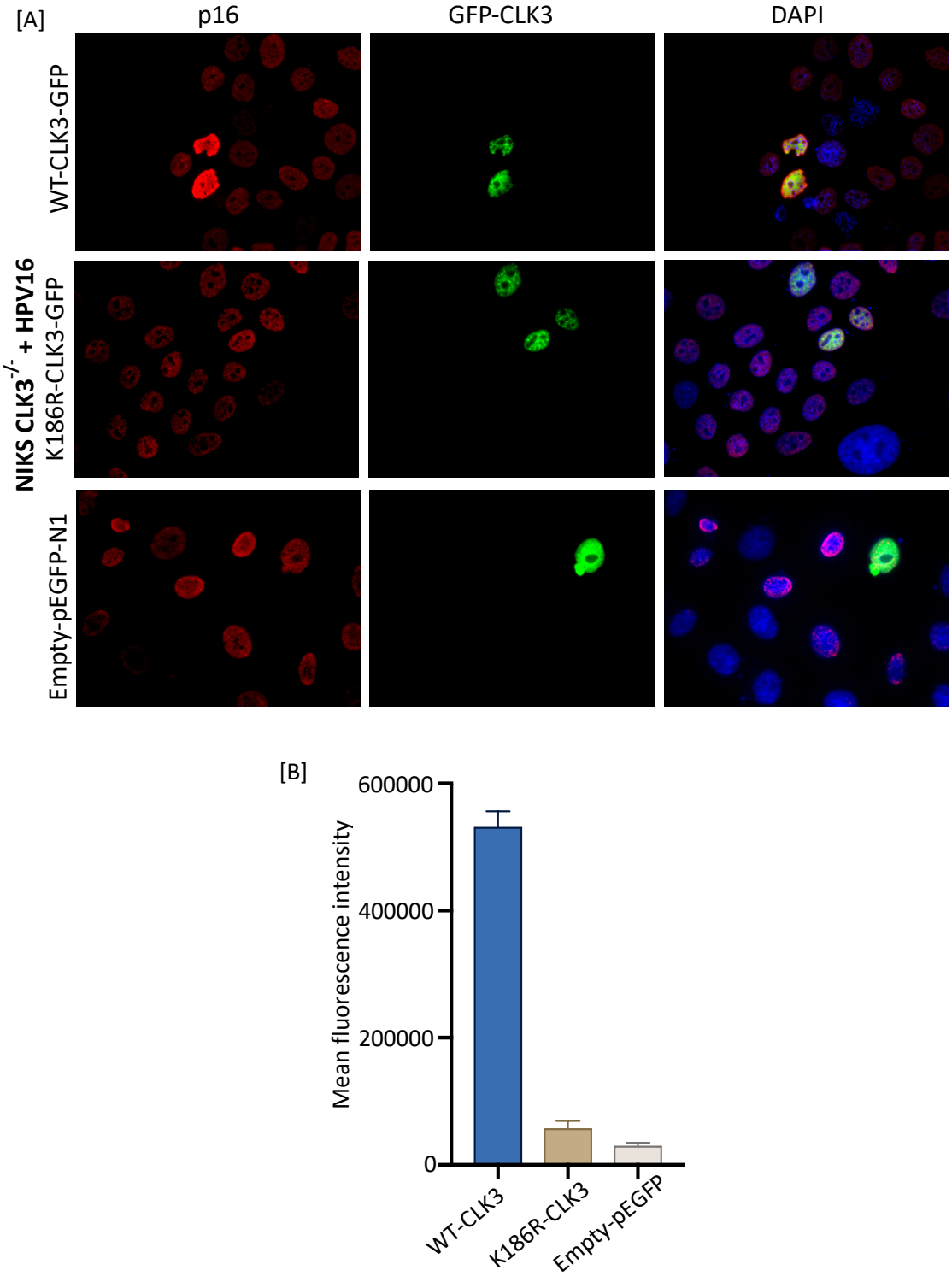



**Figure 3.5.7 p16 IF staining of NIKS (+HPV16) transfected with CLK3**

(A) IF staining of WT and CLK3$^{-/-}$ NIKS with p16 antibody (red) following 48hrs of transfection of WT-CLK3-GFP, K186R-CLK3-GFP, and Empty-pEGFP-N1 (control) plasmids. Cells expressing the plasmid are green (GFP), DAPI staining indicates the nuclei (blue). Microscopy images were at 60X magnification. (B) Image J analysis was used to quantify the fluorescence, with mean fluorescence intensity of 10 cells being represented by each bar.

## 3.6 Discussion

The main aim of this chapter was to perform CLK3 rescue experiments to observe the effects of CLK3 overexpression on HPV16 E6/E7. The use of CRISPR to knockout CLK3 in NIKS with integrated HPV16 URR E6/E7 aided in confirming two key elements: one is that cells were able to survive and proliferate without CLK3, making it a desirable candidate for targeted therapy, and second is that CRISPR allowed for total loss of CLK3 expression in cells and it was possible to observe the striking difference in E6/E7 and p16 mRNA levels, as well as E7 and p16 protein levels resulting from this.

Multiple attempts were made at overexpressing CLK3. Initially, CO-CLK3 was transfected into various HPV positive cancer cell lines using the pEGFP plasmid as a vector. Following antibiotic selection, no overexpression of CLK3 was observed. It is likely that prolonged expression of CLK3 was not tolerated in the cells, which is also true for the k186R mutant. This would suggest that CLK3 has additional, uncharacterised roles independent of its kinase function. Following various unsuccessful attempts at creating stable transfected cell lines using FACS sorting and electroporation, attempts were made to overexpress endogenous levels of CLK3. This was successful, with overexpression of CLK3 being observed at mRNA and protein levels. Chemically induced hypoxic conditions with DFO resulted in the overexpression of E6/E7 mRNA levels in two HPV16 positive cervical cancer cell lines. Previous studies where oxygen levels were maintained at 1% demonstrated a loss of E6/E7 mRNA (Hoppe-Seyler, et al., 2017, Bossler, et al., 2019). The differences observed here could result from the different cellular responses that occur under 1% vs 0.1% (or DFO) $O_2$

conditions. $CoCl_2$ treatment is understood to mimic a hypoxia response that is similar to 1% $O_2$ concentrations (Reviewed in Muñoz-Sánchez and Chánez-Cárdenas, 2018). There were mimimal changes in CLK3 and E6/E7 levels at this condition. This supports the notion that upregultion of CLK3 was causing the increased levels of E6/E7 following DFO treatment. There is a noticeable decrease in ribonucleotide reductase activity between 0.05-0.5% $O_2$ levels, resulting in reduced DNA synthesis and cell cycling (Ebbesen, et al., 2004). At 0.1% $O_2$ concentration, protein accumulation is greatly diminished, partly as a result of reduced phosphorylation of the eIF2R inititation factor, and partly due to an increase in protein degredation (Kraggerud, et al., 1995). In addition, mRNA translation is greatly supressed, along with ATP production. In addition, as previously discussed, severe hypoxic conditions lead to p53-dependent upregulation of various genes, one of which is CLK3. It is likely that other genes being upregulated could also affect E6/E7 levels, leading to the difference in E6/E7 levels between previous two studies and the observations made here.

This was followed by transient transfection of GFP-WT-CLK3 in CLK3 KO NIKS with HPV16 URR E6/E7, followed by IF staining with HPV16 E7 and p16 antibodies. Increased levels of HPV E7 and p16 were observed following CLK3 overexpression, further validating the relationship between CLK3 and E6/E7. Importantly, no overexpression of E7 or p16 were observed in cells transfected with the mutant kinase-inactive CLK3, highlighting the importance of the kinase function in regulating E6/E7. This was also repeated with CLK3 KO NIKS stably transfected with the complete HPV16 genome, and similar results were

observed with both E7 and p16 staining. The next chapter will focus on the interactions between CLK3 and E6/E7, to further understand the role of CLK3 in E6/E7 regulation.

# Chapter 4

# Investigating the mechanism by which CLK3 regulates E6/E7 expression

# 4 Investigating the mechanism by which CLK3 regulates E6/E7 expression

## 4.1 Introduction

Date presented in the previous chapter demonstrate the importance of CLK3 in maintaining E6/E7 expression in NIKS. Our next objective was to identify the CLK3 substrate(s) and downstream mechanisms involved.

As discussed in chapter 1, CLK3 is known to phosphorylate SR proteins (Howell, et al., 1991, Colwill, et al., 1996, Nayler, et al., 1997, Bullock, et al., 2009) and to increase the stability of c-MYC by activating the ubiquitin-specific protease, USP-13 (Zhou, et al., 2020). SRFS2 has been implicated in the control of E6/E7 expression through stabilisation of the E6/E7 mRNA (MacFarlane, et al., 2015) but previous nascent RNA labelling experiments in our laboratory suggested the loss of E6/E7 mRNA following CLK3 knockdown in HPV-positive cancer cell lines was due to a decrease in the rate of E6/E7 transcription, rather than mRNA stability (Williams, 2021). We hypothesized that given SRSF2's known role in transcription elongation through the release of paused RNA polymerase II (Lin, et al., 2008, Ji, et al., 2013), it might act to increase E6/E7 transcription (in a CLK3-dependent manner) in addition to its previously-described effects on mRNA stability.

c-MYC controls gene expression at multiple levels: it acts as a transcription factor, binding to E-box elements in target promoters with its heterodimerisation partner (MAX) (Blackwood, et al., 1991, E-box ref); it is required for the release of RNA polymerase II

following promoter-proximal pausing (Rahl, et al., 2011), therefore playing a role in transcription elongation and it is also involved in the upregulation of mRNA translation (Cole and Cowling, 2008, Elkon, et al, 2015). We identified a putative E-box sequence within the HPV16 URR (Supplementary Figure 8.1.2), and therefore also tested the hypothesis that CLK3 acts via c-MYC to directly increase the activity of the promotor, resulting in increased E6/E7 transcription.

A third possibility is that CLK3 acts via an as-yet unknown substrate to mediate its effects on E6/E7 expression. Figure 4.1.1 summarises the three main possibilities we initially considered concerning the mechanism by which CLK3 regulates HPV oncogene expression. The experiments described in this chapter were focused on testing and refining these hypotheses.

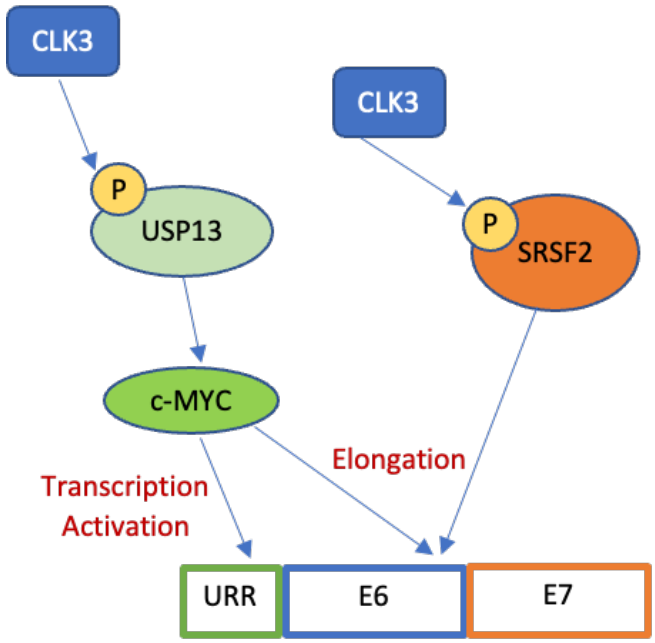



**Figure 4.1.1 Hypothesised pathway of CLK3 interaction with E6/E7**

## 4.2 CLK3 phosphorylates and redistributes SRSF2

Firstly, an experiment was designed to confirm the relationship between CLK3 and SRSF2. The pEGFP-N1 plasmids containing the WT-CLK3 and K186R-CLK3 inserts were transfected into various HPV positive cancer cell lines and stained with SRSF2 antibody 48hrs of post-transfection. The aim of this experiment was to determine the functionality of the expressed kinase, as well as the effect of CLK3 overexpression on SRSF2. Previously, studies have established the significance of the kinase activity of CLKs in phosphorylating and redistributing SR proteins (Colwill, et al., 1996, Nayler, et al., 1997, Bullock, et al., 2009). This study attempts to validate the previous observations, focusing on the importance of the kinase function of CLK3 on SRSF2.

SRSF2, like other SR proteins, is stored in nuclear speckles within the nucleus (Misteli, et al., 1997). CLKs are known to phosphorylate the RS2 domain of SR proteins and dissolve the nuclear speckles, redistributing the SR-proteins to the site of transcription (Colwill, et al., 1996). The staining pattern of SRSF2 seen in 93VU147T and SiHa cells confirms their existence as speckles within the nucleus. Both cell lines display a complete dissolution of the nuclear speckles when transfected with the WT-CLK3-GFP, confirming the expression of a functional GFP fusion protein following transfection of the pEGFP-N1 plasmid (Fig 4.2.1). Equally important are the observations made when the cells were transfected with the K186R-CLK3-GFP plasmid, where the SRSF2 speckles remain intact (Fig 4.2.1 B, D). This

confirms the need for the CLK3 kinase function for the redistribution of SRSF2, as seen previously with other CLK family members.

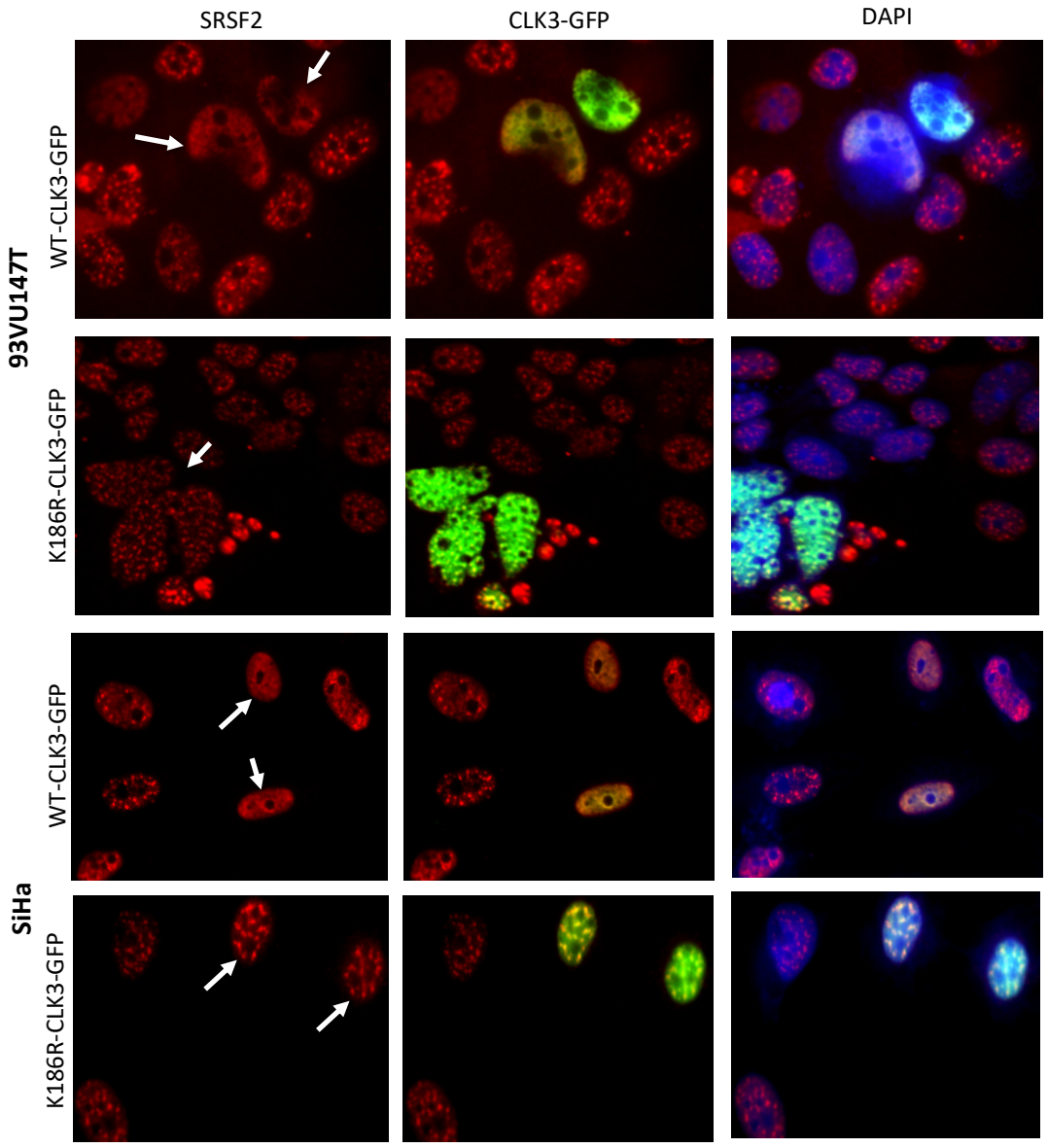



**Figure 4.2.1 Phosphorylation and redistribution of SRSF2 by CLK3**
IF staining of 93VU147T and SiHa cell lines with SRSF2 antibody following transient transfection of pEGFP-N1-CLK3-GFP and pEGFP-K186R-CLK3-GFP for 48hrs. SRSF2 staining is red, cells expressing the plasmid are green (marked with arrow), and DAPI staining indicates the nuclei (blue). Microscopy images were at 60 X magnification.

Moving forward, to determine if the loss of CLK3 had any effect on the expression of SRSF2, western blot and qRT-PCR analysis was completed using the WT and CLK3 KO NIKS. It was hypothesised that there will be minimal difference in SRSF2 mRNA levels, and most likely on protein level as well. As expected, qRT-PCR analysis suggested no significant difference in the mRNA levels of SRSF2 between the $CLK3^{-/-}$ NIKS and the WT NIKS, and no difference in SRSF2 protein levels were observed from the western blot (Figure 4.2.2 A, B). CLK3 was believed to be only involved in post-translational modifications of SRSF2, therefore the loss of CLK3 is likely to affect the phosphorylation of SRSF2, and quite possibly it speckling patterns within the nucleus. As expected (McFarlane et al.,), NIKS stably expressing HPV16 E6/E7 displayed increased SRSF2 expression (Figure 4.2.2 B) and the reduced levels of SRSF2 in $CLK^{-/-}$-E6/E7 NIKS compared to CLK3 WT E6/E7 NIKS are likely due to the reduced E6/E7 expression in the absence of CLK3 (Figure Chapter 3).

To investigate this, IF staining of WT and CLK3 KO NIKS was carried out with SRSF2 antibody. The staining patterns revealed minimal difference, with Image J quantification showing no significant difference in mean fluorescence intensity (Fig 4.2.2 C, D). In addition, there was also no significant difference in the number/intensity of SRSF2 speckles between the WT and $CLK3^{-/-}$ cells. The cells were then stained with phospho-SRSF2 specific antibody to identify any difference in the phosphorylation level. However, IF staining and Image J analysis revealed no significant difference in the phosphorylation levels of SRSF2 between the WT and CLK3 KO NIKS.

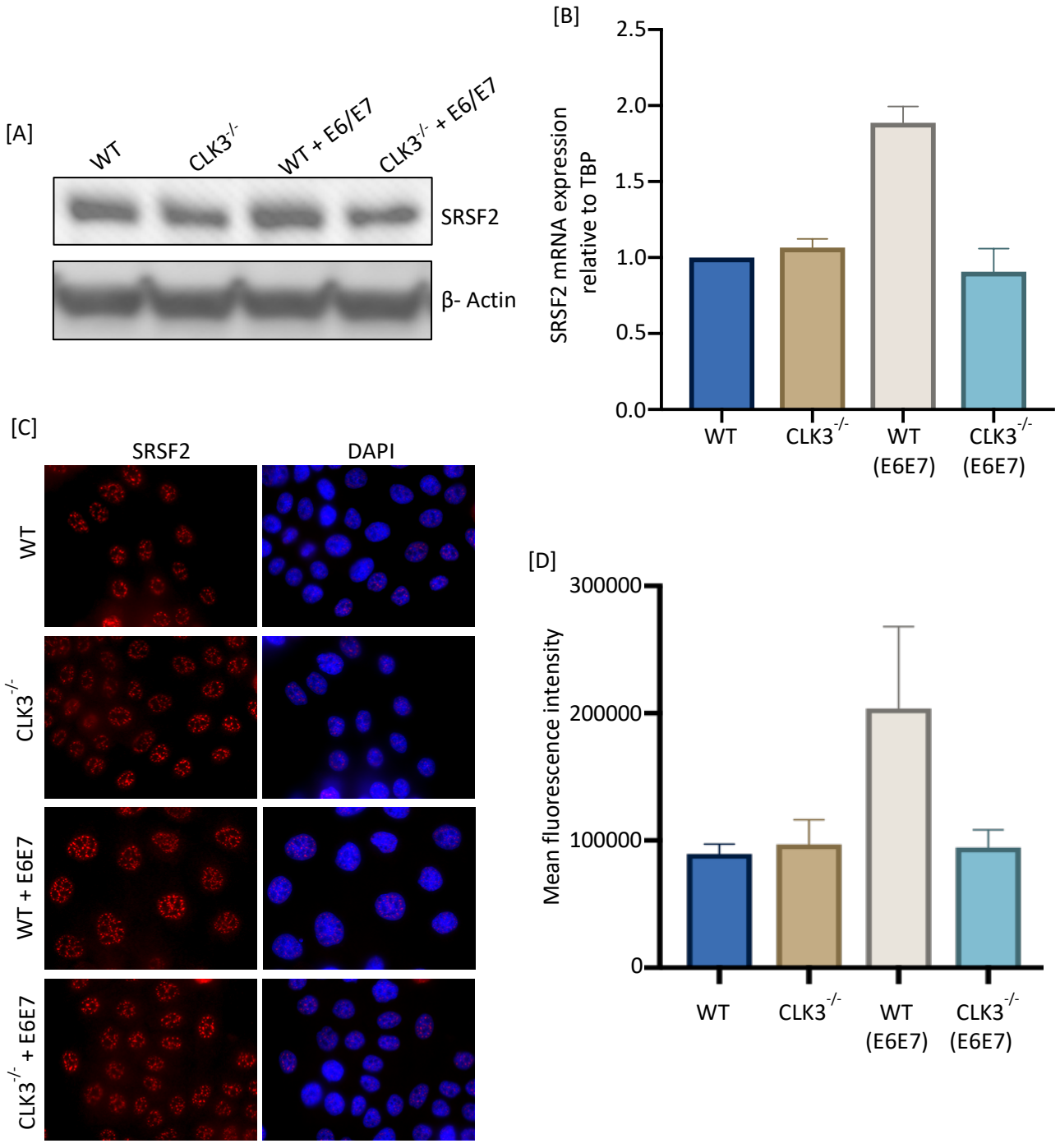



**Figure 4.2.2 CLK3 KO effect on SRSF2**

A) Western blot. 40µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of SRSF2 is approximately 35kDa. β-Actin was used as a loading control. B) qRT-PCR analysis of SRSF2 mRNA relative to TBP in wildtype and CLK3 KO NIKS with URR E6/E7 C) IF staining of SRSF2 in NIKS, NIKS CLK3$^{-/-}$, NIKS + E6/E7, NIKS CLK3$^{-/-}$ + E6/E7. SRSF2 staining is in red, and DAPI staining indicates the nuclei (blue). Microscopy images were at 60 X magnification. D) Image J quantification of SRSF2 staining, with the mean fluorescence intensity of 10 cells being represented by each bar.

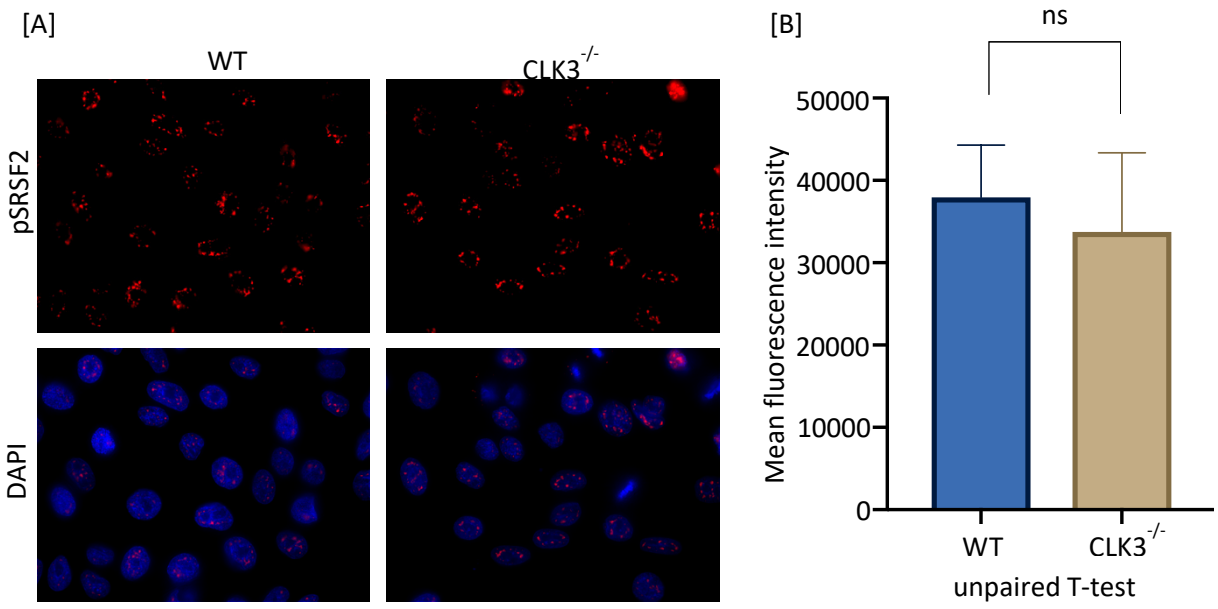



**Figure 4.2.3 CLK3 KO effect on pSRSF2**
A), IF staining of pSRSF2 in NIKS, and CLK3$^{-/-}$ NIKS$^{-}$, with pSRSF2 staining in red, and DAPI staining in blue. Microscopy images were at 60 X magnification. B) Image J quantification of pSRSF2 staining, with the mean fluorescence intensity of 10 cells being represented by each bar.

## 4.3 Loss of CLK3 results in loss of c-MYC

Having observed no clear differences in SRSF2 phosphorylation or localisation in WT versus CLK3$^{-/-}$ NIKS, we next investigated whether the stabilisation of c-MYC by CLK3 recently reported in cholangiocarcinoma cells (Zhou, et al., 2020) also occurs in NIKS. Western blot analysis of WT and CLK3$^{-/-}$ NIKS confirmed a significant loss of c-MYC protein (Figure 4.3.1, A) in the absence of CLK3. In addition, qRT-PCR analysis revealed a significant increase in c-MYC mRNA levels in the CLK3 KO cell lines, compared to the WT (Figure 4.3.1, B), consistent with a role for CLK3 in upregulating c-MYC at the protein level, as reported in the previous study (Zhou, et al., 2020). IF staining added further strength to this, with less c-MYC staining being observed in the CLK3 KO cell line, and Image J quantification confirmed the significant difference (Figure 4.3.1 C, D).

