2008-01-01

Evaluation of Regulation of Late Gene Expression in Human Papillomavirus-infected Cervical Epithelia

Joanna Fay
Technological University Dublin

Evaluation of Regulation of Late Gene Expression in Human Papillomavirus-Infected Cervical Epithelia

A thesis submitted for the degree of Doctor of Philosophy

by

Joanna Fay B.Sc. (Biochemistry and Molecular Biology)

Supervisors:
Dr. Helen Lambkin
Prof. Stefan Schwartz
Ms. Brid Ann Ryan

School of Biological Sciences,
Dublin Institute of Technology

January 2008

Abstract

Cervical cancer is second only to breast cancer as the most common cancer among women worldwide and is caused by Human Papillomavirus (HPV). The aim of this study was to evaluate the role of RNA-binding proteins in the expression of HPV in cervical cells and clinical samples. Firstly, an audit was performed on cervical biopsy samples taken by laser loop excision over 5 years in the National Maternity Hospital, Holles Street, Dublin, Ireland. The presence of HPV and HPV-16 was evaluated on fifty of these samples and the distribution of multiple RNA-binding proteins, including SR proteins and hnRNPs, was evaluated and related to differentiation (cytokeratins, syndecan) and proliferation-associated (PCNA, p16INK4a) biomarkers. The samples consisted of histologically normal cervical epithelium, HPV-induced low-grade and high-grade pre-malignant lesions and cervical cancers. Low expression of these proteins was detected in basal epithelial cells of normal cervix, with increased expression in intermediate layers, and a lack of expression in the upper superficial layer. Expression of all RNA-binding proteins increased in neoplastic lesions and highest expression was demonstrated in cervical cancers. The expression profile of the RNA-binding proteins was similar to PCNA expression in normal and lesional cervical epithelia, but was not associated with HPV-16 status, p16INK4a or differentiation markers.

The effect of a number of these proteins on transcription of the late HPV genes was evaluated in modified HeLa cell lines (pBEL and pBELM) by transfection studies, with effects evaluated using real-time PCR and Northern blot analyses. It was determined that adenoviral E4orf4, splicing proteins SRp30c and PTB as well as HPV-16 E2 can induce the

expression of the late HPV genes and these were shown to alter the ratio of expression of HPV L1 mRNAs, demonstrating that these proteins may play a role in the differential expression of proteins during the HPV life-cycle.

I certify that this thesis which I now submit for examination for the award of Doctor of Philosophy, is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for an award in any other Institute or University.

The work reported on in this thesis conforms to the principles and requirements of the Institute's guidelines for ethics in research.

The Institute has permission to keep, to lend or to copy this thesis in whole or in part, on condition that any such use of the material of the thesis be duly acknowledged.

Signature Joanna Fay Date 20-5-05

For my family

There is no such thing as a self-made man. We are made up of thousands of others. Everyone who has ever done a kind deed for us, or spoken one word of encouragement to us, has entered into the makeup of our character and our thoughts, as well as our success.

George Matthew Adams

Acknowledgements

I would like to thank all the staff in the School of Biological Sciences for their help during this project. I would especially like to thank Ms. Brid Ann Ryan, whose links with Uppsala University provided me with the opportunity to collaborate with Prof. Stefan Schwartz.

I would like to thank my supervisors, Dr. Helen Lambkin and Prof. Stefan Schwartz, for their guidance and support throughout the project. I would also like to thank Dr. Fergus Ryan for all his help in just about every area of molecular biology!

I would like to thank all the technicians, especially Kathleen and Daireen, who (very patiently) answered my every question and were always available for advice.

To all the post-grads, past and present, both in Lab 230 and B11:3, in particular Leanne, Hilary, Ros, Dan, Barbara Monika, Anna, Klaartje, Dave and Paul, thanks for making the lab such a great place to work, with your help the last 4 years have flown by.

To all my friends and family, especially my Mum and sisters, thank you for always being there and on occasion providing the necessary distraction from research, reminding me that there was life outside the lab!

Abbreviations

Abbreviation	Meaning
ACE	AC-rich element
AP-1	activator protein 1
AAs	amino acids
Ad E4orf4	adenoviral E4orf4
APC	anaphase-promoting complex
APC/C	anaphase-promoting complex/cyclosome
ARE	AU-rich element
ATP	adenosine 5'-triphosphate
BPV	bovine papillomavirus
CBC	cap binding complex
CDK	cyclin-dependant kinases
CF I	cleavage factor I
CF II	cleavage factor II
CIN	cervical intraepithelial neoplasia
CIS	carcinoma in situ
CK	cytokeratin
CK2	casin kinase 2
CMV	cytomegalovirus
CPSF	cleavage and polyadenylation specificity factor
CRPV	cotton-tail rabbit papillomavirus
CS	chondroitin sulphate
CstF	cleavage stimulating factor
CstF-64	cleavage stimulating factor, 64kDa subunit
CT	threshold cycle
CTD	carboxy-terminal domain
DSE	down-stream element
DTT	dithiothreitol
E6-AP	E6 associated protein
EGFR	epithelial growth factor
ESE	exonic splicing enhancer
ESS	exonic splicing silencer
FFPE	formalin-fixed, paraffin-embedded
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GCR	glucocorticoid receptor
GIN	glandular intraepithelial neoplasia
GMP	guanosine mono-phosphate
HAT	histone acetlytransferase
HCII	hybrid capture II
HDAC	histone deacetylases
hnRNP	heterogeneous ribonucleoprotein
HPV	human papillomavirus
HS	heparan sulphate

HSPG	heparan sulphate proteoglycan
IHC	immunohistochemistry
ISE	intronic splicing enhancer
ISS	intronic splicing silencer
kDa	kilodalton
LCR	long control region
LLETZ	large loop excision of the transformation zone
NMH	National Maternity Hospital
NRE	negative regulatory element
ORF	open reading frame
PAB II	poly(A) binding protein II
pAE	early polyadenylation signal
pAL	late polyadenylation signal
PAP	poly(A) polymerase
PCBP 1	poly r(C) binding protein I
PCBP 2	poly r(C) binding protein II
PCNA	proliferating cell nuclear antigen
PDZ domain	PSD-95/discs large/zonacculdens 1 domain
poly(A)	polyadenylation
PP2A	protein phosphatase 2
PTB	polypyrimidine tract binding protein
PV	papillomavirus
RBD	RNA binding domain
RNA pol II	RNA polymerase II
RRM	RNA recognition motif
RS domain	arginine – serine domain
SA	splice acceptor
SCC	squamous cell carcinoma
SD	splice donor
SFM	serum free medium
snRNP	small nuclear ribonucleoprotein
SR protein	serine – arginine rich protein
TGFβ	transforming growth factor β
Tm	melting temperature
TZ	transformation zone
U2AF65	U2 auxiliary factor, 65kDa subunit
URR	upstream regulatory region
UTR	untranslated region
VLP	virus-like particles
YB-1	Y-box protein 1

Table of Contents

Table of Figures

1. Chapter 1 – Introduction

1.1 Virology

The diverse microbial world of bacteria, fungi and protozoa had been established by the latter half of the 19th century [1]. In 1840 Jacob Henle was the first to hypothesise that infectious agents existed that were too small to be observed with the light microscope [2]. This remained a hypothesis for 50 years until 1892, when Dimitrii Ivanowsky demonstrated that even after a filtration step that blocked the passage of bacteria, the sap from a tobacco plant infected with tobacco mosaic disease retained its infectious properties. As no organism could be cultured from the infected sap, the agent failed to satisfy Koch's postulates, a set of criteria to identify if an organism was responsible for causing a specific disease. Ivanowsky firmly believed that this inability to culture an organism from the sap was a result of some technical failure in his experiment [2]. This apparent failure was explained by the work of Martinus Beijerinck who demonstrated that the infectious agent could only reproduce itself in living tissue and not in the cell free sap of the plant [2]. Over the next decade many scientists conducted work in this area, Friedrich Loeffler and Paul Frosch were the first to identify an infectious filterable agent from animals (foot and mouth disease) in 1898 [1] and Walter Reed identified the first from humans (yellow fever) in 1901 [2]. At first the agent was termed *contagium vivium fluidum* (contagious living liquid) but this evolved to the term virus from the Latin word for poison [1].

1.2 Papillomavirus

The viral nature of cutaneous warts was first demonstrated by Ciuffo in 1907 when he demonstrated the transmission of human warts using cell free filtrates [3]. Richard Scope was the first to describe a papillomavirus (PV) as the responsible agent for papillomas in the cotton-tail rabbit when he identified cotton-tail rabbit papillomavirus (CRPV) in 1933 [4].

Papillomavirus virions were first visualised in 1949 using electron microscpy [5] but due to the absence of a tissue culturing system for the papillomaviruses little research was carried out until the 1970s when the first papillomavirus genome was cloned in bacteria [6]. This allowed the analysis of the molecular biology of these viruses. It was also about this time that it was first suggested that human papillomaviruses (HPV) may play a role in the aetiology of cervical cancer. Meisels and Fortin were the first to report that koilocytic atypias in cervical smears may have been as a result of HPV infection [5]. In the early 1980s Harald zur Hausen and colleagues were the first to detect HPV DNA in genital warts; HPV-6 and HPV-11 were the first types to be isolated [5]. In 1983 zur Hausen isolated HPV-16 and a year later HPV-18 directly from cervical cancer biopsies [7]. It took nearly a decade after these findings were published for the causal role of HPV in cervical cancer to be accepted.

Papillomaviruses are members of the Papillomaviridae family [8], they are a large group of small DNA viruses that can induce warts in a number of higher vertebrates including humans, cattle, rabbits, sheep, dogs and some avian species including parrots, and chaffinches [9], although each papillomavirus is species-specific. Their genome is a closed circular, double stranded DNA molecule which is complexed with histones. The DNA is

then condensed into nucleosomes and is encapsidated in an icosahedral virion with a diameter of 55 nanometers [6].

Papillomavirus nomenclature divides the viruses in this family into genus, species, types, subtypes and variants. There are 16 genera in the Papillomavirus family, each genus is identified by a letter in the Greek alphabet, alpha-papillomaviruses, beta-papillomaviruses etc. [8]. Each genus groups together species that are phylogenetically related but may be biologically diverse. Five genera contain only human, ape and monkey papillomaviruses, while the remaining 11 genera contain papillomaviruses which infect other mammals and birds. The alpha-papillomaviruses contain the clinically most important HPVs, those associated with genital and mucosal lesions [8]. The different genera share less than 60% nucleotide homology in a product expressed in the latter stages of the viral lifecycle, the late gene 1 (L1) [10].

There are 44 species divided among the 16 genera, species share between 60-70% nucleotide identity in L1. Each species is identified by a number with each genera containing a varying number of species, the alpha-papillomaviruses contain 15 species while the epsilon- and zeta-papillomaviruses contain only one [10].

The types contained in a species will not only be phylogenetically related but will also share biological and pathological properties. Differing types are generally named for the host species of the virus, for example bovine papillomavirus (BPV) and cotton-tail rabbit papillomavirus (CRPV). When more than one distinct isolate of papillomavirus occurs in a single host, each isolate is distinguished by a number, for example human papillomavirus type 1 (HPV-1) and human papillomavirus type 2 (HPV-2). To be defined as a separate

type papillomaviruses must have a L1 coding sequence that shares less than 90% sequence homology with any other known type [8].

The term subtype refers to an isolate whose L1 sequence is between 90-98% homologus with a known type. Only three HPV isolates fulfil this criterion as subtypes, they are, HPV-46 (subtype of HPV-20), HPV-55 (subtype of HPV-44) and HPV-64 (subtype of HPV-34) [8].

A variant shares more than 98% sequence identity with another papillomavirus type [10]. Variants show maximal divergence when they have been isolated from ethnic groups that have evolved for a long period of time with little or no contact.

Various papillomavirus types infect many different species, but due to their medical importance, HPVs are the most extensively studied papillomavirus types.

1.3 Human Papillomavirus and Cervical Cancer

There are an estimated 10 million cases of cancer diagnosed around the world every year [11]. Cervical cancer is the second most common cancer in women worldwide after breast cancer [11]. The World Health Organisation estimated that in 2005 over 500,000 new cases of cervical cancer were diagnosed and that nearly 260,000 women died from the disease [12]. The burden of disease is not spread equally around the world with 90% of the new diagnoses and 95% of the deaths occurring in developing countries [12].

Cervical cancer arises in the epithelium lining of the cervix. The two major morphological types are squamous cell carcinoma (SCC) and adenocarcinoma, they account for approximately 85% and 10% of cervical cancers respectively [13]. The remaining 5% of cervical cancers are adenosquamous in origin or are other rare tumour types [13].

Squamous cell carcinomas generally arise from metaplastic squamous mucosa that occurs in the region of the transformation zone of the cervix [14]. The incidence of squamous cell carcinoma has been in decline in recent years, probably as a result of organised cervical screening programmes [13]. Adenocarcinoma arises from the glandular epithelium of the endocervix [14]. The incidence of this type of cervical cancer has been increasing, mainly in young women, possibly because screening is less effective for detecting this type of cancer [13]. Cervical cancer develops from precursor lesions called cervical intraepithelial neoplasia (CIN), it can take between 10 and 20 years for cervical cancer to develop [12, 13]. These precursor lesions can be detected by screening programmes many years before they become cancerous, allowing easier treatment and better outcomes. It has been shown that in countries that have established systematic screening programmes the number of deaths due to cervical cancer has been reduced by 70% or more [15].

After the discovery by Harald zur Hausen and colleagues of the presence of HPV DNA in cervical biopsies in the early 1980s a great quantity of research has been carried out in this area. HPV DNA has been detected in 99% of cervical cancers and 94% of CIN cases [16]. HPV was the first virus to be classed as carcinogenic to humans by the World Health Organisation in 1995, and it is recognised that HPV has a causal role in the development of cervical cancer [17, 18]. Of the approximate 100 types of HPV that have been isolated, 40 have been shown to preferentially infect genital mucosa [19], all of these types are contained within the alpha-papillomavirus genus [10]. These 40 types have been divided into 2 risk groups depending on their association with the development of cervical abnormalities, including cervical cancer. There are 13 types in the high risk group, which is also known as the carcinogenic group, they are listed in Table 1.1. These HPV types have

been shown to have a high association with the development of cancer of the cervix, and there is also evidence to support a role for some HPV types in the development of cancers other than in the cervix, for example HPV-16 and HPV-18 in cancer of the vulva, vagina, penis and oral cavity [20]. The second risk group, the probably carcinogenic group, contains 5 HPV types for which there is limited data associating these types to cervical cancer [19, 20].

Risk Group	HPV Types
High-risk	16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59 and66
Probably high-risk	26, 53, 68, 73 and 82

Table 1.1 Risk groups for HPV types

Although there is an approximate 75% lifetime risk for acquiring a HPV infection [11] and it has been deemed to be the cause of cervical cancer, only a small number of women who contract a HPV infection will ever develop cancer. This is due to the fact that most women will clear the infection within 12 to 18 months after infection [21]. Only a small proportion of those infected will not clear it and develop a persistent infection [15, 16]. Out the 13 high risk HPV types, the 8 most common types (HPV-16, -18, -45, -31, -33, -52, -58 and –35) cause 90% of cervical cancers [19, 22]. HPV-16 has been found to be very carcinogenic with a 40% risk of developing CIN 3 after 5 years of persistent infection [23]. If no screening occurred it is estimated that cervical cancer would occur in approximately 1% of women who acquire HPV infection [15].

1.4 HPV Vaccine Development

With the knowledge that cervical cancer is a result of a viral infection, researchers in the early 1990s started to look at ways of preventing this type of cancer through a prophylactic vaccine. Two companies, Merck and GlaxoSmithKline (GSK), have each developed separate vaccines that they hope will reduce the rate of cervical cancer due to HPV [24]. Both vaccines are based on the fact that the HPV L1 protein is highly immunogenic and also will spontaneously form into virus-like particles (VLPs) [24]. These VLPs when injected into the patient (3 intramuscular injections over 6 months) result in an effective immune response [25]. As the VLPs are HPV type-specific and they generate very little cross-reactivity with other types, it was necessary to produce vaccines that would protect against multiple types [24]. The GSK vaccine, called Cervarix, is bivalent and offers protection against HPV-16 and HPV-18 [26]. Merck produced a quadrivalent vaccine, called Gardasil, against HPV-16, HPV-18, HPV-6 and HPV-11, this vaccine will not only protect against cervical cancer (HPV-16 and HPV-18), it will offer protection against genital warts (HPV-6 and HPV-11) [26]. Clinical trials have proven that both of the vaccines are very effective at not only reducing persistent infection of the types vaccinated against, but also reducing CIN 2 and CIN 3 caused by HPV-16 and HPV-18 [27]. The Food and Drug Administration in the United States licensed the Merck quadrivalent vaccine for use in June 2006 [27]. It has also been approved for use in a number of countries including Australia, Canada and New Zealand [25]. In the United States it has been recommended in that girls aged between 11 and 12 be vaccinated [25].

While the effectiveness of these vaccines has been shown in the clinical trials that have been conducted, a number of questions still remain to be answered. The length of duration

of protection that the vaccine offers or if booster vaccines will be required is not yet known [28]. The longest follow-up that has taken place so far is 53 months for Cervarix and 36 months for Gardasil, at these time points both vaccines still offered good protection but a much longer follow-up is required to establish duration of protection [25]. Gardasil has been shown to be safe and immunogenic in 9-15 year old boys and girls and Cervarix in 10-14 year old girls but trials are ongoing to test if this is true for women over 26 (for both vaccines) and in young men for Gardasil [25]. Further studies are needed to establish if these vaccines can be used in infants and immunosuppressed individuals, what the optimal age of vaccination is, if the vaccine will be effective against established infection, if other high-risk types will become more common in the place of HPV-16 and HPV-18, and if it will be effective in males [27, 28].

It is hoped that these vaccines will reduce the occurrence of cervical cancer by approximately 70% [29]. But screening programmes will need to continue as it will take 10 to 15 years for the vaccinated individuals to reach screening age and also to detect cervical cancer caused by HPV types other than HPV-16 and HPV-18.

1.5 Human Papillomavirus Life Cycle

HPVs are strictly epitheliotrophic [4] and have two distinct phases in their life cycle. The first, called the non-productive phase of the life cycle, involves the establishment of the viral genome in the host cell, and viral replication [30]. The second, or productive phase, involves the high level amplification of the viral genome, the production of the viral capsid proteins and the formation of the virion [30].