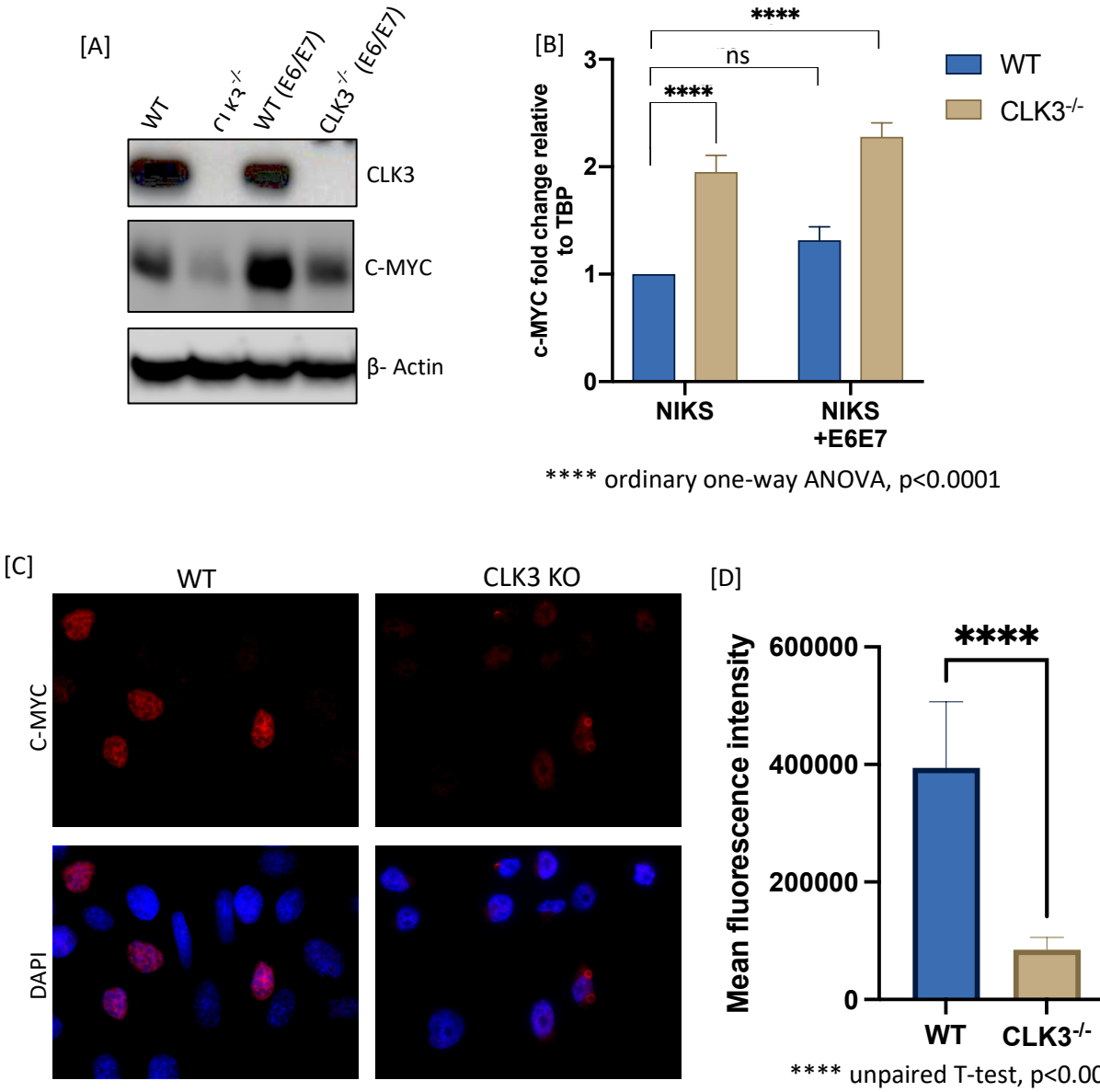



**Figure 4.3.1 CRISPR-Cas9 knockout of CLK3 in NIKS results in loss of c-MYC**

A) Western blot of WT and CLK3 KO NIKS with and without URR E6/E7. 50µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of CLK3 is approximately 59kDa, Molecular weight of c-MYC is approximately 62kDa. β-Actin was used as a loading control. B) qRT-PCR analysis of c-MYC mRNA relative to TBP in wildtype and CLK3 KO NIKS and NIKS with URR E6/E7. C) IF staining of c-MYC in NIKS, and NIKS CLK3 KO. Microscopy images were at 60 X magnification. D) Image J quantification of c-MYC staining, with the mean fluorescence intensity of 10 cells being represented by each bar.

Furthermore, a tetracycline inducible shRNA system, developed by Luke Williams, was used to knockdown CLK3 in SiHa cells. Knockdown was induced with treatment of 1µg/µl of doxycycline, and protein analysis revealed a significant loss of c-MYC over 72hrs (fig 4.3.2 A). qRT-PCR analysis revealed no significant difference in c-MYC mRNA levels, confirming that, as in NIKS, CLK3 was specifically acting on c-MYC at the protein level.

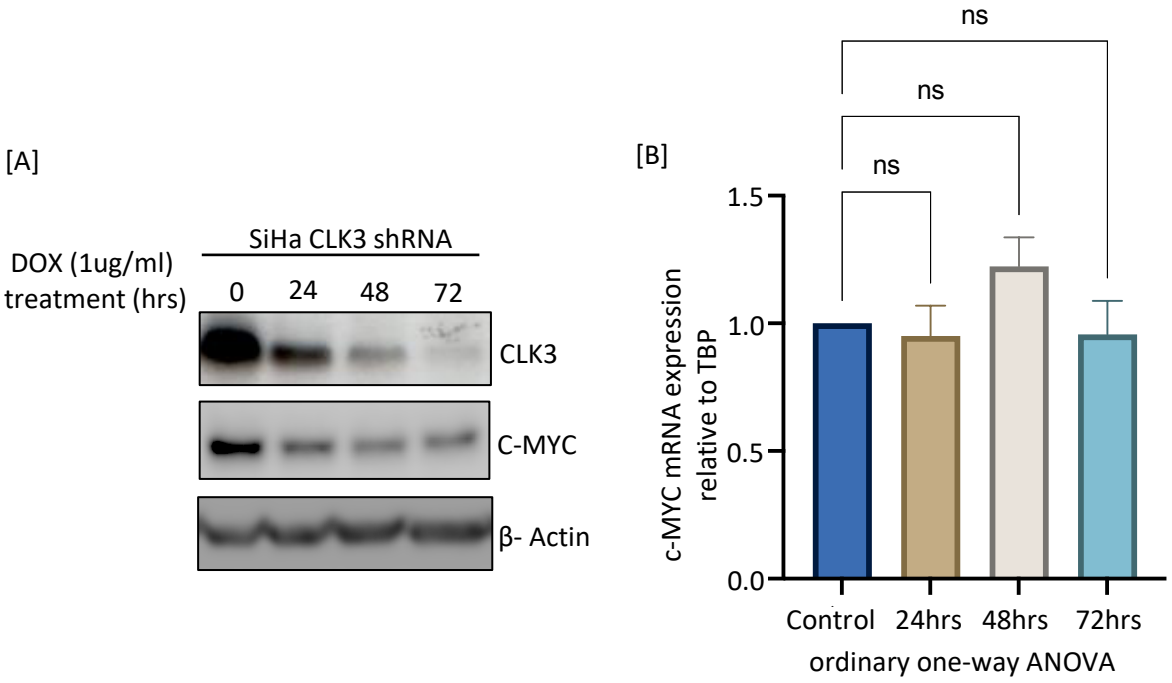



**Figure 4.3.2 shRNA knockdown of CLK in SiHa results in loss of c-MYC**
A) Western blot of shRNA knockdown of CLK3 in SiHa over 72hrs, induced by treatment with 1µg/µl of doxycycline. 40µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of CLK3 is approximately 59kDa, Molecular weight of c-MYC is approximately 62kDa. β-Actin was used as a loading control. B) qRT-PCR analysis of c-MYC mRNA relative to TBP in SiHa over 72hrs following doxycycline treatment to induce CLK3 shRNA knockdown (This was only repeated twice). shNT siRNA transduced SiHa were used as control.

Following these observations, it was necessary to investigate whether overexpression of CLK3 will result in upregulation of c-MYC. WT-CLK3 and K186R-mutant CLK3 were transfected into CLK3 KO NIKS, following which the cells were stained with a c-

MYC specific antibody. Markedly stronger intensity for c-MYC was observed specifically in those cells expressing the WT-CLK3, and not the K186R-mutant CLK3, which highlighted the importance of the kinase function in c-MYC regulation. Additional repeats are required to further validate this and quantify the difference in fluorescence intensity with Image J software.

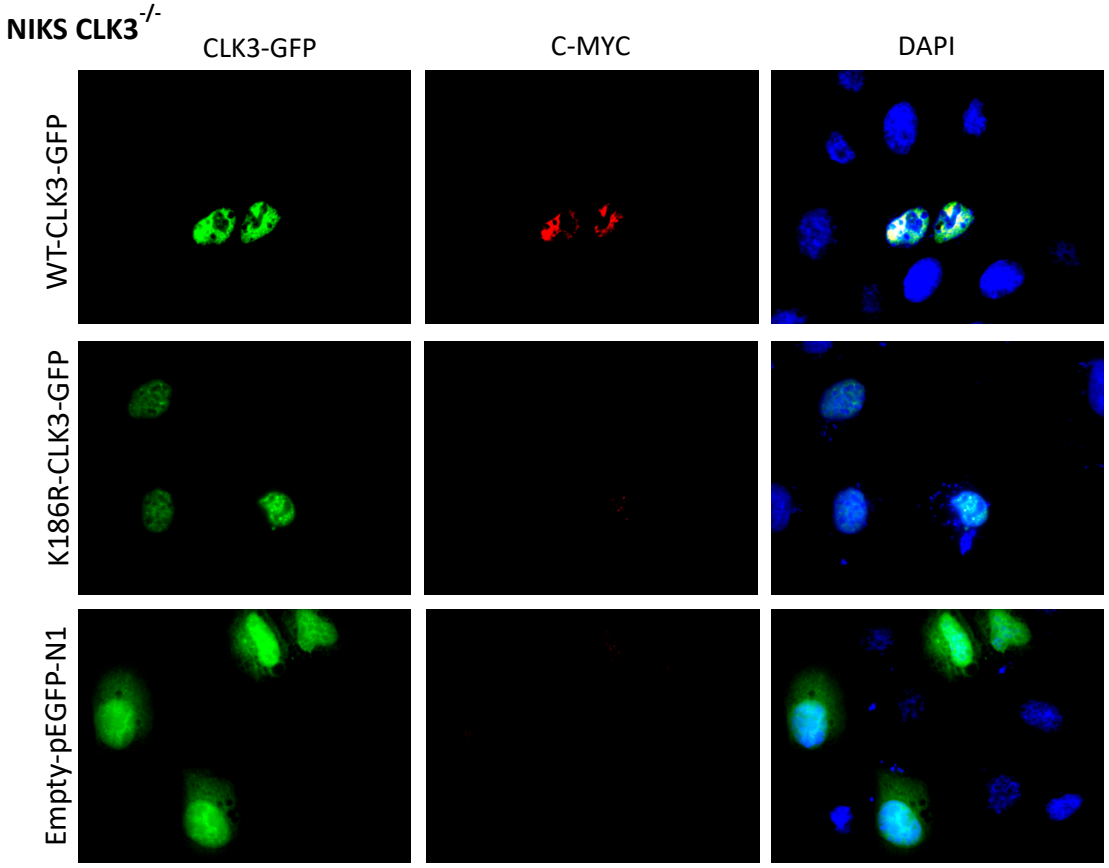



**Figure 4.3.3 Overexpression of CLK3 results in upregulation of c-MYC**
IF staining of CLK3$^{-/-}$ NIKS with c-MYC antibody following transient transfection of pEGFP-N1-CLK3-GFP, pEGFP-N1-K186R-CLK3-GFP or pEGFP-N1-Empty (control) plasmids for 48hrs. c-MYC is stained red, cells expressing the plasmid are green (GFP), DAPI staining indicates the nuclei (blue). Microscopy images were at 60X magnification.

## 4.4 Overexpression of c-MYC upregulates HPV16 E7 protein.

Following confirmation of loss of c-MYC upon CLK3 loss, it was necessary to determine if c-MYC was involved in controlling E6/E7 expression downstream of CLK3, as hypothesised. Using the in-situ approach developed in chapter 3, c-MYC-GFP was transfected into CLK3$^{-/-}$ NIKS harbouring the HPV16 URR E6/E7 construct and IF staining with HPV16 E7 antibody was carried out to investigate the effect of c-MYC on E7 expression. We used CLK3 KO NIKS due to the low basal level of c-MYC and E7 expression in these cells compared to WT NIKS. Increased expression of E7 was observed in cells overexpressing c-MYC-GFP, revealing a novel role for c-MYC in upregulating E7 expression (Figure 4.4.1, A). The plasmid was also transfected in CLK3$^{-/-}$ NIKS without E6/E7 and stained with E7 antibody to confirm the signal was specific to E7, and Empty-pEGFP-N1 plasmid was used as negative control (Figure 4.4.1, A). Image J quantification was carried out to quantify the effect of c-MYC on E7 expression (Figure 4.4.1, B). It was not possible to study the effect of c-MYC overexpression on E6 protein, as we have not been able to find an E6 antibody that performs sufficiently for E6 detection by IF in these cells.

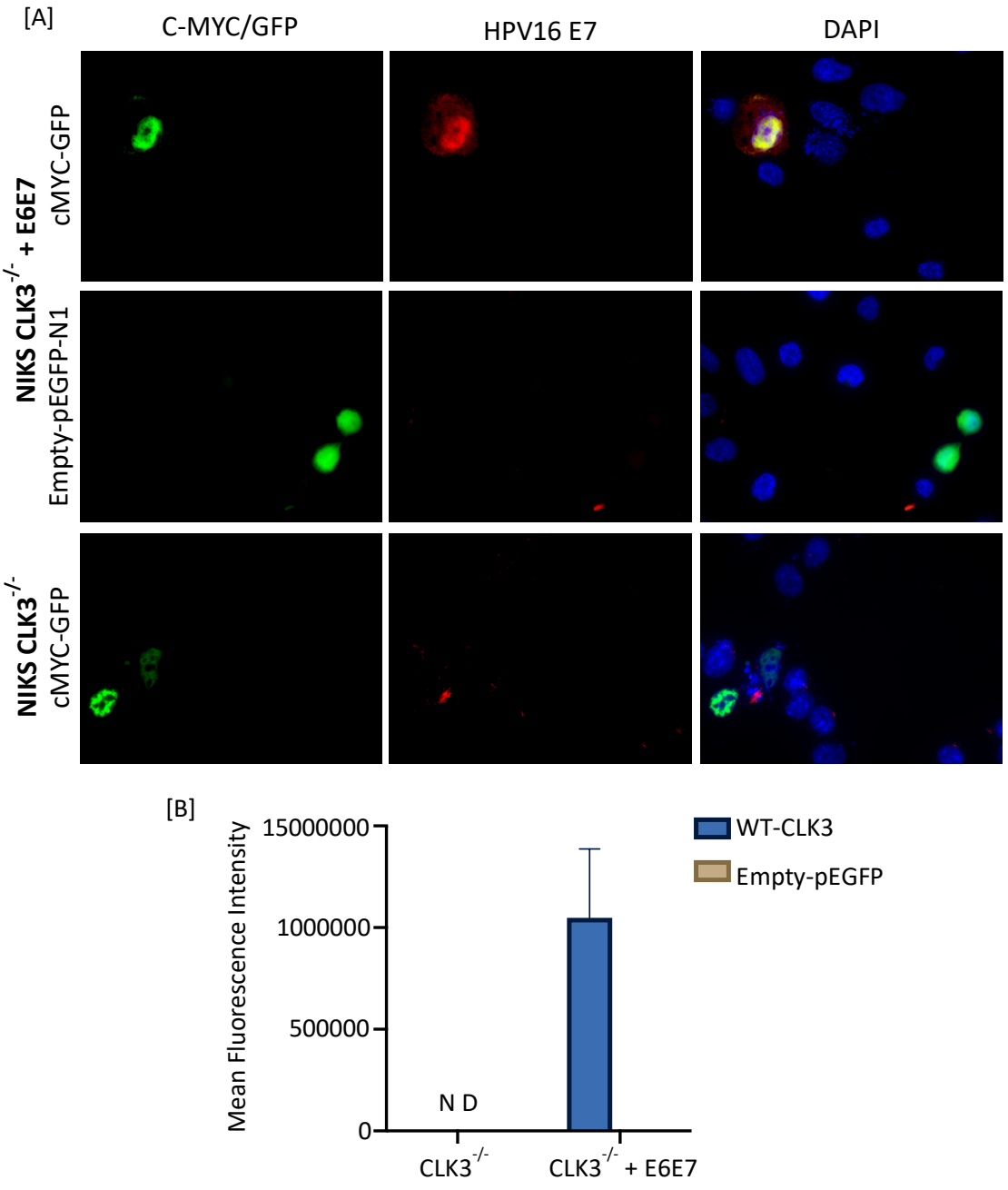



**Figure 4.4.1 c-MYC overexpression results in overexpression of HPV16 E7**
A) IF staining of CLK3$^{-/-}$ NIKS, with and without HPV16 URR E6/E7, with HPV16 E7 antibody (red), following transient transfection of c-MYC-GFP or Empty-pEGFP-N1 plasmids (green). DAPI staining confirms presence of nuclei (blue). Microscopy images were at 60 X magnification. C) Image J quantification of E7 staining, with the mean fluorescence intensity of 10 cells being represented by each bar. E7 Antibody produces minimal background staining; therefore, no bars are present for control and CLK3$^{-/-}$ cell line.

## 4.5 siRNA knockdown of c-MYC downregulates HPV16 E7 protein

Having confirmed the upregulation of HPV16 E7 protein upon c-MYC overexpression, it was necessary to determine the effect of c-MYC on E6/E7 mRNA. Our working hypothesis was that CLK3 exerts its effect on E7 protein expression by increasing transcription of the E6/E7 mRNA. It was therefore predicted that loss of c-MYC would also primarily effect E6/E7 mRNA. To determine this, siRNA knockdown of c-MYC was carried out in two HPV16-positive cancer cell lines (SiHa and CaSki) and NIKS-HPV16 URR E6/E7. ONTARGETplus siRNAs were used as they are chemically modified with 2'-O-methyl ribosyl substitution at position 2 of the guide strand to reduce off target degradation of mRNA by RISC. Successful knockdown of c-MYC was confirmed following protein extraction at 48hrs and 72hrs post siRNA transfection, and loss of HPV16 E7 protein was observed following c-MYC siRNA knockdown (figure 4.5.1). This was followed by RNA extraction and qRT-PCR analysis to determine the effect of c-MYC knockdown on E6/E7 mRNA (figure 4.5.2). Contrary to our hypothesis, we observed an increase E6E7 mRNA levels upon c-MYC knockdown, suggesting that the effects of c-MYC on E7 protein expression may occur at the post-transcriptional level. As expected, both SiHa and CaSki cell lines displayed loss of viability at 96 hours post-transfection with c-MYC siRNA.

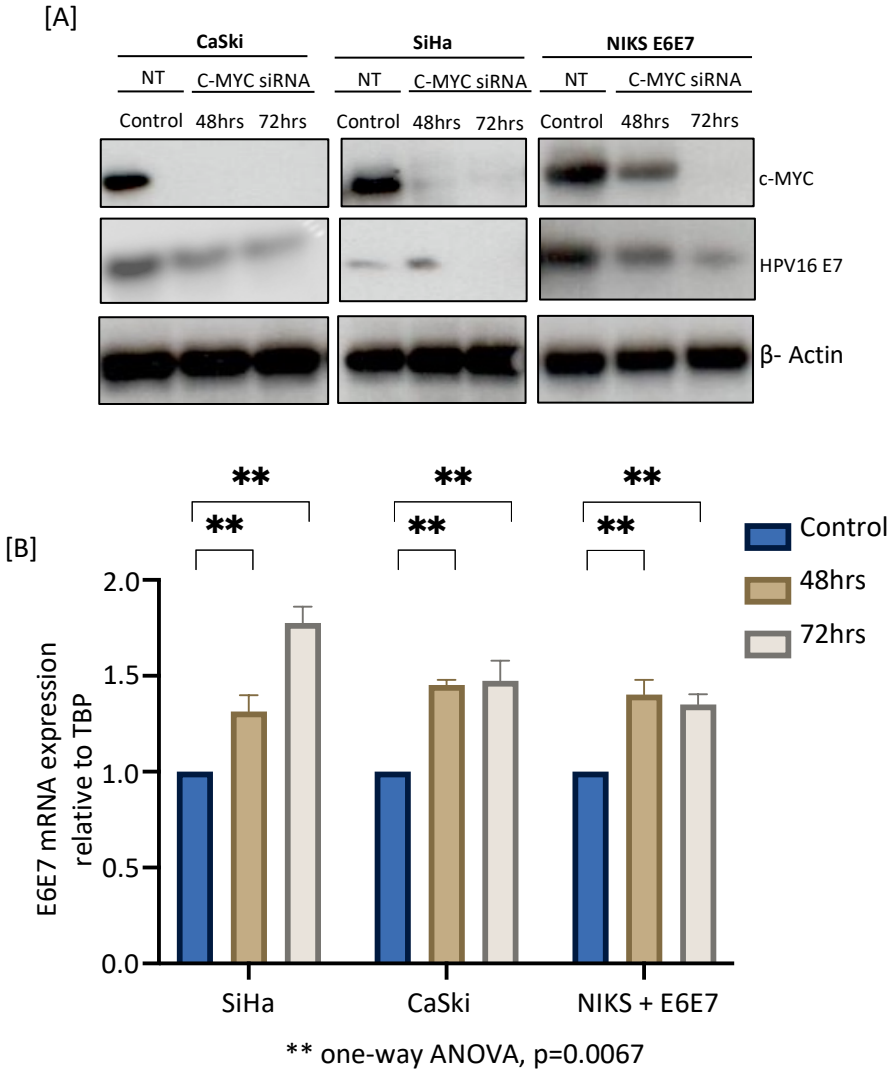



**Figure 4.5.1 siRNA knockdown of c-MYC results in downregulation of HPV16 E7**
A) Western blot analysis following ONTARGETplus siRNA knockdown of c-MYC in SiHa, CaSki and NIKS (with E6/E7). siRNA transfection was completed over 72hrs. non-targeting control siRNA was used as control. 40µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of c-MYC is approximately 62kDa, molecular weight of HPV16 E7 is approximately 18kDa. β-Actin was used as a loading control. B) qRT-PCR analysis of E6/E7 mRNA relative to TBP following siRNA knockdown of c-MYC. E6/E7 levels are represented as fold change compared to control.

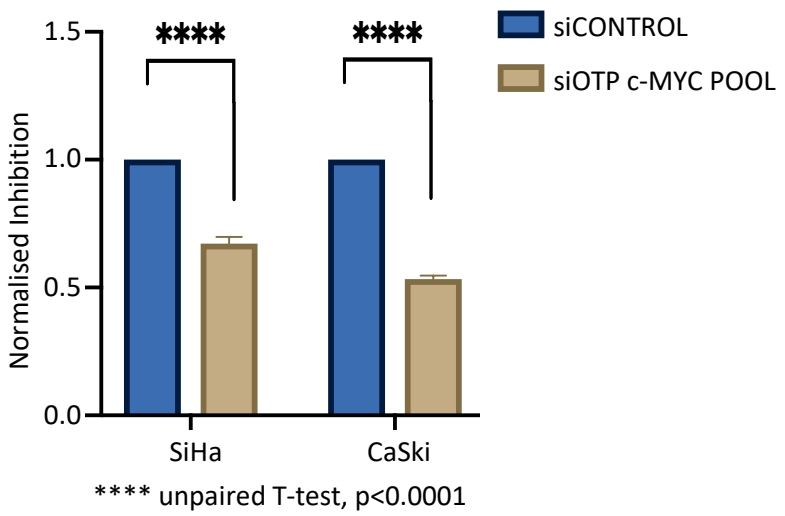



**Figure 4.5.2 siRNA knockdown of c-MYC results in loss of cell viability**
Cell viability assay was performed over five days in CaSki and SiHa to determine the effect of siRNA c-MYC knockdown. siRNA targeting PLK1 was used as positive control and non-targeting siRNA was used as negative control, and MTS assay was performed to determine cell viability on day five.

## 4.6 c-MYC does not act on the promoter of E6/E7

To our surprise, c-MYC siRNA knockdown experiments suggested c-MYC was only involved in regulating HPV16 E7 protein but not E6/E7 mRNA. To further validate this and to investigate the effect of CLK3 on the E6/E7 promoter, a luciferase assay was designed. The HPV16 URR from the HPV16-positive UM-SCC047 HNSCC cell line was cloned into the pGL3 basic plasmid, upstream of the firefly luciferase coding sequence (Fig 4.6.1 A). This plasmid was initially co-transfected into WT and CLK3$^{-/-}$ NIKS along with pcDNA-RLuc8 plasmid renilla Luciferase control plasmid and the Dual luciferase reporter assay system (Promega) was used to quantify the luciferase assay. The firefly luciferase activity was normalised against the renilla control plasmid to account for any difference in transfection efficiency. It was possible to observe a decrease of approximately 50% in firefly luciferase

activity in the CLK3$^{-/-}$ NIKS, compared to WT NIKS (Fig 4.6.1 B), consistent without previous observations (Chapter 3 and Williams, 2021) implicating a role for CLK3 in upregulation of E6/E7 transcription and suggesting that this effect is mediated via the URR.

This was followed by an experiment to investigate whether CLK3 was acting on the promoter via c-MYC. To improve the signal-noise ratio of the reporter assays, the URR-pGL3 and pcDNA-RLuc8 plasmids were co-transfected into the highly-transfectable HEK-293T cell line, along with the c-MYC-GFP plasmid. In addition, WT-CLK3, K186R-CLK3, and empty pEGFP-N1 plasmids were also transfected simultaneously to further validate CLK3 involvement. As shown in Figure 4.6.1 C, no significant difference in the luciferase activity was observed in cells overexpressing the MYC-GFP plasmid, suggesting no involvement for c-MYC in E6/E7 promoter activation. This is also true for cells overexpressing the K186-CLK3 and pEGFP-N1 empty plasmids. However, an approx. 3-fold increase in luciferase activity was observed in cells transfected with WT-CLK3, further validating the involvement of CLK3 and its kinase activity in E6/E7 promoter activation.