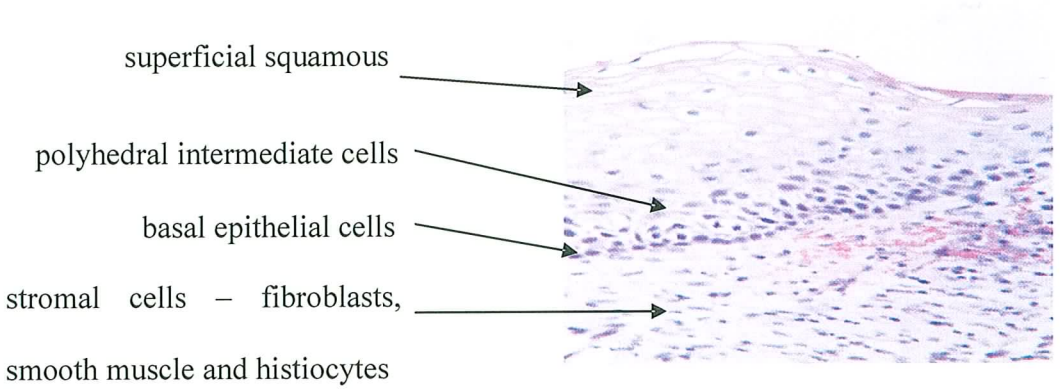


Figure 1.1 Cervical stratified squamous epithelium

Haematoxylin and eosin stained cervical stratified squamous epithelium with epithelial layers indicated

Although proliferating cells are infected by HPV, viral particles are formed exclusively in terminally differentiating keratinocytes of the infected epithelium [31]. For infection to occur HPV needs to access the basal cell layer of the epithelium (Fig. 1.1), this possibly occurs through the site of an abrasion or micro-lesion [31]. After gaining entry to a cell, possibly after attachment via heparan sulphate, and entry via an as yet unknown receptor [32], the viral genome remains as extrachromosomal elements, called episomes, in the nucleus. At this early stage of infection the virus uses the replication machinery of the host cell to establish itself in low copy number, generally 50-100 copies per cell [33]. The DNA is generally replicated once per cell cycle in a bidirectional mode [34]. This is the non-productive phase of the viral lifecycle and occurs in the basal layer of the epithelium. Only the early gene products, which are under the control of the early promoter, are expressed at this stage (Fig. 1.2). As these cells divide, the viral episomes are divided equally between the daughter cells, some of these daughter cells migrate upwards away from the basal layer and start to undergo differentiation. The remaining infected cells stay in the basal layer and

so provide a reservoir of viral DNA for further cell divisions [33]. Due to the fact that the highly immunogenic virion particles are not produced in the basal layer, the immune response is not alerted, and so these infected cells can continue to proliferate [16]. Uninfected epithelial cells exit the cell cycle after they leave the basal layer and start to undergo terminal differentiation. HPV infected cells continue to cycle in the upper epithelial layers due to the action of some of the viral proteins. As the infected epithelial cells start to undergo differentiation the late viral promoter is activated and the late genes are expressed. This is the productive phase of the life cycle and it occurs in the parabasal and superficial epithelial layers (Fig. 1.1 and 1.2). The viral genome is packaged into the capsid forming the infectious particles. HPVs are non-lytic viruses, the newly formed viruses are not released from the cell until it reaches the surface of the epithelium and the terminally differentiated cells are sloughed off [31]. The release of the virion is possibly aided by the action of one of the early HPV proteins, the E4 protein, in disturbing keratin integrity [7].

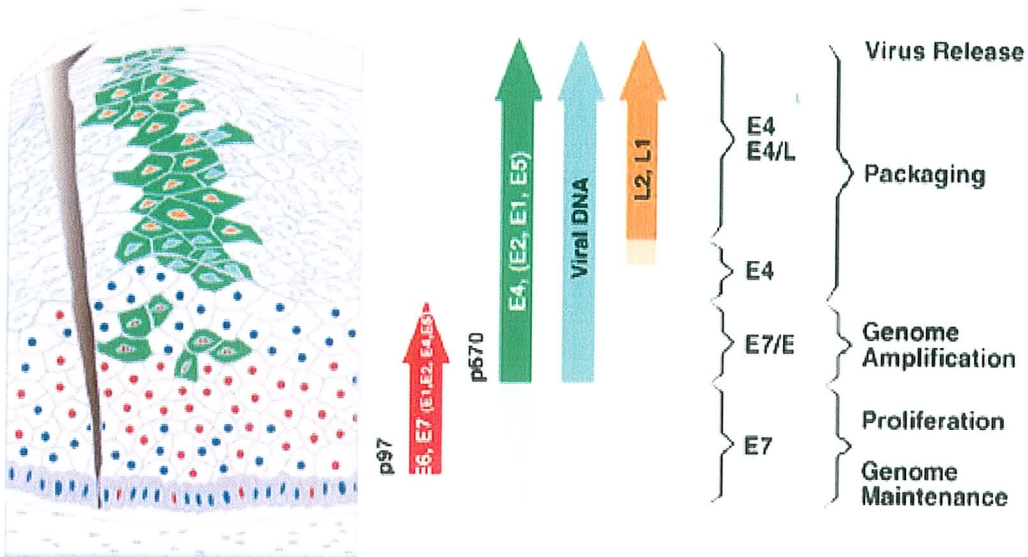


Figure 1.2 Diagrammatic representation of cervical epithelium indicating HPV-16 gene expression in infected epithelial cells

Diagram modified from [31]. Red nuclei represent expression of E6 and E7, which are required for S-phase entry. Green cells represent up-regulation of the late promoter (p670) increasing expression of the viral replication proteins (E1, E2, E4 and E5). Light blue nuclei represent amplification of the viral genome. Yellow nuclei represent production of the viral capsid proteins (L1 and L2).

1.6 Human Papillomavirus Genome

The genome of HPV is approximately 8000 nucleotides long [4]. HPV genomes usually contain eight open reading frames (ORFs) [16]. Generally only one of the strands in the double helix is transcriptionally active and so transcription occurs in one direction only [5] producing polycistronic mRNAs [35]. HPV genomes can be divided into three regions, a long control region (LCR) or upstream regulatory region (URR), an early region (E) and a late region (L) (Fig. 1.3). The LCR covers about 10% of the genome and contains binding sites for a number of viral proteins and also host cellular transcription factors that can

positively or negatively control the transcription of the viral genes and the replication of the viral genome [36]. The early region covers over 50% of the genome and can contain up to eight ORFs, but most HPVs only contain E1, E2, E4, E5, E6 and E7. HPV-31 is the only type to have been proven to contain a protein encoding E8 ORF [37]. All of the early genes are expressed in the non-productive phase of the viral life cycle. E4, although located in the early region, is also expressed during the productive stage of the life cycle [3]. These genes contribute to the regulation of viral transcription and viral DNA replication and also affect the growth of host cells [36]. The late region contains two ORFs, L1 and L2. These genes encode the two structural proteins which make up the viral capsid. These late genes are only expressed in the productive stages of the viral life cycle [35]. The HPV genome generally contains two major promoters, the early promoter that controls the expression of the early genes and a differentiation dependant late promoter that controls the expression of the late genes [9, 32]. Two polyadenylation signals are also present in the HPV genome. The early polyadenylation signal (pAE) is located between the early and late regions and during the non-productive stage of the viral life cycle almost all transcripts are polyadenylated at this site [38, 39]. The late polyadenylation signal (pAL) is located after the L1 coding region and is used during the productive stage of the life cycle to produce the viral capsid proteins [38].

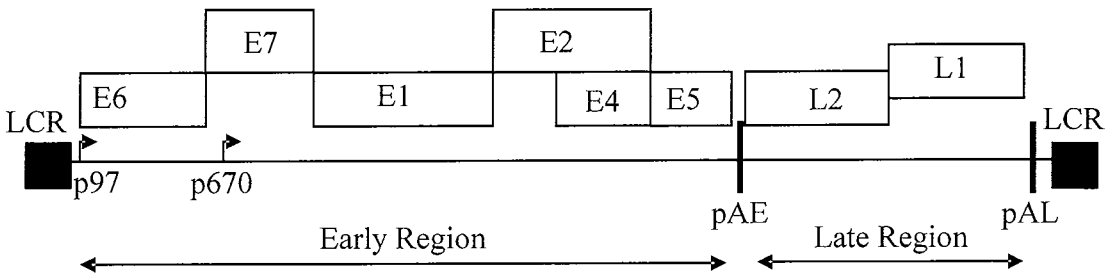


Figure 1.3 HPV-16 genome

The genome is 8kb long and contains 8 ORFs (represented by white boxes) and a long control region (represented by black boxes). The location of the early promoter p97 and the late promoter p670 are indicated by arrows. Also shown are the early (pAE) and late (pAL) polyadenylation signals.

1.7 HPV Promoters and Proteins

The six major ORFs in the early region of the HPV genome that are transcribed in the non productive phase of the viral life cycle, originate from the early promoter located in the E6 ORF at p97 in HPV-16 and p99 in HPV-31 [38]. The early promoter is a potent promoter and is tightly controlled by cis-elements that are located in the LCR. The late gene transcripts originate from sequences located in E7 ORF located at p670 in HPV-16 and p742 in HPV-31 [38]. This promoter only becomes active in differentiated cells [40].

1.7.1 E1 Protein

The E1 ORF is the largest ORF in the HPV genome and it encodes one of the most conserved HPV proteins [41]. The E1 protein varies in size depending on the HPV type, from 68-85kDa [42]. E1 has been shown to be an essential replication protein and is the only HPV protein that has been shown to have enzymatic activity; it can function as an ATPase and as an ATP-dependant helicase. The helicase activity is required to catalyse the

unwinding of the DNA at the origin of replication [41]. E1 is expressed throughout the life cycle but is only expressed at high levels when the late promoter is activated [35]. E1 has been shown to bind the α subunit of DNA polymerase, recruiting the replication machinery to the viral origin of replication [36, 41]. The E1 protein initiates replication by binding to AT-rich regions within the LCR, binding affinity is weak but it is strengthened by forming complexes with locally bound E2 [43]. Once E1 binds the viral DNA it forms a hexomeric ring structure around the DNA molecule. E1 can associate with cyclin/CDK complexes, these complexes can phosphorylate E1 which is necessary to maintain its nuclear localisation and provide a mechanism of regulating its activity [35]. It has also been shown that E1 can bind to human topoisomerase I through its DNA binding domain and its C-terminal domain [44]. This stimulates its activity and prevents the overwinding of DNA during replication [44].

1.7.2 E2 Protein

The HPV E2 ORF codes for a protein 48-50 kilodalton (kDa) in size, that has functions in both replication and transcription [32, 42]. The E2 protein forms dimeric complexes that bind to HPV DNA in the LCR through β-barrel structures. E2 binds DNA at a palindromic recognition sequence $ACCN_6GGT$ [41]. The LCR of high risk HPVs contain 4 E2 binding sites, 3 of these sites flank the E1 recognition sequence increasing the affinity of the E1 DNA binding [35]. E2 can act as either an activator or a repressor of transcription from the early promoter. At low levels it is an activator of early gene expression and at high levels it acts as a repressor by interfering with the binding of transcription factors such as TFIID and Sp1 [32]. E2 is a major regulator of E6 and E7 expression but when the late promoter is

activated transcriptional repression by E2 is overcome and leads to an increase in the level of E1 and E2 expression [35]. E2 may also alter the structure of chromatin through interaction with histone acetlytransferases (HATs), this affects the expression of genes from the early promoter [35]. HPV-18 E2 has been shown to be anti-proliferative by inducing cell cycle arrest during G1, and also may be pro-apoptotic possibly through the activation of caspase 8 and the activation of the extrinsic apoptotic pathway [45]. Apoptosis is a genetically determined programme that leads to the induction of caspase-activated deoxyribonucleases [46]. The E2 ORF is frequently disrupted during integration, this allows over-expression of E6 and E7 to occur.

1.7.3 E4 Protein

The E4 ORF codes for a protein between 10-44kDa depending on the HPV type [42], and completely overlaps the E2 ORF, but it is read in a different reading frame [9]. The E4 ORF contains no start codon and so is a splice product that includes the first 5 amino acids (AAs) of E1 (which contains the start codon) and the E4 ORF, it is therefore referred to as E1^E4 [32]. It is the most highly expressed of all the HPV proteins with highest levels occurring in the suprabasal layers [35]. E1^E4 is expressed throughout the life cycle and is localised primarily in the cytoplasm. Expression of E1^E4 occurs before L1 expression but in co-ordination with genome amplification. E1^E4 has been shown to be associated with cytokeratins through a N-terminal LLXLL motif [35], (where L = leucine, X = any amino acid). This binding possibly leads to the collapse of the cytokeratin network but only a limited amount of collapse has been seen in lesions [35]. It has also been shown that E1^E4 can interact with E4-DEAD box protein through its C-terminal domain. As DEAD box

proteins are generally involved in the regulation of mRNA stability this may suggest a role for E1^E4 in post-transcriptional regulation [47].

1.7.4 E5 Protein

HPV E5 is a small, approximately 14kDa, hydrophobic membrane associated protein [35, 42, 43]. It contains 3 membrane helices with short regions at the N- and C-terminal that extend beyond the membrane. E5 is mainly distributed in the golgi apparatus, endoplasmic reticulum and occasionally the plasma membrane [43]. HPV has been shown to have weak transforming ability but along with E6 and E7 it can increase the overall transforming ability of the virus [35]. As it is expressed in very low levels in infected cells most of the information about the E5 protein comes from over-expression studies which may affect the overall result [35]. HPV-16 E5 has been shown to activate the epithelial growth factor receptor (EGFR) which can lead to the over-expression of a number of proto-oncogenes including c-jun and c-fos [48]. Heterodimerisation of these two proteins forms the activator protein 1 (AP-1) which can not only stimulate mitosis but through AP-1 binding sites in the LCR can increase the expression of E6 and E7 [49].

1.7.5 E6 Protein

The E6 protein is approximately 18kDa in size and contains 2 zinc finger-like motifs [9, 42]. It is one of the earliest proteins expressed in the HPV life cycle and has been shown to be localised to both the cytoplasm and the nucleus [50]. Like all viruses, HPV can only replicate within its host cell using the cellular replication machinery, consequently HPV

requires the epithelial cell to remain replication competent after it leaves the basal layer. E6 helps to alter the cellular environment to allow viral replication in the intermediate epithelial layers [50]. E6 from high-risk HPV types (high-risk E6) is an anti-apoptotic protein through its interaction with p53 but no functional interaction between E6 from low-risk HPV types (low-risk E6) and p53 has been observed [9]. p53 is the so called 'guardian of the genome', and functions as a sequence specific DNA binding protein and acts as a transcription factor that up-regulates the expression genes involved in cell cycle arrest and apoptosis [35]. When DNA damage occurs, p53 transcriptionally activates the expression of a number of regulators of cell cycle progression including $p21^{WAF1/CIP1}$ gene. p21 inhibits some of the cyclin-dependant kinases that are necessary for the progression from G1 to S phase, allowing DNA damage to be repaired. If the damage is too great to be repaired p53 can induce the death of the cell through apoptosis.

In order to prevent this and to keep the cell cycling, high risk E6 has been shown to indirectly interact with p53 through a cellular ubiquitin-ligase, the E6-associated protein (E6-AP) [51]. This complex of E6 and E6-AP then binds the p53 protein [52]. E6-AP-p53 recruits a complex of enzymes called the ubiquitin complex which ubiquitinates lysines on p53 initiating its proteolysis by the proteosome [53]. With p53 degraded by E6, the normal growth arrest signals of the G1/S phase boundary are bypassed [54]. It has also been shown that in cells that express E6 *in vivo*, the stability of p53 is reduced, its half-life decreases from approximately 3 hours to 20 minutes [55].

HPV-16 E6 has been shown to increase the activity of the enzyme telomerase [56]. Telomerase is a multisubunit enzyme responsible for replicating telomeric DNA present at the ends of chromosomes [54]. In most somatic cells the telomerase enzyme is not active

and so the telomeres present at the ends of the chromosomes gradually shorten over successive DNA replications. It is possible that once the telomeres reach a certain shortened point, a natural pathway leading to senescence and cell death is activated [57]. In contrast to this, most tumour cells maintain the length of their telomeres through telomerase activity. The telomerase enzyme has a protein core TERT which catlyses reverse transcription, and an RNA molecule (TERC) which serves as a template for synthesis [56]. The regulation of telomerase occurs at the level of the TERT subunit, as this is the limiting factor. E6 can up-regulate the transcription of TERT from its promoter, possibly trough the binding of transcription factors Myc and Sp1 [32].

HPV-16 and HPV-18 E6 proteins contain a carboxyl terminal PSD-95/discs large/zonaocculdens 1 (PDZ) domain [58]. A PDZ domain is an approximately 90 AA protein-protein interaction domain. It is most often found within cellular proteins that are located at areas of cell-cell contact such as tight junctions in epithelial cells, possibly aiding in signal transduction [32]. E6 can bind to a number of cellular proteins that contain PDZ domains including, hDlg, the human homologue of the *Drosophila melanogaster* tumour suppressor protein discs large, MUPP1, a multi-PDZ domain containing protein, and hScrib, the human homologue of the *D. Melanogaster* tumour suppressor scribble [54]. The binding of these proteins to HPV-16 and HPV-18 E6 proteins results in E6-AP-mediated ubiquitation and proteolysis [35]. This may contribute to the promotion of unregulated cell proliferation as the degradation of these signalling molecules prevents the transmission of the signal to downstream effectors. Mutations within the PDZ domain of E6 results in the removal of its transforming abilities [59].

1.7.6 E7 Protein

E7 is made is approximately 10kDa in size, and is a zinc binding phosphoprotein [5, 42]. It binds zinc through the carboxyl terminal domain which contains two copies of a Cys - X - X - Cys domain, similar to the zinc binding domain of E6 [43]. E7 proteins dimerise through this zinc finger motif [60, 61]. E7 is phosphorylated at two serine residues at the amino terminal domain by casin kinase 2 (CK2) [62]. E7 has a short half-life (approximately 30 minutes) and is rapidly degraded by the proteosome [63]. The protein is mainly located in the nucleus of the host cell.

E7 is believed to cause the post-mitotic suprabasal cells to become permissive for DNA synthesis [64]. It contributes to creating a favourable environment for DNA amplification during the productive stage of the HPV lifecycle [65]. This is achieved mainly through the interaction between E7 and members of the Rb tumour repressor family. Interaction with these proteins facilitates the progression of the cell cycle into S phase [66]. pRb, p107 and p130 are all members of this family of proteins and they all contain binding sites for HPV E7 [58]. These proteins all regulate transcription of genes with promoters that are dependant on E2F activation.

pRb is the most studied member of this family of proteins. Its state of phosphorylation is regulated through the cell cycle [66]. At the G1/S boundary it is phosphorylated at multiple serine residues by cyclin-dependent kinases allowing progression into S phase. In its hypophosphorylated form pRb is active and binds the E2F transcription factor [54]. The sequestering of E2F by pRb represses transcription from genes that have promoters containing E2F sites, therefore inhibiting transcription [67, 68]. Many proteins that are required for DNA synthesis, such as DNA polymerase α, have E2F-dependent promoters

and so are transcribed in a cell cycle-dependent manner due to the regulation of pRb and E2F [3]. It is also believed that pRb plays an important role in cell cycle exit following differentiation [54]. Deregulation of this control mechanism is an important factor in the ability of E7 to transform cells. E7 binds the hypophosphorylated form of pRb, inactivating it and targeting it for degradation via the ubiquitin-dependent protein degradation pathway. The degradation of pRb is essential to efficiently overcome cell cycle arrest and hence contribute to uncontrolled cell growth [69, 70]. Low-risk E7 can also bind pRb but at a ten-fold lower efficiency when compared with high-risk E7 [9] and degradation of pRb mainly occurs with high risk E7 proteins [35].

E7 interacts with inhibitors of cyclin-dependent kinases such as p21 and p27 [58]. p21 is a very important mediator in the inhibition of cyclin-dependent kinases during keratinocyte differentiation [71]. Hence the sequestering of p21 by E7 may be critical in allowing the continued replication of the viral genome in differentiated squamous epithelial cells where otherwise replication of DNA would no longer occur. It can be said that E7 helps to uncouple the withdrawal from the cell cycle and the process of differentiation [64]. E7 may be involved in overcoming transforming growth factor β (TGFβ) growth arrest through the binding of p27, a possible mediator in TGFβ regulation [72]. E7 also interacts with class I histone deacetylases (HDAC). HDACs are involved in the regulation of transcription through the remodelling of chromatin making it unavailable for transcription. E7 binds to HDACs indirectly through Mi2β, this interaction has been shown to specifically increase the levels of E2F-mediated transcription in differentiating cells driving the cells into S phase [35].

It has been found that E7 can cause abnormal chromosomal duplication resulting in multipolar and abnormal mitosis. Although E6 can amplify the effect, it is E7 that plays the major role in inducing these chromosomal-related mitotic disturbances [58]. Abnormal centrosome duplication induced by E7 rapidly results in genomic instability and is one of the hallmarks of a cancer cell [73].

1.7.7 L1 Protein

The L1 protein is the major capsid protein, it is one of the most conserved proteins in the HPV genome and is approximately 55kDa in size [41]. Three hundred and sixty L1 monomers assemble into 72 pentameric capsomers that together form the viral capsid [35]. L1 capsomers are formed in the cytoplasm and are transported into the nucleus through the nuclear pore by a nuclear adaptor/receptor Kapα2β [35]. The virion containing the viral DNA is then assembled in the nucleus. L1 proteins can on their own form into virus-like particles (VLPs) and immunisation with these particles can induce an immune response to a future challenge with the virus [24]. L1 can bind to cell surface hepran sulphates, this is thought to be the site of viral attachment to the host cell [74].

1.7.8 L2 Protein

The L2 protein is the minor capsid protein and is approximately 50 kDa in size [41]. There are 12 copies of L2 in the viral capsid, possibly associated with the 12 penta-associated L1 capsomers [74]. The exact function of L2 is unknown but it is thought to be involved in the initial stages of infection [74]. The N-terminal region of the protein either protrudes from the capsid or lies in an exposed site in the capsid as its external loop is able to interact with

cellular receptors [74]. L2 can also bind DNA through charged amino acids in the first 12 AAs of the N-terminal region. This binding may be involved in the introduction of the viral DNA into the capsid [41]. The production of both L1 and L2 is dependent on epithelial differentiation [75] and their transcription is initiated from the late promoter located in the E7 ORF [33].

1.8 Regulation of HPV Gene Expression

The expression of the HPV genes is complex involving regulation at a number of levels including at the level of transcription and at different RNA processing stages.

1.8.1 Regulation of HPV Transcription

Transcription of HPV occurs in only one direction as only one DNA strand is active [5]. HPV mRNAs are generally polycistronic [30], with many containing 3 or more messages on one mRNA [32]. Two major promoters, regulated by both cellular and viral transcription factors, direct the expression of the HPV genes [36]. The early promoter is generally located upstream of the E6 coding region and is regulated by the binding of transcription factors to elements that are located in the LCR [35]. For example, TFIID has been shown to bind to TATA boxes located approximately 30bp upstream from the early start site [32] and HPV-16 LCR contains 4 binding sites for the viral E2 protein [5]. Binding sites for a number of other transcription factors have also been identified including Oct-1, SP-1 and YY-1 [5, 9, 35]. Some of the enhancer elements identified in the LCR are thought to bind keratinocytic-specific transcription factors thereby being involved in the tissue tropism of the virus [9, 35]. The late promoter is located within the E7 coding region and is only active

in terminally differentiating keratinocytes [9]. The late proteins L1 and L2 are transcribed from this site as well as the E1^E4 protein. It is thought to be regulated by the binding of differentiation-specific cellular factors but these factors have not been identified [32].

1.8.2 RNA Processing

Messenger RNA (mRNA) is produced in the nucleus by RNA polymerase II (RNA pol II) before being transported to the cytoplasm to act as a template for protein synthesis [76]. The initial product of RNA pol II is pre-mRNA which must undergo several processing steps before it is considered functional RNA [76]. All eukaryotic mRNAs contain a 5' cap, most will contain a 3' poly(A) tail and many will also undergo splicing (Fig. 1.4). These processes have previously been termed as post-transcriptional events, but it has been reported that these processes are mainly carried out co-transcriptionally [77]. While some of these modifications may continue after the termination of transcription it has been shown that the processing events can be stimulated by RNA pol II [78]. Many of the trans-acting factors that are required for the processing events bind to the carboxy-terminal domain (CTD) of the RNA polymerase [79].

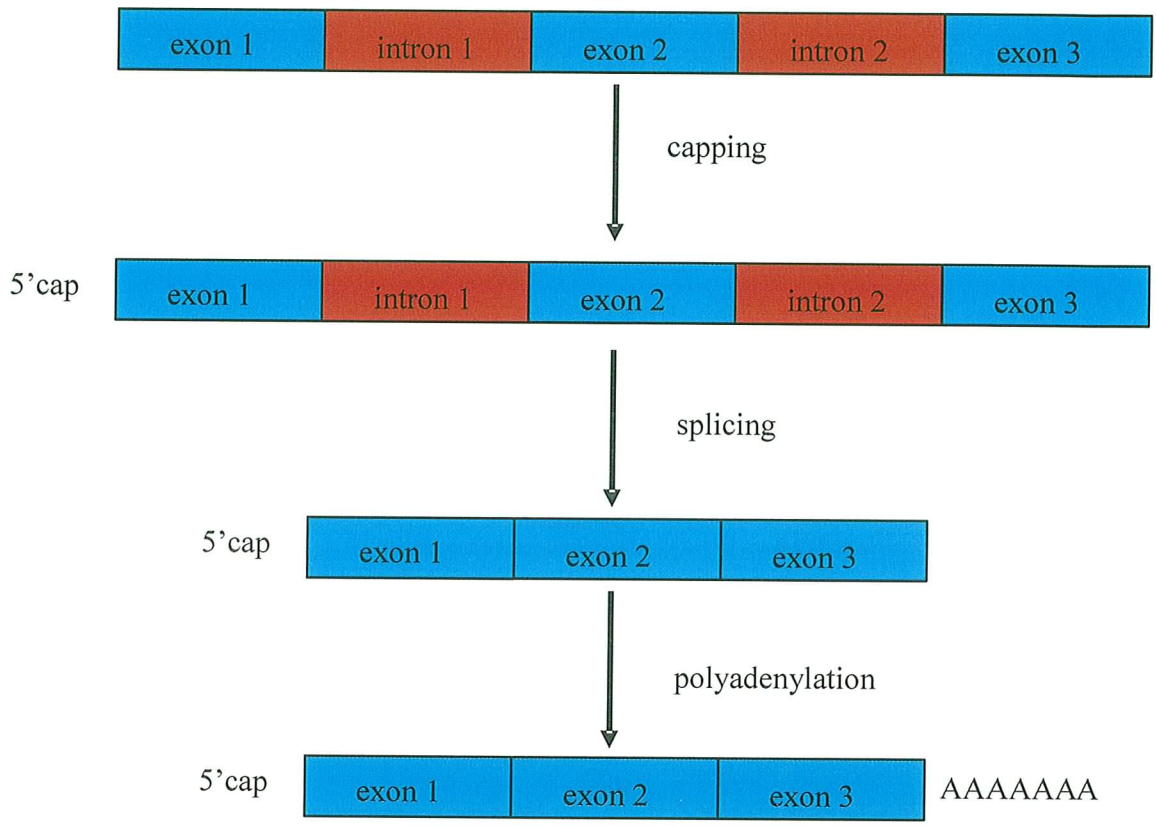


Figure 1.4 RNA Processing

The three RNA processing reactions, the addition of a 5' methyguanosine cap (5' cap), the removal of intronic sequences by splicing and the addition of a 3' poly(A) tail (AAAAAAA), required to produce mature messenger RNA that can be exported from the nucleus and used as a template for translation into protein.

(i) Capping

After 20-30 nucleotides of the nascent RNA have been produced RNA pol II pauses allowing the 3 step capping reaction to occur [80]. Firstly the 5' triphosphate end of the transcript is hydrolysed to form a diphosphate end by RNA triphosphotase. The diphosphate end is then capped with guanosine mono-phosphate (GMP) by RNA guanylyltransferase [81]. Guanine methyltransferase subsequently methylates the 5' guanine base via the nitrogen at position 7 of the purine ring, forming a 7-methylguanosine

cap. The cap is then bound by a nuclear cap binding complex (CBC) consisting of a 20kDa and an 80kDa protein [80]. The 5' cap complex has a number of functions - it protects the RNA from 5'→ 3' exonuclease attack, it can stimulate the splicing of the first intron, it can enhance translation and can facilitate transport from the nucleus to the cytoplasm [77, 78, 81, 82]. Capping occurs co-transcriptionally, as RNA pol II switches from initiation of transcription to processive elongation of the RNA, the CTD (a heptad amino acid sequence repeated 52 times in mammals [77]), is hyperphosphorylated at the second serine of the heptad [78]. The hpyerphosphorylated CTD has affinity for the two human capping enzymes, guanylyltransferase and guanine methyltransferase [78].

(ii) Splicing

Nearly all mammalian genes are interspersed by non-coding regions called introns [77]. Before translation into proteins can occur these non-coding sequences must be removed and the coding sequences (exons) placed side by side, this process is called splicing. A number of consensus sequences are required for splicing to occur, the 5' splice site or splice donor (SD), the branch point sequence, the polypyrimidine tract and the 3' splice site or splice acceptor (SA) [83]. The SD consists of the sequence AG | GURAGU where | marks the exon-intron boundary, the branch point consists of the sequence YNYURAC, which occurs about 100 nucleotides upstream form the SA, and the SA which consists of the sequence YAG | N [80]. The splicing reaction is a two step process carried out by the spliceosome. First the 2' OH group of the adenosine located in the branch point sequence promotes cleavage of the SD through a transesterification reaction resulting in the cleavage of the phosphodiester bond at the SD, this results in a free 5' exon and a lariat shaped intron. The 3' OH group at the end of the free exon attacks the phosphodiester bond at the SA resulting

in the joining of the exons and the release of the lariat [76]. The spliceosome is made up of small nuclear ribonucleoproteins (snRNPs) and non snRNP protein factors [84]. Spliceosome assembly starts with the recruitment of the U1 snRNP to the 5' splice site and the U2 snRNP to the branch point sequence [76]. Extrinsic factors, such as members of the SR protein family, help to stabilise the interactions of the snRNPs and the RNA (61). Once the U1 and U2 snRNPs have bound, this complex interacts with a pre-formed U5-U4-U6 tri-snRNP to form the immature spliceosome which undergoes arrangements that result in the formation of the fully formed spliceosome [85]. Splicing factors are recruited rapidly to the nascent RNA and many introns are removed co-transcriptionally and the process often continues post-transcriptionally [79] (Fig. 1.4).

(iii) Polyadenylation

The 3' ends of most eukaryotic genes are defined not by the termination of transcription but by a number of signal sequences that result in the cleavage and addition of up to 250 adenosine residues to the pre-mRNA [76]. The most conserved of these sequences is the hexanucleotide AAUAAA that is generally located between 10 to 30 nucleotides upstream of the site of polyadenylation [86, 87]. The polyadenylation (poly(A)) site itself is often immediately preceded by a CA dinucleotide and followed 10-20 nucleotides later by a U- or GU-rich downstream element (DSE) [80, 83]. The polyadenylation signal sequence and the DSE are both bound by multisubunit complexes. First cleavage and polyadenylation specificity factor (CPSF) binds to the AAUAAA sequence, then 3 proteins, cleavage stimulation factor (CstF), cleavage factor I (CFI) and cleavage factor II (CFII) bind to the CPSF-RNA complex [86]. The interaction between the CstF 64kDa subunit (CstF-64) and the DSE stabilises the complex [86]. Poly(A) polymerase (PAP) binds to this complex

before cleavage of the RNA by CFI and CFII occurs allowing rapid addition of the adenosine residues to occur once the RNA has been cleaved [77]. PAP adds the first 12 adenosine residues of the tail slowly, poly(A) binding protein II (PABII) binds to this short tail enhancing the activity of PAP allowing the rapid addition of adenosines until the tail reaches a length of approximately 200 residues [77, 86]. Some of the possible functions of the poly(A) tail include, facilitation of transport from the nucleus to the cytoplasm, stabilisation of the mRNA molecules in the cytoplasm and facilitating translation by permitting enhanced recognition of the mRNA by the ribosomal machinery [76, 83]. The termination of transcription requires the presence of a functional poly(A) site and usually occurs several hundred nucleotides downstream from the poly(A) site [77]. After cleavage occurs the RNA on the downstream side of the cleavage site remains attached to the RNA pol II but is unstable and its degradation is linked to the termination of transcription [79].

1.9 Alternative Processing

Before the human genome had been sequenced it had been suggested by examination of the proteome that humans had between 80,000 and 100,000 genes [83]. When the Human Genome Project was completed it was found that humans have approximately 20,000 – 25,000 genes [88]. Alternative processing of pre-mRNA has been used to explain the large difference between the number of proteins produced and the number of genes possessed. The two main mechanisms of alternative processing are alternative splicing and alternative polyadenylation (Fig. 1.5).

1.9.1 Alternative Splicing

It is now thought that between 60-80% of human genes are spliced in at least one exon [80], and is the most prevalent mechanism for generating proteomic diversity. Enhancer and silencer regions are short conserved sequences located in introns or exons that can stimulate or inhibit the use of weak splice sites [89]. Enhancers and silencers require the binding of specific regulatory proteins to act upon local splice sites. These regulatory proteins include serine/arginine rich (SR) proteins that generally bind enhancers and stimulate splicing, and heterogeneous nuclear ribonucleoproteins (hnRNPs) that generally bind to silencers inhibiting splicing [90]. There are a number of models of alternative splicing including constitutive, cassette exon, alternative SD, alternative SA, intron retention and mutually exclusive exons [89].

1.9.2 Alternative Polyadenylation

Over half of all human genes have been shown to contain more than one poly(A) site in the 3'untranslated region (UTR) [38]. An example of alternative polyadenylation occurs in the IgM gene, two forms of the IgM protein can exist - a membrane bound form and a secreted form [77]. The IgM pre-mRNA contains two poly(A) sites one of which is a weak site located within an intron. Early in an infection when the membrane form of IgM is produced this internal poly(A) site is repressed by an upstream SD and is removed by splicing. In plasma cells when the secreted form of IgM is produced there are higher levels of CstF-64 allowing recognition and use of the weak poly(A), this out-competes the splicing process and the shorted secreted form of the IgM protein is produced [77].

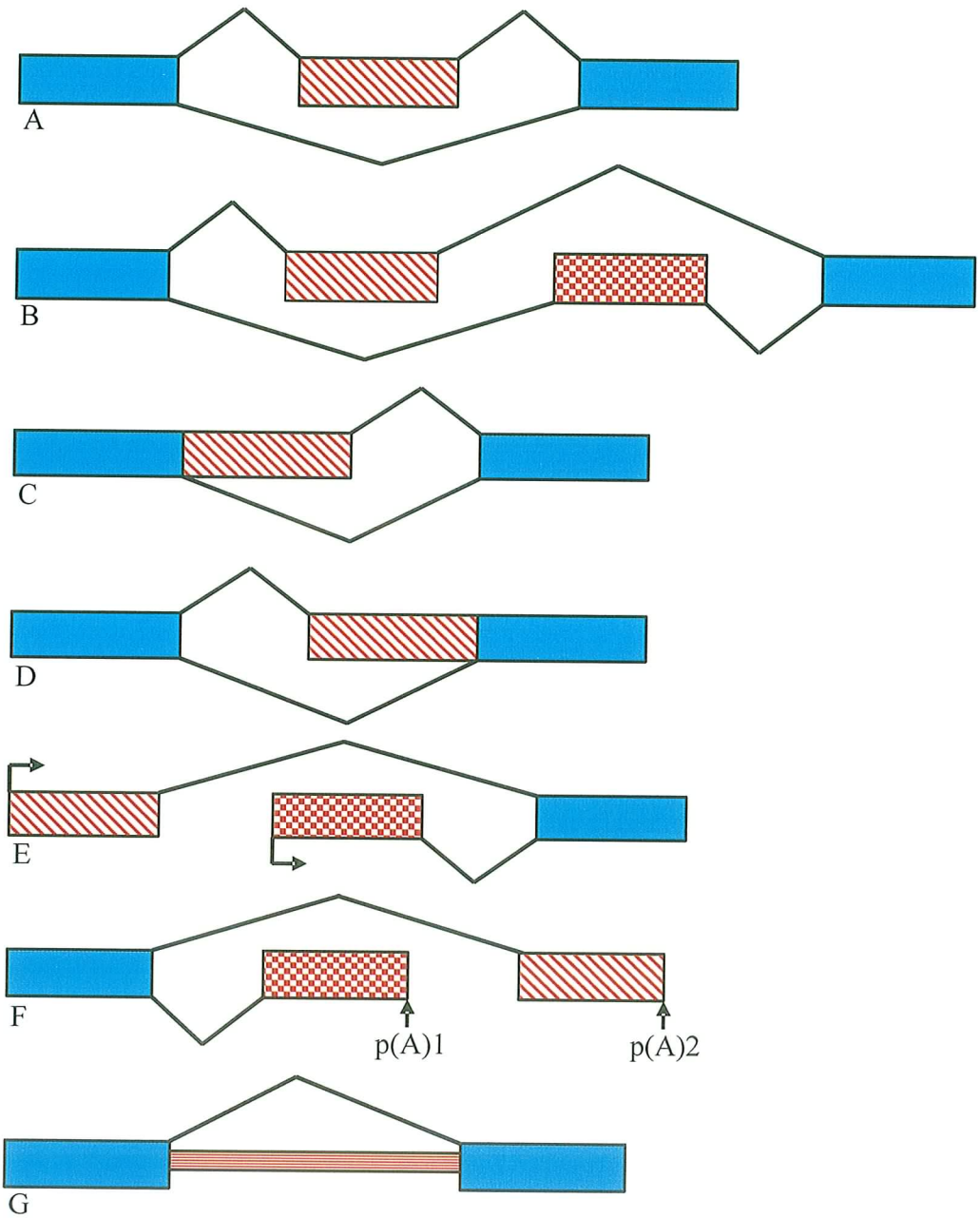


Figure 1.5 Patterns of alternative processing
The 7 main patterns of alternative splicing. Blue boxes indicate a constitutive exon that is present in all mature mRNAs, exons that may or may not be present are in red. A. Cassette exon, the middle exon can either be included or excluded from the mRNA. B. Mutually exclusive exons, when two or more adjacent cassette exons are spliced in such a way that only one cassette exon is included in the mRNA at any one time. C. Alternative 5' splice site, when a choice of 5' splice site is available it allows the possible lengthening of an exon. D. Alternative 3' splice site, when a choice of 3' splice site is available it allows the possible shortening of an exon. E. Alternative promoters, allows a switch in the first included exon. F. Alternative polyadenylation signals, allows a switch in the last included exon. G. A retained intron, an intronic sequence between the exons can be either excluded from or retained in the mRNA.