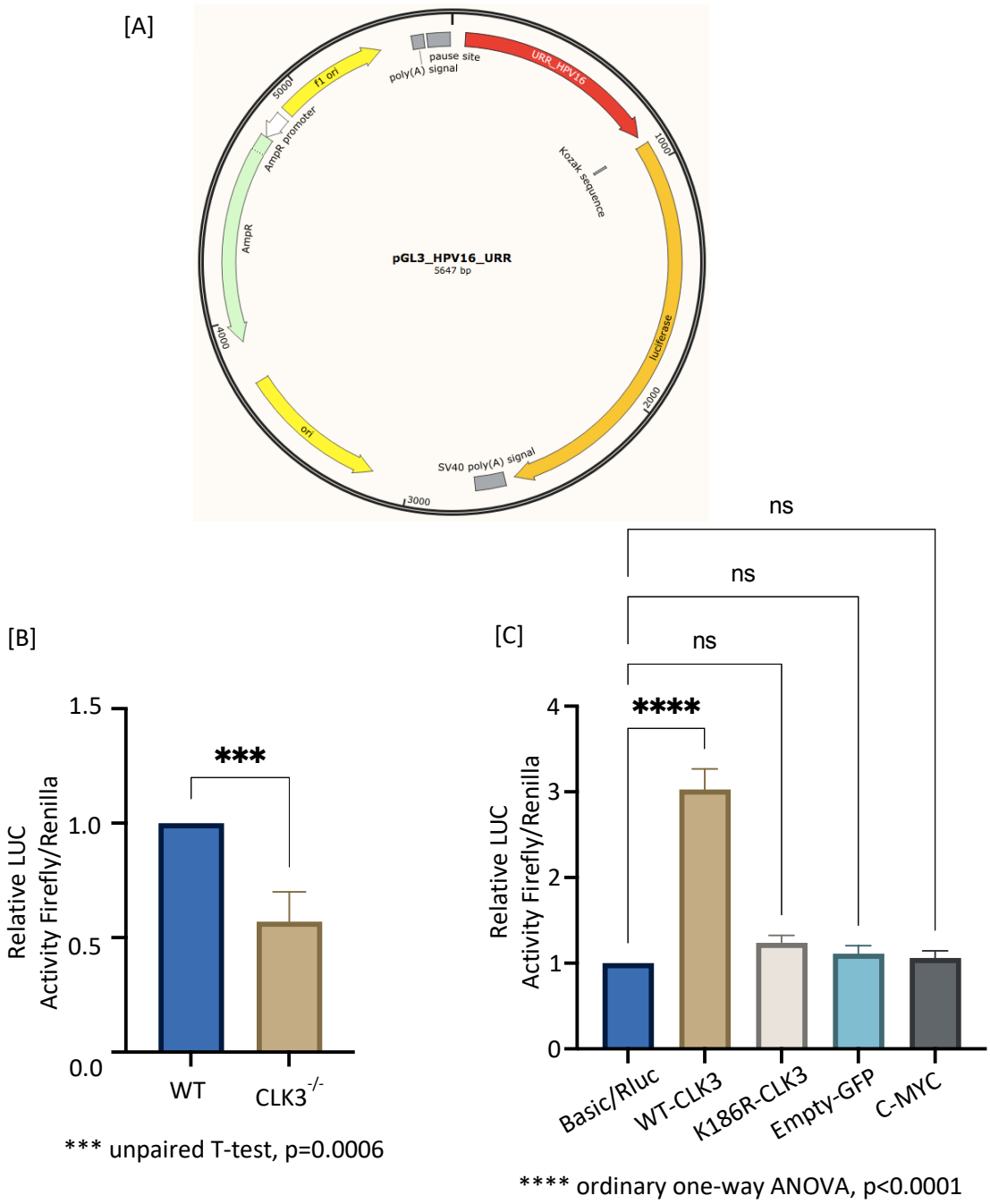



**Figure 4.6.1 Luciferase reporter assay**

A) Map of URR-pGL3 plasmid. B) WT and CLK3$^{-/-}$ NIKS luciferase activity, 48hrs post-transfection with URR-pGL3 and pcDNA-RLu8 plasmids. The firefly luciferase activity was measured following normalisation against renilla luciferase activity to account for difference in transfection efficiency. C) Luciferase activity in HEK293T cells following overexpression of either WT-CLK3, K186R-CLK3, Empty-pEGFP-N1, or MYC-GFP plasmids. Basic-pGL3 plasmid was used as control to account for background luciferase activity.

## 4.7 CLK3 inhibition reduces c-MYC and HPV16 E6/E7 expression

Over the last two decades, structurally diverse CLK inhibitors have been described in the literature, with the majority of them being published in recent years. most of the reported CLK inhibitors interact relatively weakly with CLK3—only 4 out of 24 compounds described in Moyano et al exhibit $IC_{50}$/ $K_d$ values towards CLK3 < 100 nM (Moyano, et al., 2020). In collaboration with Kinsensus Ltd, following a screen of a 633-compound kinase inhibitor library (Selleck) against CLK3, we identified a cyclin-dependent kinase inhibitor, R547 (KIN100250-002) with activity against CLK3 (Fig 4.7.1). Initially R547 was tested in NIKS with the HPV16 E6/E7 to determine if c-MYC and HPV16 E7 protein levels were affected by the inhibitor. From the western blot, it is possible to observe a significant loss of c-MYC and HPV16 E7 upon 3hrs of inhibition (Fig 4.7.2). Inhibition over 6hrs resulted in increased levels of c-MYC and E7, possibly as a result of the inhibition of other CDKs.

[A] [B]

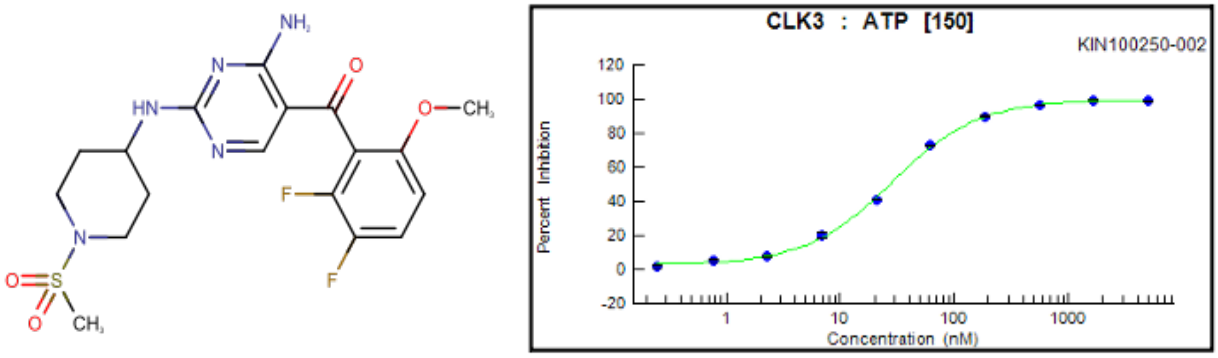



**Figure 4.7.1 R547 structure and 10-point $IC_{50}$ titration for CLK3 inhibition**
A) structure of R547 (KIN100250-002), and B) 10-point $IC_{50}$ titration for CLK3 inhibition (ZLYTE assay, ThermoFisher SelectScreen Services).

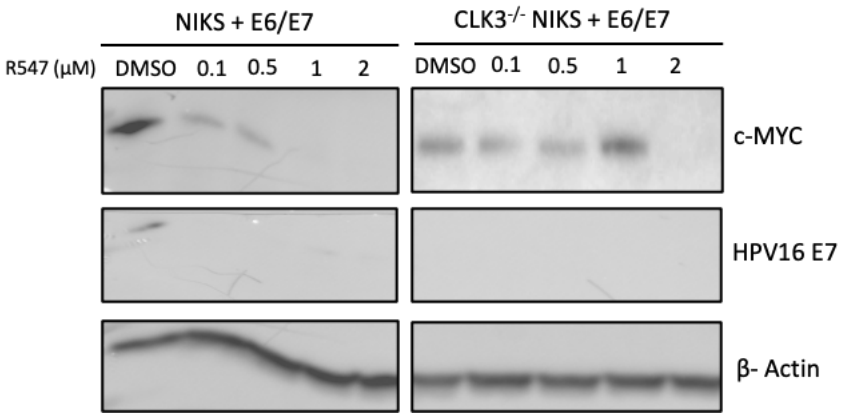



**Figure 4.7.2 R547 treatment of NIKS**
Western blot following 3hrs treatment of NIKS and CLK3$^{-/-}$ NIKS with HPV16 URR E6/E7 with R547. 40μg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of c-MYC is approximately 62kDa, molecular weight of HPV16 E7 is approximately 18kDa. β-Actin was used as a loading control.

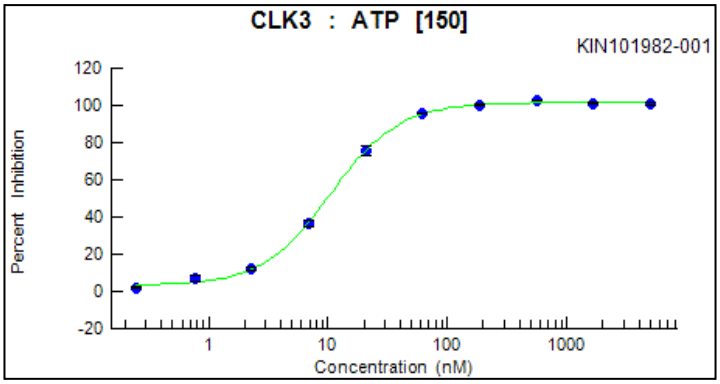



**Figure 4.7.3 Compound KIN101982-001**
10-point $IC_{50}$ titration for CLK3 inhibition for compound KIN101982-001 (ZLYTE assay, ThermoFisher SelectScreen Services).

In work conducted by Dr Clive McCarthy (Kinsensus Ltd), the crystal structure of CLK3 was analysed to inform modification of R547 to improve its potency for CLK3. The newly modified compound (KIN101982-001) $IC_{50}$ was determined to be 10nM, significantly more potent against CLK3 (Fig 4.7.3). Significant loss of c-MYC and HPV16 E7 was observed with 10nM treatment for 3hrs (Fig 4.7.4, A). In addition, significant downregulation of E6/E7 and p16 mRNA levels were observed with the same concentration following 3hrs of inhibition (Fig 4.7.4, B). $CLK3^{-/-}$ NIKS with HPV16 URR E6E7 was used as an important control to demonstrate the effects of the inhibitors were indeed due to the inhibition of CLK3. The selectivity of KIN101982-001 for CLK3 over other kinases remains to be determined but it is anticipated that multiple modifications will be necessary to gain a selective inhibitor.

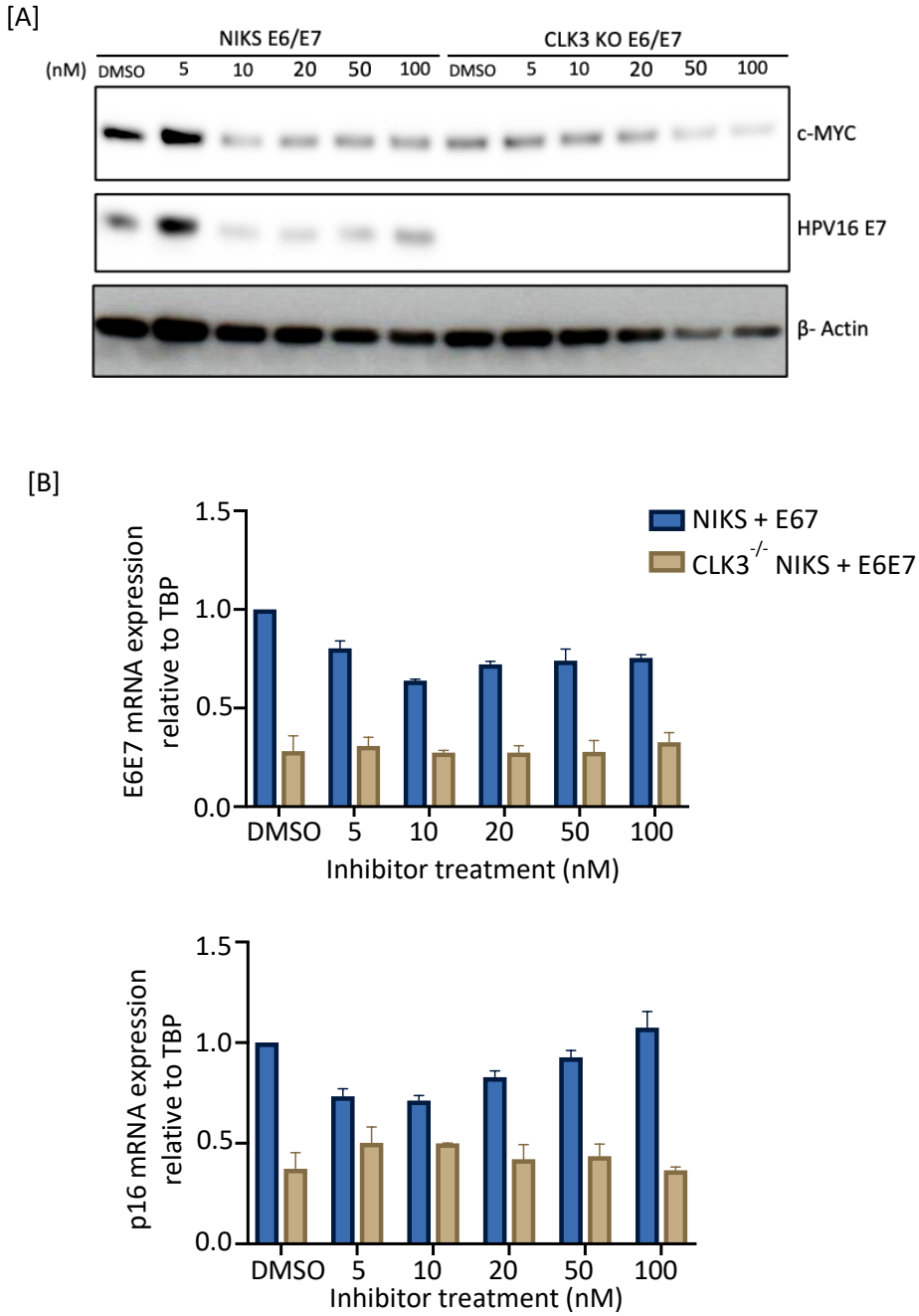



**Figure 4.7.4 KIN101982-001 treatment of NIKS**

A) Western blot following 3hrs treatment of NIKS and CLK3$^{-/-}$ NIKS with HPV16 URR E6/E7 with KIN101982-001. 40µg of lysate was loaded in each well and separated by SDS page and probed with primary antibodies overnight. Molecular weight of c-MYC is approximately 62kDa, molecular weight of HPV16 E7 is approximately 18kDa. β-Actin was used as a loading control. B) qRT-PCR analysis of E6/E7 and p16 mRNA relative to TBP following 3hrs of treatment with KIN101982-001. E6/E7 and p16 levels are represented as fold change compared to control.

## 4.8 Discussion

The focus of this chapter was to determine how CLK3 controls E6/E7 expression. As a known substrate of CLK3 and having been indicated in E6/E7 mRNA control previously, SRSF2 was initially considered as a possible mediatory protein in CLK3-E6/E7 regulation. However, following various studies, we concluded that loss of CLK3 in NIKS had no significant effect on SRSF2 expression or its phosphorylation state. It is possible that other members of the CLK family are able to maintain SRSF2 levels in the absence of CLK3 as SRSF2 is also a known target of other CLKs.

Following this, our focus turned to c-MYC. The stability of c-MYC is known to be controlled via CLK3 phosphorylation of USP13 in cholangiocarcinoma cells. We therefore hypothesised that CLK3 was controlling E6/E7 mRNA expression via c-MYC, as it is a known transcription factor of wide array of genes. In addition, the E-box sequence CACGTG was identified within the HPV16 URR promoter sequence, and we hypothesised that c-MYC controls E6/E7 mRNA levels by activating the promoter via E-box binding. We were able to confirm that loss of CLK3 resulted in a significant loss of c-MYC in NIKS, as well as SiHa cell lines. In addition, CLK3 overexpression was found to upregulate the level of c-MYC, which was confirmed by IF staining. Following this, c-MYC was overexpressed in HPV16 URR-E6/E7-NIKS and IF staining confirmed upregulation of E7 protein. siRNA knockdown of c-MYC in SiHa, CaSki, and NIKS with HPV16 URR E6/E7 further confirmed a role of c-MYC in E7 protein regulation, with loss of E7 being observed at 48hrs and 72hrs post-transfection of c-MYC siRNA. As expected, cell viability assay following c-MYC siRNA knockdown in all

three cell lines resulted in significant loss of viability by 96 hours post-transfection (similar to effects seen upon CLK3 knockdown (Williams, 2021)). Surprisingly, qRT-PCR analysis provided evidence of c-MYC regulation of HPV16 E6/E7 only occurring at protein level. This was further validated with a luciferase assay, which confirmed no role for c-MYC in E6/E7 promotor activation. However, the assay also provided strong evidence of CLK3 involvement in promoter activation, with significant difference in luciferase activity being observed in WT NIKS compared to the $CLK3^{-/-}$ NIKS, as well as increased luciferase activity in HEK293T cells overexpressing the WT-CLK3 plasmid but not the kinase-inactive K186R mutant. The experiments completed in this chapter revealed two novel findings: (1) the involvement of CLK3 in E6E7 promoter (HPV16 URR) activation to initiate transcription and (2) the involvement of c-MYC in the upregulation of E7 protein levels. Interestingly this latter effect appears to be at the post-transcriptional level, suggesting CLK3 upregulates transcription of E6/E7 in a c-MYC-independent fashion and E7 protein levels (possibly via increased translation or protein stabilisation) in a c-MYC-dependent fashion. This model is discussed further in chapter 6.

With our collaborators, it was possible to carry out an initial CLK3 inhibitor study. The CDK inhibitor R547 was identified as a CLK3 inhibitor from a 633-compound library, and initial experiments were promising, with loss of c-MYC and HPV16 E7 proteins being used as markers to confirm loss of CLK3 activity in NIKS. R547 was further modified to increase the selectivity and potency towards CLK3, and the $IC_{50}$ was significantly improved to 10nM. The new compound (KIN101982-001) showed great potential, with significant loss of c-MYC

and HPV16 E7 being observed with 10nM treatment. qRT-PCR analysis further validated this, with significant loss of E6/E7 and p16 mRNA being observed following inhibition with the same dose.

# Chapter 5

# Expression and purification of recombinant CLK3

# 5 Expression and purification of recombinant CLK3

## 5.1 Introduction

Chapter three confirmed a key role for CLK3 in controlling the transcription of HPV16 E6/E7. The aim of chapter four was to identify how CLK3 interacted with E6/E7, and it was initially hypothesised that CLK3 was acting via either c-MYC or SRSF2 to control the transcription rate of E6/E7. However, data suggested no significant role for SRSF2 or c-MYC in controlling E6/E7 mRNA transcription, although further experiments are required to probe the role of SR proteins in this function (see chapter 6). It is therefore possible that CLK3 acts via other SR proteins or unknown substrates of the kinase to control HPV16 E6/E7 transcription. We were therefore interested to identify other potential substrates of CLK3, which will be the main aim of this chapter. When treating CLK3 as a molecular target for potential therapeutic use against HPV-associated cancers, it is necessary to confirm if any additional signalling pathways and cellular functions will be affected. Therefore, identifying other substrates of CLK3 was considered the next significant step.

Kinase substrate identification has been completed using various methods including the use of microarrays, phage display, mass spectrometry and bioinformatics prediction of substrates (Xue, et al., 2013). The advantage of using a mass-spectrometry based assays compared to other methods is that it not only identifies phosphoproteins, but also reveals their phosphorylation sites. In addition, it is also possible to identify the substrates by detecting them whilst they interact with the kinase (Amano, et al., 2010). However, phosphorylation is an event that is usually considered to be transient, as post-

phosphorylation, the interaction between the kinase and the substrate is quickly interrupted, allowing the kinase to perform further phosphorylation. It is therefore difficult to catch the protein-substrate interactions, and usually results in a high-false positive rate.

One method that was preferred when identifying phosphoproteins is the use of a cell lysate as a protein pool for a kinase reaction, followed by mass spectrometric analysis. One of the main challenges with this approach was the level of background phosphorylation that will be present in the lysate samples due to other kinases. To overcome this, a method introduced by Shokat's group known as Analogue-Sensitive Kinase Allele (ASKA) technology was selected for this study (Shah, et al., 1997. Previously, ASKA has been successfully used to identify substrates of v-Src, PERK, CDK1, CDK2 and several other kinases (Shah, et al., 1997, Maas, et al., 2014, Blethrow, et al., 2008, Chi, et al., 2008). The ASKA technology has the potential to be utilised across all kinases as the basic principles of the method revolve around modifying a conserved region within all kinases known as the gatekeeper domain. This region is found within the ATP-binding pocket, and can be modified to create a larger ATP-binding pocket, allowing the mutated kinase to utilise analogue forms of ATP. This modification involves replacing a 'bulky' amino acid, that is usually found at the gatekeeper domain, with a small one such as glycine or alanine. This mutant form is able to use analogue forms of ATP, which cannot be used by wild-type kinases found within cells, to phosphorylate its substrates. Using this technique, substrates of the kinase can be elegantly labelled using ATP analogues such as ATP-γ-S, which adds a thiophosphate group to the phosphorylated substrates. Mass spectrometry analysis can be used to reveal the identity

of the protein corresponding to the tagged peptide. Figure 5.1 displays the schematic diagram of the ASKA methodology.

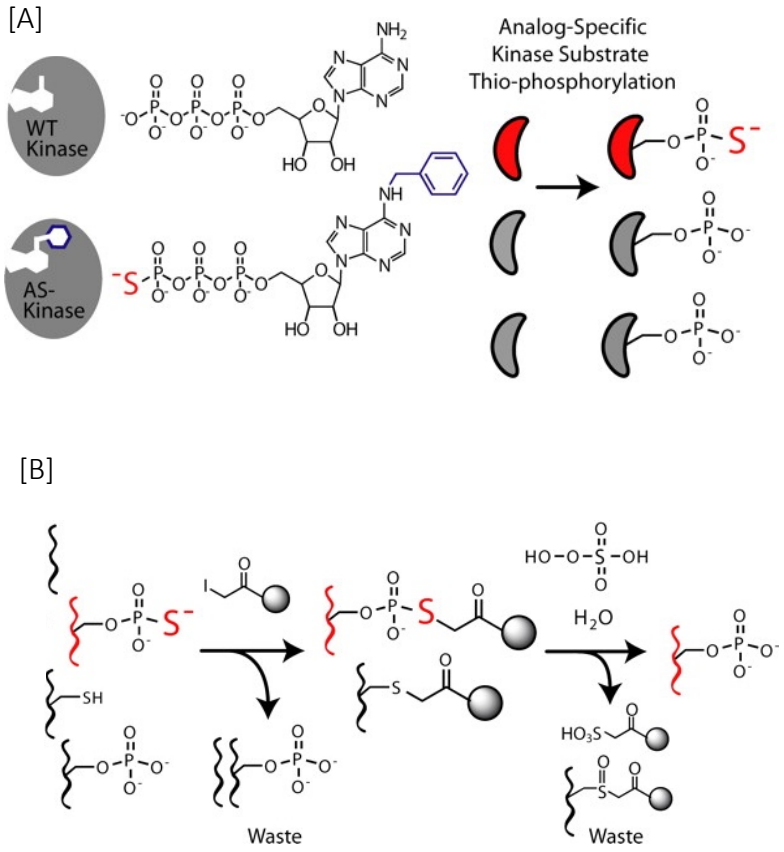



**Figure 5.1.1 Method of protein kinase identification**
A) Kinase is engineered to accept an ATP analog, with a modification at the gatekeeper domain. ATP analog, which bears a transferable thiophosphate group (red) is used to perform a kinase reaction with the mutant kinase. This results in specific tagging of the substrates of the kinase with a thiophosphate group. B) The protein mixture containing the tagged substates is digested to peptides, and the products are allowed to react with iodoacetyl-agarose. Thiophosphate groups react to form covalent bonds, and unbound peptides are washed away. The bound peptides are liberated by oxidation-promoted hydrolysis of the Sulphur-phosphorous bond. The thiophosphoryl sulfur atom is replaced with oxygen in this step (Blethrow, et al., 2008).

## 5.2 Identifying the gatekeeper domain of CLK3

The gatekeeper residue lines the bottom of the kinase active site and generally consists of a large “bulky” amino acid. Previously, studies have aligned gatekeeper residues of known kinases to identify the residue (Liu, et al., 1999, Schnabel, et al., 2018), and the same approach was utilised here. In general, the bulky side chain is either Thr, Ile, Leu, Met, or Phe. The structure of CLK3 was aligned with CDK1 and CDK2 structures using PyMOL and the gatekeeper domain of the kinase was identified as a phenylalanine at position 236 (fig 5.2E), with assistance from Prof Mark Wass (School of Biosciences, University of Kent).

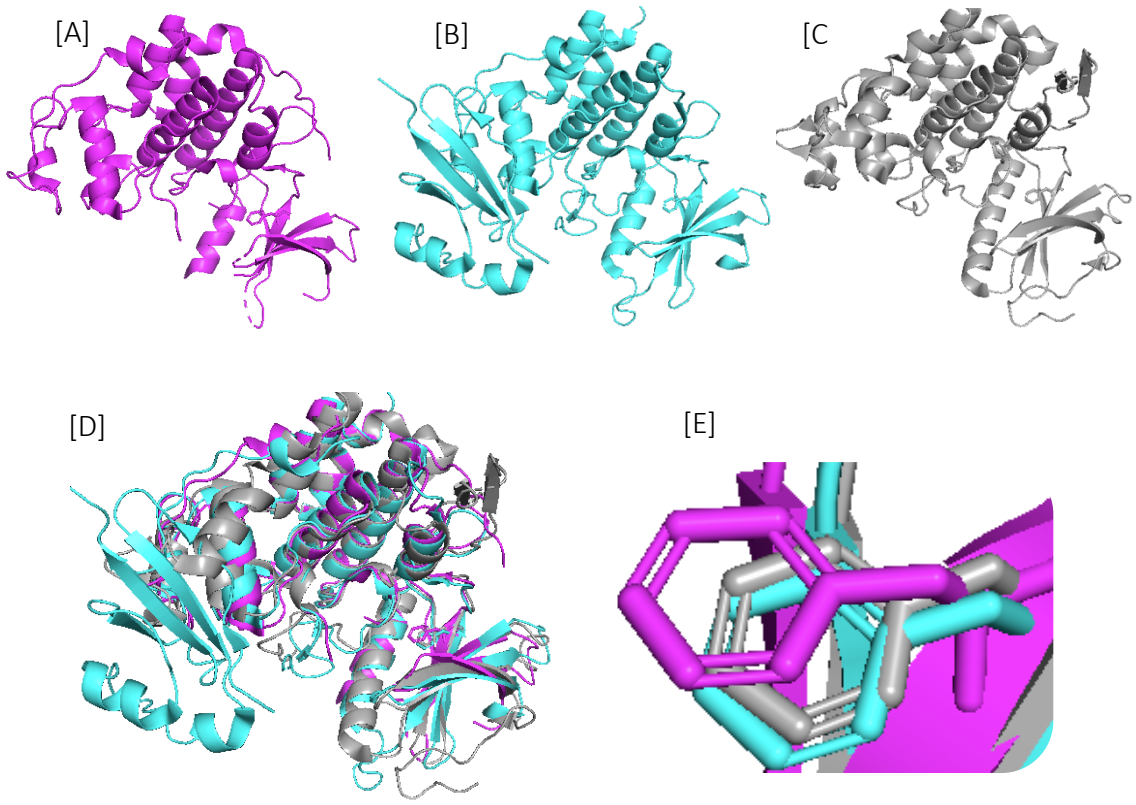



**Figure 5.2.1 Gatekeeper domain of CLK3**
Structural comparison of CLK3 [C] with CDK1 [A] and CDK2 [B] using PYMOL to identify the gatekeeper domain of CLK3. [D] structural alignment of all three kinases, and [E] gatekeeper domain. CDK1 and CDK2 gatekeeper domains were previously determined at position 80 as phenylalanine (Strambi, et al., 2013). From PYMOL analysis, CLK3 gatekeeper domain was identified as phenylalanine at position 236.

## 5.3 Plasmid design for CLK3 protein expression

The next step was to generate mutant CLK3 protein containing the modified gatekeeper domain. Commonly, the bulky residue is replaced with a small amino acid such as glycine or alanine to create the altered ATP-binding site. Codon-optimised wild-type CLK3 was used to create the mutants, and primers were designed to carry out two-step site-directed mutagenesis using PCR. Following mutagenesis, the inserts were cloned into a pET-28 (+) 6xHis-3xFLAG plasmid (a gift from Dr Chris Mulligan, School of Biosciences, University of Kent). In total, four different plasmids were made: wild-type CLK3, kinase inactive CLK3, mutant with F236A mutation, and mutant with F236G mutation. Successful cloning, sequencing errors and correct orientation of the inserts was confirmed by restriction digest, colony PCR, and Sanger sequencing of the plasmids.

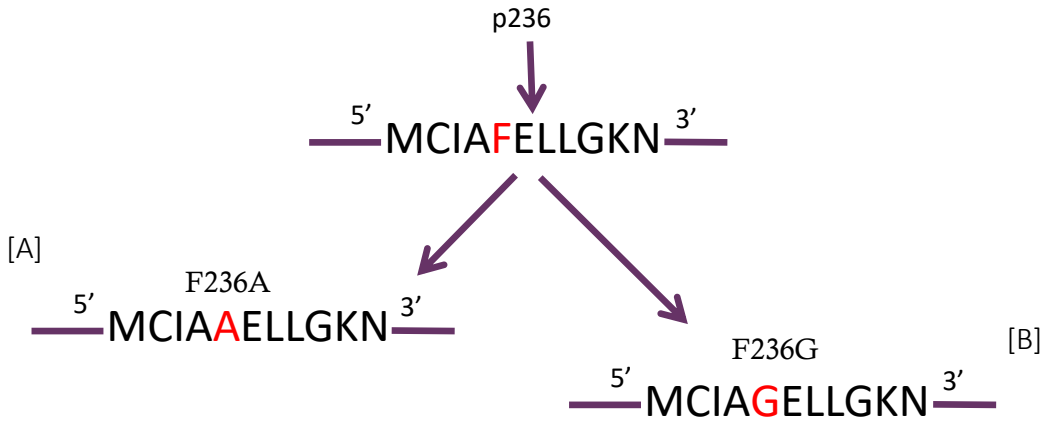



**Figure 5.3.1 Site-directed mutagenesis of CLK3**
Schematic diagram of site directed mutagenesis carried out to create two mutants, one with an alanine mutation (A), and one with a glycine mutation (B).

## 5.4 CLK3 expression and purification

Initially, CLK3 expression was attempted in competent BL21 (DE3) *E. coli* using the newly constructed pET-28(+) CO-CLK3 plasmid. Figure 5.4.1 represents a diagram of the plasmid construct. Western blot was carried out and FLAG-tag primary monoclonal antibody was used to confirm if CLK3 expression had occurred (Fig 5.4.2). From multiple repeats with varied temperature and IPTG concentrations, no evidence of CLK3 expression was observed.

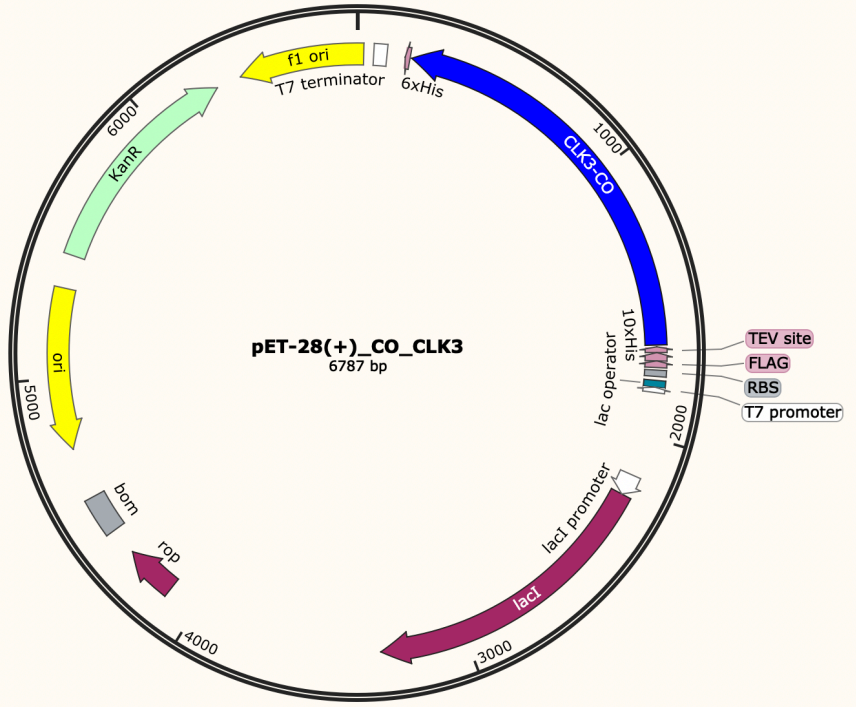



**Figure 5.4.1 Plasmid map of pET-28(+)-CLK3**
Modified pET-28(+) CLK3 plasmid. The plasmid was previously modified with the addition of 10x His tag, FLAG tag and a TEV site. Blue region is the CLK3 gene insert cloned in, Red is the Lac repressor sequence, responsible for inhibiting gene expression, which is reversed in the presence of IPTG.

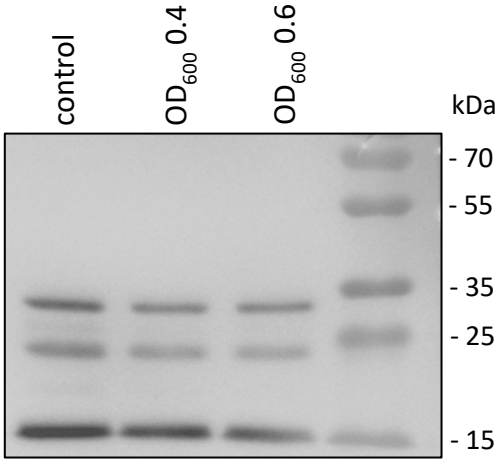



**Figure 5.4.2 CLK3 expression in BL21 (DE3) E. coli**
Western blot analysis following CLK3 expression attempt. FLAG-tag primary monoclonal antibody was used for protein detection. Cells were boiled in sample buffer at 95°C for 10mins and ran on SDS-PAGE gel. PageRuler plus prestained protein ladder was used as a molecular marker. No expression of CLK3 observed.

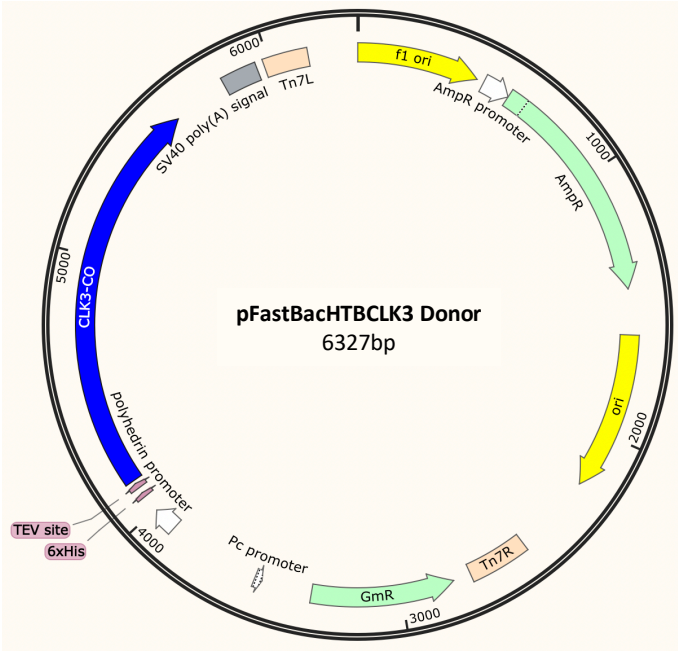



**Figure 5.4.3 Plasmid map of pFastBacHTB-CLK3**
Schematic diagram of pFastBacHTB-CLK3 donor plasmid. Gene insert is in blue, with gentamycin acetyltransferase and ß-lactamase sequences in green.

Following this, CLK3 expression was attempted using the Baculovirus system. A schematic diagram of the pFastBacHTB-CLK3 donor plasmid can be found in figure 5.43. The empty donor plasmid, along with the DH10Bac *E. coli* were a kind gift from Dr Chris Toseland (University of Sheffield). The general steps involved in CLK3 expression using this system is displayed on figure 5.4.4. All three mutants (K186R, F136G, F136A) of CLK3, along with the WT kinase were successfully expressed, and Coomassie staining was performed to confirm expression following P2 viral stock collection (fig 5.4.5, A). Following expression, purification of WT-CLK3 was carried out. Figure 5.4.5, B displays a western blot with the purified recombinant WT-CLK3. The next step was to express and purify the mutant constructs. It was not possible to make any further progression with purification of the mutants as the labs were shut down due to the COVID pandemic. Unfortunately, time constraints prevented any further experiments.

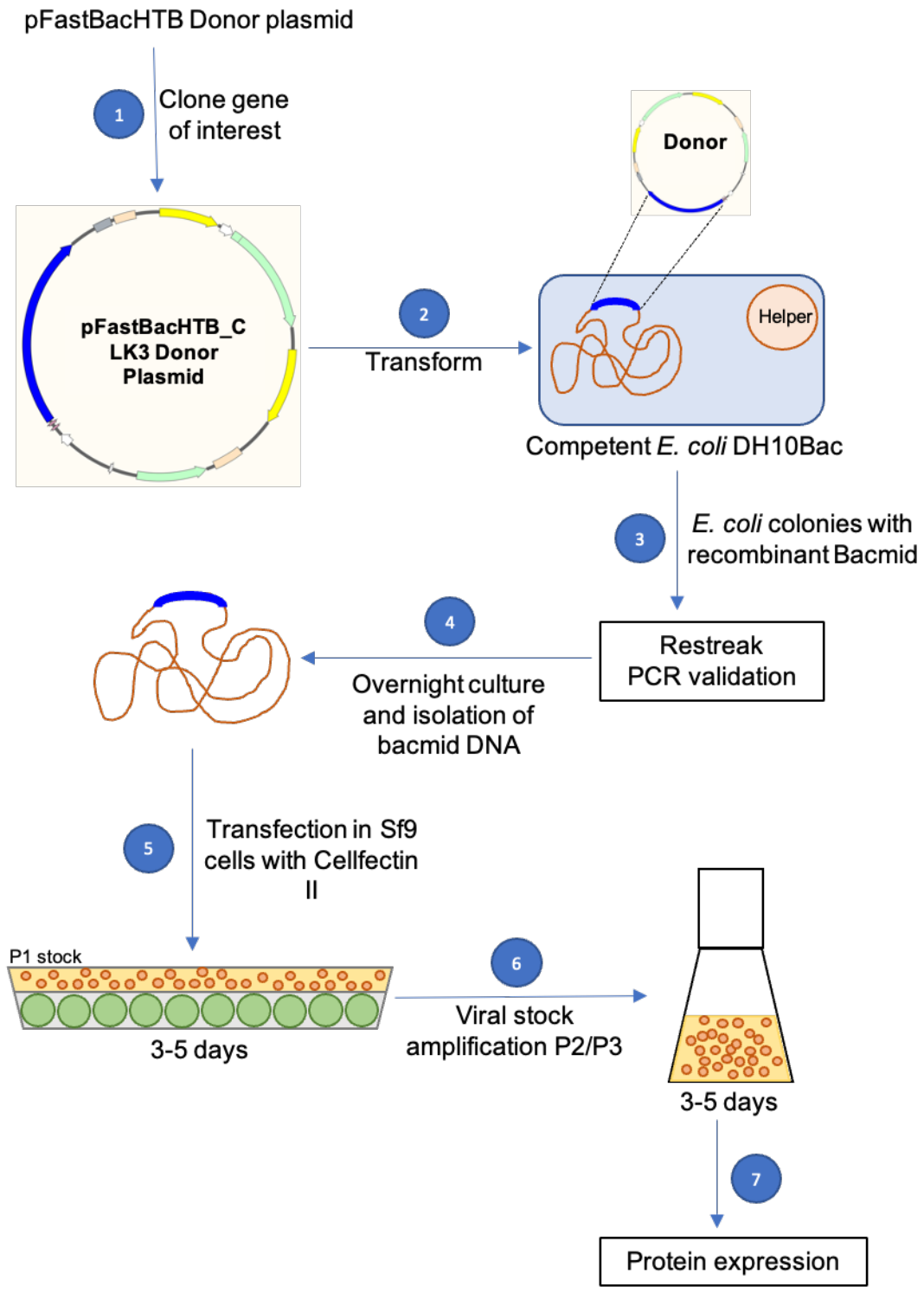



**Figure 5.4.4 The baculovirus system**
Schematic diagram detailing the baculovirus expression system as a flow chart. Sf9 cells are represented in green and viral particles are represented in red.

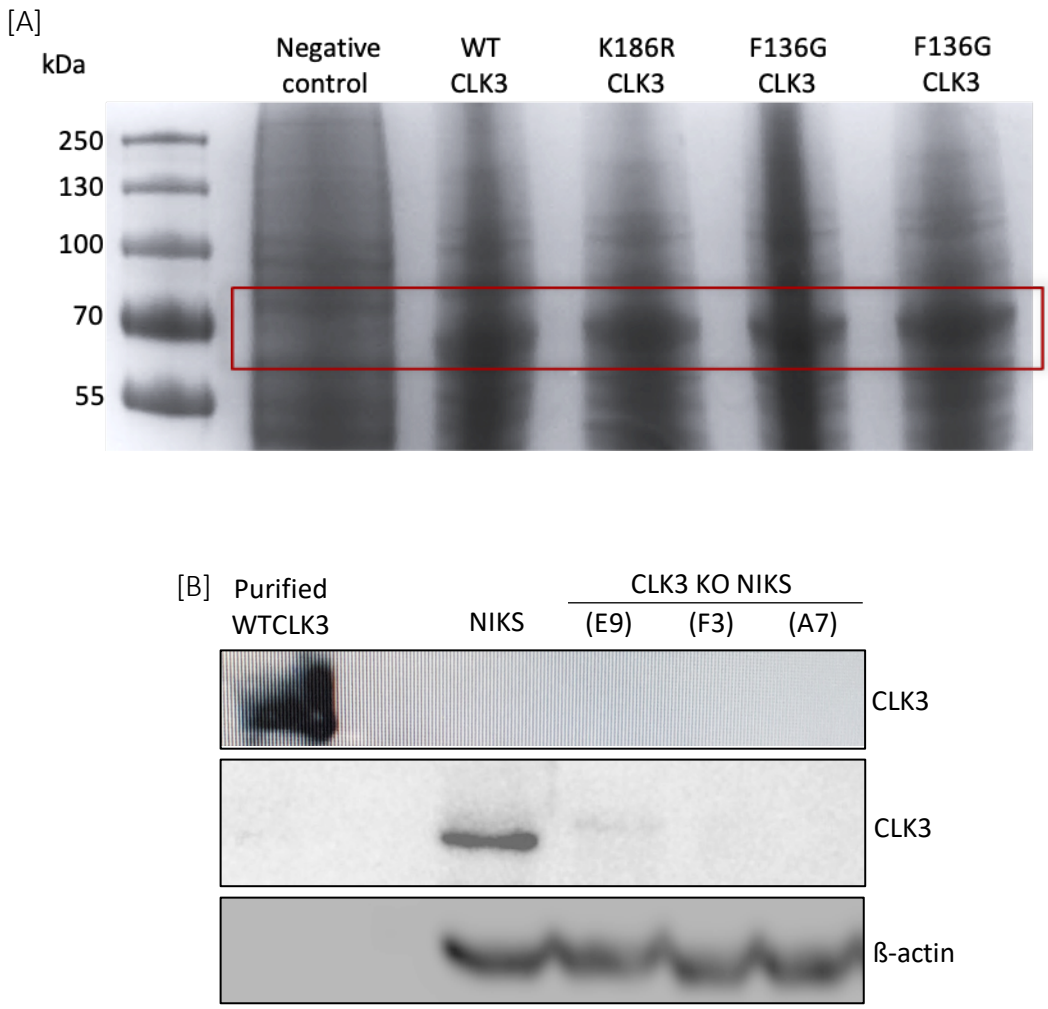



**Figure 5.4.5 Expression and purification of CLK3**
(A) Coomassie staining of SDS-page gel showing expression of all CLK3 constructs. Negative control was Sf9 cells without any bacmid. Cells were boiled in sample buffer at 95°C for 10mins and ran on SDS-PAGE gel. Coomassie staining was followed. (B) Western blot of WT-CLK3 following successful purification. NIKS cells were used as positive control, and three CLK3 KO clones (E9, F3, A7) were used as negative control. 1µl of purified CLK3 was used, with 50µg of lysate for the controls. Blot was incubated with CLK3-primary antibody overnight. β-Actin was used as a loading control. PageRuler plus prestained protein ladder was used as a molecular marker.

## 5.5 Discussion

The aim of this chapter was to identify unknown substrates of CLK3. The first step involved in this was to design, express and purify the WT CLK3 kinase, as well as three mutants, with the following mutations: K186R, F226A and F226G. The K186R substitution mutation produces a mutant protein with an inactive kinase function. Both F226A and F226G are substitutions to replace the bulky phenylalanine with either an alanine or glycine to produce a modified ATP binding pocket that would allow analogue forms of ATP to bind. All four versions of CLK3 were successfully designed and expressed using the baculovirus expression system, and the WT CLK3 was successfully purified. However, it was not possible to make any further progression to complete the objective of this chapter. This was due to two main factors; the lab group collaborating with us at the University of Kent during this experiment relocated, therefore we no longer had access to the baculovirus expression system to complete the purification of the three mutant CLK3s. In addition, due to the COVID pandemic, lab access was significantly limited for several months, which also prevented further experiments.

Previously, SRSF1 lacking the RS2 domain was demonstrated to still have the ability to relocate to the nucleus, however was unable to undergo further phosphorylation or re-disperse to the site of transcription. SRSF1 Is a known target of CLK3, therefore we purchased the 22 amino acids long RS2 domain peptide to be used as a substrate for validating the kinase function of all four CLK3 proteins following purification. After this is

confirmed, the step will be to determine if the mutants with the modified gatekeeper domain can utilise the analogue form of ATP ($N^6$-(benzyl)ATP-γ-S) to phosphorylate the peptide. It will also be important to confirm that WT and K186R kinases are not able to use the analogue form to phosphorylate its substrate. Following confirmation, the modified CLK3, along with the analogue ATP will be added to the cell lysate (from $CLK3^{-/-}$ NIKS) in vitro. The mixture will then be digested to form peptides which will then be incubated with iodoacetyl-agarose resin, resulting in covalent capture of the thiophosphate tagged peptides. Thiophosphorylation is resistant to dephosphorylation, thus increasing the sensitivity of this method. The resin will then be treated with Oxone, a peroxide agent, to oxidise the sulphur atom and release the phophopeptides, which can then be eluted and analysed using mass spectrometry to identify not only the substrates of CLK3, but also the region that is phosphorylated. Mass spectrometry has excellent sensitivity and resolution to identify a single protein, however analysis of phospho-proteins is typically more difficult due to the weaker signal exerted by the phosphopeptides, and it can also be difficult to observe signal from low-abundance phospho-proteins, especially when there is high-abundance of unphosphorylated proteins. Thiophosphorylation allows us to bypass these difficulties.

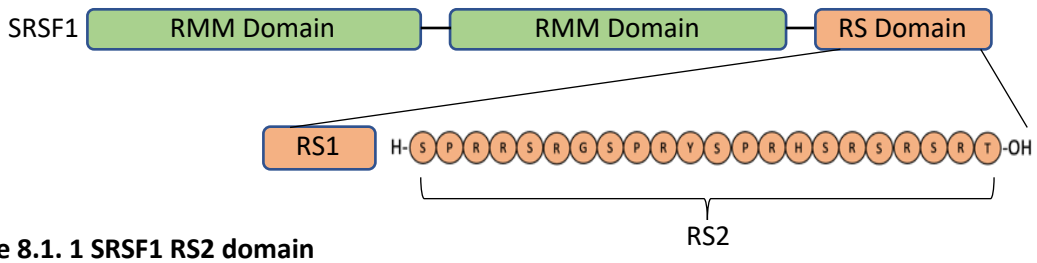



**Figure 8.1. 1 SRSF1 RS2 domain**
Diagram of SRSF1 structure. 22 amino acids long RS2 domain was purchased from Severnbiotech.

# 6 General Discussion

The results presented in this thesis show that: i) CRISPR knockout of CLK3 in NIKS with stably transduced HPV16 E6/E7 results in total loss of detectable E7 protein and significant loss of E6/E7 mRNA; ii) overexpression of CLK3 in CLK3 KO NIKS results in upregulation of HPV16 E7 in a kinase dependent manner; iii) CRISPR knockout of CLK3 in NIKS and shRNA knockdown of CLK3 in SiHa results in loss of c-MYC protein; iv) overexpression of c-MYC results in upregulation of HPV16 E7; v) siRNA knockdown of c-MYC results in significant loss of HPV16 E7 exclusively at the protein level; vi) inhibition of CLK3 with novel potent inhibitor KIN101982-001 results in loss of c-MYC and HPV16 E7 at protein level and E6/E7 and at mRNA level.

In previous work from my supervisor's lab, evidence was presented identifying a dependence in CLK3 for the viability of HPV-associated cancer cell lines, and the loss of HPV E6/E7 mRNA and E7 protein following siRNA and shRNA knockdown of CLK3 in multiple HPV positive cervical and head and neck cancer cell lines (Williams, 2021). As expected, the loss of E6/E7 transcription was accompanied by reactivation of the p53 and pRB pathways, and cells were observed to enter senescence.

To build on these observations CLK3 expression was deleted in NIKS through the generation of biallelic frameshift mutations in the first coding exon using CRISPR/Cas9, Part of the HPV16 genome consisting of the URR containing the p97 promoter, together with the E6/E7 coding region was cloned from the 93VU147T head and neck cancer cell line and

was stably transduced into both WT and CLK3 KO NIKS using a lentivirus. It is important to note that the NIKS did not enter senescence following loss of CLK3, and in fact, they showed no significant difference in proliferation rate compared to the WT NIKS. Early evidence showed stronger loss of HPV16 E7 protein and E6/E7 mRNA in CLK3 knockout NIKS than in the cancer cell lines transfected with CLK3 siRNA, consistent with the complete loss of CLK3 achieved in the KO NIKS. However, the completion of CLK3 add-back 'rescue' experiments to further validate the functional importance of CLK3 in E6/E7 regulation and to test the requirement for CLK3 kinase activity presented a major obstacle.

it was difficult to produce retroviral or lentiviral particles from both WT and kinase inactive (K186R) mutant plasmids. It was possible that CLK3 was somehow interfering with the viral packaging, preventing any viable particles from being made. CLKs have been previously shown to regulate HIV-1 gene expression, where inhibition of CLKs resulted in the failure of HIV-1 replication (Wong, et al., 2011). Following this, CLK3 overexpression was attempted using pEGFP-N1 plasmid in various HPV-associated cancer cell lines and NIKS, followed by long term selection with G418 sulfate. We observed that cells were shutting down the expression of both WT and K186R mutant CLK3 over time while GFP was continuously expressed successfully from the empty pEGFP-N1 plasmid. Studies have demonstrated that overexpression of proteins and enzymes can be harmful to cells, by overloading specific pathways, disrupting regulation, or aggregating together (Vavouri, et al, 2009, Tang and Amon, 2013, Makanae, et al., 2013). Overexpressing any protein long term can exhaust cellular resources. It is likely that overexpressing CLK3 for a long period

had toxic effects to the cell, by perhaps increasing the global mRNA transcription, leading to exhausting the cellular resources, and resulting in apoptosis. Transient transfection was attempted to avoid the possible toxic effects, however, due to the poor transfection efficiency (10-15%), it was not possible to observe any significant overexpression of CLK3 when analysing the entire cell population by qRT-PCR or western blotting. Luke Williams attempted electroporation transfection of pEGFP-CO-CLK3 using the MaxCyte system and was able to achieve much higher efficiency. However, 24hrs following FACS sorting, complete loss of CO-CLK3-GFP expression was observed, despite strong expression of GFP from the cells containing the empty vector.

It was possible to increase endogenous levels of CLK3, both at mRNA and protein level, by growing cells in hypoxic conditions or treating them with mimetic agents. However, we were unable to observe the direct effect of CLK3 rescue here, due to the global effect of hypoxia on gene expression. We observed a loss of HPV16 E7 protein when hypoxia was induced, however we were able to observe an increase in E6/E7 mRNA, which could have resulted due to the upregulation of CLK3. It will be necessary to investigate how E6E7 mRNA levels vary in CLK3 KO NIKS compared to the WT. This will further clarify if the upregulation of E6E7 mRNA was due to increased levels of CLK3. It was possible that other confounding factors were affecting the protein expression of E7. Interestingly, HIF-1α, the major TF responsible for hypoxic effects has been demonstrated to functionally counteract c-MYC activity by direct interaction and induction of Mxil (inhibits c-MYC via interactions with Max (Corn, et al., 2005, Zhang, et al., 2007). We have demonstrated the role of c-MYC in HPV16

E7 expression with loss of E7 being observed following siRNA knockdown of c-MYC. It is therefore possible that we observe a loss of HPV16 E7 protein following hypoxia due to the inhibition of c-MYC functions..