1.9.3 Alternative Processing in HPV

Like many viruses, HPV has a small genome, only 8kb long, and in order to code for all proteins that are required efficient use must be made of the small space. In eukaryotic cells most genes have their own promoter and the expression of each gene can be regulated through this promoter. Most HPV types contain only two promoters, one that controls the expression of early genes (the early promoter) and one that controls the expression of the late genes (the late promoter) [38] (Fig. 1.3). This means that HPV must use a different method for regulating the expression of the individual genes. Almost all papillomavirus types produce polycistronic mRNAs that contain a number of suboptimal splice sites allowing it to undergo alternative splicing [41]. Most papillomavirus types also contain more than one polyadenylation site, usually one at the end of the early coding region (pAE) and one at the end of the late coding region (pAL). The use of both alternative splicing and alternative polyadenylation allows HPV to differentially express all the genes necessary for the continuation of its life cycle.

In HPV-16 there are 11 splice sites, 10 of which are located in the early region. It has been shown that from the early promoter to the early polyadenylation signal 14 different mRNA species can be produced [38]. Overall 20 different mRNA species could be produced from the HPV-16 genome (Fig. 1.6)

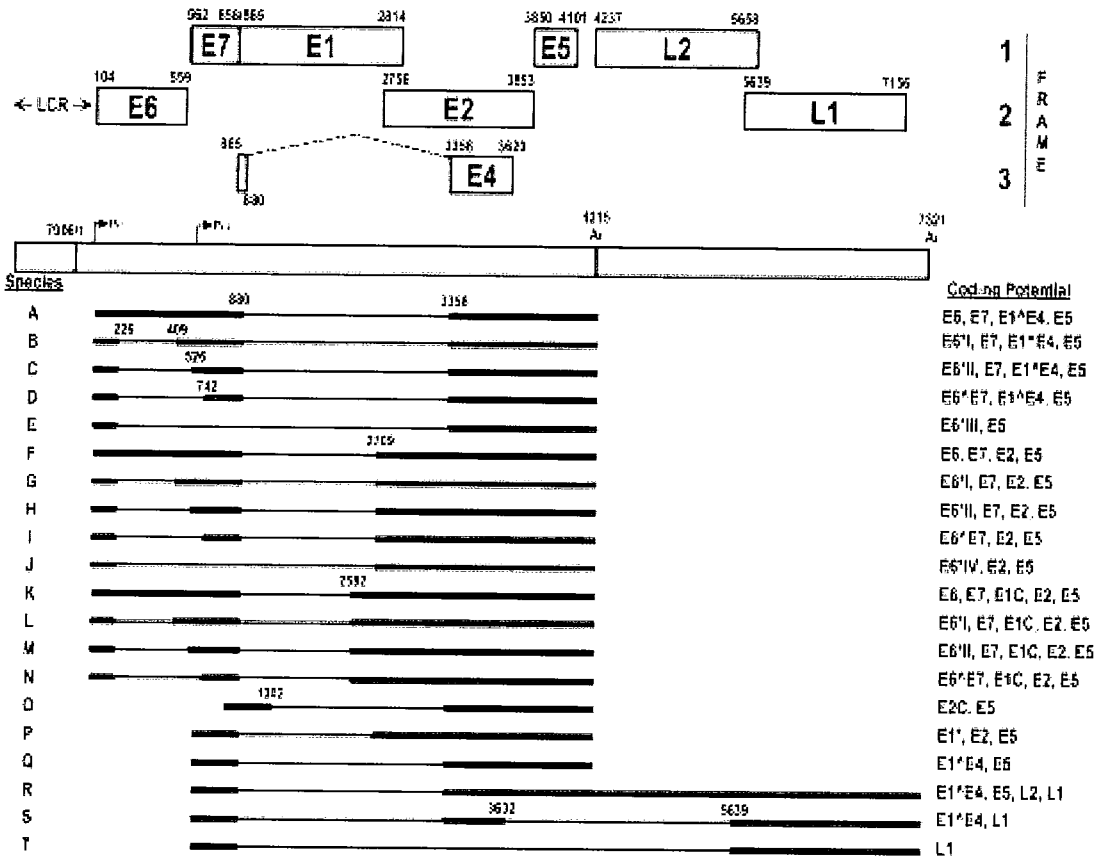


Figure 1.6 HPV-16 genome and transcription map

Schematic diagram of the HPV-16 genome (from [38]) showing the 3 separate reading frames the DNA can be read in and the 20 different RNA transcripts that could be produced by alternative splicing and polyadenylation. The coding potential of each mRNA is shown on the right.

Alternative splicing is especially important to the E6, E7, E4 and E2 proteins. E6 and E7 proteins are coded for on a bicistronic mRNA which contains one SD and three SAs. There is only a small gap between the E6 stop codon and E7 start codon, leading to low translation efficiency of E7 from full length E6E7 pre-mRNA. The splice sites in this region allow 4 mRNA species to be produced, E6*I, E6*II, E6^E7 and when no splicing occurs E6E7 [38]. The unspliced mRNA will produce E6 protein, while both E6*I and

E6*II disrupt the E6 coding region but lead to an increased efficiency in E7 protein production [41, 91]. The E6^E7 mRNA disrupts both the E6 and E7 coding regions possibly allowing an increase in translation efficiency for coding regions that occur downstream.

The E4 coding region lies entirely within the E2 coding region, although in a different frame, this means that splicing is essential for its production. The first 5 AAs of E4 are coded for by the start of the E1 coding region, the mRNA is then spliced from SD880 to SA3358 where the E4 coding region begins. E2 can be formed by using the SA2709 instead. Use of the SA3358 also has a role in the blocking of the production of the late transcripts. When SA3358 is used it encourages transcripts to be polyadenylated at the early poly(A) signal, preventing late gene transcription [92]. Therefore the use of SA3358 must be regulated in order to produce E4 and block late transcripts but it also must be flexible to allow production of E2 [92].

Splicing and polyadenylation is also important within the late region of HPV-16. As the 3' end of L2 and the 5' end of L1 overlap, alternative splicing is required for L1 protein production. The primary transcripts that are produced from the late promoter share a common SD (SD880) but will use either a proximal or distal SA. Use of the proximal SA (SA 3358) will result in polyadenylation at the early poly(A) signal and the production of E4, whereas the use of the distal SA (SA5639) located in the late region, splices out the pAE and results in the production in L1 [38]. During the non-productive phase of the HPV-16 lifecycle the mRNAs that are produced are polyadenylated almost exclusively at the early poly(A) signal thereby helping to regulate early gene expression. Differentiation of the epithelium leads to approximately 50% read-through of this site in HPV-31 allowing

the production of L2 [93]. While there are some co-transcriptional events that help to regulate the expression of the late genes, post-transcriptional events are also involved. This is indicated by the detection of L1 and L2 mRNA in undifferentiated epithelial cells, but these proteins can only be detected in terminally differentiated cells [38]. In this way, through the use of 11 splice sites and 2 polyadenylation signals, HPV-16 can regulate the expression of all of its proteins during its life-cycle (Fig. 1.7).

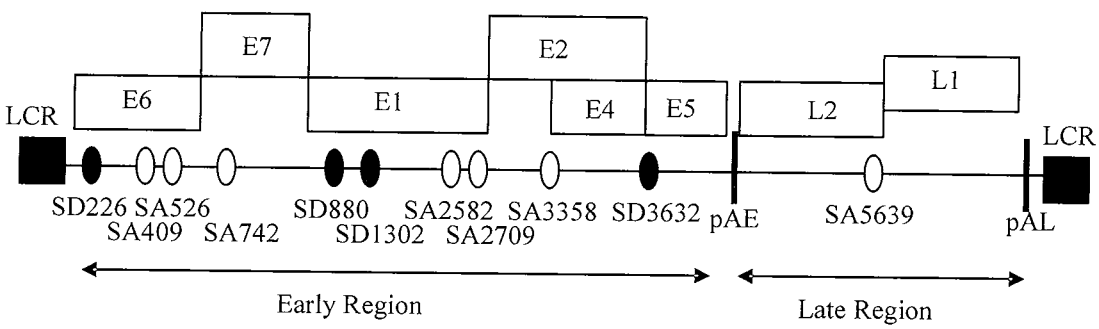


Figure 1.7 HPV-16 genome showing the 11 alternative splice sites and 2 polyadenylation sites

Schematic diagram showing the locations of the splice donors (SD) (black circles) and splice acceptors (SA) (white circles) within the HPV-16 genome. Also shown are the early (pAE) and late (pAL) polyadenylation signals and long control region (LCR).

With the regulation of gene expression playing such an important role in the viral life cycle, the knowledge of how it occurs is essential to further our understanding of this virus. This study aimed to examine the expression of proteins involved in the regulation of splicing in patients with cervical abnormalities and also to study the effect of some of these proteins on HPV late gene expression *in vitro*.

2. Chapter 2 – Audit of laser-loop excised cervical biopsies in the National Maternity Hospital, 1989 - 1993

2.1 Introduction

2.1.1 The Cervix

The cervix is the cylindrical lower part of the uterus and at its lower end it protrudes into the vagina [94]. There is a constricted opening at each end of the cervix, called the os. The internal os is located at the upper end of the cervix and is the opening of the cervix inside the uterine cavity [95] (Fig. 2.1). The external os is located at the lower end of the cervix and opens into the vagina [96] (Fig. 2.1). The cervix contains only a few muscle fibres, it is mainly made up of dense connective tissue (nearly 85%) [94]. The endocervical canal is the passageway between the external os and the uterine cavity (Fig. 2.1). It varies in width and length. The lining of the endocervical canal differs completely from the rest of the uterine endometrium as it contains large branched glands [95]. The endocervix is lined with a single layer of mucous-secreting simple columnar epithelium. The columnar epithelium is gradually replaced by metaplastic squamous epithelium, the area where this occurs is called the transformation zone (TZ) or squamo-columner junction [96]. The TZ is the area of the cervical epithelium most likely to be affected by disease [95]. The ectocervix is the portion of the cervix that projects into the vagina and has a convex surface that has an epithelial lining [97]. The epithelium of the ectocervix is non-keratinised stratified squamous

epithelium which is continuous below with the squamous epithelium lining of the vagina [98].

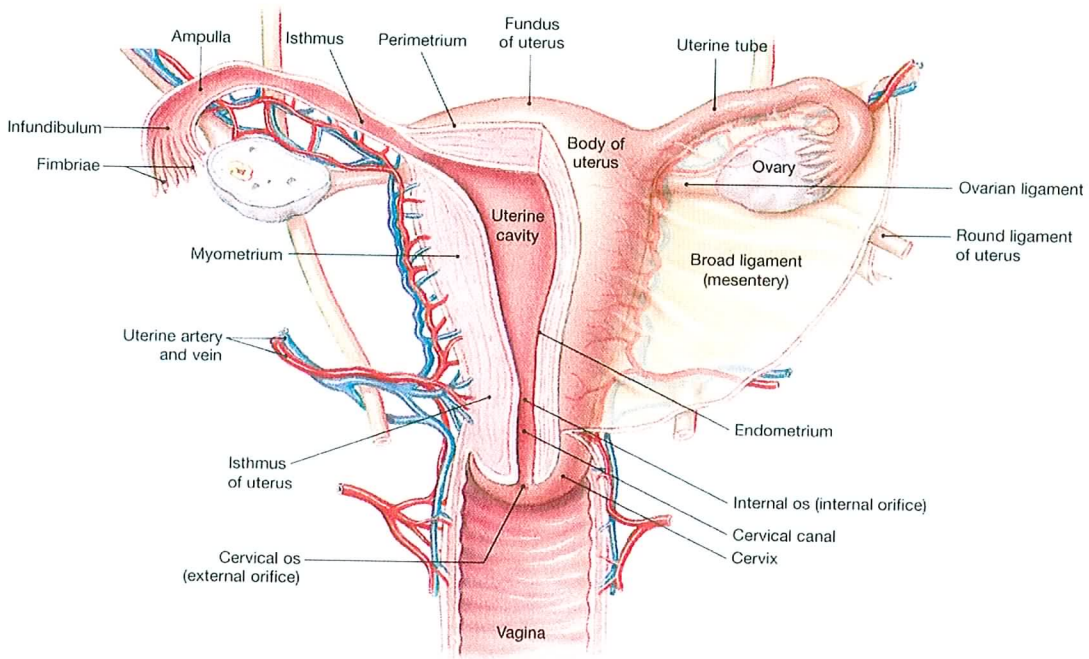


Figure 2.1 The uterus

Diagram of the uterus (from [99]) indicating location of the cervix, the internal and external os and the cervical canal.

2.1.2 Cervical Abnormalities due to HPV Infection

The most common result of HPV infection of the cervix is condyloma acuminata (genital warts) [100]. These are papillary lesions that are generally caused by non-oncogenic forms of HPV (HPV-6, HPV-11). These lesions contain characteristic cellular changes including multinucleation and koilocytic atypia. The cellular changes are generally confined to the superficial layers of the epithelium [98].

Infection with high-risk HPV types can lead to the development of abnormal neoplastic cervical cells. These cells lose their regular stratified pattern, have a high nuclear to cytoplasmic ratio, show a variation in shape, and have an increased mitotic activity [95]. Up until the 1960s these non-invasive intraepithelial neoplasias of the cervix had a 4 tiered terminology consisting of mild dysplasia, moderate dysplasia, severe dysplasia and carcinoma *in situ* (CIS), and it was believed that only CIS was a pre-cursor to cervical cancer [101]. By the 1970s it became evident that all dysplasias had the ability to progress to cancer and a new nomenclature for cervical lesions was introduced [101]. Three grades of cervical intraepithelial neoplasia (CIN) were categorised with CIN 1 the lowest and CIN 3 the highest grade of neoplasia (Fig. 2.2). CIN is recognised by disturbances of cellular maturation and stratification as well as by the presence of cytological atypia [98]. In CIN 1 lesions the maturation of the cytoplasm occurs in the superficial two thirds of the epithelium. Cytoplasmic maturation in CIN 2 occurs in the upper third of the epithelium, while in CIN 3, maturation of the cytoplasm is minimal or completely absent [98]. If cytoplasmic maturation occurs at all in CIN 3 it will occur in the upper third of the epithelium.

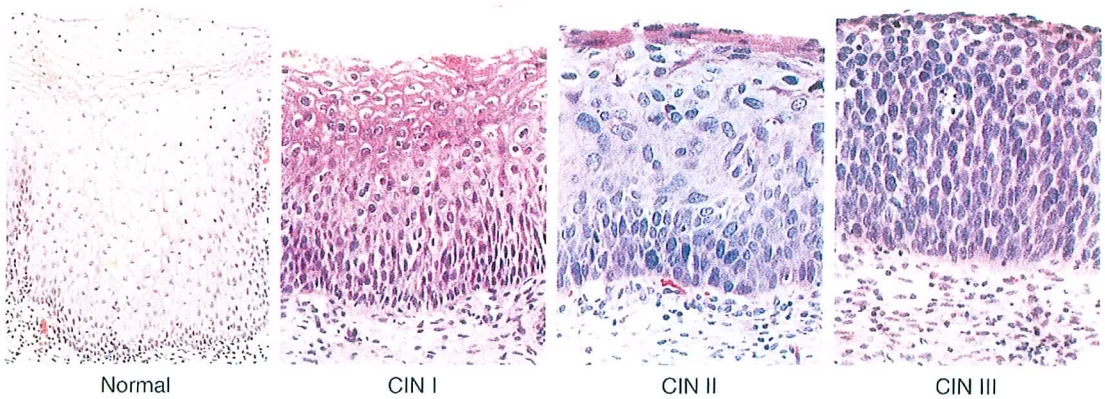


Figure 2.2 CIN grades 1 – 3 with normal squamous epithelium as a comparison

Normal – normal squamous epithelium, CIN 1 – cervical intraepithelial neoplasia grade 1, with koilocytotic atypia, CIN 2 – with progressive atypia in all layers, CIN 3 – with diffuse atypia in all layers and loss of maturation (from [102]).

The 3 grades of CIN lesions also have different likelihoods of both regression to a lower lesion and progression to a higher lesion. CIN 1 has a 60% chance of regression, CIN 2, 40% and CIN 3, 33% [101]. CIN 1 has a 30% risk of becoming a persistent lesion, a 10% chance of progression to CIN 3 and 1% chance of progression to invasion [101]. CIN 2 has a 40% chance of persistence, 20% chance of progression to CIN 3 and a 5% chance of progression to invasion. CIN 3 has a 56% chance of persistence and less than 12% chance of progression to invasion [98, 101, 103]. It is now recognised that high-grade CIN lesions (CIN 2 or CIN 3) can often develop quickly after initial infection with high-risk HPV types, and the these lesions may not be preceded by a CIN 1 lesion [101].

Generally CIN lesions are asymptomatic and are not readily visible to the naked eye, diagnosis of a lesion is generally made using a cervical smear and microscopic examination (PAP smear) [104, 105]. If the lesion is not detected and cervical cancer develops it can

spread locally into the cervical stroma, the paracervix and parametrial tissues, the body of the uterus, the vagina and late in the course of disease to the bladder and the rectum [98].

2.1.3 LLETZ as a Treatment for CIN

Large loop excision of the transformation zone (LLETZ, also known as loop electrosurgical excision procedure, LEEP) is a method of removing the transformation zone of the cervix as a treatment for CIN [106-108]. A loop diathermy method for the diagnosis and treatment of CIN was first described by Cartier [107-109]. In the early 1980s this method was modified slightly and introduced to the UK by Prendiville and Cullimore [107, 108]. LLETZ has a major advantage over destructive techniques such as laser or cold coagulation, as the tissue that is removed by the LLETZ technique can be examined histologically allowing the diagnosis of micro-invasive or invasive lesions [109, 110]. LLETZ also has advantages over other excisional techniques such as knife cone biopsy, as it does not require a general anaesthetic and it is associated with lower levels of complications [110]. For these reasons LLETZ has become one of the most frequently used treatments for CIN. Success rates for the use of LLETZ as a treatment range from 60% - 95% [106].