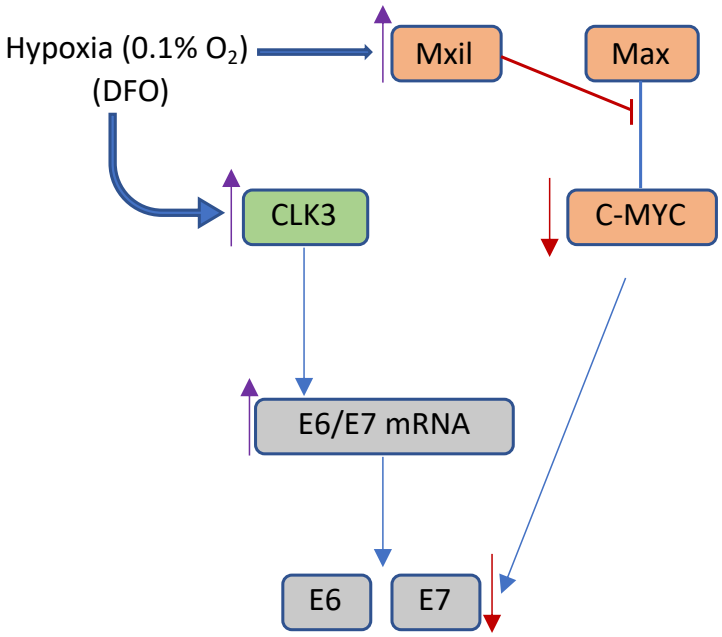



**Figure 8.1. 2 Theorised effect of severe hypoxia on E6/E7 via CLK3 and c-MYC**
Under severe hypoxic conditions (or with DFO treatment), CLK3 levels are upregulated. Hypoxia is also known to interfere with c-MYC functions by inducing Mxil to inhibit c-MYC:Max heterodimer complex formation. We therefore suspect that increased levels of CLK3 results in the increase of E6/E7 mRNA, and loss of c-MYC functions leads to either decreased stability or reduced translation of E7 (and possibly E6).

Due to the pleiotropic effects of hypoxia on cells and the large number of genes regulated under hypoxic conditions, it was not possible to definitively demonstrate the effect of CLK3 on E6/E7 using this approach. Transient transfection of WT and K186R mutant CLK3 plasmids in CLK3 KO NIKS, followed by immunostaining of the cells with p16 or HPV16 E7 specific antibodies enabled us to observe the direct overexpression effect of CLK3 in a

clean system with no endogenous expression of the kinase. It was possible to observe the upregulation of E7 and p16 following overexpression of WT CLK3, but not with K186R mutant. This was an important experiment as it not only validated our previous findings from loss-of-function experiments, but also highlighted the significance of the kinase function of CLK3 in regulating E6/E7. It suggested that a CLK3 kinase inhibitor could be utilized to regulate E6/E7 levels and therefore provided proof-of-principle for attempting to target CLK3 in HPV-associated cancers.

After we established the relationship between CLK3 and E6/E7, our focus turned to establishing the mechanism by which CLK3 was regulating E6/E7. As discussed previously, SRSF2 was considered as the obvious intermediate protein linking CLK3 and E6/E7 together. SRSF2 is a target of CLK3 and has been shown to modulate the expression of HPV E6/E7 (McFarlane, et al., 2015). We therefore hypothesised that CLK3 was acting via SRSF2 in HPV positive cancer cells. Initially, WT-CLK3 and K186R mutant CLK3 were transfected into various HPV positive cancer cells, and the cells were stained with SRSF2 specific antibody. The SRSF2 speckles were found to dissolve following overexpression of only WT CLK3, but not the kinase inactive mutant. Following this, SRSF2 expression was investigated in NIKS, with no significant difference in phosphor-SRSF2, or total SRSF2 (protein and mRNA) apparent between WT and CLK3 KO lines. It was possible that other members of the CLK family were contributing to account for the lack of CLK3 and correct the difference in SRSF2 phosphorylation. Numerous studies have demonstrated the redundancy among the CLK family to phosphorylate various SR proteins (Howell, et al., 1991, Colwill, et al., 1996, Nayler,

et al., 1997, Bullock, et al., 2009). It was also possible that other SR proteins were involved in contributing towards CLK3 regulation of E6/E7. It is therefore necessary to further investigate the expression/phosphorylation status of other SR proteins, particularly focusing on those previously implicated in HPV E6/E7 modulation (SRSF1-4).

As previously stated, CLK3 was demonstrated to regulate the stability of c-MYC in CCA cells via phosphorylation of USP13 (zhou, et al., 2020). Given the significance of c-MYC in a wide array of cancers and its role as a TF, we hypothesised that CLK3 could be acting via c-MYC to increase E6/E7 levels. We first established that CLK3 regulates the expression of c-MYC in NIKS and SiHa cells, with loss of c-MYC protein being observed following either CRISPR knockout (NIKS) or shRNA knockdown (SiHa) of CLK3. We also demonstrated that overexpression of WT-CLK3 but not K186R-CLK3 in NIKS resulted in the upregulation of c-MYC, although further independent repeats of the overexpression experiment are needed to conclusively state this. Following this, we overexpressed c-MYC in NIKS with stably transduced E6/E7 and observed an upregulation of E7. siRNA knockdown of c-MYC in CaSki, SiHa, and NIKS further validated this, with loss of HPV16 E7 being observed across all three cell lines. To our knowledge, we are the first to observe a role for c-MYC in the regulation of E7 expression, thus we have revealed a novel role for this much-studied proto-oncogene.

Interestingly, my results suggest c-MYC acts exclusively to modulate either E7 stability or translation (either directly or indirectly, via the transcriptional activation of an intermediary gene), with no loss of E6/E7 mRNA being observed following siRNA

knockdown. This suggests that CLK3 upregulates E6/E7 mRNA transcription through an unknown pathway, possibly via SR proteins, and E7 protein via c-MYC. Figure 5.5.1 helps visualise this theory. We only investigated HPV16 positive cells, therefore further investigation is required to determine the role of c-MYC in regulating E7 expression in other HPV types. It is also important to note that due the lack of a HPV16 E6 specific antibody, it was not possible to investigate the relationship between c-MYC and E6, and therefore currently, we can only conclude that c-MYC is involved in modulating the translation/stability of HPV16 E7.

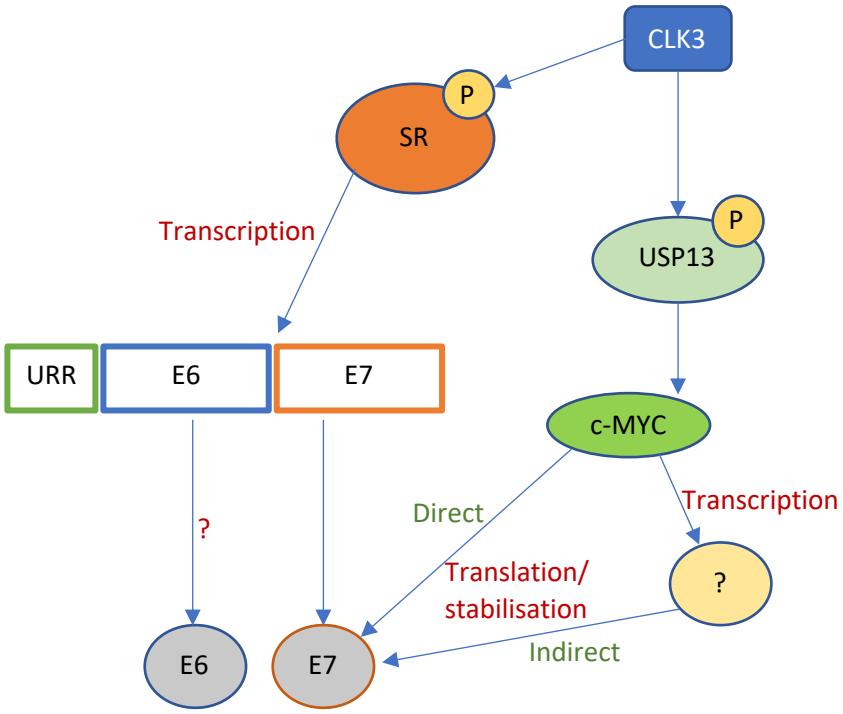



**Figure 5.5.1 CLK3 controls E6/E7 at both transcription and translation/stabilisation**
CLK3 is known to phosphorylate SR proteins and USP13. From the evidence gathered during this study, we theorise two modes of action for CLK3 on E6/E7 mRNA. CLK3 is thought to phosphorylate one or more of the SR proteins and redistribute them to the site of mRNA transcription, where these specific SR proteins are thought to be involved in the transcription of E6/E7 mRNA. Additionally, CLK3 phosphorylates USP13, which then stabilises c-MYC, which is thought to be involved in the translation of E6/E7. We have demonstrated that c-MYC plays a direct role in either HPV16 E7

protein stability or translation. It is also possible that c-MYC acts indirectly on E7 via transcription of another unknown protein. Whether c-MYC regulates E6 protein levels remains unknown.

Early studies have demonstrated that E7 protein undergoes various post-translational modifications, including phosphorylation by casein kinase 2 in the CR2 domain (Firzlaff, et al., 1989, Barbosa, et al., 1990), and ubiquitination by UBE2L3 and Cullin 1/Skp2-containing E3 ubiquitin kinase (Oh, et al., 2004). It is possible for c-MYC to be involved indirectly in E7 protein regulation via the transcriptional activation of an intermediate protein, which then could act on E7 translation or stability.

c-MYC has been implicated in a wide range of cancers. In fact, the oncogene is deregulated in over 50% of all human cancers, and this deregulation is commonly associated with poor prognosis and unfavorable patient survival. More importantly, upregulation of c-MYC has been noted in HPV E6E7-transfected cells as well as patients with the cancer, where it has been found as a negative prognostic marker for survival (kübler, et al., 2015). As a TF rather than an enzyme, c-MYC, lacks a specific active site into which small molecule inhibitors can bind, which typically makes it difficult to target and inhibit its functions using strategies analogous to those that have been so successful for targeting kinases in cancer treatment. In addition, c-MYC is an intracellular protein, c-MYC cannot be targeted using antibodies. Other strategies have been explored for targeting MYC indirectly. Inhibition of BRD4, which controls MYC transcription, with JQ1 showed potent anti-cancer effects in multiple haemopoietic cancer cell lines and pancreatic ductal adenocarcinoma (PDAC) cell line exhibiting c-MYC overexpression (Mertz, et al., 2011, Mazur, et al., 2015). Inhibition of

the PI3K|AKT|mTOR pathway, which is involved in regulating MYC mRNA translation, showed remarkable efficacy in MYC-driven cancer cell lines including neuroblastoma, small-cell lung carcinoma, and breast cancer (Yu, et al., 2001, Frost, et al., 2004, Chapuis, et al., 2010). More recently, inhibition of USP7, which directly binds and stabilizes N-MYC, with P22077 significantly suppressed growth of N-MYC amplified neuroblastoma in a xenograft model (Tavana, et al., 2016). In addition to evidence from previous studies in CCA, we have shown here that loss of CLK3 can lead to the loss of c-MYC in HPV16-associated cancer cell lines and NIKS. It is important to expand further and investigate the effect of CLK3 in other cancer cell lines, potentially focusing on those in which c-MYC is overexpressed, or in which it is known to play a key functional role. Targeting CLK3 to destabilise c-MYC could prove to be a potential therapy against a wide array of cancers.

Despite the availability of effective vaccines, HPV still remains a global burden, especially in third world countries. The overall incidence of cervical cancer has reduced thanks to the effective screening and vaccination programmes implemented in many countries, however cervical cancer remains the fourth most common cause of cancer among women and fourth leading cause of deaths worldwide. 84% of the estimated 528,000 new cases were diagnosed in less developed regions, where cervical cancer ranks second, only after breast cancer. Improved accessibility of the vaccine, increased screening, and better education of cervical cancer in developing/third world countries are the only way to reduce the global HPV burden. It is also essential to identify treatments to improve the outcome for those with recurrent or metastatic HPV-associated cancers.

In collaboration with Kinsensus Ltd, we identified a CDK inhibitor, R547, from a 633-compound library, with high potency against CLK3 (28.3nM). R547 was then further modified to increase the potency and potentially the selectivity towards CLK3, generating KIN101982-001, which has an $IC_{50}$ of 10nM for CLK3 but for which we don't yet have selectivity data. Early data from protein and RNA analysis of NIKS stably transduced with E6/E7 treated with KIN101982-001 seems promising. Significant loss of c-MYC and HPV16 E7 were observed at protein level and significant loss of E6/E7 and p16 mRNA following treatment for 3 hours was apparent from qRT-PCR analysis. It will be important to examine the effect of the inhibitor in HPV-associated cancer cell lines, and to determine whether the cells enter senescence as expected. However, this compound is still expected hit multiple CDKs, therefore further modifications will be necessary to improve the selectivity against CLK3 before any effects on cell viability can be attributed to CLK3 inhibition.

Our lab has demonstrated that CLK3 knockdown results in cell cycle arrest followed by senescence in SiHa and CaSki cancer cell lines. We will be conducting shRNA knockdown of CLK3 followed by CLK3 re-expression (in shCLK3 SiHa and CaSki) to assess whether cells are able to escape senescence. In addition, we will use senolytic drugs Dasatinib (approved anti-cancer drug) and Navitoclax to test whether these could be combined with CLK3 inhibition to eliminate senescent tumour cells.

Doxycycline controlled Tet system will be implemented in a subcutaneous xenograft model to further characterise CLK3 in a preclinical model. The SiHa cells transfected with shRNA targeting CLK3 (same as the ones used in chapter 3.1) will be used along with shNT SiHa control cells. The tumour would be allowed to grow to a palpable size (e.g., 200mm$^3$) before CLK3 knockdown is initiated by adding doxycycline in the drinking water of the mice. In addition, and given out findings on the importance of c-MYC for the regulation of E7 by CLK3, it might be interesting to use a recently published mouse model to further study the effects of CLK3 inhibition. This model was created following transfection of the AMES-16 plasmids (constitutively active AKT, c-MYC, HPV16-E6/E7 with a luciferase reporter gene, and Sleeping Beauty transposase (SB100)) into the reproductive tract of C57BL/6 mice and has been demonstrated as an improved preclinical model to cervical carcinoma (Henkle, et al., 2021). Current preclinical models for cervical cancer cell lines have multiple drawbacks; i) they are not derived from anatomically relevant tissue, ii) they do not accurately predict clinical outcomes of novel treatment, iii) and they do not model clinical progression through precancerous lesions. This novel model will look to improve on this by i) locally expressing HPV E6E7 in the reproductive tract of mice, ii) being able to monitor tumour progression from HSIL to SCC, iii) presenting the opportunity to use multiple high-risk HPV types, iv) and being accompanied by a mutated PI3K/AKT pathway, a common occurrence in cervical cancers.

Currently, CLK3 is known to phosphorylate SR proteins and USP13. To further characterise the kinase, and increase our understanding of the relationship between CLK3

and E6/E7, we attempted to identify other unknown substrates of the kinase. Shokat's method of ASKA technology was considered as the method to determine any unknown substrates of CLK3. CLK3 was successfully modified to create two mutant kinases with a modified gatekeeper domain that can use analogue forms of ATP to tag its targets for mass-spectrometry analysis. We initially faced a major obstacle when attempting to express the kinase, and had no success using BL21 (DH3) *E. coli* for the expression. We then opted to use the baculovirus expression system (Sf9 cells) in collaboration with the Toseland group at the University of Kent. The system had multiple failures due to severe contamination issues, however, we were able to successfully express the WT and K186R kinase dead CLK3, along with the two kinases with the modified gatekeeper domain (F236A, F236G). We also had success purifying the WT kinase, however, as a result of the relocation of our collaborators, and due to the COVID pandemic, we lacked the resources to complete the large-scale expression necessary for purification of the remaining three mutants.

To conclude, the data presented here demonstrate that CLK3 regulates E6/E7 mRNA expression, and E7 protein via c-MYC. Our findings suggest that CLK3 represents a potential novel therapeutic target for HPV-driven cancers. Currently, there are no treatments for HPV infections and CLK3 inhibitors also presents an opportunity as treatment to eliminate HPV infections, perhaps being considered as a topical ointment for the site of infection. Finally, CLK3 also has the potential to be targeted in a wide array of cancers/diseases where c-MYC is overexpressed or known to play a functional role.

# 7 References

Agger, K., Cloos, P.A.C. and Rudkjaer, L. (2009). The H3K27me3 demethylase JMJD3 contributes to the activation of the INK4A-ARF locus in response to oncogene- and stress-induced senescence. *Genes & Development*, 23(10), pp.1171–1176.

Akagi, K., Li, J. and Broutian, T.R. (2013). Genome-wide analysis of HPV integration in human cancers reveals recurrent, focal genomic instability. *Genome Research*, 24(2), pp.185–199.
Aksoy, P., Gottschalk, E.Y. and Meneses, P.I. (2017). HPV entry into cells. *Mutation research*, 772, pp.13–22.

Alcorta, D.A., Xiong, Y. and Phelps, D. (1996). Involvement of the cyclin-dependent kinase inhibitor p16 (INK4a) in replicative senescence of normal human fibroblasts. *Proceedings of the National Academy of Sciences*, 93(24), pp.13742–13747.

Allen-Hoffmann, B.L., Schlosser, S.J. and Ivarie, C.A.R. (2000). Normal Growth and Differentiation in a Spontaneously Immortalized Near-Diploid Human Keratinocyte Cell Line, NIKS. *Journal of Investigative Dermatology*, 114(3), pp.444–455.

Alvarez-Salas, L.M., Cullinan, A.E. and Siwkowski, A. (1998). Inhibition of HPV-16 E6/E7 immortalization of normal keratinocytes by hairpin ribozymes. *Proceedings of the National Academy of Sciences*, 95(3), pp.1189–1194.

Amano, M., Tsumura, Y. and Taki, K. (2010). A proteomic approach for comprehensively screening substrates of protein kinases such as Rho-kinase. *PloS One*, 5(1), p.e8704.

Araki, S., Dairiki, R. and Nakayama, Y. (2015). Inhibitors of CLK Protein Kinases Suppress Cell Growth and Induce Apoptosis by Modulating Pre-mRNA Splicing. *PLoS ONE*, 10(1).

Arbeit, J.M., Münger, K. and Howley, P.M. (1993). Neuroepithelial carcinomas in mice transgenic with human papillomavirus type 16 E6/E7 ORFs. *The American Journal of Pathology*, 142(4), pp.1187–1197.

Ashrafi, G.H., Brown, D.R. and Fife, K.H. (2006). Down-regulation of MHC class I is a property common to papillomavirus E5 proteins. *Virus research*, 120(1–2), pp.208–11.

Aubol, B.E., Plocinik, R.M. and Hagopian, J.C. (2013). Partitioning RS Domain Phosphorylation in an SR Protein through the CLK and SRPK Protein Kinases. *Journal of molecular biology*, 425(16).

B, Y., R, F. and K, L. (1994). The Doa Locus Encodes a Member of a New Protein Kinase Family and Is Essential for Eye and Embryonic Development in Drosophila Melanogaster. *Genes & development*.

Baker, C.C., Phelps, W.C. and Lindgren, V. (1987). Structural and transcriptional analysis of human papillomavirus type 16 sequences in cervical carcinoma cell lines. *Journal of Virology*, 61(4), pp.962–971.

Banks, L., Pim, D. and Thomas, M. (2012). Human tumour viruses and the deregulation of cell polarity in cancer. *Nature Reviews Cancer*, 12(12), pp.877–886.

Babu, N., Pinto, S.M. and Biswas, M. (2020). Phosphoproteomic analysis identifies CLK1 as a novel therapeutic target in gastric cancer. *Gastric Cancer*, 23(5), pp.796–810.

Bachmann, M.F., Rohrer, U.H. and Kündig, T.M. (1993). The influence of antigen organization on B cell responsiveness. *Science (New York, N.Y.)*, 262(5138), pp.1448–1451.

Barbosa, M.S., Edmonds, C. and Fisher, C. (1990). The region of the HPV E7 oncoprotein homologous to adenovirus E1a and Sv40 large T antigen contains separate domains for Rb binding and casein kinase II phosphorylation. *The EMBO Journal*, 9(1), pp.153–160.

Barradas, M., Anderton, E. and Acosta, J.C. (2009). Histone demethylase JMJD3 contributes to epigenetic control of INK4a/ARF by oncogenic RAS. *Genes & Development*, 23(10), pp.1177–1182.

Bayo, S., Bosch, F.X. and de Sanjosé, S. (2002). Risk factors of invasive cervical cancer in Mali. *International Journal of Epidemiology*, 31(1), pp.202–209.

Bergant Marušič, M., Ozbun, M.A. and Campos, S.K. (2012). Human papillomavirus L2 facilitates viral escape from late endosomes via sorting nexin 17. *Traffic (Copenhagen, Denmark)*, 13(3), pp.455–467.

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

Bhullar, K.S., Lagarón, N.O. and McGowan, E.M. (2018). Kinase-targeted cancer therapies: progress, challenges and future directions. *Molecular Cancer*, 17(1).

Blackwell TK, Kretzner L, Blackwood EM, Eisenman RN, Weintraub H. Sequence-specific DNA binding by the c-MYC protein. *Science*. 1990;250(4984):1149–51.

Blackwell TK, Huang J, Ma A, Kretzner L, Alt FW, Eisenman RN, et al. Binding of MYC proteins to canonical and noncanonical DNA sequences. *Mol Cell Biol*. 1993;13(9):5216–24.

Blackwood EM, Eisenman RN. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with MYC. *Science*. 1991;251(4998):1211–7.

Blethrow, J.D., Glavy, J.S. and Morgan, D.O. (2008). Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. *Proceedings of the National Academy of Sciences*, 105(5), pp.1442–1447.

Bossler, F., Kuhn, B.J. and Günther, T. (2019). Repression of Human Papillomavirus Oncogene Expression under Hypoxia Is Mediated by PI3K/mTORC2/AKT Signaling. *mBio*, 10(1).

Bowler, E., Porazinski, S. and Uzor, S. (2018). Hypoxia leads to significant changes in alternative splicing and elevated expression of CLK splice factor kinases in PC3 prostate cancer cells. *BMC Cancer*, 18(1).

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 Research*, 56(20), pp.4620–4624.

Bravo, I.G., de Sanjosé, S. and Gottschling, M. (2010). The clinical importance of understanding the evolution of papillomaviruses. *Trends in Microbiology*, 18(10), pp.432–438.

Brotherton, J.M.L. (2019). Impact of HPV vaccination: Achievements and future challenges. *Papillomavirus Research*, 7, pp.138–140.

Brown, J.M. (1999). The hypoxic cell: a target for selective cancer therapy--eighteenth Bruce F. Cain Memorial Award lecture. *Cancer Research*, 59(23), pp.5863–5870.

Bullock, A.N., Das, S. and Debreczeni, J.É. (2009). Kinase Domain Insertions Define Distinct Roles of CLK Kinases in SR Protein Phosphorylation. *Structure(London, England:1993)*, 17(3–2), pp.352–362.

Burd, E. (2003). Human Papillomavirus and Cervical Cancer. *Clinical Microbiology Reviews*, 16(1), pp.1–17.

Butz, K., Denk, C. and Ullmann, A. (2000). Induction of apoptosis in human papillomaviruspositive cancer cells by peptide aptamers targeting the viral E6 oncoprotein. *Proceedings of the National Academy of Sciences*, 97(12), pp.6693–6697.

Butz, K., Ristriani, T. and Hengstermann, A. (2003). siRNA targeting of the viral E6 oncogene efficiently kills human papillomavirus-positive cancer cells. *Oncogene*, 22(38), pp.5938–5945.

Cáceres, 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 & Development*, 12(1), pp.55–66.

Campbell, J., Ryan, C.J. and Lord, C.J. (2018). Identifying Genetic Dependencies in Cancer by Analyzing siRNA Screens in Tumor Cell Line Panels. *Methods in Molecular Biology*, 1711, pp.83–99.

Carroll, M., Ohno-Jones, S. and Tamura, S. (1997). CGP 57148, a tyrosine kinase inhibitor, inhibits the growth of cells expressing BCR-ABL, TEL-ABL, and TEL-PDGFR fusion proteins. *Blood*, 90(12), pp.4947–4952.

Castro-Muñoz, L.J., Manzo-Merino, J. and Muñoz-Bello, J.O. (2019). The Human Papillomavirus (HPV) E1 protein regulates the expression of cellular genes involved in immune response. *Scientific Reports*, 9(1).

Center for Drug Evaluation and Research (2021). Oncology (Cancer) / Hematologic Malignancies Approval Notifications. *FDA*. [online] Available at: https://www.fda.gov/drugs/resources-information-approved-drugs/oncology-cancer-hematologic-malignancies-approval-notifications.

Chang, J.-L., Tsao, Y.-P. and Liu, D.-W. (2001). The expression of HPV-16 E5 protein in squamous neoplastic changes in the uterine cervix. *Journal of Biomedical Science*, 8(2), pp.206–213.

Chapuis, N., Tamburini, J., Green, A., et al. (2010) Dual inhibition of PI3K and mTORC1/2 signaling by NVP-BEZ235 as a new therapeutic strategy for acute myeloid leukemia. *Clin. Cancer Res.*16, 5424–5435.

Chatterjee, A. (2014). The next generation of HPV vaccines: nonavalent vaccine V503 on the horizon. *Expert Review of Vaccines*, 13(11), pp.1279–1290.

Chaturvedi, A.K., Anderson, W.F. and Lortet-Tieulent, J. (2013). Worldwide Trends in Incidence Rates for Oral Cavity and Oropharyngeal Cancers. *Journal of Clinical Oncology*, 31(36), pp.4550–4559.

Chaturvedi, A.K., Engels, E.A. and Anderson, W.F. (2008). Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *Journal of clinical oncology: official journal of the American Society of Clinical Oncology*, 26(4), pp.612–9.

Chen, J., Xue, Y. and Poidinger, M. (2014). Mapping of HPV transcripts in four human cervical lesions using RNAseq suggests quantitative rearrangements during carcinogenic progression. *Virology*, 462–463, pp.14–24.

Chen, Z., Schiffman, M. and Herrero, R. (2007). Human papillomavirus (HPV) types 101 and 103 isolated from cervicovaginal cells lack an E6 open reading frame (ORF) and are related to gamma-papillomaviruses. *Virology*, 360(2), pp.447–453.

Chi, Y., Welcker, M. and Hizli, A.A. (2008). Identification of CDK2 substrates in human cell lysates. *Genome Biology*, 9(10), p.R149.

Choe, J., Vaillancourt, P. and Stenlund, A. (1989). Bovine papillomavirus type 1 encodes two forms of a transcriptional repressor: structural and functional analysis of new viral cDNAs. *Journal of Virology*, 63(4), pp.1743–1755.

Ciesielska, U., Nowińska, K. and Podhorska-Okołów, M. (2012). The role of human papillomavirus in the malignant transformation of cervix epithelial cells and the importance

of vaccination against this virus. *Advances in Clinical and Experimental Medicine: Official Organ Wroclaw Medical University*, 21(2), pp.235–244.

Cohen-Eliav, M., Golan-Gerstl, R. and Siegfried, Z. (2013). The splicing factor SRSF6 is amplified and is an oncoprotein in lung and colon cancers. *The Journal of Pathology*, 229(4), pp.630–639.

Cole, M.D. and Cowling, V.H. (2008). Transcription-independent functions of MYC: regulation of translation and DNA replication. *Nature Reviews Molecular Cell Biology*, 9(10), pp.810–815.

Coller HA, Grandori C, Tamayo P, Colbert T, Lander ES, Eisenman RN, et al. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. *Proc Natl Acad Sci U S A.* 2000;97:3260–3265.

Comerford, S.A., Maika, S.D. and Laimins, L.A. (1995). E6 and E7 expression from the HPV 18 LCR: development of genital hyperplasia and neoplasia in transgenic mice. *Oncogene*, 10(3), pp.587–597.

Colombo, N., Dubot, C. and Lorusso, D. (2021). Pembrolizumab for Persistent, Recurrent, or Metastatic Cervical Cancer. *New England Journal of Medicine*.

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

Corkery, D.P., Holly, A.C. and Lahsaee, S. (2015). Connecting the speckles: Splicing kinases and their role in tumorigenesis and treatment response. *Nucleus (Austin, Tex.)*, 6(4), pp.279–288.

Corn, P.G., Ricci, M.S. and Scata, K.A. (2005). Mxi1 is induced by hypoxia in a HIF-1–dependent manner and protects cells from c-Myc-induced apoptosis. *Cancer Biology & Therapy*, 4(11), pp.1285–1294.

Crook, T., Vousden, K.H. and Tidy, J.A. (1991). Degradation of p53 can be targeted by HPV E6 sequences distinct from those required for p53 binding and trans-activation. *Cell*, 67(3), pp.547–556.

Culp, T.D., Budgeon, L.R. and Marinkovich, M.P. (2006). Keratinocyte-Secreted Laminin 5 Can Function as a Transient Receptor for Human Papillomaviruses by Binding Virions and Transferring Them to Adjacent Cells. *Journal of Virology*, 80(18), pp.8940–8950.

Curado, M.P. and Boyle, P. (2013). Epidemiology of head and neck squamous cell carcinoma not related to tobacco or alcohol. *Current opinion in oncology*, 25(3), pp.229–34.

Dang, C.V., Le, A. and Gao, P. (2009). MYC-Induced Cancer Cell Energy Metabolism and Therapeutic Opportunities. *Clinical Cancer Research*, 15(21), pp.6479–6483.

Danos, O., Katinka, M. and Yaniv, M. (1982). Human papillomavirus 1a complete DNA sequence: a novel type of genome organization among papovaviridae. *The EMBO journal*, 1(2), pp.231–236.

Das, R., Yu, J. and Zhang, Z. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Molecular Cell*, 26(6), pp.867–881.

Das, S. and Krainer, A.R. (2014). Emerging functions of SRSF1, splicing factor and oncoprotein, in RNA metabolism and cancer. *Molecular cancer research: MCR*, 12(9), pp.1195–1204.

Davies, R., Hicks, R. and Crook, T. (1993). Human papillomavirus type 16 E7 associates with a histone H1 kinase and with p107 through sequences necessary for transformation. *Journal of Virology*, 67(5), pp.2521–2528.

Davy, C.E., Jackson, D.J. and Wang, Q. (2002). Identification of a G2 Arrest Domain in the E1ΛE4 Protein of Human Papillomavirus Type 16. *Journal of Virology*, 76(19), pp.9806–9818.

Day, P.M., Roden, R.B.S. and Lowy, D.R. (1998). The Papillomavirus Minor Capsid Protein, L2, Induces Localization of the Major Capsid Protein, L1, and the Viral Transcription/Replication Protein, E2, to PML Oncogenic Domains. *Journal of Virology*, 72(1), pp.142–150.

Dominguez Sanchez, J. M. 'Role of CLK3 in HIF1a driven cardiac hypertrophy. ETH Zurich. Phd thesis. 2017.

Doorbar, J. (2013). The E4 protein; structure, function and patterns of expression. *Virology*, 445(1–2), pp.80–98.

Doorbar, J., Egawa, N. and Griffin, H. (2015). Human papillomavirus molecular biology and disease association. *Reviews in Medical Virology*, 25(2), pp.2–23.
Doorbar, J., Quint, W. and Banks, L. (2012). The Biology and Life-Cycle of Human Papillomaviruses. *Vaccine*, 30, pp.F55–F70.

Dowhanick, J.J., McBride, A.A. and Howley, P.M. (1995). Suppression of cellular proliferation by the papillomavirus E2 protein. *Journal of Virology*, 69(12), pp.7791–7799.

D'Souza, G., Agrawal, Y. and Halpern, J. (2009). Oral Sexual Behaviors Associated with Prevalent Oral Human Papillomavirus Infection. *The Journal of Infectious Diseases*, 199(9), pp.1263–1269.

Duncan, P.I., Howell, B.W. and Marius, R.M. (1995). Alternative splicing of STY, a nuclear dual specificity kinase. *The Journal of Biological Chemistry*, 270(37), pp.21524–21531.

Duncan, P.I., Stojdl, D.F. and Marius, R.M. (1997). In vivo regulation of alternative pre-mRNA splicing by the Clk1 protein kinase. *Molecular and Cellular Biology*, 17(10), pp.5996–6001.

Duncan, P.I., Stojdl, D.F. and Marius, R.M. (1998). The Clk2 and Clk3 dual-specificity protein kinases regulate the intranuclear distribution of SR proteins and influence pre-mRNA splicing. *Experimental Cell Research*, 241(2), pp.300–308.

Durst, M., Gissmann, L. and Ikenberg, H. (1983). A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proceedings of the National Academy of Sciences*, 80(12), pp.3812–3815.

Dyson, N., Guida, P. and Münger, K. (1992). Homologous sequences in adenovirus E1A and human papillomavirus E7 proteins mediate interaction with the same set of cellular proteins. *Journal of Virology*, 66(12), pp.6893–6902.

Ebbesen, P., Eckardt, K.-U. and Ciampor, F. (2004). Linking measured intercellular oxygen concentration to human cell functions. *Acta Oncologica*, 43(6), pp.598–600.

Edmond, V., Moysan, E. and Khochbin, S. (2011). Acetylation and Phosphorylation of SRSF2 Control Cell Fate Decision in Response to Cisplatin. *The EMBO journal*, 30(3), pp.510–23.

El Marabti, E. and Younis, I. (2018). The Cancer Spliceome: Reprograming of Alternative Splicing in Cancer. *Frontiers in Molecular Biosciences*, 5(80).

Elkon, R., Loayza-Puch, F. and Korkmaz, G. (2015). Myc coordinates transcription and translation to enhance transformation and suppress invasiveness. *EMBO reports*, 16(12), pp.1723–1736.

Epstein, A.C., Gleadle, J.M. and McNeill, L.A. (2001). C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell*, 107(1), pp.43–54.

Erkelenz, S., Mueller, W. and Evans, M. (2013). Position-dependent Splicing Activation and Repression by SR and hnRNP Proteins Rely on Common Mechanisms. *RNA (New York, N.Y.)*, 19(1).

Escobar-Hoyos, L., Knorr, K. and Abdel-Wahab, O. (2019). Aberrant RNA Splicing in Cancer. *Annual Review of Cancer Biology*, 3(1), pp.167–185.

Fan, L., Iyer, J. and Zhu, S. (2001). Inhibition of N-MYC expression and induction of apoptosis by iron chelation in human neuroblastoma cells. *Cancer Research*, 61(3), pp.1073–1079.

Ferlay, J., Soerjomataram, I. and Dikshit, R. (2014). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, 136(5), pp.E359–E386.

Firzlaff, J.M., Galloway, D.A. and Eisenman, R.N. (1989). The E7 protein of human papillomavirus type 16 is phosphorylated by casein kinase II. *The New Biologist*, 1(1), pp.44–53.

Flores, E., Allen-Hoffmann, B.L. and Lee, D. (1999). Establishment of the Human Papillomavirus Type 16 (HPV-16) Life Cycle in an Immortalized Human Foreskin Keratinocyte Cell Line. *Virology*, 262(2), pp.344–354.

Fregoso, O.I., Das, S. and Akerman, M. (2013). Splicing-factor oncoprotein SRSF1 stabilizes p53 via RPL5 and induces cellular senescence. *Molecular Cell*, 50(1), pp.56–66.

Frost, P., Moatamed, F., Hoang, B., et al. (2004). In vivo antitumor effects of the mTOR inhibitor CCI-779 against human multiple myeloma cells in a xenograft model. *Blood* 104, 4181–4187.

Fu, X.D. and Maniatis, T. (1992). The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site. *Proceedings of the National Academy of Sciences of the United States of America*, 89(5), pp.1725–1729.

Funk, J.O., Waga, S. and Harry, J.B. (1997). Inhibition of CDK activity and PCNA-dependent DNA replication by p21 is blocked by interaction with the HPV-16 E7 oncoprotein. *Genes & Development*, 11(16), pp.2090–2100.

Gadducci, A., Guerrieri, M.E. and Greco, C. (2013). Tissue biomarkers as prognostic variables of cervical cancer. *Critical reviews in oncology/hematology*, 86(2), pp.104–29.

Ganti, K., Broniarczyk, J. and Manoubi, W. (2015). The Human Papillomavirus E6 PDZ Binding Motif: From Life Cycle to Malignancy. *Viruses*, 7(7), pp.3530–3551.

Ghosh, G. and Adams, J.A. (2011). Phosphorylation mechanism and structure of serine-arginine protein kinases. *The FEBS journal*, 278(4), pp.587–97.

Gillison, M.L., Koch, W., Capone, R. (2000). Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *JNCI*, 92 (9). pp.709-720.

Gillison, M.L., D'Souza, G. and Westra, W. (2008). Distinct risk factor profiles for human papillomavirus type 16-positive and human papillomavirus type 16-negative head and neck cancers. *Journal of the National Cancer Institute*, 100(6), pp.407–20.

Giroglou, T., Florin, L. and Schafer, F. (2001). Human Papillomavirus Infection Requires Cell Surface Heparan Sulfate. *Journal of Virology*, 75(3), pp.1565–1570.

Giuliano, A.R., Tortolero-Luna, G. and Ferrer, E. (2008). Epidemiology of Human Papillomavirus Infection in Men, Cancers other than Cervical and Benign Conditions. *Vaccine*, 26, pp.K17–K28.

Goodwin, E.C., Yang, E. and Lee, C.J. (2000). Rapid induction of senescence in human cervical carcinoma cells. *Proceedings of the National Academy of Sciences of the United States of America*, 97(20), pp.10978–10983.

Gottschling, M., Bravo, I.G. and Schulz, E. (2011). Modular organizations of novel cetacean papillomaviruses. *Molecular Phylogenetics and Evolution*, 59(1), pp.34–42.

Graham, K. and Unger, E. (2018). Overcoming tumor hypoxia as a barrier to radiotherapy, chemotherapy and immunotherapy in cancer treatment. *International Journal of Nanomedicine*, Volume 13, pp.6049–6058.

Graham, S.V. and Faizo, A.A.A. (2017). Control of human papillomavirus gene expression by alternative splicing. *Virus Research*, 231, pp.83–95.

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

Griffin, H., Wu, Z. and Marnane, R. (2012). E4 antibodies facilitate detection and type-assignment of active HPV infection in cervical disease. *PloS One*, 7(12), p.e49974.

Grm, H.S. and Banks, L. (2004). Degradation of hDlg and MAGIs by human papillomavirus E6 is E6-AP-independent. *The Journal of General Virology*, 85(Pt 10), pp.2815–2819.

Gross, G. and Pfister, H. (2004). Role of human papillomavirus in penile cancer, penile intraepithelial squamous cell neoplasias and in genital warts. *Medical Microbiology and Immunology*, 193(1), pp.35–44.

Guo, M., Song, L.-P. and Jiang, Y. (2006). Hypoxia-mimetic agents desferrioxamine and cobalt chloride induce leukemic cell apoptosis through different hypoxia-inducible factor-1alpha independent mechanisms. *Apoptosis: An International Journal on Programmed Cell Death*, 11(1), pp.67–77.

Hanahan, D. and Weinberg, Robert A. (2011). Hallmarks of Cancer: The Next Generation. *Cell*, 144(5), pp.646–674.

Hanes, J., von der Kammer, H. and Klaudiny, J. (1994). Characterization by cDNA cloning of two new human protein kinases. Evidence by sequence comparison of a new family of mammalian protein kinases. *Journal of Molecular Biology*, 244(5), pp.665–672.

Harper, D.M. and DeMars, L.R. (2017). HPV vaccines – A review of the first decade. *Gynecologic Oncology*, 146(1), pp.196–204.

Hauschild, A., Grob, J.-J. and Demidov, L.V. (2012). Dabrafenib in BRAF-mutated metastatic melanoma: a multicentre, open-label, phase 3 randomised controlled trial. *The Lancet*, 380(9839), pp.358–365.

Hawley-Nelson, P., Vousden, K.H., Hubbert, N.L. and Lowy, D.R. (1989). HPV16 E6 and E7 proteins cooperate to immortalize human foreskin keratinocytes. *The EMBO Journal*, 8(12), pp.3905–3910.

He, W., Staples, D. and Smith, C. (2003). Direct activation of cyclin-dependent kinase 2 by human papillomavirus E7. *Journal of Virology*, 77(19), pp.10566–10574.

Heck, D.V., Yee, C.L. and Howley, P.M. (1992). Efficiency of binding the retinoblastoma protein correlates with the transforming capacity of the E7 oncoproteins of the human papillomaviruses. *Proceedings of the National Academy of Sciences of the United States of America*, 89(10), pp.4442–4446.

Henkle, T.R., Lam, B. and Kung, Y.J. (2021). Development of a Novel Mouse Model of Spontaneous High-Risk HPVE6/E7–Expressing Carcinoma in the Cervicovaginal Tract. *Cancer Research*, 81(17), pp.4560–4569.

Herrero, R., Quint, W. and Hildesheim, A. (2013). Reduced Prevalence of Oral Human Papillomavirus (HPV) 4 Years after Bivalent HPV Vaccination in a Randomized Clinical Trial in Costa Rica. *PLoS ONE*, 8(7), p.e68329.

Hirota, K. and Semenza, G.L. (2005). Regulation of hypoxia-inducible factor 1 by prolyl and asparaginyl hydroxylases. *Biochemical and Biophysical Research Communications*, 338(1), pp.610–616.

Höckel, M., Schlenger, K. and Knoop, C. (1991). Oxygenation of carcinomas of the uterine cervix: evaluation by computerized O2 tension measurements. *Cancer Research*, 51(22), pp.6098–6102.

Höckel, M. and Vaupel, P. (2001). Tumor Hypoxia: Definitions and Current Clinical, Biologic, and Molecular Aspects. *JNCI: Journal of the National Cancer Institute*, 93(4), pp.266–276.

Hoppe-Seyler, K., Bossler, F. and Lohrey, C. (2017). Induction of dormancy in hypoxic human papillomavirus-positive cancer cells. *Proceedings of the National Academy of Sciences*, [online] 114(6), pp.E990–E998.

Horner, S.M., DeFilippis, R.A. and Manuelidis, L. (2004). Repression of the Human Papillomavirus E6 Gene Initiates p53-Dependent, Telomerase-Independent Senescence and Apoptosis in HeLa Cervical Carcinoma Cells. *Journal of Virology*, 78(8), pp.4063–4073.

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

Hu, Z., Zhu, D. and Wang, W. (2015). Genome-wide profiling of HPV integration in cervical cancer identifies clustered genomic hot spots and a potential microhomology-mediated integration mechanism. *Nature Genetics*, 47(2), pp.158–163.

Huang, H., Yu, J. and Yu, D. (2014). Effects of chemically induced hypoxia on in vitro expression of hypoxia inducible factor-lα, vascular endothelial growth factor, aggrecanase-1, and tissue inhibitor of metalloproteinase-3 in rat mandibular condylar chondrocytes. *Journal of Oral & Facial Pain and Headache*, 28(3), pp.269–276.

Huibregtse, J.M., Scheffner, M. and Howley, P.M. (1993). Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins. *Molecular and Cellular Biology*, 13(8), pp.4918–4927.

Hwang, E.S., Riese, D.J. and Settleman, J. (1993). Inhibition of cervical carcinoma cell line proliferation by the introduction of a bovine papillomavirus regulatory gene. *Journal of Virology*, 67(7), pp.3720–3729.

IARC Working Group on the Evaluation of Carcinogenic Risk to Humans. Biological Agents. Lyon (FR): International Agency for Research on Cancer; 2012, (IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, No. 100B.) Human Papillomaviruses.

Imai, Y., Matsushima, Y. and Sugimura, T. (1991). Purification and characterization of human papillomavirus type 16 E7 protein with preferential binding capacity to the underphosphorylated form of retinoblastoma gene product. *Journal of Virology*, 65(9), pp.4966–4972.

Ivan, M., Kondo, K. and Yang, H. (2001). HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science (New York, N.Y.)*, 292(5516), pp.464–468.

Iwai, K., Yaguchi, M. and Nishimura, K. (2018). Anti-tumor efficacy of a novel CLK inhibitor via targeting RNA splicing and MYC-dependent vulnerability. *EMBO molecular medicine*, 10(6).

Jaakkola, P., Mole, D.R. and Tian, Y.-M. (2001). Targeting of HIF-α to the von Hippel-Lindau Ubiquitylation Complex by O2-Regulated Prolyl Hydroxylation. *Science*, 292(5516), pp.468–472.

Jabbar, S.F., Abrams, L. and Glick, A. (2009). Persistence of High-Grade Cervical Dysplasia and Cervical Cancer Requires the Continuous Expression of the Human Papillomavirus Type 16 E7 Oncogene. *Cancer Research*, 69(10), pp.4407–4414.

Jang, M.K., Shen, K. and McBride, A.A. (2014). Papillomavirus Genomes Associate with BRD4 to Replicate at Fragile Sites in the Host Genome. *PLoS Pathogens*, 10(5), p.e1004117.

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. *Journal of Virology*, 69(5), pp.2989–2997.

Ji, X., Zhou, Y. and Pandit, S. (2013). SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell*, 153(4), pp.855–868.

Jia, R., Li, C. and McCoy, J.P. (2010). SRp20 is a proto-oncogene critical for cell proliferation and tumor induction and maintenance. *International Journal of Biological Sciences*, 6(7), pp.806–826.

Jiang, M. and Milner, J. (2002). Selective silencing of viral gene expression in HPV-positive human cervical carcinoma cells treated with siRNA, a primer of RNA interference. *Oncogene*, 21(39), pp.6041–6048.

Johung, K., Goodwin, E.C. and DiMaio, D. (2007). Human papillomavirus E7 repression in cervical carcinoma cells initiates a transcriptional cascade driven by the retinoblastoma family, resulting in senescence. *Journal of Virology*, 81(5), pp.2102–2116.

Joyce, J.G., Tung, J.-S. and Przysiecki, C.T. (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. *Journal of Biological Chemistry*, 274(9), pp.5810–5822.

Kajitani, N., Satsuka, A. and Kawate, A. (2012). Productive Lifecycle of Human Papillomaviruses that Depends Upon Squamous Epithelial Differentiation. *Frontiers in Microbiology*, 3(152).

Kastenhuber, E.R. and Lowe, S.W. (2017). Putting p53 in Context. *Cell*, 170(6), pp.1062–1078.

Kirnbauer, R., Booy, F. and Cheng, N. (1992). Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. *Proceedings of the National Academy of Sciences*, 89(24), pp.12180–12184.

Klingenberg, B., Hafkamp, H.C. and Haesevoets, A. (2010). p16INK4A overexpression is frequently detected in tumour-free tonsil tissue without association with HPV. *Histopathology*, 56(7), pp.957–967.

Kraggerud, S.M., Sandvik, J.A. and Pettersen, E.O. (1995). Regulation of protein synthesis in human cells exposed to extreme hypoxia. *Anticancer Research*, 15(3), pp.683–686.

Kreimer, A.R., González, P. and Katki, H.A. (2011). Efficacy of a bivalent HPV 16/18 vaccine against anal HPV 16/18 infection among young women: a nested analysis within the Costa Rica Vaccine Trial. *The Lancet Oncology*, 12(9), pp.862–870.

Kübler, K., Heinenberg, S. and Rudlowski, C. (2014). c-myc copy number gain is a powerful prognosticator of disease outcome in cervical dysplasia. *Oncotarget*, 6(2), pp.825–835.

Kuroyanagi, N., Onogi, H. and Wakabayashi, T. (1998). Novel SR-protein-specific kinase, SRPK2, disassembles nuclear speckles. *Biochemical and Biophysical Research Communications*, 242(2), pp.357–364.

Lambert, P.F., Spalholz, B.A. and Howley, P.M. (1987). A transcriptional repressor encoded by BPV-1 shares a common carboxy-terminal domain with the E2 transactivator. *Cell*, 50(1), pp.69–78.

Laniosz, V., Dabydeen, S.A. and Havens, M.A. (2009). Human papillomavirus type 16 infection of human keratinocytes requires clathrin and caveolin-1 and is brefeldin a sensitive. *Journal of Virology*, 83(16), pp.8221–8232.

Lassen, P., Eriksen, J.G. and Hamilton-Dutoit, S. (2009). Effect of HPV-Associated p16INK4A Expression on Response to Radiotherapy and Survival in Squamous Cell Carcinoma of the Head and Neck. *Journal of Clinical Oncology*, 27(12), pp.1992–1998.

Lee, D., Kim, H.-Z. and Jeong, K.W. (2002). Human papillomavirus E2 down-regulates the human telomerase reverse transcriptase promoter. *The Journal of biological chemistry*, 277(31), pp.27748–56.

Leszczynska, K.B., Foskolou, I.P. and Abraham, A.G. (2015). Hypoxia-induced p53 modulates both apoptosis and radiosensitivity via AKT. *The Journal of Clinical Investigation*, 125(6), pp.2385–2398.

Lew, J., Qi, Z. and Huang, Q.Q. (1995). Structure, function, and regulation of neuronal Cdc2-like protein kinase. *Neurobiology of Aging*, 16(3), pp.263–268; discussion 268-270.

Lewis, BC., Shim, H., Li, Q., et al. Identification of putative c-MYC-responsive genes: characterization of rcl, a novel growth-related gene. *Mol Cell Biol.* 1997;17:4967–4978.

Li, X., Wang, J. and Manley, J. (2005). Loss of splicing factor ASF/SF2 induces G2 cell cycle arrest and apoptosis, but inhibits internucleosomal DNA fragmentation. *Genes & Development*, 19(22), pp.2705–2714.

Lin, C., Lovén, J., Rahl, P., et al. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* 151, 56–67.

Lin, S., Coutinho-Mansfield, G. and Wang, D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. *Nature structural & molecular biology*, 15(8), pp.819–826.

Lison, D., De Boeck, M. and Verougstraete, V. (2001). Update on the genotoxicity and carcinogenicity of cobalt compounds. *Occupational and Environmental Medicine*, 58(10), pp.619–625.

Liu, D., ZHANG, X.-X. and XI, B.-X. (2014). Sine oculis homeobox homolog 1 promotes DNA replication and cell proliferation in cervical cancer. *International Journal of Oncology*, 45(3), pp.1232–1240.

Liu, Y., Bishop, A. and Witucki, L. (1999). Structural basis for selective inhibition of Src family kinases by PP1. *Chemistry & Biology*, 6(9), pp.671–678.

Liu, Y., Gao, F. and Song, W. (2017). Periostin contributes to arsenic trioxide resistance in hepatocellular carcinoma cells under hypoxia. *Biomedicine & Pharmacotherapy*, 88, pp.342–348.

Loening, A.M., Fenn, T.D. and Wu, A.M. (2006). Consensus guided mutagenesis of Renilla luciferase yields enhanced stability and light output. *Protein Engineering Design and Selection*, 19(9), pp.391–400.

Long, J.C. and Caceres, J.F. (2009). The SR protein family of splicing factors: master regulators of gene expression. *The Biochemical journal*, 417(1), pp.15–27.

Loomis, R.J., Naoe, Y. and Parker, J.B. (2009). Chromatin Binding of SRp20 and ASF/SF2 and Dissociation from Mitotic Chromosomes Is Modulated by Histone H3 Serine 10 Phosphorylation. *Molecular cell*, 33(4), pp.450–461.

Maas, N.L., Singh, N. and Diehl, J.A. (2014). Generation and characterization of an analog-sensitive PERK allele. *Cancer Biology & Therapy*, 15(8), pp.1106–1111.

Maemondo, M., Inoue, A. and Kobayashi, K. (2010). Gefitinib or Chemotherapy for Non–Small-Cell Lung Cancer with Mutated EGFR. *New England Journal of Medicine*, 362(25), pp.2380–2388.

Maglennon, G.A., McIntosh, P. and Doorbar, J. (2011). Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression. *Virology*, 414(2), pp.153–63.

Mahey, S., Kumar, R. and Arora, R. (2016). Effect of cobalt(II) chloride hexahydrate on some human cancer cell lines. *SpringerPlus*, 5(1).

Mak, R., Van Renterghem, L. and Cuvelier, C. (2004). Cervical smears and human papillomavirus typing in sex workers. *Sexually Transmitted Infections*, 80(2), pp.118–120.

Makanae. K., Kintaka. R., and Makino. T., (2013) Identification of dosage-sensitive genes in Saccharomyces cerevisiae using the genetic tug-of-war method Genome Research 23:300–311.

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

de Martel, C., Plummer, M. and Vignat, J. (2017). Worldwide burden of cancer attributable to HPV by site, country and HPV type. *International Journal of Cancer*, 141(4), pp.664–670.

Martín Moyano, P., Němec, V. and Paruch, K. (2020). Cdc-Like Kinases (CLKs): Biology, Chemical Probes, and Therapeutic Potential. *International Journal of Molecular Sciences*, 21(20), p.7549.

Mazur, P., Herner, A., Mello, S., et al. (2015) Combined inhibition of BET family proteins and histone deacetylases as a potential epigenetics-based therapy for pancreatic ductal adenocarcinoma. *Nat. Med.* 21, 1163–1171.

McArthur, G.A., Chapman, P.B. and Robert, C. (2014). Safety and efficacy of vemurafenib in BRAFV600E and BRAFV600K mutation-positive melanoma (BRIM-3): extended follow-up of a phase 3, randomised, open-label study. *The Lancet Oncology*, 15(3), pp.323–332.

McBride, A.A., Romanczuk, H. and Howley, P.M. (1991). The papillomavirus E2 regulatory proteins. *Journal of Biological Chemistry*, 266(28), pp.18411–18414.

McBride, A.A. (2008). Chapter 4 Replication and Partitioning of Papillomavirus Genomes. *Advances in Virus Research*, pp.155–205.

McBride, A.A. (2013). The Papillomavirus E2 proteins. *Virology*, 445(1–2), pp.57–79.

McFarlane, M., MacDonald, A.I., Stevenson, A. and Graham, S.V. (2015). Human Papillomavirus 16 Oncoprotein Expression Is Controlled by the Cellular Splicing Factor SRSF2 (SC35). *Journal of Virology*, 89(10), pp.5276–5287.