2.1.4 Detection of HPV

There are a number of different techniques that can be used to detect HPV in samples. Many are based on the direct detection of HPV nucleic acid using a hybridisation method. Southern and northern blots can be used to detect HPV DNA or RNA in a solid phase [111]. These methods are time-consuming and require large amounts of highly pure and

ideally full-size nucleic acid molecules, so generally only fresh or fresh-frozen samples can be used [4, 111-113]. *In situ* hybridisation can be used to detect both HPV DNA and RNA in tissue samples using labelled probes [114]. As the signal is localised to the HPV nucleic acid, the morphology of the area can be examined and also it can be established if the virus has integrated into the host genome [114, 115]. The signal can be amplified to increase the sensitivity of this method, although in order to detect specific HPV types specific probes and multiple reactions would be required [113]. The only Federal and Drug Administration approved test for the detection of HPV is the Hybrid Capture 2 (HCII) test [115]. This is a solution-based hybridisation method that uses two RNA probe cocktails, one that contains probes for 5 low-risk HPV types, and one that contains probes for 13 high-risk types, to detect the presence of HPV [111, 113]. The HPV DNA-RNA hybrid is immobilised on the wall of the microtiter plate before a luminescent product is generated through a series of reactions. As the intensity of the signal is proportional to the amount of target that was present in the sample this method is a semi-quantitative method for HPV detection [116]. This method is also useful as samples taken for cervical smears that are contained in PreserveCyt solution can be used, but a specific HPV type cannot be identified in a sample using the HCII method [114, 115].

2.1.5 PCR Detection of HPV

Since its invention in 1983 by Kary Mullis [117], PCR has become a very common tool in molecular biology and is the most widely used target amplification method for the detection of HPV DNA [113]. As PCR amplifies DNA, it allows testing of samples with low levels of tissue or cells, poor quality DNA and samples with low levels of virus [115]. From the

earliest days of PCR DNA from formalin-fixed, paraffin-embedded samples were used to assess the presence of HPV DNA [118]. PCR reactions with small amplification targets (approx. 200bp) work best with these samples [112]. For the detection of HPV DNA there are two main types of PCR primers, type-specific and consensus [115]. Type-specific PCR uses a set of primers that is specific for only one distinct type of HPV. In order to detect multiple types in one sample or to asses the presence of different HPV types in a group of samples, multiple type-specific PCR reactions have to be carried out. Another method to detect the presence of HPV DNA is to use consensus or general primers that can detect a broad-spectrum of HPV types with one primer set. These primers will target a conserved region of the HPV genome, generally L1 but sometimes E1 [113]. For HPV detection there are three types of general primers. The first type of primer set is composed of one forward and one reverse primer aimed at a conserved region. These primers will only be completely complementary to one or two types so a low annealing temperature is used to compensate. An example of this type is the GP5+/GP6+ primers [113]. The second type is a set of forward and reverse primers that contain one or more degenerate base pairs that compensate for the sequence variance between different HPV types. The MY09/MY11 primer pair is an example of this type of general primer [113]. The third type of general primer pairs involves the use of a combination of a number of distinct forward and reverse primers, aimed at the same position of the viral genome. The PGMY and SPF_{10} primers are an example of these primers [113]. The GP5+/GP6+ and the MY09/MY11 primer pairs are the most commonly used primers for general HPV detection [119, 120]. These primers are also useful as they can be used in a nested reaction, a PCR in which 2 different primer sets, one set that amplifies a region located between the first primer pair, are used to increase the

sensitivity of the PCR. The GP5+/GP6+ primers amplify a 150bp region that is located within the 450bp region amplified by the MY09/MY11 primers. Nested PCR with these primers increases the detection rate of HPV and a recent study reported that 38.8% of samples that were negative for HPV DNA by MY09/MY11 PCR alone were positive when the nested PCR was used [121].

2.1.6 Aims:

- to audit LLETZ biopsies over a 5 year period
- to establish HPV-16 prevalence in LLETZ excised cervical samples

Objectives

- collection of a cohort of tissue samples with varying grades of CIN and SCC
- identification of samples containing HPV DNA
- detection of HPV-16 prevalence in a cohort of samples
- evaluate the association between HPV-16 and neoplasia in these samples

2.2 Methods

2.2.1 Sample Collection

Ethical approval was received from the National Maternity Hospital (NMH) to collect patient information and tissue blocks for all LLETZ samples sent to the Pathology laboratory between 1989 and 1993. Ethical approval was also granted for the collection of information and blocks for 30 Wertheim's hysterectomy patients (invasive carcinoma of the cervix) between 1989 and 1995. Laboratory day books (a log of all samples received) were examined and the laboratory number for all LLETZ samples was recorded. With the laboratory number the original request forms were identified and any clinically relevant information, such as previous history, was recorded. The number of blocks that the sample had been divided into was also recorded (this ranged from 1 to 13, the majority had been divided into 6 blocks) so the correct number of blocks could be found. The tissue blocks for all 961 LLETZ samples and the 30 Wertheim's hysterectomy samples were sourced from the archival storage area.

2.2.2 Identification of Presence of Lesion

As all tissue blocks had been previously sectioned by the NMH laboratory, blocks needed to be re-sectioned to determine if the diagnosed lesion was still present. Also as there was generally more than one block per sample the exact block or blocks that contained the lesion needed to be identified. 20 cases were randomly selected for each grade of lesion (no CIN present, CIN 1, CIN 2, CIN 3 and SCC), totalling 100 cases, and were sectioned and stained with a haematoxylin and eosin stain. 5μm tissue sections were cut, floated onto a slide and melted for 4 hours at 65°C. The sections were then de-waxed for 6 minutes each

in xylene (BDH) (two baths), absolute alcohol (Merck) (two baths) and spirit (appendix I), before being placed in water. Sections were placed in Mayer's haematoxylin (appendix I) for 30 seconds to 1 minute. The sections were then washed in warm water for 1 minute before being placed in 1% eosin (appendix I) for 3 minutes. Sections were washed in running water for 1 to 2 minutes before being de-hydrated by being placed in spirit, absolute alcohol (two baths) and xylene (two baths) for 6 minutes each. Sections were mounted with DPX (BDH) and left to dry on a flat surface. On re-evaluation of morphology, 50 cases were selected for further evaluation of HPV status and RNA binding protein expression (Chapter 3).

2.2.3 DNA Extraction

DNA was extracted from all samples using the Qiagen DNA Mini kit tissue protocol. A change to the beginning of the protocol for paraffin-embedded samples suggested by the manufacturer was included. Two 5μm sections were cut from each block and were placed into a sterile eppendorf tube. To prevent contamination between blocks a different blade was used for each block, the microtome was cleaned with 70% alcohol (appendix I) between each block and gloves were worn at all times. To remove the paraffin wax 1200μl of xylene (BDH) was added to the tube and it was vortexed vigorously. The sample was centrifuged at full speed for 5 minutes at room temperature. The supernatant was removed and 1200μl 100% ethanol (Merck) was added to the tube. The contents of the tube were mixed gently by vortexing before centrifuging at full speed for 5 minutes at room temperature. The supernatant was removed, again being careful not to disturb the pellet. The ethanol wash was repeated once. After the supernatant was removed from the second

wash the open tube was incubated at 37°C for 10 to 15 minutes until all the ethanol had evaporated. The pellet was re-suspended in 180µl Buffer ATL (Qiagen) and 20µl Proteinase K (Qiagen) was added. The tube was incubated at 56°C for 5 days, additional Proteinase K (20µl) was added twice each day and the tube was vortexed after each addition. After the 5 day incubation the tube was centrifuged briefly to remove any drops from inside the lid before 200µl Buffer AL (Qiagen) was added. The tube was pulse-vortexed for 15 seconds to mix the contents and was then incubated at 70°C for 10 minutes, again the tube was briefly centrifuged to remove any drops from the lid. 200µl 100% ethanol (Merck) was added and was mixed by pulse vortexing for 15 seconds, the tube was centrifuged briefly. The entire sample in the tube was applied to the QIAamp spin column (Qiagen) and was centrifuged at 8000rpm for 1 minute. The spin column was placed in a new collection tube and the tube containing the filtrate was discarded. 500µl Buffer AW1 (Qiagen) was added to the spin column and it was centrifuged at 8000rpm for 1 minute. The spin column was placed in a new collection tube and the tube containing the filtrate was discarded. 500µl Buffer AW2 (Qiagen) was added to the spin column and it was centrifuged at 14000rpm for 3 minutes. The spin column was placed in a new collection tube and the tube containing the filtrate was discarded. The spin column was centrifuged at 14000rpm for 1 minute to prevent any carryover of Buffer AW2. The spin column was placed into a sterile 1.5ml eppendorf tube and the collection tube containing the filtrate was discarded. 50µl Buffer AE (Qiagen) was added to the spin column and was incubated at room temperature for 5 minutes. The column was centrifuged at 8000rpm for 1 minute. A further 50µl Buffer AE was added to the column, it was incubated at room temperature and was centrifuged at 8000rpm for 1 minute. The sample was stored at -20°C until required.

2.2.4 PCR Amplification of Cervical Biopsy DNA

A β-globin PCR was carried out using the PC03/PC04 primer pair to establish quality of DNA extracted, while HPV DNA was detected using the MY09/MY11 and GP5+/GP6+ primer pairs. Type specific HPV-16 DNA was detected using a nested PCR with the HPV-16O and HPV-16I primer pairs. Primer sequences, magnesium concentrations and size of amplified product are detailed in Table 2.1.

Primer	Primer Sequence 5'→3'	$MgCl_2$ Concentration	Primer Concentration	Product Size
PC03	acacaactgtgttcactagc	3mmol/l	0.5μmol/l	110bp
PC04	caacttcatccacgttcacc			
MY09	cgtccaaaaggaaactgatc	2mmol/l	5μmol/l	450bp
MY11	gcacagggacataacaatgg			
GP5+	tttgttactgtggtagatactac	2mmol/l	50μmol/l	150bp
GP6+	gaaaaataaactgtaaatcatat			
HPV-16O(A)	caacaaaccgttgtgtgattt	2mmol/l	0.25μmol/l	192bp
HPV-16O(B)	tccatgcatgattacagctggg			
HPV-16I(A)	tcaaaagccactgtgtcctg	2.5mmol/l	2.5μmol/l	119bp
HPV-16I(B)	cgtgttcttgatgatctgca			

Table 2.1 Primers used for DNA quality assessment and HPV detection

Primer sequences, magnesium concentrations, primer concentration and product size for primers used in this chapter. HPV-16O – outer primers for nested PCR, HPV-16I – inner primers for nested PCR.

All PCR reactions were carried out in a final volume of 25μl containing 1x reaction buffer (200mmol/l Tris-HCl (pH8.4), 500mmol/l KCl) (Invitrogen), 2mmol/l dNTPs (Invitrogen), 1U Platinum Taq polymerase (Invitrogen) and 5μl template DNA. $MgCl_2$ (Invitrogen) and primer concentrations are detailed in Table 2.1. The PCR protocol was as follows: an initial 5 minute denaturation step at 95°C, followed by a cycle of 30 seconds at 95°C, 1 minute at 55°C for β-globin primers, 56°C for MY09/MY11 primers or 42°C for GP5+/GP6+ primers, 1 minute at 72°C and a final elongation step of 5 minutes at 72°C. 40 cycles were used for

all PCRs. The nested HPV-16 type-specific PCR was carried out as above but only 3µl template DNA was used. The nested PCR protocol was as follows: an initial 5 minute denaturation step at 95°C, followed by a cycle of 30 seconds at 95°C, 1 minute at 59°C, 1 minute at 72°C and a final elongation step of 5 minutes at 72°C, 20 cycles were used for the 1st round and 30 for the second round of the nested HPV-16 PCR. All PCR products were analysed on a 2% agarose gel (appendix I) stained with ethidium bromide (0.5µg/ml ethidium bromide (Fluka) dissolved in gel). Gels were run at 120V for 1 hour.

2.2.5 Controls for PCR

Positive controls were included for all PCRs. For β-globin and MY09/MY11 – GP5+/GP6+ nested PCR, DNA extracted from HeLa cells (a HPV-18 positive cervical cancer cell line) was used as a positive control. For nested HPV-16 PCR, DNA extracted from SiHa cells (a HPV-16 positive cervical cancer cell line) was used as a positive control. A negative control for the PCR reaction, containing no DNA, was also always included.

2.3 Results

2.3.1 Audit of LLETZ Samples Received By NMH Pathology Laboratory, 1989-1993

All LLETZ samples that were received in the pathology laboratory of the National Maternity Hospital during the 5 years covered by ethical approval were recorded. Details of diagnosis and any other comments made on the original request form or at diagnosis were documented. A total of 961 LLETZ samples were received in the laboratory (Table 2.2). The samples were graded by the National Maternity Hospital pathologist as, no cervical intraepithelial neoplasia (CIN) present in the sample, CIN 1, CIN 2, CIN 3 or where there was multiple CIN grades present in the sample CIN 1-2, CIN 2-3 and CIN 1-3.

Of the 961 LLETZ samples, 210 (21.9%) samples were diagnosed as having no CIN present in the sample, 164 (17.1%) were diagnosed with CIN 1, 369 (38.4%) with CIN 2 (CIN 2 and CIN1-2, Table 2.2) and 205 (21.3%) with CIN 3 (CIN 3, CIN 2-3 and CIN 1-3, Table 2.2). The total number of patients diagnosed with CIN during this time frame was 742, which is 77.2% of patients who were treated by LLETZ between 1989 and 1993. Glandular intraepithelial neoplasia (GIN) was rarely diagnosed in these samples with only 5 cases over the 5 year period. Only 1 case of squamous cell carcinoma (SCC) was found in the LLETZ samples within the timeframe of the ethical approval. This required a change in the ethical approval to include 30 Wertheim's hysterectomy samples between 1989 and 1995 to use as SCC samples.

	1989	1990	1991	1992	1993	Total
No CIN	15	84	25	45	41	210
CIN 1	17	73	26	22	26	164
CIN 2	12	87	46	45	44	234
CIN 3	5	13	14	18	11	61
CIN 1-2	12	68	27	17	11	135
CIN 2-3	7	54	13	16	17	107
CIN 1-3	5	16	7	4	5	37
CIN grade not specified	0	0	1	0	3	4
GIN (* = with CIN 3)	0	1	2	1*	1*	5
Adenocarcinoma	0	0	0	0	1	1
SCC	0	0	0	1	0	1
No cervical tissue present	0	0	0	1	0	1
Full thickness dysplasia	0	0	1	0	0	1
Total number of samples	73	396	162	170	160	961

Table 2.2 LLETZ Diagnosis 1989 – 1993

CIN – cervical intraepithelial neoplasia, GIN – glandular intraepithelial neoplasia, SCC – squamous cell carcinoma.

When the data on the age of patients was examined it was found that of the 948 diagnosed with no CIN, CIN 1, CIN 2, CIN 3 and SCC (total 978), 956 had age details given (Table 2.3). The mean age for the no CIN and CIN 3 group was the same at 34.2 years. The CIN 1 and CIN 2 groups were slightly lower at 31.4 years and 32.2 years respectively. The SCC group had the highest mean age at 48.8 years. Overall the patients undergoing LLETZ treatment (excluding Wertheim's hysterectomy patients) ranged in age from 17 to 67 years with a mean patient age of 33 years.

	No. of Cases (with age data)	Mean Age	Range (years)	Min Age	Max Age
No CIN	210 (208)	34.2	48	19	67
CIN 1	164 (163)	31.4	40	18	58
CIN 2	369 (359)	32.2	49	17	66
CIN 3	205 (196)	34.2	41	20	61
SCC	30 (30)	48.8	39	28	67

Table 2.3 Age data for groups no CIN, CIN1, CIN 2, CIN 3 and SCC

CIN = cervical intraepithelial neoplasia, SCC = squamous cell carcinoma.

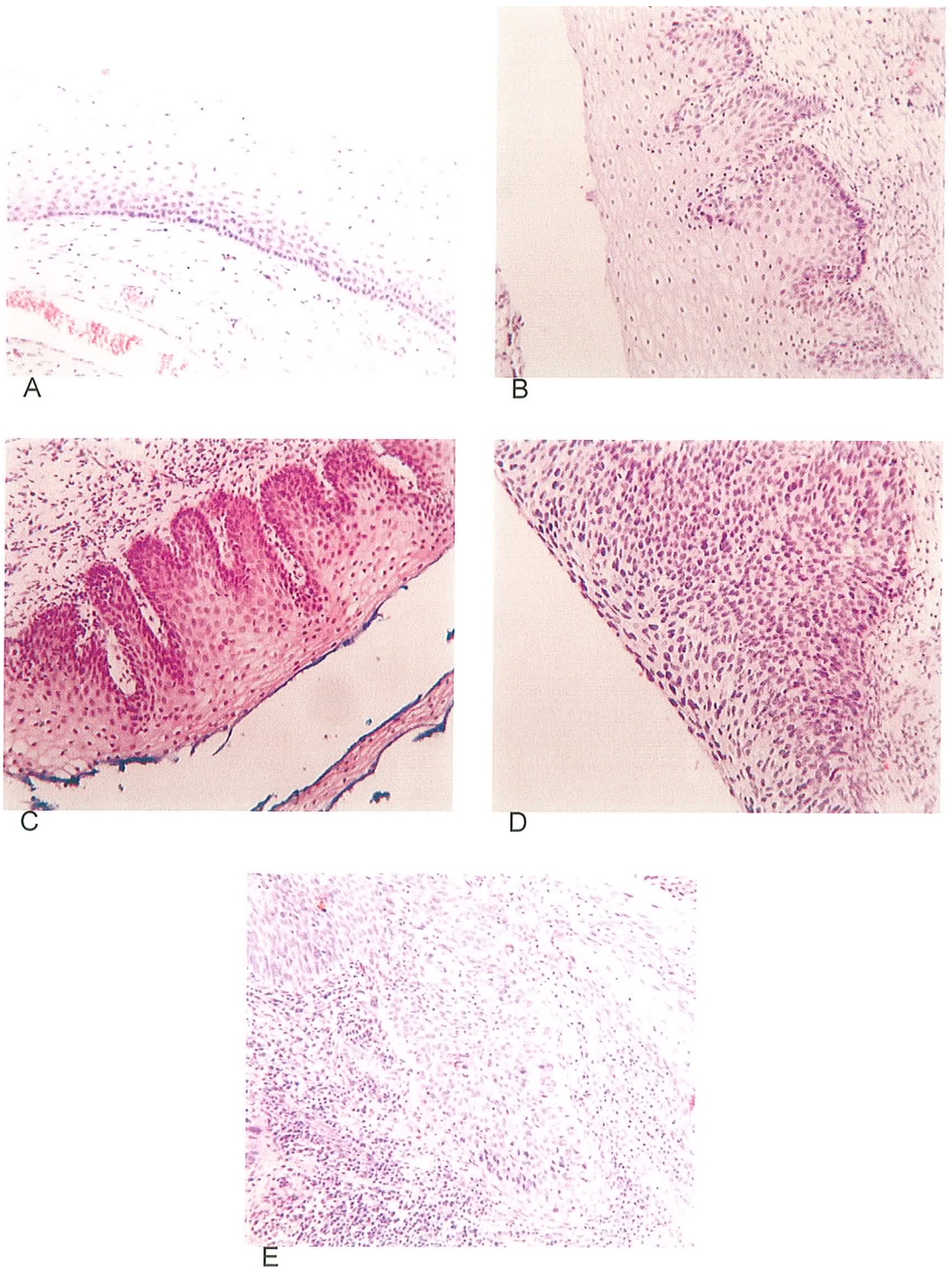


Figure 2.3 Cervical epithelium with varying grades of pre-malignant and malignant neoplastic lesions

Cervical epithelium stained with haematoxylin and eosin. A - no CIN present, B - CIN 1, C - CIN 2, D - CIN 3, E - SCC. CIN = cervical intraepithelial neoplasia, SCC = squamous cell carcinoma. Magnification 400x.