McLaughlin-Drubin, M.E., Crum, C.P. and Münger, K. (2011). Human papillomavirus E7 oncoprotein induces KDM6A and KDM6B histone demethylase expression and causes epigenetic reprogramming. *Proceedings of the National Academy of Sciences*, 108(5), pp.2130–2135.

Menssen A, Hermeking H. (2002) Characterization of the c-MYC-regulated transcriptome by SAGE: identification and analysis of c-MYC target genes. *Proc Natl Acad Sci U S A.* 99:6274–6279.

Mertz, J., Conery, A., Bryant, B., et al. (2011) Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc. Natl Acad. Sci. USA*108, 16669–16674.

Misteli, T., Cáceres, J.F. and Clement, J.Q. (1998). Serine phosphorylation of SR proteins is required for their recruitment to sites of transcription in vivo. *The Journal of Cell Biology*, 143(2), pp.297–307.

Mo, S., Ji, X. and Fu, X.-D. (2013). Unique role of SRSF2 in transcription activation and diverse functions of the SR and hnRNP proteins in gene expression regulation. *Transcription*, 4(5), pp.251–259.

Mole, S., Milligan, S.G. and Graham, S.V. (2008). Human Papillomavirus Type 16 E2 Protein Transcriptionally Activates the Promoter of a Key Cellular Splicing Factor, SF2/ASF. *Journal of Virology*, 83(1), pp.357–367.

Moody, C.A. and Laimins, L.A. (2010). Human papillomavirus oncoproteins: pathways to transformation. *Nature Reviews Cancer*, 10(8), pp.550–560.

Mooren, J.J., Gültekin, S.E. and Straetmans, J.M.J.A.A. (2014). P16(INK4A) immunostaining is a strong indicator for high-risk-HPV-associated oropharyngeal carcinomas and dysplasias, but is unreliable to predict low-risk-HPV-infection in head and neck papillomas and laryngeal dysplasias. *International Journal of Cancer*, 134(9), pp.2108–2117.

MRC Protein Phosphorylation and Ubiquitylation Unit. (n.d.). *List of clinically approved kinase inhibitors | MRC PPU*. [online] Available at: https://www.ppu.mrc.ac.uk/list-clinically-approved-kinase-inhibitors [Accessed 20 Oct. 2021].

Münger, K., Basile, J.R. and Duensing, S. (2001). Biological activities and molecular targets of the human papillomavirus E7 oncoprotein. *Oncogene*, 20(54), pp.7888–7898.

Munger, K., Gwin, T.K. and McLaughlin-Drubin, M. (2013). p16 in HPV-associated cancers. *Oncotarget*, 4(11), pp.1864–1865.

Münger, K., Phelps, W.C. and Bubb, V. (1989). The E6 and E7 genes of the human papillomavirus type 16 together are necessary and sufficient for transformation of primary human keratinocytes. *Journal of Virology*, 63(10), pp.4417–4421.

Muñoz-Sánchez, J. and Chánez-Cárdenas, M.E. (2018). The use of cobalt chloride as a chemical hypoxia model. *Journal of Applied Toxicology*, 39(4), pp.556–570.

Muraki, M., Ohkawara, B. and Hosoya, T. (2004). Manipulation of alternative splicing by a newly developed inhibitor of Clks. *The Journal of Biological Chemistry*, 279(23), pp.24246–24254.

Murphy, B.J., Laderoute, K.R. and Chin, R.J. (1994). Metallothionein IIA is up-regulated by hypoxia in human A431 squamous carcinoma cells. *Cancer Research*, 54(22), pp.5808–5810.

Nayler, O., Stamm, S. and Ullrich, A. (1997). Characterization and comparison of four serine- and arginine-rich (SR) protein kinases. *Biochemical Journal*, 326(3), pp.693–700.

Nees, M., van Wijngaarden, E. and Bakos, E. (1998). Identification of novel molecular markers which correlate with HPV-induced tumor progression. *Oncogene*, 16(19), pp.2447–2458.

Nguyen, C.L. and Münger, K. (2008). Direct association of the HPV16 E7 oncoprotein with cyclin A/CDK2 and cyclin E/CDK2 complexes. *Virology*, 380(1), pp.21–25.
Nie, Z., Hu, G., Wei, G., et al. (2012) c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* 151, 68–79.

Nobre, R.J., Herráez-Hernández, E. and Fei, J.-W. (2009). E7 Oncoprotein of Novel Human Papillomavirus Type 108 Lacking the E6 Gene Induces Dysplasia in Organotypic Keratinocyte Cultures. *Journal of Virology*, 83(7), pp.2907–2916.
Oh, K.-J., Kalinina, A. and Wang, J. (2004). The papillomavirus E7 oncoprotein is ubiquitinated by UbcH7 and Cullin 1- and Skp2-containing E3 ligase. *Journal of Virology*, 78(10), pp.5338–5346.

Papasaikas, P. and Valcárcel, J. (2016). The Spliceosome: The Ultimate RNA Chaperone and Sculptor. *Trends in Biochemical Sciences*, 41(1), pp.33–45.

Parfenov, M., Pedamallu, C.S. and Gehlenborg, N. (2014). Characterization of HPV and host genome interactions in primary head and neck cancers. *Proceedings of the National Academy of Sciences*, [online] 111(43), pp.15544–15549.

Park, S.Y., Piao, Y. and Thomas, C. (2016). Cdc2-like kinase 2 is a key regulator of the cell cycle via FOXO3a/p27 in glioblastoma. *Oncotarget*, 7(18), pp.26793–26805.

PaVE (n.d.). *PaVE: Papilloma virus genome database*. [online] pave.niaid.nih.gov. Available at: https://pave.niaid.nih.gov/#search/search_database [Accessed 20 Oct. 2021].

Pedroza-Saavedra, A., Lam, E.W.-F. and Esquivel-Guadarrama, F. (2010). The human papillomavirus type 16 E5 oncoprotein synergizes with EGF-receptor signaling to enhance cell cycle progression and the down-regulation of p27Kip1. *Virology*, 400(1), pp.44–52.

Peter, M., Rosty, C., and Couturier. (2006). *MYC* activation associated with the integration of HPV DNA at the *MYC* locus in genital tumours. *Oncogene*. 25, 5985-5993 (2006).

Perna D., Faga, G., and Verrecchia A. (2012). Genome-wide mapping of Myc binding and gene regulation in serum-stimulated fibroblasts. *Oncogene*, 31(13):1695–709.

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

Pim, D., Broniarczyk, J. and Bergant, M. (2015). A Novel PDZ Domain Interaction Mediates the Binding between Human Papillomavirus 16 L2 and Sorting Nexin 27 and Modulates Virion Trafficking. *Journal of Virology*, 89(20), pp.10145–10155.

Pinidis, P., Tsikouras, P. and Iatrakis, G. (2016). Human Papilloma Virus' Life Cycle and Carcinogenesis. *Mædica*, 11(1), pp.48–54.

Rahl, P.B., Lin, C.Y. and Seila, A.C. (2010). c-Myc Regulates Transcriptional Pause Release. *Cell*, 141(3), pp.432–445.

Rayess, H., Wang, M.B. and Srivatsan, E.S. (2011). Cellular senescence and tumor suppressor gene p16. *International Journal of Cancer*, 130(8), pp.1715–1725.

Reuschenbach, M., Huebbers, C.U. and Prigge, E.-S. (2015). Methylation status of HPV16 E2-binding sites classifies subtypes of HPV-associated oropharyngeal cancers. *Cancer*, 121(12), pp.1966–1976.

Richards, R.M., Lowy, D.R. and Schiller, J.T. (2006). Cleavage of the papillomavirus minor capsid protein, L2, at a furin consensus site is necessary for infection. *Proceedings of the National Academy of Sciences*, 103(5), pp.1522–1527.

Sahab, Z., Sudarshan, S.R. and Liu, X. (2012). Quantitative Measurement of Human Papillomavirus Type 16 E5 Oncoprotein Levels in Epithelial Cell Lines by Mass Spectrometry. *Journal of Virology*, 86(17), pp.9465–9473.

Sano, T., Oyama, T. and Kashiwabara, K. (1998). Expression Status of p16 Protein Is Associated with Human Papillomavirus Oncogenic Potential in Cervical and Genital Lesions. *The American Journal of Pathology*, 153(6), pp.1741–1748.

Sarastry, R. 'use of CRISPR-cas9 genome editing to delete CLK3 in human keratinocytes' MSc Thesis, University College London.

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

Schelhaas, M., Shah, B., and Holzer, M. (2012). Entry of human papillomavirus type 16 by actin-dependent, clathrin- and lipid raft- independent endocytosis. *PLoS Pathog*. 8 (4), e1002657.

Schnabel, J., Hombach, P. and Waksman, T. (2018). A chemical genetic approach to engineer phototropin kinases for substrate labeling. *Journal of Biological Chemistry*, 293(15), pp.5613–5623.

Selinka, H.-C., Giroglou, T. and Sapp, M. (2002). Analysis of the infectious entry pathway of human papillomavirus type 33 pseudovirions. *Virology*, 299(2), pp.279–287.

Sequist, L.V., Yang, J.C.-H. and Yamamoto, N. (2013). Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 31(27), pp.3327–3334.

Shah, K., Liu, Y. and Deirmengian, C. (1997). Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. *Proceedings of the National Academy of Sciences of the United States of America*, 94(8), pp.3565–3570.

Shah, M., Nunes, M.R. and Stearns, V. (2018). CDK4/6 Inhibitors: Game Changers in the Management of Hormone Receptor–Positive Advanced Breast Cancer? *Oncology (Williston Park, N.Y.)*, 32(5), pp.216–222.

Shamloo, B. and Usluer, S. (2019). p21 in Cancer Research. *Cancers*, 11(8), p.1178.

Shukla, S., Mahata, S. and Shishodia, G. (2014). Physical state & copy number of high risk human papillomavirus type 16 DNA in progression of cervical cancer. *The Indian Journal of Medical Research*, 139(4), pp.531–543.

Simonart, T., Degraef, C. and Andrei, G. (2000). Iron chelators inhibit the growth and induce the apoptosis of Kaposi's sarcoma cells and of their putative endothelial precursors. *The Journal of Investigative Dermatology*, 115(5), pp.893–900.

Smalley Rumfield, C., Roller, N., Pellom, S.T., Schlom, J. and Jochems, C. (2020). Therapeutic Vaccines for HPV-Associated Malignancies. *ImmunoTargets and Therapy*, 9, pp.167–200.

Sondka, Z., Bamford, S. and Cole, C.G. (2018). The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. *Nature Reviews. Cancer*, 18(11), pp.696–705.

Songyang, Z., Fanning, A. and Cantley, L. (1997). Recognition of Unique Carboxyl-Terminal Motifs by Distinct PDZ Domains. *Science*, 275(5296), pp.73–77.

Srebrow, A. and Kornblihtt, A.R. (2006). The connection between splicing and cancer. *Journal of Cell Science*, 119(13), pp.2635–2641.

Stanley, M. (2010). HPV - immune response to infection and vaccination. *Infectious Agents and Cancer*, 5(1).

Steele, C., Sacks, P.G. and Adler-Storthz, K. (1992). Effect on cancer cells of plasmids that express antisense RNA of human papillomavirus type 18. *Cancer Research*, 52(17), pp.4706–4711.

Steger, G. and Corbach, S. (1997). Dose-dependent regulation of the early promoter of human papillomavirus type 18 by the viral E2 protein. *Journal of virology*, 71(1), pp.50–8.

Sterling, J.C., Handfield-Jones, S. and Hudson, P.M. (2001). Guidelines for the management of cutaneous warts. *British Journal of Dermatology*, 144(1), pp.4–11.

Stern DA, R. (1982). Fatti statistici relativi alle malattie cancerose. *Giornale per servire ai progressi della Patologia e della Terapia*, 2:507-17.

Stevens, H., Rector, A. and Bertelsen, M.F. (2008). Novel papillomavirus isolated from the oral mucosa of a polar bear does not cluster with other papillomaviruses of carnivores. *Veterinary Microbiology*, 129(1), pp.108–116.

Stickeler, E., Kittrell, F. and Medina, D. (1999). Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. *Oncogene*, 18(24), pp.3574–3582.

Straight, S.W., Hinkle, P.M. and Jewers, R.J. (1993). The E5 oncoprotein of human papillomavirus type 16 transforms fibroblasts and effects the downregulation of the epidermal growth factor receptor in keratinocytes. *Journal of virology*, 67(8), pp.4521–32.

Strambi, A., Mori, M. and Rossi, M. (2013). Structure Prediction and Validation of the ERK8 Kinase Domain. *PLoS ONE*, 8(1), p.e52011.

Syrjänen, K., Syrjänen, S. and Lamberg, M. (1983). Morphological and immunohistochemical evidence suggesting human papillomavirus (HPV) involvement in oral squamous cell carcinogenesis. *International Journal of Oral Surgery*, 12(6), pp.418–424.

Syrjänen, S., Rautava, J. and Syrjänen, K. (2016). HPV in Head and Neck Cancer—30 Years of History. *HPV Infection in Head and Neck Cancer*, pp.3–25.

Tam, B.Y., Chiu, K. and Chung, H. (2020). The CLK inhibitor SM08502 induces anti-tumor activity and reduces Wnt pathway gene expression in gastrointestinal cancer models. *Cancer Letters*, 473, pp.186–197.

Tang, S., Tao, M. and McCoy, J.P. (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. *Journal of Virology*, 80(9), pp.4249–4263.

Tang, YC., Amon, A., (2013) Gene copy-number alterations: A cost-benefit analysis Cell 152:394–405.

Tavana, O., Li, D., Dai., C., et al. (2016). HAUSP deubiquitinates and stabilizes N-Myc in neuroblastoma. *Nat. Med.* 22, 1180–1186.

Teicher, B.A., Holden, S.A. and al-Achi, A. (1990). Classification of antineoplastic treatments by their differential toxicity toward putative oxygenated and hypoxic tumor subpopulations in vivo in the FSaIIC murine fibrosarcoma. *Cancer Research*, 50(11), pp.3339–3344.

Thomas, M., Laura, R. and Hepner, K. (2002). Oncogenic human papillomavirus E6 proteins target the MAGI-2 and MAGI-3 proteins for degradation. *Oncogene*, 21(33), pp.5088–5096.

Thomas, M., Massimi, P. and Navarro, C. (2005). The hScrib/Dlg apico-basal control complex is differentially targeted by HPV-16 and HPV-18 E6 proteins. *Oncogene*, 24(41), pp.6222–6230.

Tsao, Y.P., Li, L.Y. and Tsai, T.C. (1996). Human papillomavirus type 11 and 16 E5 represses p21(WafI/SdiI/CipI) gene expression in fibroblasts and keratinocytes. *Journal of virology*, 70(11), pp.7535–9.

Van Doorslaer, K., Ould M'hamed Ould Sidi, A. and Zanier, K. (2009). Identification of Unusual E6 and E7 Proteins within Avian Papillomaviruses: Cellular Localization, Biophysical Characterization, and Phylogenetic Analysis. *Journal of Virology*, 83(17), pp.8759–8770.

Vaupel, P., Höckel, M. and Mayer, A. (2007). Detection and characterization of tumor hypoxia using pO2 histography. *Antioxidants & Redox Signaling*, 9(8), pp.1221–1235.

Vaupel, P., Mayer, A. and Höckel, M. (2004). *Tumor Hypoxia and Malignant Progression*. ScienceDirect.

Vavouri, T., Semple, JI., and Garcia-Verdugo, R. (2009) Intrinsic protein disorder and interaction promiscuity are widely associated with dosage sensitivity Cell 138:198–208.

Velazquez-Dones, A., Hagopian, J.C. and Ma, C.-T. (2005). Mass Spectrometric and Kinetic Analysis of ASF/SF2 Phosphorylation by SRPK1 and Clk/Sty. *Journal of Biological Chemistry*, 280(50), pp.41761–41768.

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

Vinokurova, S., Wentzensen, N. and Kraus, I. (2008). Type-Dependent Integration Frequency of Human Papillomavirus Genomes in Cervical Lesions. *Cancer Research*, 68(1), pp.307–313.

von Knebel Doeberitz, M., Rittmüller, C. and zur Hausen, H. (1992). Inhibition of tumorigenicity of cervical cancer cells in nude mice by HPV E6-E7 anti-sense RNA. *International Journal of Cancer*, 51(5), pp.831–834.

Watson, R.A. (2005). Human Papillomavirus: Confronting the Epidemic—A Urologist's Perspective. *Reviews in Urology*, 7(3), pp.135–144.

Wechsler, E.I., Wang, Q. and Roberts, I. (2012). Reconstruction of Human Papillomavirus Type 16-Mediated Early-Stage Neoplasia Implicates E6/E7 Deregulation and the Loss of Contact Inhibition in Neoplastic Progression. *Journal of Virology*, 86(11), pp.6358–6364.

Weinstein, I.B. and Joe, A. (2008). Oncogene Addiction. *Cancer Research*, 68(9), pp.3077–3080.

Wetherill, L., Holmes, K., and Verow., M. (2012). High-risk human papillomavirus E5 oncoprotein displays channel-forming activity sensititve to small-molecule inhibitors. *J Virol*. 86 (9), pp 5341-5351.

White, E.A., Kramer, R.E. and Tan, M.J.A. (2012). Comprehensive Analysis of Host Cellular Interactions with Human Papillomavirus E6 Proteins Identifies New E6 Binding Partners and Reflects Viral Diversity. *Journal of Virology*, 86(24), pp.13174–13186.

White, E.A., Sowa, M.E. and Tan, M.J.A. (2012). Systematic identification of interactions between host cell proteins and E7 oncoproteins from diverse human papillomaviruses. *Proceedings of the National Academy of Sciences of the United States of America*, 109(5), pp.E260–E267.

Williams, L. R. 'Cdc like kinase 3 regulation of human papillomavirus oncogene expression'. PhD thesis, University College London.

Woodman, C.B.J., Collins, S.I. and Young, L.S. (2007). The natural history of cervical HPV infection: unresolved issues. *Nature Reviews Cancer*, 7(1), pp.11–22.

Wong, R., Balachandran, A. and Mao, A.Y. (2011). Differential effect of CLK SR Kinases on HIV-1 gene expression: potential novel targets for therapy. *Retrovirology*, 8(1), p.47.

Woodham, A.W., Da Silva, D.M. and Skeate, J.G. (2012). The S100A10 subunit of the annexin A2 heterotetramer facilitates L2-mediated human papillomavirus infection. *PloS One*, 7(8), p.e43519.

Wouters, A., Pauwels, B. and Lardon, F. (2007). Review: implications of in vitro research on the effect of radiotherapy and chemotherapy under hypoxic conditions. *The Oncologist*, 12(6), pp.690–712.

Wu, D. and Yotnda, P. (2011). Induction and Testing of Hypoxia in Cell Culture. *Journal of Visualized Experiments*, [online] 54(e2899), pp.1–4.

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 & Development*, 11(3), pp.334–344.

Xie, X., Piao, L. and Bullock, B.N. (2014). Targeting HPV16 E6-p300 interaction reactivates p53 and inhibits the tumorigenicity of HPV-positive head and neck squamous cell carcinoma. *Oncogene*, 33(8), pp.1037–1046.

Xue, L. and Tao, W.A. (2013). Current technologies to identify protein kinase substrates in high throughput. *Frontiers in biology*, 8(2), pp.216–227.

Yajid, A.I., Zakariah, M.A. and Zin, A.A.M. (2017). Potential Role of E4 Protein in Human Papillomavirus Screening: a Review. *Asian Pacific Journal of Cancer Prevention : APJCP*, 18(2), pp.315–319.

Yang, L., Mohr, I. and Fouts, E. (1993). The E1 protein of bovine papilloma virus 1 is an ATP-dependent DNA helicase. *Proceedings of the National Academy of Sciences*, 90(11), pp.5086–5090.

Yeakley, J.M., Tronchère, H. and Olesen, J. (1999). Phosphorylation Regulates In Vivo Interaction and Molecular Targeting of Serine/Arginine-rich Pre-mRNA Splicing Factors. *The Journal of Cell Biology*, 145(3), pp.447–455.

Yildiz, I.Z., Usubütün, A. and Firat, P. (2007). Efficiency of immunohistochemical p16 expression and HPV typing in cervical squamous intraepithelial lesion grading and review of the p16 literature. *Pathology, Research and Practice*, 203(6), pp.445–449.

Yim, E.-K. and Park, J.-S. (2005). The Role of HPV E6 and E7 Oncoproteins in HPV-associated Cervical Carcinogenesis. *Cancer Research and Treatment*, 37(6), p.319.

Yoshida, T., Kim, J.H. and Carver, K. (2015). CLK2 Is an Oncogenic Kinase and Splicing Regulator in Breast Cancer. *Cancer Research*, 75(7), pp.1516–1526.

Yu, K., Toral-Barza, L., Discafani, C., et al. (2001) mTOR, a novel target in breast cancer: the effect of CCI-779, an mTOR inhibitor, in preclinical models of breast cancer. *Endocr. Relat. Cancer* 8, 249–258.

Yun, C. and Fu, X. (2000). Conserved SR Protein Kinase Functions in Nuclear Import and Its Action Is Counteracted by Arginine Methylation in SaccharoMYCes Cerevisiae. *The Journal of cell biology*, 150(4), pp.707–18.

Zerfass-Thome, K., Zwerschke, W. and Mannhardt, B. (1996). Inactivation of the cdk inhibitor p27KIP1 by the human papillomavirus type 16 E7 oncoprotein. *Oncogene*, 13(11), pp.2323–2330.

Zhang, B., Li, P. and Wang, E. (2003). The E5 protein of human papillomavirus type 16 perturbs MHC class II antigen maturation in human foreskin keratinocytes treated with interferon-γ. *Virology*, 310(1), pp.100–108.

Zhang, H., Gao, P. and Fukuda, R. (2007). HIF-1 inhibits mitochondrial biogenesis and cellular respiration in VHL-deficient renal cell carcinoma by repression of C-MYC activity. *Cancer Cell*, 11(5), pp.407–420.

Zhang, H.S., Postigo, A.A. and Dean, D.C. (1999). Active Transcriptional Repression by the Rb–E2F Complex Mediates G1 Arrest Triggered by p16INK4a, TGFβ, and Contact Inhibition. *Cell*, 97(1), pp.53–61.

Zhang, W., Kazakov, T. and Popa, A. (2014). Vesicular Trafficking of Incoming Human Papillomavirus 16 to the Golgi Apparatus and Endoplasmic Reticulum Requires γ-Secretase Activity. *mBio*, 5(5).

Zhang, Z., Yan, J. and Chang, Y. (2011). Hypoxia Inducible Factor-1 as a Target for Neurodegenerative Diseases. *Current medicinal chemistry*, 18(28), pp.4335–4343.

Zhou, C., Wu, Y.-L. and Chen, G. (2011). Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. *The Lancet Oncology*, 12(8), pp.735–742.

Zhou, Q., Lin, M. and Feng, X. (2020). Targeting CLK3 inhibits the progression of cholangiocarcinoma by reprogramming nucleotide metabolism. *The Journal of Experimental Medicine*, 217(8).

Zhou, Z. and Fu, X.-D. (2013). Regulation of Splicing by SR proteins and SR Protein-Specific Kinases. *Chromosoma*, 122(3), pp.191–207.

Zhu, D., Xu, S. and Deyanat-Yazdi, G. (2018). Synthetic Lethal Strategy Identifies a Potent and Selective TTK and CLK1/2 Inhibitor for Treatment of Triple-Negative Breast Cancer with a Compromised G1–S Checkpoint. *Molecular Cancer Therapeutics*, 17(8), pp.1727–1738.

Zu, M., Li, C. and Fang, J.-S. (2015). Drug Discovery of Host CLK1 Inhibitors for Influenza Treatment. *Molecules*, 20(11), pp.19735–19747.