2.3.2 HPV DNA Detection

From the 100 samples examined using the H&E stain (Fig. 2.3), 50 samples (10 for each grade of lesion) were selected for the detection of HPV DNA. The 50 samples were given an identifier in the form of a patient number. Patient numbers and the grade of lesion present in the sample as well as patient age and any previous history, are shown in Table 2.4.

Cases	Cytomorphology	Average Age (Range)	Previous History	
			Abnormal Smear	CIN on Biopsy
1 – 10	No CIN	38.5 (32 – 48)	10	5
11 – 20	CIN 1	28.5 (21 – 46)	7	5
21 – 30	CIN 2	27.3 (20 – 38)	9	3
31 – 40	CIN 3	35.7 (24 – 45)	9	3
41 – 50	SCC	49.1 (28 – 67)	-	-

Table 2.4 Summary of age of patients and cytomorphology results

DNA was extracted from the 50 blocks and to confirm the presence of DNA a PCR was carried out on all extracted DNA using the PC03/PC04 β-globin primers. These primers amplify a 110bp region of the β-globin gene. Amplifiable DNA was extracted from all samples (Fig. 2.4 and Appendix II).

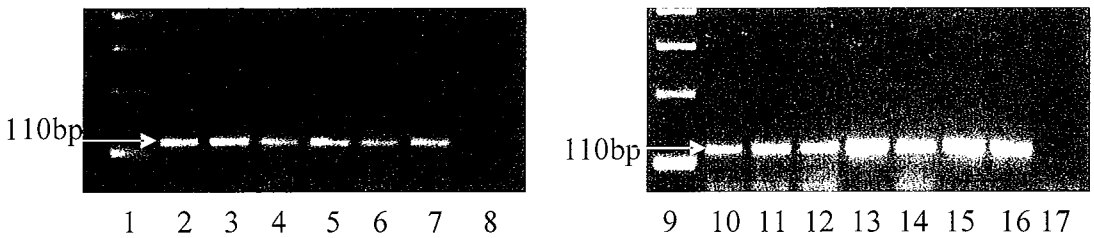


Figure 2.4 PCR amplification of the β-globin gene using the PC03/PC04 primers

Lane 1 100bp DNA ladder, lane 2 positive control, lanes 3 – 7 patients 1 – 5, lane 8 negative control, lane 9 100bp DNA ladder, lane 10 positive control, lanes 11 – 16 patients 6 – 11, lane 17 negative control. PCR products run on a 1% agarose gel.

To detect HPV DNA two primer sets were used to amplify regions within the L1 gene. As this is the most conserved region of the HPV genome it allows detection of multiple HPV types with one set of primers. As the GP5+/GP6+ primers lie within the area amplified by the MY09/MY11, these two sets of primers can be used in a nested PCR, amplifying a 150pb region of the HPV genome. The nested PCR was run on all 50 samples (Fig. 2.5 and Appendix III) and 49 out of the 50 samples (98%) indicated the presence of HPV DNA.

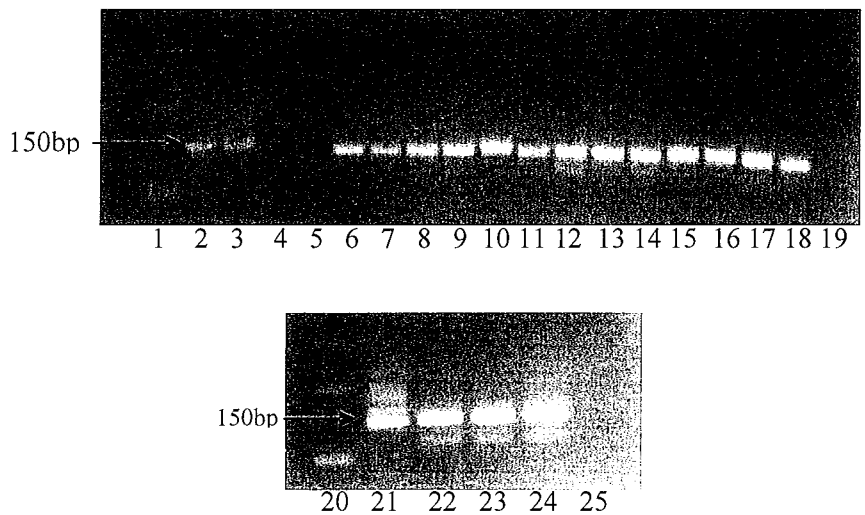


Figure 2.5 Nested PCR amplification of the L1 HPV gene using the MY09/MY11 primers in combination with the GP5+/GP6+ primers

Lane 1 100bp DNA ladder, lane 2 positive control, lane 3 patient 9, lanes 4 and 5 – skipped, lanes 6 – 18 patients 10 – 22, lane 19 negative control, lane 20 100bp DNA ladder, lane 21 positive control, lanes 22 – 24 patients 48 – 50, lane 25 negative control. PCR products run on a 1% agarose gel.

To investigate how many of the HPV positive samples were HPV-16 positive, a PCR was carried out on a selection of samples using HPV-16 type-specific primers. As a nested PCR had to be run in order to detect HPV DNA in all samples, primers were designed to amplify a region outside of the type-specific amplified region. These newly designed primers, in combination with the HPV-16 type-specific primers, could then be used as a nested PCR. The outer primers were optimized for magnesium (Fig. 2.6). The outer primers amplify a 192bp region and the inner primers amplify a 119bp region of the HPV-16 genome.

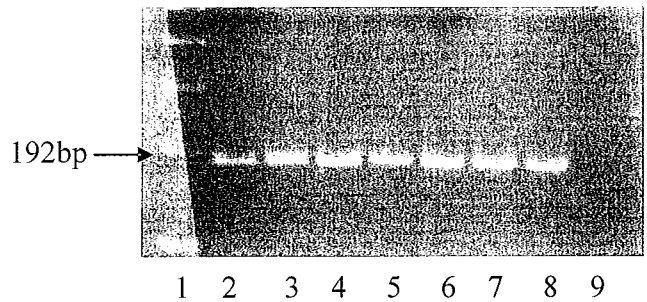


Figure 2.6 Magnesium titration for HPV-16 outer primers

Lane 1 100bp DNA ladder, lane 2 1 mmol/l magnesium, lane 3 1.5 mmol/l magnesium, lane 4 2 mmol/l magnesium, lane 5 2.5 mmol/l magnesium, lane 6 3 mmol/l magnesium, lane 7 3.5 mmol/l magnesium, lane 8 4 mmol/l magnesium, lane 9 negative control. PCR products run on a 1% agarose gel.

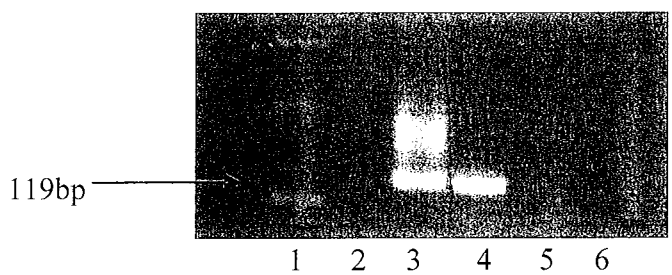


Figure 2.7 Nested HPV-16 PCR using HPV-16O and HPV-16I primer pairs

Lane 1 100bp DNA ladder, lane 2 positive control for round 1 of the nested PCR, lane 3 positive control for both rounds of the nested PCR, lane 4 patient 43, lane 5 negative control for round 1 of the nested PCR, lane 6 negative control for both rounds of the nested PCR. Products run on a 1% agarose gel.

Once the conditions had been optimised (2mmol/l magnesium, 20 cycles for round 1 and 40 cycles for round 2 and 1µl of DNA added for the second round of the PCR) (Fig. 2.7) the nested HPV-16 PCR was applied to the 50 samples. (Fig. 2.8 and Appendix IV).

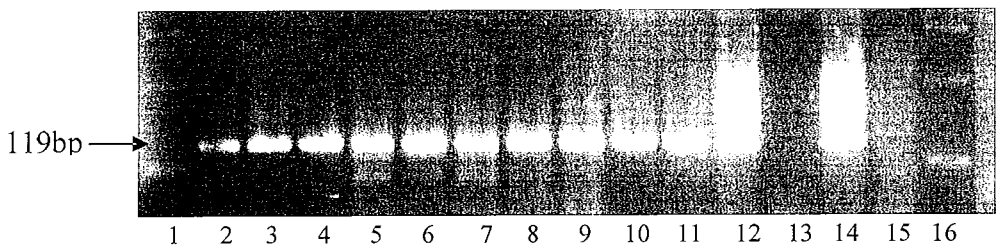


Figure 2.8 Nested HPV-16 PCR using HPV-16O and HPV-16 I primer pairs

Lane 1 100bp DNA ladder, lane 2 positive control, lanes 3 – 4 patients 24 - 25, lane 5 – 9 patient 31 – 35, lane 10 – 14 patients 41 – 45, 15 negative control, lane 16 100bp DNA ladder. PCR products run on a 1% agarose gel.

It was determined that 27 samples (54%) contained HPV-16 DNA, and in 26 of these cases CIN was present. As the CIN grade increased from CIN 1 to CIN 3 the number of HPV-16 positive samples also increased: from 5 CIN 1 to 7 CIN 3. 8 out of the 10 SCC samples contained HPV-16 DNA. A summary of the HPV PCR results is shown in Table 2.5.

Cytomorphology (number of cases)	HPV Positive	HPV-16 Positive
No CIN (10)	9	1
CIN 1 (10)	10	5*
CIN2 (10)	10	6
CIN 3 (10)	10	7
SCC (10)	10	8

Table 2.5 Summary of patient diagnosis, HPV PCR and HPV-16 PCR results

* = one case weakly positive, CIN = cervical intraepithelial neoplasia, SCC = squamous cell carcinoma.

2.3.3 Key Findings

- 961 LLETZ biopsies were received by the laboratory in the National Maternity Hospital during the 5 year period 1989 – 1993
- 21.9% of these cases were categorised as histologically normal
- 77.2% were diagnosed with CIN (4 patients - CIN grade was not specified)
    - 17.1% CIN 1
    - 38.4% CIIN 2
    - 21.3% CIN 3
- 98% of a 50 patient sample cohort were HPV positive (nested MY-GP PCR)
- 54% of the HPV positive patients were HPV-16 positive

2.4 Discussion

2.4.1 LLETZ samples With No CIN Present

LLETZ was introduced as a method of treatment for CIN in the National Maternity Hospital in 1989. Samples in this study were collected from this time for five years.

Over the period covered by this study there was a large number of samples that were treated for a high-grade cervical lesion using the LLETZ method but upon histological examination no lesion was present. 21.85% of all samples examined in the laboratory during this time were diagnosed as having no CIN present. This is higher than the number of patients diagnosed with CIN 1 (17.06%). The negative result for a LLETZ sample could be due to a number of factors, the lesion may have spontaneously regressed, there may be small foci of dysplasia that are removed during excision but not sampled during sectioning and also the failure to remove the dysplasia [106]. Other studies, that examined cases of negative LLETZ samples in their centres, reported that between 14% and 34% of samples had no lesion present [106, 122-124]. The study by Livasy et al. also found that patients with a negative LLETZ who had a recurrent high-grade dysplasia after treatment did so within a short period of time (less than one year) [106]. This indicates that all patients treated for CIN using the LLETZ method, including those with a negative result, must be followed up closely to rule out recurrence of disease. In this study the majority of the 10 cases with no CIN present were positive for HPV, but only 1 case was positive for HPV-16.

2.4.2 Age profile in LLETZ Cases

The 978 cases that had data on patient age were grouped into 5 groups on the basis of diagnosis, no CIN present, CIN 1, CIN 2, CIN 3 and SCC. The age range for each group was similar with the no CIN present, CIN 1, CIN 2 and CIN 3 groups having a range of 17-67 years (Table 2.3). The CIN 2 group had the largest range (49 years, 17-66 years) of these 4 groups. The no CIN present group had an age range of 48 years (19-67 years). The CIN 1 had an age range of 40 years and CIN 3 had an age range of 41 years, (18-58 years and 20-61 years respectively). The CIN 1 and CIN 2 groups had the lowest mean ages of 31.4 years and 32.2 years respectively. The no CIN present and CIN 3 groups had a mean age of 2 years more than this at 34.2 years for both groups. The mean age in the SCC group was 48.8 years, this is 12 years more than any other group including the CIN 3 group. The age range for the SCC group was 28-67 years. The 28 year old patient is the youngest patient in this group, the next youngest patients are 13 years older at 41 years old. 28 years old is very young for the development of cervical cancer and may indicate a relatively short time for progression from pre-cancerous lesions to the development of cancer in this patient.

When the patient ages are compared there is some statistical significance between the groups. The CIN 1 group shows a significant difference ($p = < 0.001$) with the CIN 3 group. There is also a significant difference ($p < 0.001$) between the SCC group and the 4 other groups, no CIN, CIN 1, CIN 2, and CIN 3. The difference in mean age between the CIN 1 group when compared to the CIN 3 group is a result of the time interval between initial infection and persistence of the virus and the development of high-grade CIN. It has been reported that the length of time between initial infection and the development of CIN 3 can be between 7 to 15 years [23]. Also the difference between the SCC group and the

other groups is due to the long period of time that cervical cancer takes to develop. According to Moscicki et al. [23] women with invasive cervical cancer tend to be on average, more than 10 years older than women with CIN 3 and in North America the median age for the diagnosis of cervical cancer is 47 years [125].

2.4.3 Detection of HPV

When DNA was extracted from the group of 50 samples it was examined using PCR to determine the presence of HPV. It was found that general primer sets MY09/MY11 and GP5+/GP6+ on their own could not amplify HPV DNA from any of the samples. This could be due to low amount of HPV DNA present in the sample, the lesion generally in all samples except the SCC samples, covered only a small area, so the number of cells containing HPV DNA in each sample was very small. When using a nested PCR technique this study found that 98% of the patients were HPV positive. 9 out of the 10 samples diagnosed with no CIN present were HPV positive. While these are both high figures it has to be remembered that all of these patients had previous abnormal smears and most had previous abnormal biopsies. Globally, between 8% and 22% of women with normal smear tests have been shown to be HPV positive [126-128], therefore in patients with cervical abnormalities a much higher presence of HPV would be expected.

After a persistent HPV-16 infection the chance of having a CIN 3 lesion after 5 years is 40% [23]. It has been established that women in Europe are more likely to become infected with HPV-16 than any other high-risk HPV type when compared to women in other parts of the world [129]. The presence of HPV-16 at the time of treatment for CIN also increases the chance of the recurrence of the disease when compared to other high-risk HPV types

[130]. When the 49 HPV positive samples were examined for the presence of HPV-16 it was found that 55.1% were HPV-16 positive. This is similar to the results of a study by Mann et al., who found that 58% of HPV positive patients were infected with HPV-16 [131]. In the current study only one sample with no CIN present was HPV-16 positive. The percentage of HPV-16 positive cases in CIN 1, CIN 2 and CIN 3 was 50%, 60% and 70% respectively. In the group of SCC patients 80% were HPV-16 positive. Although the figures for the number of HPV-16 positive cases in the high-grade lesions is slightly higher than expected, they correlate to other studies that examined the presence of high-risk HPV in LLETZ biopsies. In studies by van Ham et al. [132] and Sarian et al. [133] it was reported that 94% and 92% respectively, of high-grade lesions (CIN 2 and CIN 3) excised by LLETZ biopsy were positive for high-risk HPV. The increase in the number of HPV-16 positive patients correlates with increasing CIN grade. This indicates a high level of oncogenic potential for this HPV type. Overall there is no statistically significant difference in age between those women who are HPV-16 positive and HPV-16 negative.

3. Chapter 3 – Analysis of expression of RNA-binding proteins in normal and neoplastic cervical epithelium

3.1 Introduction

3.1.1 Splicing

Nearly all mammalian genes are interspersed by non-coding regions called introns [77]. Before translation into proteins can occur these non-coding sequences must be removed and the coding sequences (exons) placed side by side, this process is called splicing. This process is regulated to ensure the correct coding sequences are retained in the mRNA. RNA-binding proteins such as SR proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs) are involved in the regulation of splicing.

3.1.2 Splicing Regulatory Proteins

(i) SR Proteins

SR proteins are part of a family of closely related essential splicing factors which regulate splicing [134]. Eleven SR proteins have been characterised in humans including SRp75, SRp55, SRp40, SRp30, SRp20, ASF/SF2, SC35, 9G8 [38]. Generally SR proteins are made up of one or two RNA binding domains (RBD), which are located at the N-terminus, and arginine – serine (RS) dipeptide repeats of different lengths located at the C-terminus, called the RS domain [90]. SR proteins function in both constitutive and alternative splicing events and have been shown to act in various stages of spliceosome assembly [90] (Fig. 3.1). Interactions between the SR proteins and two other components of the

spliceosome (U1-70K and U2AF^{65}) is thought to help bring the 5' and 3' splice sites together [134]. SR proteins can bind to exonic splicing enhancers (ESE) and stimulate the splicing out of an upstream intron [80] (Fig. 3.1). SR proteins are themselves controlled during the cell cycle through phosphorylation and dephosphorylation of serine residues located within the RS domain of the protein [80, 135]. This allows tight control to be placed over alternative splicing. SR protein ASF/SF2 has been shown to interact indirectly with the HPV-16 negative regulatory element (NRE) and possibly helps to regulate the processing of late HPV mRNA transcripts [136].

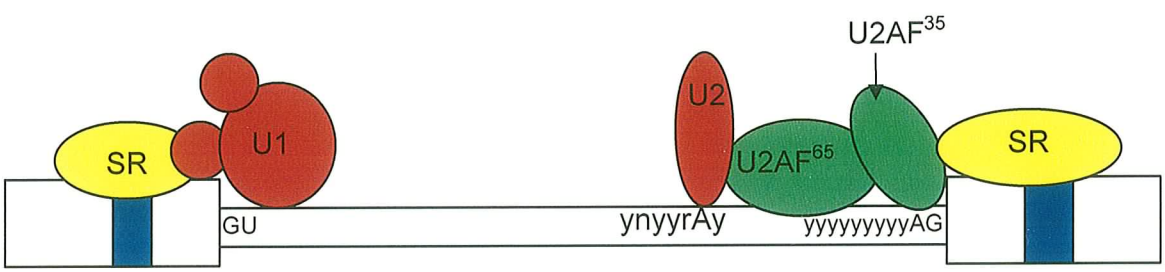


Figure 3.1 Positive regulation of alternative splicing by SR proteins

Blue bars indicate splicing enhancers, SR proteins bind to enhancers and help to recruit members of the spliceosome (U1, and via U2AF^{35} and U2AF^{65}, U2) to allow splicing to occur. GU = splice donor sequence, ynyyrAy = branch point sequence, yyyyyyyyy = polypyrimidine tract, AG = splice acceptor sequence.

(ii) Heterogeneous nuclear ribonucleoproteins (hnRNPs)

hnRNPs are abundant nucleoplasmic pre-mRNA binding proteins that have important roles in the biogenesis of mRNA. Many hnRNPs shuttle between the nucleus and the cytoplasm and are involved in the regulation of splicing, mRNA polyadenylation and mRNA export [137]. Approximately 30 major types of hnRNPs have been identified in human cells

including hnRNP A1, hnRNP C and hnRNP H [138]. The are identified by their association with unspliced mRNA precursors and are not a single family of related proteins [90].

hnRNP A1

hnRNP A1 binds to exonic or intronic splicing silencer elements repressing the splicing of certain alternatively spliced exons [139]. It has been shown that hnRNP A1 binds to multiple sites within the L1 coding region, binding of hnRNP A1 inhibits the use of the late 3' splice site [140].

hnRNP C1/C2

hnRNP C1/C2 has been shown to bind the polypyrimidine tract of Hepatitis C virus and may play a role in the initiation and/or regulation of the viral lifecycle [141]. hnRNP C1/C2 has also been shown to interact with an upstream enhancer element present in HPV-16 and helps to increase the polyadenylation at the early polyadenylation signal (pAE) [142].

hnRNP H

hnRNP H contains 3 RNA-binding domains that recognise G-rich elements [90]. Up-regulation of expression is seen in some tumour tissues that normally express low cytoplasmic levels [143]. hnRNP H has been shown to bind within the L2 coding region of HPV-16, this binding correlates with efficient use of the early polyadenylation signal [144].

hnRNP I (Polypyrimidine Tract Binding Protein (PTB))

In mammals PTB acts as a key splicing suppressor [90, 145]. It generally binds to intronic sequences that mediate splicing repression [80]. In bovine papillomavirus type 1 (BPV-1) it binds an exonic splicing silencer and represses the use of a suboptimal 3' splice site, it competes with $U2AF^{65}$ for binding sites, which may explain this inhibitory action [146]. PTB binds to the early untranslated region in HPV-16 and affects the use of the early polyadenylation signal, it is unknown if binding stimulates or inhibits the use of this signal [142].

(iii) $U2AF^{65}$

This protein is an essential splicing factor that recognises the 3' splice donor region and binds to the polypyrimidine tract [80]. Binding of $U2AF^{65}$ to the polypyrimidine tract helps to promote U2 snRNP recruitment to the spliceosome possibly by stabilising the interaction between U2 snRNA and the branch point sequence [147]. $U2AF^{65}$ has also been shown to stabilise the binding of U1 snRNA to weak SD sites within the HPV-16 NRE [148].

(iv) HuR

HuR is a 36kDa polypeptide that contains 3 RNA binding domains [149]. Two of these RNA recognition motifs (RRMs) recognise AU-rich elements (AREs) that help to regulate the lifespan of the mRNA that they are located on [149]. Overexpression of HuR leads to the stabilisation of mRNAs containing AREs *in vitro,* therefore prolonging the life of the mRNA [149]. HuR binds to the late untranslated region of HPV-1 [150, 151], HPV-16

[148] and HPV-31 [152], it is thought that binding inhibits late gene expression but its exact functional role in the HPV lifecycle is as yet unknown.

3.1.3 Markers of Differentiation, Proliferation and Viral Infection

(i) Cytokeratins

Cytokeratins are structural proteins of epithelial cells [153]. Approximately 50 cytokeratin genes have been identified and these are expressed in a tissue-specific manner [154]. They can be divided into two groups based on molecular weight and isoelectric points [153]. Type 1 cytokeratins are acidic and have low molecular weights and type 2 cytokeratins are basic or neutral and have high molecular weights [154]. Type 1 and type 2 can polymerise to from hetero-polymeric intermediate filaments [153]. Intermediate filaments are abundant in stratified epithelium especially in the supra-basal layers of the epidermis [153]. Cytokeratin 5 and cytokeratin 14 are both expressed in the highly proliferative basal cells of stratified epithelium [155]. Once these cells terminally differentiate the cytokeratin expression profile changes, with a down-regulation of the expression of cytokeratin 5 and 14 and an increase in expression of either cytokeratin 10 or 13 [154, 155].

(ii) Syndecan-1

There are four members of the Syndecan family, Syndecan-1, -2, -3 and -4 and they are all cell-surface heparan sulphate proteoglycans (HSPGs) [156, 157]. They are expressed on virtually all cell types and they are believed to have roles in cell adhesion, cell signaling and possibly act as co-receptors with integrins and cell-cell adhesion molecules [158]. They function as co-receptors by binding to ligands and increasing the ligand concentration in the

vicinity of their signaling receptors [157]. Syndecans contain 3 major domains, a short cytoplasmic domain, a single-span transmembrane domain and a extracellular domain with attachment sites for 3 or 5 heparan sulphate (HS) or chondroitin sulphate (CS) chains [159]. The HS chain allows interaction with a large group of proteins including fibroblast growth factors and vascular endothelial growth factors [159]. The extracellular domain can also be proteolytically shed as part of a regulated host response to injury or infection [157]. Syndecan-1 is highly expressed at the baso-lateral surface of epithelial cells and is an important regulator of cell-cell and cell-extracullular matrix interactions [158]. Syndecan-1 and -2 have been demonstrated to promote the attachment and internalization of HIV molecules through the extracellular HS moiety [157].

(iii) Proliferating Cell Nuclear Antigen (PCNA)

PCNA is a member of the DNA sliding clamp family of proteins [160]. It is necessary for the replication of DNA as it provides replicative polymerases with the high processivity required to duplicate the entire genome [161, 162]. It is composed of 3 identical PCNA monomers (each with two identical domains) which form a doughnut shaped homotrimer [160]. This homotrimer encircles the DNA and on its outer surface can engage in protein-protein interactions [160]. There are a number of known PCNA interacting proteins including proteins involved in DNA replication and repair (Polymerase δ and DNA ligase I), base excision repair (Polymerase β and AP-endonuclease I) and cell cycle control (p21, cyclin/CDK complexes) [160-162].

(iv) p16INK4a

p16INK4a is a member of the CDK inhibitor family of proteins [163]. It is a tumour suppressor that blocks the progression of the cell cycle and has been demonstrated to be induced in cells undergoing senescence [164]. p16INK4a binds to CDK5/CDK6 inducing an allosteric conformational change that inhibits the binding of ATP to the CDK complex [165]. Without ATP the interaction of CDK4/CDK6 and cyclin D is inhibited, which leads to pRB remaining in its hypophosphorylated state leading to cell cycle arrest at the G1 – S boundary [164, 165]. Loss of expression of p16INK4a has been reported in some neoplasms – familial melanoma, but over-expression, resulting in diffuse staining, has been detected in others – gastric adenocarcinoma, Hodgkin, and non-Hodgkin lymphomas [166]. In cervical pre-neoplastic and neoplastic lesions that are associated with HPV infection, the functional inactivation of pRB by HPV E7 results in the accumulation of p16INK4a as pRB normally inhibits the transcription of p16INK4a [166]. This results in diffuse staining that is regarded as a surrogate marker for HPV infection in high grade cervical lesions and cervical cancer [166].

(v) HPV L1

The HPV L1 protein is the major protein in the viral capsid [41]. It can cause a strong immune response in the host and therefore is used as the basis of both HPV vaccines [24]. L1 is expressed from the late promoter located within the E7 coding region and is produced in the differentiated epithelial layers [32, 36]. It is only in these upper layers that the infectious viral particles can be detected [31]. Immunohistochemistry using antibodies

directed against L1 have been used to detect its presence in cells from both cervical smears and cervical tissue sections [167-170].

3.1.3 Regulation of HPV protein expression by RNA-binding proteins

Persistence of HPV infection is seen as necessary, but not in itself sufficient, for the development of cervical cancer [21]. Persistence is characterised by the expression of the early HPV proteins, E1 E2, E4, E5, E6 and E7. Expression of E6 and E7 in particular is necessary to drive proliferation in the infected cell, and is required for progression to high-grade lesions and cancer [5, 58]. Expression of the HPV late genes, L1 and L2, is inhibited in high-grade lesions and cancer, allowing the persistently infected cell to remain undetected by the immune system of the host. Therefore, HPV persistence and progression to cancer requires high expression of some HPV genes and either no expression or low levels of expression of other HPV genes. Regulation of HPV gene expression occurs both at the level of transcription and at the level of RNA processing, for example RNA splicing, polyadenylation, stability and transport [38, 171]. These processes are controlled by RNA-binding proteins, which bind regulatory RNA elements and enhance or silence a particular RNA processing signal [90, 172, 173]. RNA-binding proteins, for example heterogeneous nuclear ribonucleoproteins (hnRNPs) or serine/arginine-rich (SR) proteins, are themselves tightly regulated during the cell cycle in order to keep a high level of control over the RNA processing mechanism [135].

It is possible that the levels of various RNA-binding proteins in host cervical epithelial cells may influence the transcription of HPV genes, or that these protein levels may themselves be affected by the presence of HPV. It has been demonstrated that a number of cellular

RNA-binding proteins interact with regulatory RNA elements on HPV mRNAs, the late mRNAs in particular [136, 140, 142, 148, 150-152, 174-179]. It has previously been shown, for example, that hnRNP C1/C2 and HuR interact with an AU-rich RNA instability element in the late 3' UTR of HPV-1 [150, 151, 174], CstF64, U2AF^{65} and HuR along with ASF/SF2 bind to a negative element within the HPV-16 late UTR RNA [136, 148, 175], and that the late UTR of HPV-31 interacts with CstF64, HuR and U2AF^{65} [152]. In addition hnRNP K and poly(rC) binding proteins interact with inhibitory elements within HPV-16 L2, reducing the expression of the late genes [176]. hnRNP A1 was shown to bind to multiple splicing silencers in the HPV-16 L1 coding region, thereby inhibiting splicing of the late mRNAs [140] and hnRNP H binds to motifs within HPV-16 L2 promoting the use of the early polyadenylation signal [177]. The polypyrimidine tract binding protein (PTB) and hnRNP C1/C2 bind to a U-rich sequence in the HPV-16 early UTR that stimulates early mRNA polyadenylation [142]. In bovine papillomavirus-1 (BPV-1) it was demonstrated that a number of SR proteins bind to exonic splicing silencers contained within the BPV-1 mRNA [178, 179]. As a result of the many interactions of RNA-binding proteins with papillomavirus mRNAs, it may be speculated that expression levels of these factors may be affected in the cervical epithelium as a result of HPV infection and the development of neoplasia. In this present study RNA-binding protein expression is evaluated in normal and neoplastic cervical epithelial lesions and is correlated with HPV-16 expression.

3.1.4 Immunohistochemistry

Immunohistochemistry (IHC) allows the identification of specific components within a specimen and also allows their location to be compared to the architecture of the specimen [180]. This method is based on the specific interaction between antigens and antibodies and was first used by Coons in 1941 [181]. Coons used a fluorescently labelled antibody to detect the antigen in a tissue section [182]. The basic technique for immunohistochemistry has remained the same since its invention, antigen is stabilised within a specimen, it is incubated in the presence of antibody and is visualised using a label [180].

There are a number of conditions that are essential for the detection of an antigen using immunohistochemistry, (i) the antigen must be preserved in the tissue in the same context that it existed in before sampling, (ii) the staining must be specific and sensitive and must occur in the absence of background staining, (iii) the antibody must be well-characterised and (iv) there must be efficient labelling and detection of the antibody [183]. The way that samples can be treated, such as fixation, processing and embedding can cause antigen to be lost irretrievably, depleted, altered or masked [184]. One of the most common forms of fixation, formalin-fixation, causes inter- and intra-molecular cross-links with certain structural proteins which can mask the antigens present in that tissue [184]. In order to overcome this, methods of antigen retrieval were developed, which were first described in 1976 by Huang et al., who used a protolytic enzyme to partially digest the tissue [181]. The most popular enzymes used today include trypsin, chymotrypsin, pronase and protease [184]. Heat-mediated antigen retrieval was introduced by Shi *et al.* in 1991, their method involved heating the sections in a boiling solution of a heavy metal salt [183]. Now non-toxic buffers are used including citrate buffer and EDTA buffer. Microwave ovens and

pressure cookers can be used to heat the samples, although better results can be obtained using the pressure cooker as it results in more uniform heating and it occurs under pressure [184].

(i) Avidin – Biotin Conjugate (ABC) Method

This is a modification of the biotin – avidin method and was developed by Hsu *et al.* This method has a greatly increased sensitivity when compared to the biotin – avidin method. Again the primary or secondary antibody can be conjugated to biotin so it can be a direct or indirect method. In the indirect method first the primary antibody binds to the antigen, secondly the biotinylated secondary antibody bids to the primary. Then complexes of avidin and biotin horse-radish peroxidase conjugate bind to the biotinylated secondary. As the avidin contains more than one biotin horse-radish peroxidase it serves to bring multiple labels to the site of the antigen [181, 185]. Once an enzyme label has been localised to the site of the antigen the substrate of the enzyme is added. This will result in a product being created that allows visualisation of the complex using a light microscope. An example of this process is the reaction between the enzyme horse-radish peroxidase, its substrate hydrogen peroxide and the chromagen diaminobenzidine (Fig. 3.2).

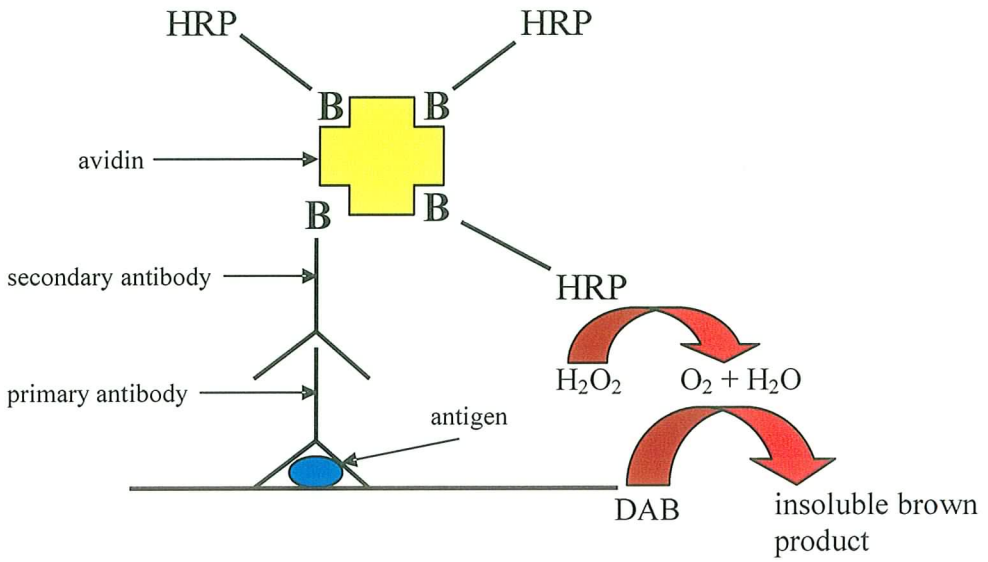


Figure 3.2 Detection of antigen using the ABC method and DAB

Secondary biotinylated antibody binds to the primary antibody, avidin-biotin-HRP complex then binds to the biotinylated antibody localising the label to the site of antigen. Once the avidin-biotin-HRP complex binds to the biotinylated secondary antibody, DAB and H_2O_2 are added leading to the development of an insoluble brown product at the location of the antigen. B = biotin, HRP = horse radish peroxidase, DAB = diaminobenzidine.

3.1.4 Application of immunohistochemistry to evaluate RNA-binding protein expression

Although the are many papers on the role of RNA-binding proteins and their involvement in the regulation of splicing, very little work has been performed using immunohistochemistry to localise their distribution in cancer cells. This present study provides an extensive evaluation and optimisation of current antibodies available for the characterisation of the expression of these RNA-binding proteins in neoplastic cells.

3.1.5 Aim

- To evaluate RNA-binding protein expression in normal and neoplastic cervical epithelium

Objectives

- To optimise immunohistochemical staining conditions for antibodies directed against RNA-binding proteins
- To relate RNA-binding protein expression to markers of proliferation and differentiation in cervical epithelium

3.2 Methods

3.2.1 Antibody Optimisation

A checkerboard optimisation technique was used to identify the optimal staining conditions for all antibodies listed in Table 3.1. Various antibody dilutions (ranged from 1:5 to 1:10000 depending on the antibody) were used with a number of different antigen retrieval techniques to assess which dilution worked best with which antigen retrieval technique. A number of different antigen retrieval techniques were used, protease enzyme treatment, microwave oven heating and pressure cooker heating. Also a number of different buffers were tested to assess which buffer gave the best results for each antibody, the buffers were 10mmol/l citrate buffer (pH 6) (appendix I), 1mmol/l EDTA buffer (pH 8) (appendix I) and Trilogy (Cell Marque).