# 8 Supplimentary Figures

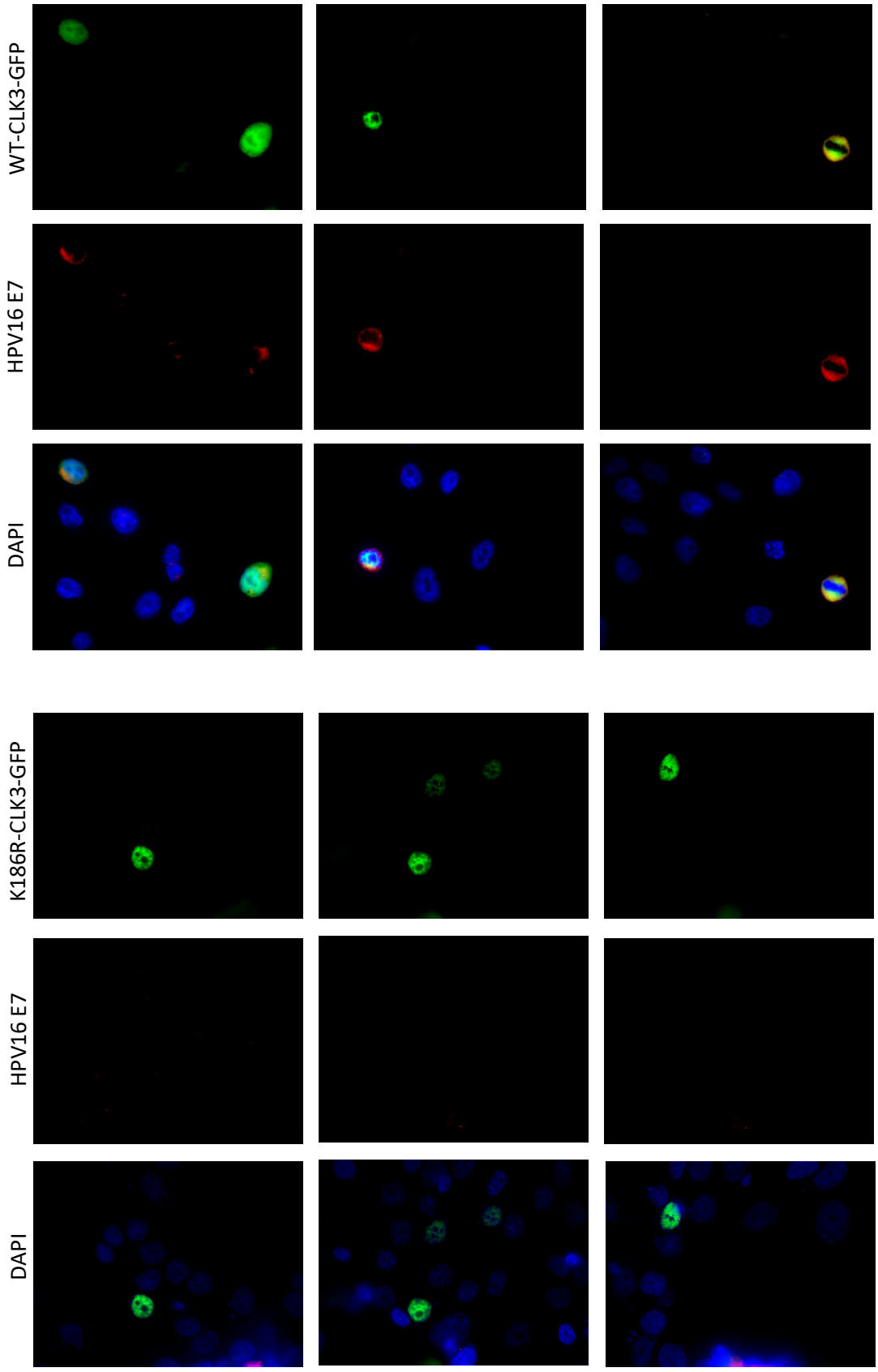
ATGCACCACTGTAAAAGATACCGCAGCCCAGAGCCTGATCCATATCTCTCTTATAGGTGGAAGCGA
AGACGATCTTACTCTCGAGAGCATGAAGGTCGGCTGCGCTACCCCTCAAGAAGGGAGCCACCTCC
TAGGAGAAGCCGCTCCCGGTCACATGATAGACTGCCCTACCAACGCAGATATCGCGAGAGAAGA
GACTCCGATACTTACAGGTGCGAGGAAAGATCACCCAGCTTCGGTGAAGATTATTACGGTCCAAG
CAGAAGCAGGCATAGGCGGCGGAGCCGGGAGAGGGGTCCTTATAGGACACGAAAGCATGCCCA
CCATTGCCACAAGCGCCGCACGCGCTCTTGCAGCTCTGCATCTTCCCGCTCACAGCAGAGCAGTAA
GCGCAGCAGCAGATCCGTGGAAGATGATAAGGAGGGGCATCTCGTGTGTAGAATTGGTGACTGG
CTCCAAGAGCGGTATGAAATTGTGGGCAATCTGGGCGAGGGAACCTTCGGTAAAGTGGTGGAGT
GTCTCGACCATGCCCGCGGCAAAAGCCAAGTGGCCTTGA**A**GATAATCAGGAATGTGGGTAAGTA
CCGCGAAGCCGCAAGACTGGAGATCAATGTGCTGAAGAAAATTAAGGAGAAAGACAAAGAAAAT
AAATTTCTCTGTGTCCTCATGTCCGATTGGTTCAATTTCCACGGGCACATGTGCATCGCCTTCGAGC
TGCTCGGAAAGAATACATTTGAATTTTTGAAAGAGAACAATTTTCAGCCCTATCCTTTGCCCCATGT
GCGGCACATGGCCTACCAACTGTGCCATGCCCTTCGGTTTCTGCATGAGAACCAGCTGACCCATAC
CGACCTGAAACCGGAGAACATCTTGTTCGTGAATTCTGAGTTCGAAACCCTCTATAATGAGCATAA
GAGCTGCGAGGAGAAATCTGTTAAGAATACCAGTATAAGAGTGGCTGACTTCGGCTCCGCTACCT
TTGACCACGAACACCACACCACAATAGTGGCCACTAGACACTATCGCCCTCCCGAAGTTATTCTGG
AGCTCGGGTGGGCCCAGCCCTGTGACGTGTGGAGTATTGGCTGCATCCTGTTCGAGTATTACAGA
GGATTTACTCTGTTTCAGACCCACGAGAATCGGGAACACCTTGTTATGATGGAGAAAATCTTGGGT
CCGATCCCTTCCCATATGATTCACCGGACAAGAAAGCAGAAGTACTTTTACAAGGGCGGACTGGT
TTGGGATGAGAACTCCAGTGATGGGCGCTACGTGAAGGAGAATTGTAAACCCCTGAAGTCTTATA
TGCTTCAGGATAGTCTGGAGCACGTCCAACTCTTTGATCTTATGCGGCGAATGCTGGAGTTCGACC
CCGCGCAGAGAATCACACTTGCTGAAGCACTGCTGCACCCCTTCTTCGCCGGACTCACACCAGAG
GAGAGATCCTTCCATACTTCTAGGAATCCTAGCAGGTGA

**Figure 8.1. 1 Codon optimized sequence of CLK3**
Adenine is mutated to Guanine for K186R kinase inactive CLK3.

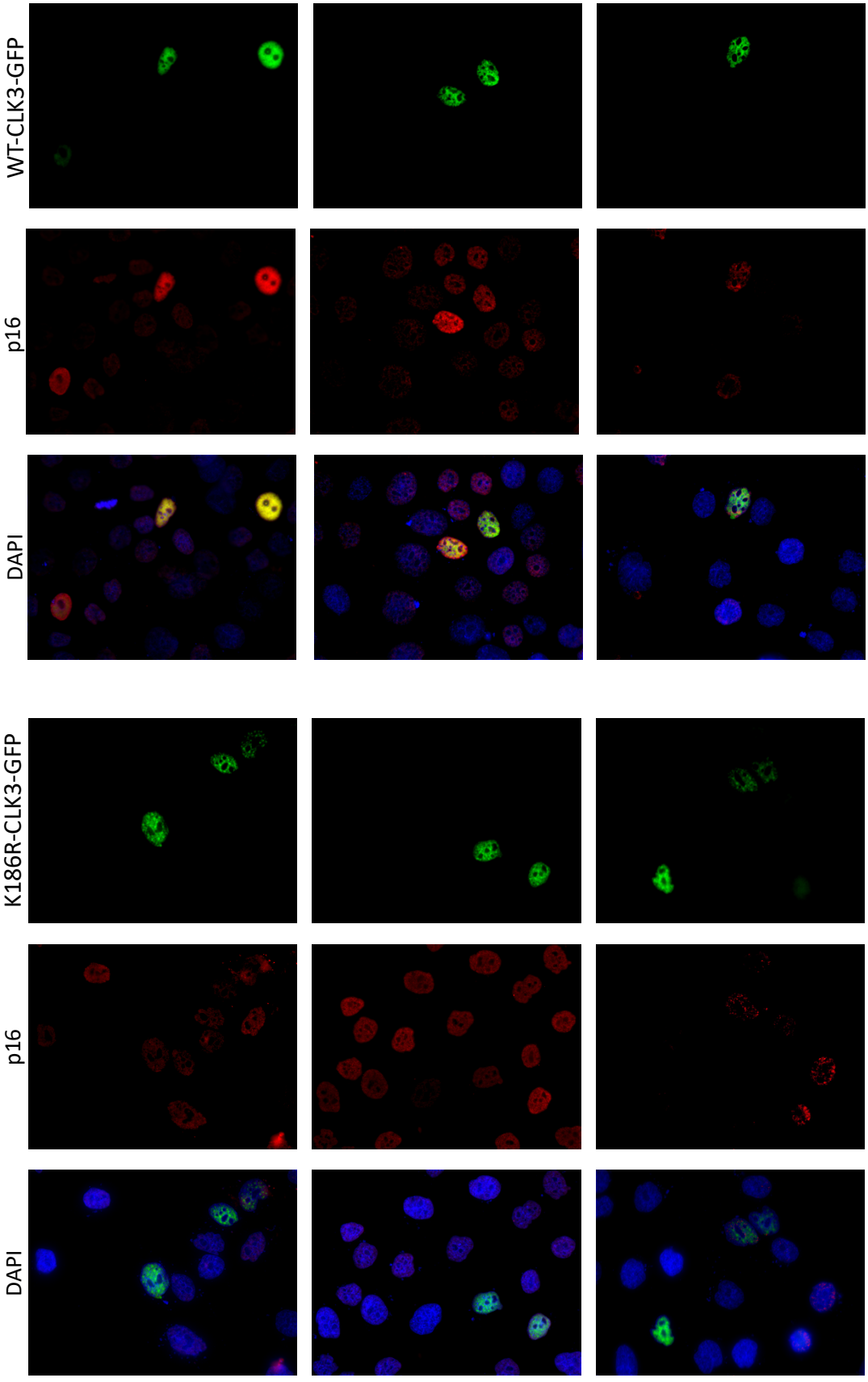
GTATGTATGTTGAATTAGTGTTGTTTGTTGTGTATATGTTTGTATGTGCTTGTATGTGCTTGTAAAT
ATTAAGTTGTATGTGTGTTTGTATGTATGGTATAATAAA**CACGTG**TGTATGTGTTTTTAAATGCTTG
TGTAACTATTGTGTCATGCAACATAAATAAACTTATTGTTTCAACACCTACTAATTGTGTTGTGGTT
ATTCATTGTATATAAACTATATTTGCTACATCCTGTTTTTGTTTTATATATACTATATTTTGTAGCGCC
AGCGGCCATTTTGTAGCTTCAACCGAATTCGGTTGCATGCTTTTTGGCACAAAATGTGTTTTTTTAA
ATAGTTCTATGTCAGCAACTATGGTTTAAACTTGTACGTTTCCTGCTTGCCATGCGTGCCAAATCCC
TGTTTTCCTGACCTGCACTGCTTGCCAACCATTCCATTGTTTTTTACACTGCACTATGTGCAACTACT
GAATCACTATGTACATTGTGTCATATAAAATAAATCACTATGCGCCAACGCCTTACATACCGCTGTT
AGGCACATATTTTTGGCTTGTTTTAACTAACCTAATTGCATATTTGGCATAAGGTTTAAACTTCTAA
GGCCAACTAAATGTCACCCTAGTTCATACATGAACTGTGTAAAGGTTAGTCATACATTGTTCATTTG
TAAAACTGCACATGGGTGTGTGCAAACCGTTTTGGGTTACACATTTACAAGCAACTTATATAATAA
TACTAAACTACAATAATTCATGTATAAAACTAAGGGCGTAACCGAAATCGGTTGAACCGAAACCG
GTTAGTATAAAAGCAGACATTTT

**Figure 8.1. 2 HPV16 URR sequence with the potential E-box binding domain highlighted**

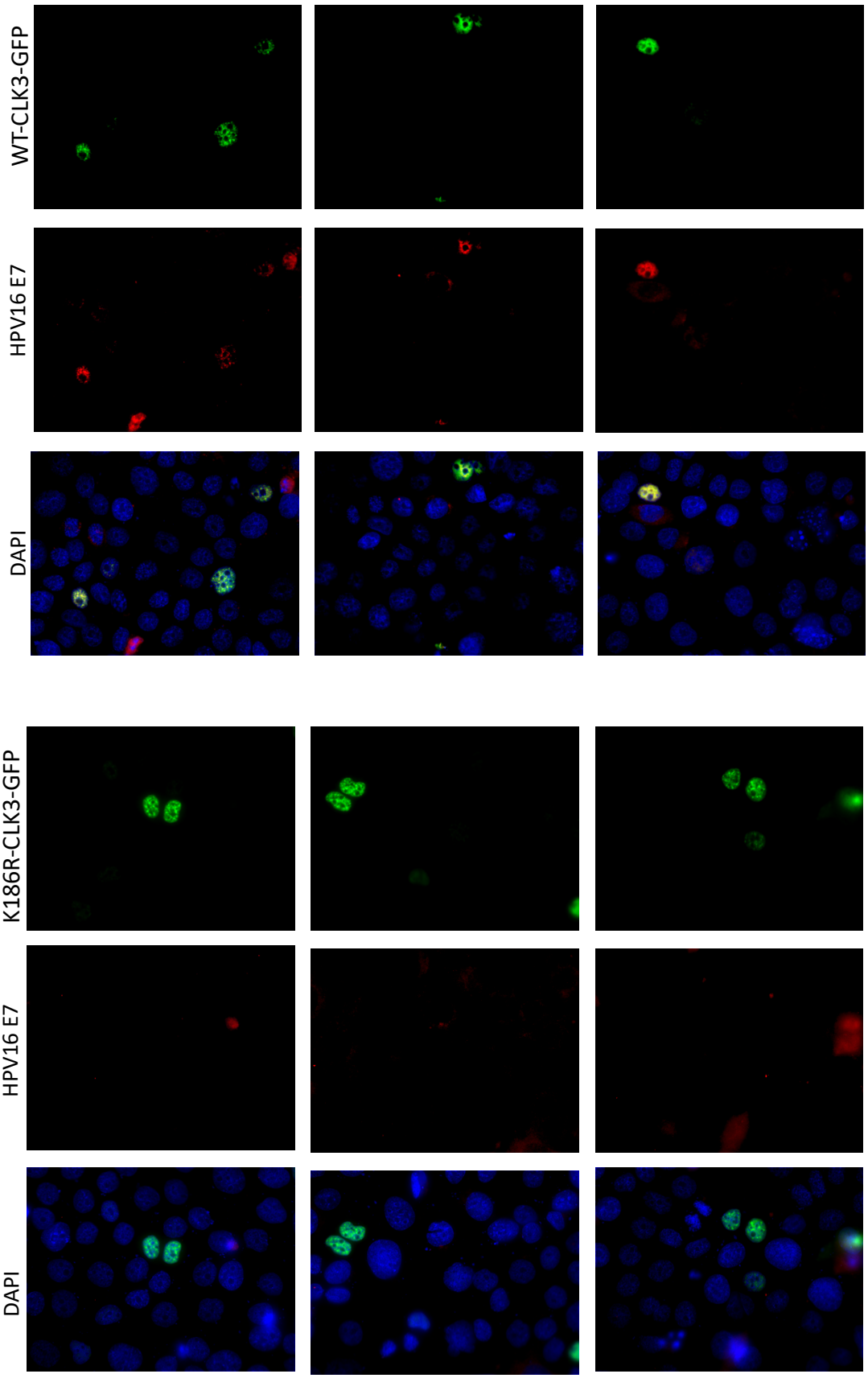
[A] **NIKS CLK3$^{-/-}$ + E6E7**



[B] **NIKS CLK3$^{-/-}$ + E6E7**



**NIKS CLK3$^{-/-}$ + HPV16**



Maxmilan Jeyakumar

**NIKS CLK3$^{-/-}$ + HPV16**

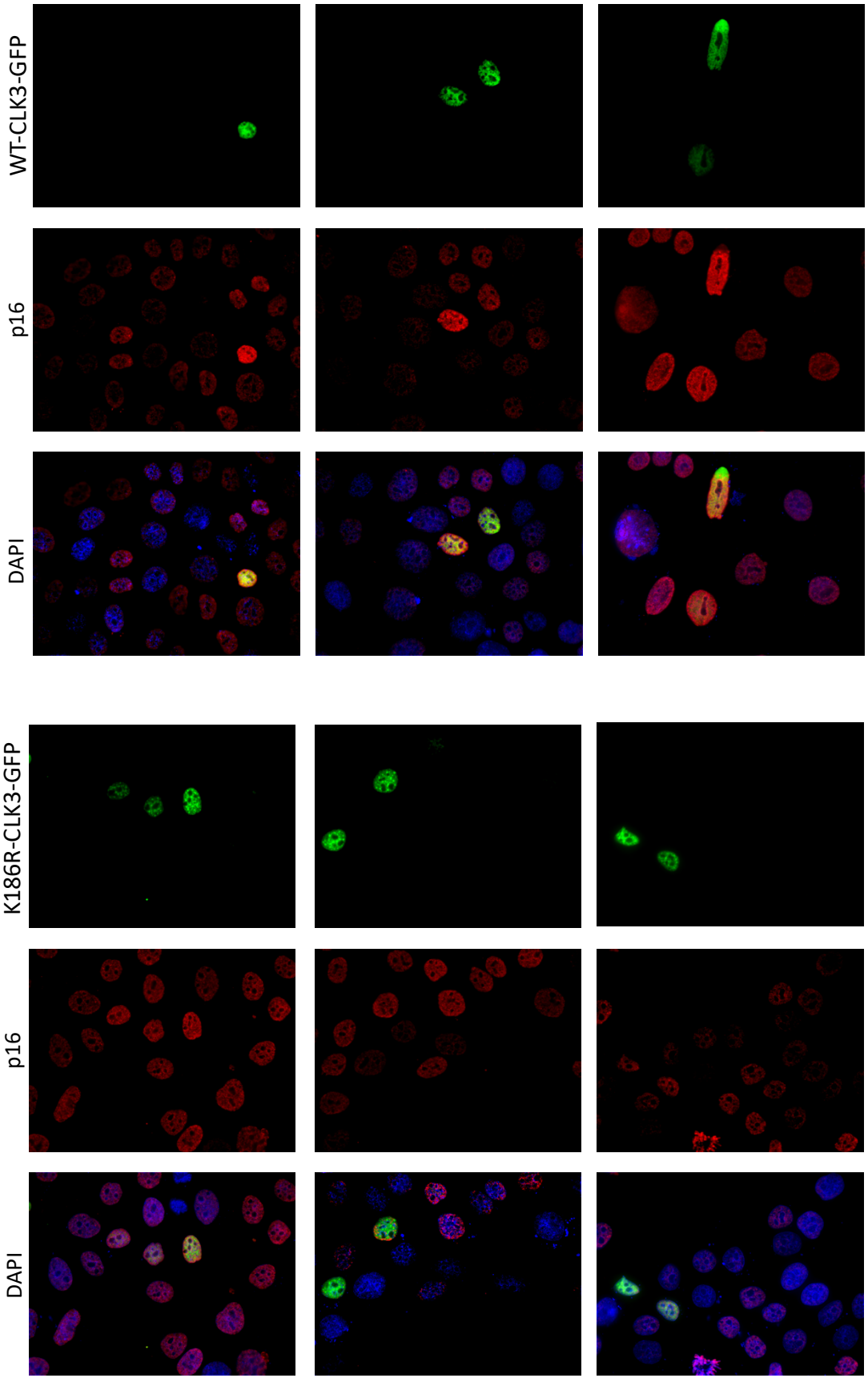
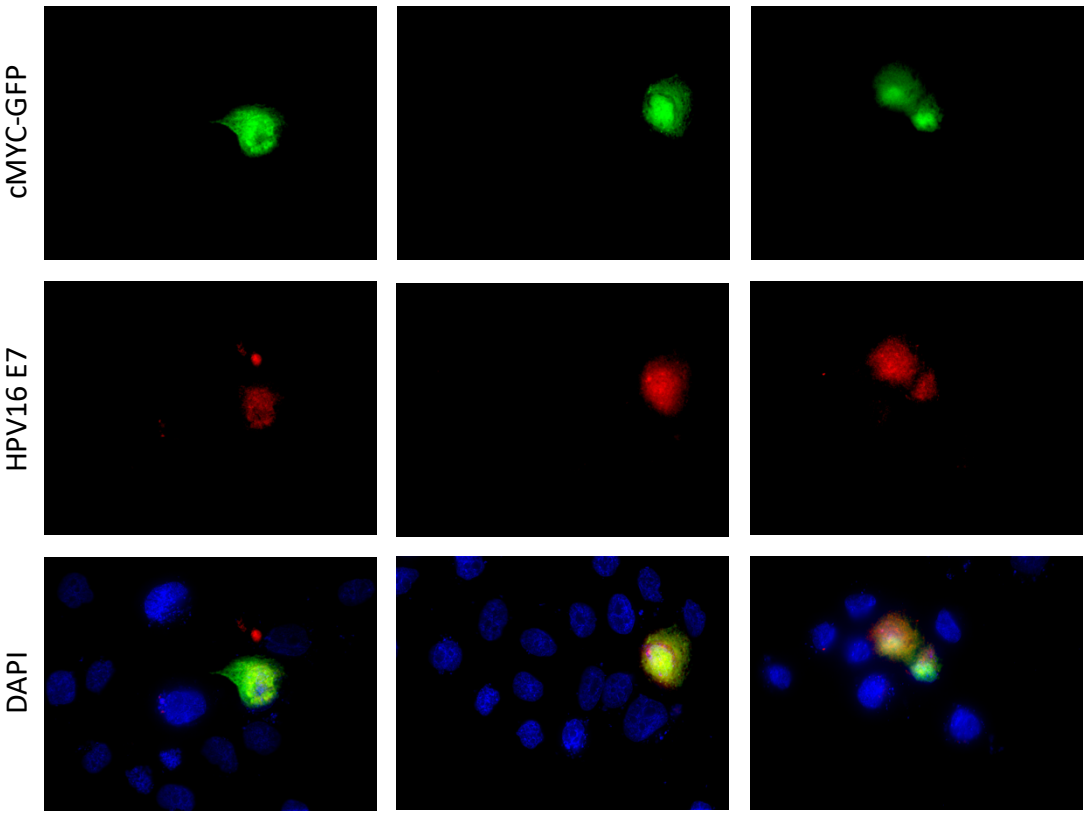
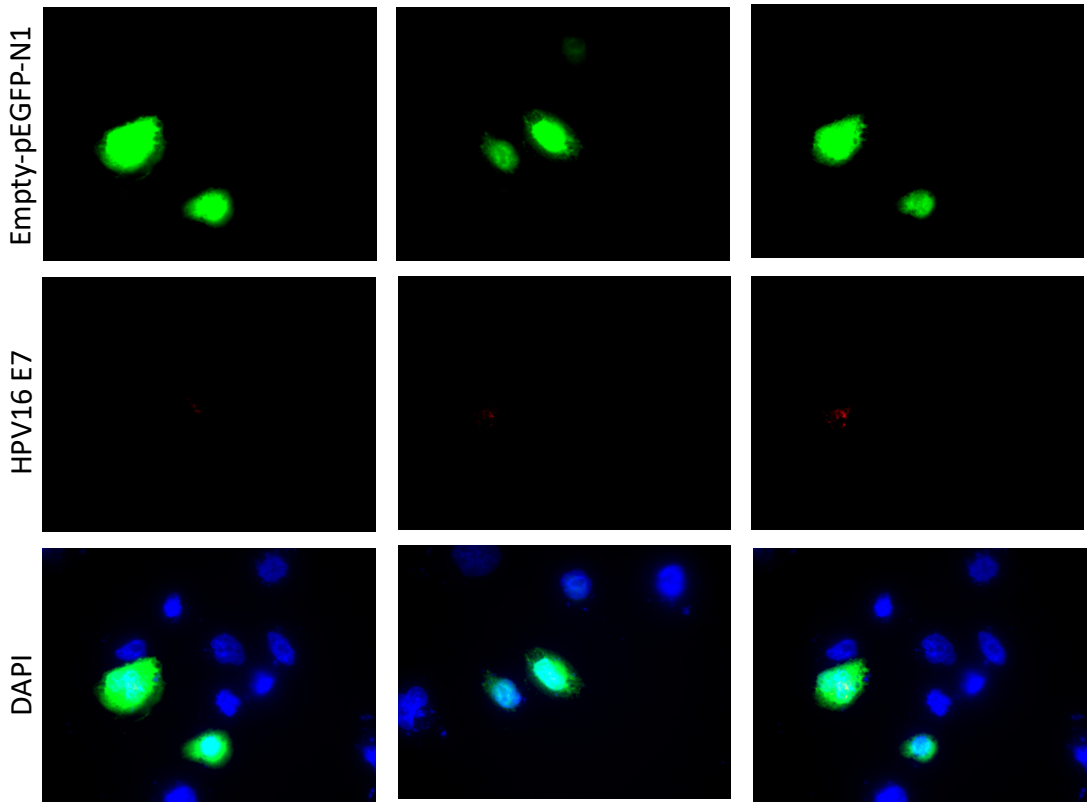


**Figure 8.1. 3 Addional microscopy images of CLK3 add back experiment in NIKS**
IF staining with HPV16 E7 or p16 antibody (red) following transient transfection of WT-CLK3-GFP and K186R-CLK3-GFP for 48hrs in $CLK3^{-/-}$ NIKS with HPV16 uRR E6/E7 and in $CLK3^{-/-}$ NIKS with the full HPV16 genome. Cells expressing the plasmid are green (GFP), and DAPI staining indicates the nuclei (blue). Microscopy images were at 60X magnification.

**NIKS CLK3$^{-/-}$ + E6E7**




**NIKS CLK3$^{-/-}$ + E6E7**



**NIKS CLK3$^{-/-}$**

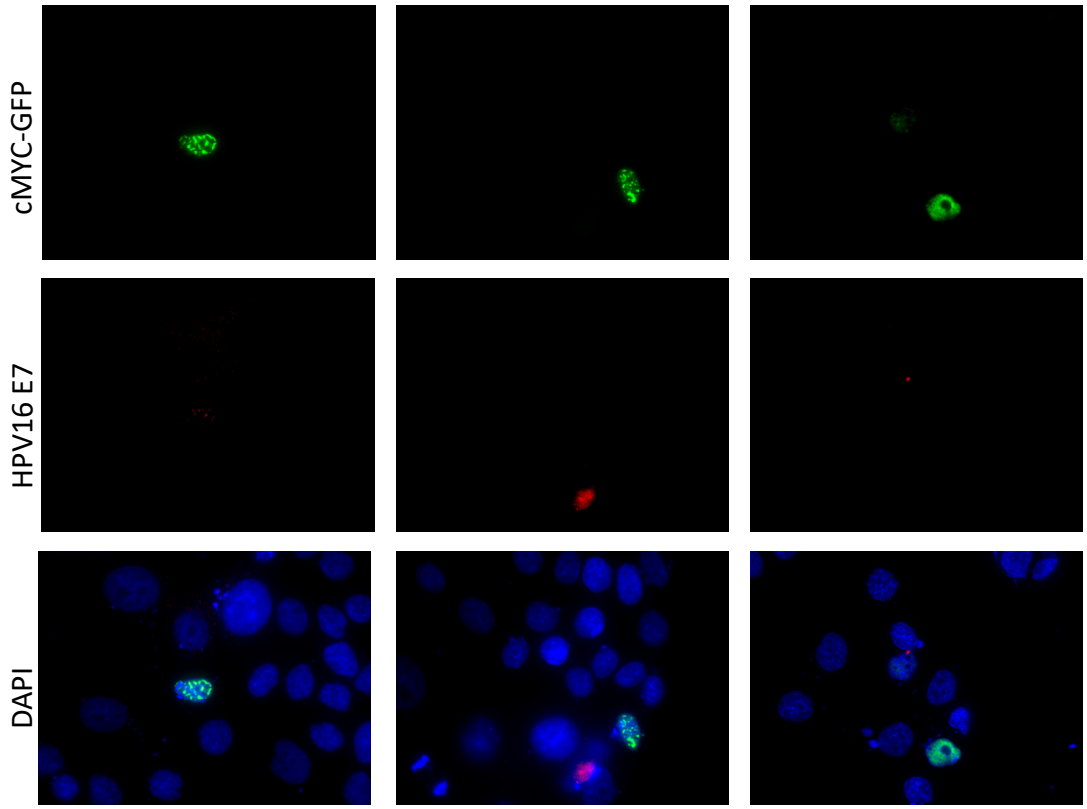



**Figure 8.1. 4 Addional microscopy images of c-MYC overexpression in NIKS**
IF staining with HPV16 E7 antibody (red) following transient transfection of cMYC-GFP in CLK3$^{-/-}$ NIKS and CLK3$^{-/-}$ NIKS with HPV16 URR E6/E7 for 48hrs. Empty-pEGFP-N1 plasmid was used as control. Cells expressing the plasmid are green (GFP), and DAPI staining indicates the nuclei (blue). Microscopy images were at 60X magnification.