Protein	Antibody	Company	Raised In
SR	1H4	Santa Cruz	mouse
9G8	P-15	Santa Cruz	goat
ASF/SF2	P-15	Santa Cruz	goat
SRp20	7B4	Santa Cruz	mouse
SC35	Y-16	Santa Cruz	goat
hnRNP A1	9H10	Abcam	mouse
hnRNP C1/C2	H-105	Santa Cruz	rabbit
hnRNP E1	E-16	Santa Cruz	goat
hnRNP E2	P-20	Santa Cruz	goat
hnRNP H	N-16	Santa Cruz	goat
hnRNP I	N-20	Santa Cruz	goat
hnRNP K	N-20	Santa Cruz	goat
HuR	19F12	Abcam	mouse
$U2AF^{65}$	N-14	Santa Cruz	goat
HPV L1	K1H8	Neomarkers	mouse
HPV-16 L1	CamVir 1	Abcam	mouse
HPV L1	Ab11777	Abcam	Rabbit
CK-5	LL002	Novocastra	mouse
CK-13	KS-1A3	Novocastra	mouse
CK-14	XM26	Novocastra	mouse
Syndecan-1 (CD-138)	MCA681H	Serotec	mouse
PCNA	Ab-1(PC10)	Neomarkers	mouse
p16INK4a	Ab-4 (16PO4)	Neomarkers	mouse

Table 3.1 Optimisation of antibodies used for staining of cervical epithelium

3.2.2 Antigen Retrieval

(i) Protease Enzyme Treatment

After dewaxing, sections were covered with 0.1% protease in PBS (appendix I) that had been pre-heated to 37°C. Sections were incubated at 37°C for 6 minutes. After incubation sections were washed in distilled water before continuing with the ABC protocol below.

(ii) Microwave Oven Treatment

Buffer (10mmol/l citrate buffer (pH 6) (appendix I) or 1mmol/l EDTA buffer (pH 8) (appendix I)) was pre-heated in a microwave oven for 5 minutes. Dewaxed sections in racks were placed into the pre-heated buffer and were heated in the microwave for 10 minutes. The sections were left in the hot buffer for a further 20 minutes before being washed in distilled water and continuing with the ABC protocol.

(iii) Pressure Cooker Treatment

400ml of water was put into the pressure cooker and a container with 500ml buffer (10mmol/l citrate buffer (pH 6) (appendix I) or 1mmol/l EDTA buffer (pH 8) (appendix I)) was placed in the bottom. The sections were placed in a rack and put into the buffer. The sections were heated on high pressure for 10 minutes. The sections were left in the hot buffer for a further 20 minutes before being washed in distilled water and continuing with the ABC protocol.

(iv) Trilogy Treatment

Trilogy (Cell Marque) is a commercially prepared buffer for dewaxing and antigen retrieval in immunohistochemistry. 400ml of water was put into the pressure cooker and two containers with 200ml Trilogy buffer (diluted 1:20 with water) were placed in the bottom. Sections, that had not been dewaxed, were placed in a rack and put into one of the containers with buffer. The sections were heated on high pressure for 15 minutes. The sections were left in the hot buffer for a further 5 minutes before the sections were transferred to the second container (hot rinse solution) of buffer. The sections were left in the hot rinse for 20 minutes before being washed in distilled water and continuing with the ABC protocol.

3.2.3 Immunohistochemistry (Avidin-Biotin Complex Immuno-Peroxidase Method)

5μm thick uterine cervical tissue sections were cut from formalin-fixed paraffin embedded tissue blocks and melted onto Superfrost Plus (VWR International) at 65°C for 2 hours. The sections were then dewaxed in xylene (BDH) and rehydrated through ethanol (Merck) to water. Antigen retrieval was performed as described above, the optimised retrieval method for each antibody is detailed in Table 3.25. Following antigen retrieval the sections were washed in distilled water and treated with 3% hydrogen peroxide in methanol (appendix I) for 10 minutes. Sections were rinsed in phosphate buffered saline pH 7.1 (PBS) (appendix I) before being stained using the Vectastain Elite ABC kit (Vector Laboratories). Normal horse serum (1:200) was applied to the sections for 5 minutes. Sections were drained and approximately 150μl of specific primary antibody (Table 3.25) was applied and incubated

at RT for 1 hour. Sections were washed with PBS for 5 min before the biotinylated secondary antibody (1:50) was applied for 15 min. Following washing with PBS the sections were covered with ABC reagent (1:50) for 15 min. Peroxidase labelling was visualised using 0.03% hydrogen peroxide (BDH) and 0.06% 2,4-diaminobenzidine (Sigma) (appendix I). Mayer's haematoxylin (appendix I) was used as a counterstain, the sections were then dehydrated in ethanol through xylene and coverslipped using DPX (BDH).

3.2.4 Controls for Immunohistochemistry

All tissues constitutively express RNA-binding proteins involved in the regulation of splicing. For positive controls skin and normal cervix were selected as both have similar histological structure. Pancreas, prostate and tonsil were also used in the optimisation of some antibodies. Negative controls for each tissue were also stained to check for background staining. All controls were formalin-fixed, paraffin-embedded tissue samples.

3.2.5 Evaluation of Cervical Epithelial Expression

The 50 formalin-fixed paraffin-embedded cases of cervical LLETZ biopsies analysed in chapter 2 for the presence of HPV were immunostained with all optimised antibodies for RNA-binding proteins, markers of differentiation, proliferation and of cervical intraepithelial neoplasia (cytokeratin-5, cytokeratin -13, cytokeratin -14, PCNA, Syndecan-1 and p16INK4a).

3.2.6 Immunohistochemistry Evaluation

No staining was graded as -, weak staining was graded +, moderate intensity staining was graded ++, and strong intensity staining was graded +++. As most RNA-binding proteins should be localised in the nucleus, cytoplasmic staining was discounted, except in the case of hnRNP E1 and hnRNP E2 as these proteins are known to shuttle from the nucleus to the cytoplasm and may give cytoplasmic staining.

3.3 Results

3.3.1 Antibody Optimisation

Many of the antibodies used had not been previously tested for use in immunohistochemistry (IHC), others had been tested for this use but had not been tested on formalin-fixed paraffin embedded samples. For all antibodies more than one tissue was used for optimisation, cervix was always included as one tissue. The checkerboard optimisation for each antibody is detailed overleaf. A summary of the optimised conditions for all antibodies is shown in Table 3.25.

(i) SR(1H4)

This mouse monoclonal had not been previously tested by the manufacturer for use in IHC. After optimisation microwave treatment with citrate buffer using a dilution of 1:80 was chosen as the optimal conditions for this antibody (Table 3.2), (Fig. 3.3). Clear epithelial nuclear staining with clean background was observed.

	Positive Control Tissue		
	Skin	Cervix	Pancreas
	Dilution (Range)	Dilution (Range)	Dilution (Range)
No Pre-treatment	NS (1:50-1:150)	NS (1:50-1:150)	NS (1:50-1:150)
Protease	1:50 (1:50-1:150)	1:80 (1:50-1:150)	NS (1:50-1:150)
Microwave (cit buffer)	1:100(1:50-1:150)	1:80 (1:50-1:150)	1:80 (1:50-1:150)
Pressure Cooker (cit buffer)	1:100(1:50-1:150)	**1:80** (1:50-1:150)	1:80 (1:50-1:150)

Table 3.2 Antibody optimisation for SR(1H4) antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(ii) 9G8

This goat polyclonal had not been previously tested by the manufacturer for use in IHC. The pressure cooker treatment with the EDTA buffer was the only method that gave any nuclear staining with cervical tissue. When 1:50 dilution of antibody was used on samples from the National Maternity Hospital using EDTA buffer in the pressure cooker no nuclear

staining was identified. This antibody was re-optimised using National Maternity Hospital cervical tissue but only very weak nuclear staining could be identified at an antibody dilution of 1:10, this dilution also gave very high background staining so lower antibody dilutions could not be tested. The 1:10 dilution was tried with a 3 hour incubation with either a 1% BSA solution (appendix I) or 0.2% Tween solution (appendix I) to attempt to reduce the background. No difference could be identified when either of these solutions was used. Therefore this antibody could not be optimised for use on formalin-fixed paraffin-embedded cervical samples (Table 3.3).

	Positive Control Tissue		
	Skin	Cervix	Pancreas
	Dilution (Range)	Dilution (Range)	Dilution (Range)
No Pre-treatment	NS (1:50-1:150)	NS (1:50-1:150)	NS (1:50-1:150)
Protease	NS (1:50-1:150)	NS (1:50-1:150)	NS (1:50-1:150)
Microwave (cit buffer)	NS (1:50-1:150)	NS (1:50-1:150)	NS (1:50-1:150)
Pressure Cooker (EDTA buffer)	1:50 (1:10-1:150)	1:50(1:10-1:150)	NS (1:10-1:150)
Pressure Cooker (Trilogy buffer)	1:50 (1:10-1:150)	NS (1:10-1:150)	NT

Table 3.3 Antibody optimisation for 9G8 antibody

Cit buffer = citrate buffer, NS = no nuclear staining present, NT = not tested.

(iii) ASF/SF2

This goat polyclonal had not been previously tested by the manufacturer for use in IHC. As no nuclear staining had been identified with the previous two antibodies when using no antigen retrieval this method was not used again for any of the antibodies chosen. Pressure cooker antigen retrieval with EDTA buffer was chosen as microwave treatment was not successful on the National Maternity Hospital LLETZ cervical biopsy samples. 1:50 antibody dilution was the optimal dilution with this method (Table 3.4), (Fig. 3.4). Clear epithelial nuclear staining with relatively clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:50-1:150)	NS (1:50-1:150)
Microwave (cit buffer)	1:50(1:10-1:150)	1:50 (1:10-1:150)
Pressure Cooker (EDTA buffer)	1:50 (1:50-1:150)	1:50 (1:50-1:150)
Pressure Cooker (Trilogy buffer)	1:50 (1:50-1:150)	**1:50** (1:50-1:150)

Table 3.4 Antibody optimisation for ASF/SF2 antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(iv) SRp20

This mouse monoclonal had been previously tested by the manufacturer for use in IHC but no information on antigen retrieval or antibody dilution was given. Pressure cooker treatment with EDTA buffer using a dilution of 1:25 was chosen as the optimal conditions for this antibody (Table 3.5). Clear epithelial nuclear staining with clean background was observed. When optimised conditions were used with the National Maternity Hospital LLETZ cervical biopsy samples this antibody did not work. No re-optimisation succeeded in staining of the LLETZ cervical biopsy samples.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:50-1:150)	NS (1:50-1:150)
Microwave (cit buffer)	NS (1:50-1:150)	NS (1:50-1:150)
Pressure Cooker (EDTA buffer)	1:50 (1:5-1:150)	1:25 (1:5-1:150)
Pressure Cooker (Trilogy buffer)	1:50 (1:10-1:50)	NS (1:10-1:50)

Table 3.5 Antibody optimisation for SRp20 antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(v) SC35

This goat polyclonal had not been previously tested by the manufacturer for use in IHC. The pressure cooker treatment with EDTA buffer was the only method that gave any nuclear staining. The 1:2 and 1:5 antibody dilutions gave weak staining and had high background staining so the 1:5 dilution was tried with a 3 hour incubation with either a 1% BSA solution (appendix I) or 0.2% Tween solution (appendix I) to attempt to reduce the background. No difference could be identified when either of these solutions was used. Therefore this antibody could not be optimised for use on formalin-fixed paraffin embedded cervical samples (Table 3.6).

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:50-1:150)	NS (1:50-1:150)
Microwave (cit buffer)	NS (1:50-1:150)	NS (1:50-1:150)
Pressure Cooker (EDTA buffer)	1:10 (1:2-1:150)	1:2 (1:2-1:150)
Pressure Cooker (Trilogy buffer)	1:10 (1:2-1:25)	NS (1:2-1:25)

Table 3.6 Antibody optimisation for SC35 antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(vi) hnRNP A1

This mouse monoclonal had not been previously tested by the manufacturer for use in IHC. When this antibody was tested on LLETZ cervical biopsy samples using pressure cooker treatment and a 1:100 antibody dilution, staining was not as good as seen when optimising the antibody. Trilogy antigen retrieval solution was used to improve staining using LLETZ cervical biopsy samples for optimisation. Pressure cooker treatment with Trilogy using an antibody dilution of 1:25 was chosen as the optimal conditions for this antibody (Table 3.7), (Fig. 3.5). Clear epithelial nuclear staining with clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:50-1:150)	NS (1:50-1:150)
Microwave (cit buffer)	1:100 (1:50-1:200)	1:100 (1:50-1:200)
Pressure Cooker (EDTA buffer)	1:150 (1:50-1:500)	1:100 (1:50-1:500)
Pressure Cooker (Trilogy buffer)	1:10 (1:10-1:150)	**1:25** (1:10-1:150)

Table 3.7 Antibody optimisation for hnRNP A1 antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(vii) hnRNP C1/C2

This rabbit polyclonal had not been previously tested by the manufacturer for use in IHC. As the microwave treatment gave strong nuclear staining with no background staining and at a high antibody dilution the pressure cooker method was not tested. Microwave treatment with citrate using an antibody dilution of 1:500 was chosen as the optimal conditions for this antibody (Table 3.8), (Fig. 3.6). Clear epithelial nuclear staining with clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	1:500 (1:50-1:700)	1:150 (1:50-1:700)
Microwave (cit buffer)	1:500 (1:50-1:700)	**1:500** (1:50-1:700)

Table 3.8 Antibody optimisation for hnRNP C1/C2 antibody

Cit buffer = citrate buffer.

(viii) hnRNP E1

This goat polyclonal had been previously tested by the manufacturer for use in IHC but no information on antigen retrieval or antibody dilution was given. As this antibody gave high background staining in all areas it could not be confirmed that the cytoplasmic staining identified was the binding of the antibody to the hnRNP E1 protein or if it was background staining. Also no stratified epithelial cells demonstrated cytoplasmic staining. Therefore this antibody was not used for any further experiments (Table 3.9).

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:40-1:150)	NS (1:40-1:150)
Microwave (cit buffer)	1:100 (1:40-1:150)	1:50 (1:40-1:150)
Pressure Cooker (EDTA buffer)	1:25 (1:25-1:150)	1:50 (1:25-1:150)

Table 3.9 Antibody optimisation for hnRNP E1 antibody

Cit buffer = citrate buffer, NS = no staining present.

(ix) hnRNP E2

This goat polyclonal had been previously tested by the manufacturer for use in IHC but no information on antigen retrieval or antibody dilution was given. As this antibody gave high background staining in all areas it could not be confirmed that the cytoplasmic staining identified was the binding of the antibody to the hnRNP E2 protein or if it was background staining. Also no stratified epithelial cells demonstrated cytoplasmic staining and no staining was identified in the control normal cervical tissue. Therefore this antibody was not used for any further experiments (Table 3.10).

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:50-1:150)	NS (1:50-1:150)
Microwave (cit buffer)	NS (1:50-1:150)	NS (1:50-1:150)
Pressure Cooker (EDTA buffer)	1:10 (1:10-1:150)	NS (1:10-1:150)

Table 3.10 Antibody optimisation for hnRNP E2 antibody

Cit buffer = citrate buffer, NS = no staining present.

(x) hnRNP H

This goat polyclonal had been previously tested by the manufacturer for use in IHC but no information on antigen retrieval or antibody dilution was given. Pressure cooker treatment with citrate using an antibody dilution of 1:400 was chosen as the optimal conditions for this antibody (Table 3.11), (Fig. 3.7). Clear epithelial nuclear staining with clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:50-1:150)	NS (1:50-1:150)
Microwave (cit buffer)	1:400 (1:50-1:700)	1:200 (1:50-1:700)
Pressure Cooker (cit buffer)	1:400 (1:50-1:700)	**1:400** (1:50-1:700)

Table 3.11 Antibody optimisation for hnRNP H antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(xi) hnRNP I

This goat polyclonal had been previously tested by the manufacturer for use in IHC but no information on antigen retrieval or antibody dilution had been given. As the microwave treatment gave strong nuclear staining with no background staining and at a high antibody dilution the pressure cooker method was not tested. Microwave treatment with citrate buffer using an antibody dilution of 1:400 was chosen as the optimal conditions for this antibody (Table 3.12), (Fig. 3.8). Clear epithelial nuclear staining with relatively clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	1:500 (1:50-1:700)	1:150 (1:50-1:700)
Microwave (cit buffer)	1:400 (1:50-1:700)	**1:400** (1:50-1:700)

Table 3.12 Antibody optimisation for hnRNP I antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(xii) hnRNP K

This goat polyclonal had been previously tested by the manufacturer for use in IHC but no information on antigen retrieval or antibody dilution had been given. Only background staining was identified at any antibody dilution using pressure cooker with EDTA. Therefore this antibody could not be optimised for use on formalin-fixed paraffin embedded cervical samples (Table 3.13).

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:25-1:150)	NS (1:25-1:150)
Microwave (cit buffer)	NS (1:25-1:150)	NS (1:25-1:150)
Pressure Cooker (EDTA buffer)	1:10 (1:10-1:150)	NS (1:10-1:150)

Table 3.13 Antibody optimisation for hnRNP K antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(xiii) HuR

This mouse monoclonal had been previously tested by the manufacturer for use in IHC using FFPE samples. A dilution range of 1:10000 to 1:100000 (0.1 – 1μg/ml) was suggested. Also heat meditated antigen retrieval was suggested (13 minutes in 10mmol/l citrate buffer pH6 followed by rapid cooling on ice). As the microwave treatment gave staining with no background staining and at a high antibody dilution the pressure cooker method was not tested. Microwave treatment with citrate buffer using an antibody dilution of 1:10000 was chosen as the optimal conditions for this antibody (Table 3.14), (Fig. 3.9). Clear epithelial nuclear staining with clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Microwave (cit buffer)	1:10,000 (1:10,000-1:100,000)	**1:10000** (1:10,000-1:100,000)

Table 3.14 Antibody optimisation for HuR antibody

Cit buffer = citrate buffer.

(xiv) U2AF^{65}

This goat polyclonal had been previously tested by the manufacturer for use in IHC but no information on antigen retrieval or antibody dilution had been given. Pressure cooker treatment with Trilogy using an antibody dilution of 1:25 was chosen as the optimal conditions for this antibody (Table 3.15), (Fig. 3.10). Clear epithelial nuclear staining with relatively clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:10-1:150)	NS (1:10-1:150)
Microwave (cit buffer)	1:10 (1:10-1:150)	NS (1:10-1:150)
Pressure Cooker (EDTA buffer)	1:50 (1:25-1:150)	1:25 (1:25-1:150)
Pressure Cooker (Trilogy buffer)	1:150 (1:10-1:150)	**1:25** (1:10-1:150)

Table 3.15 Antibody optimisation for U2AF^{65} antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(xv) HPV L1 (K1H8)

This mouse monoclonal had been previously tested by the manufacturer for use in IHC using formalin-fixed, paraffin-embedded samples. A dilution of 1:20 was suggested. Also heat meditated antigen retrieval was suggested (boiling in citrate buffer, 10mmol/l, pH 6 for 10 minutes followed by cooling at room temperature for 20 minutes). Only background staining was identified at any antibody dilution using both microwave and pressure cooker with citrate and EDTA (Table 3.16). Therefore this antibody could not be optimised for use on formalin-fixed paraffin embedded cervical samples.

	Positive Control Tissue	
	Cervix* Dilution (Range)	HeLa Cells expressing HPV-16 L1** Dilution (Range)
Microwave (cit buffer)	NS (1:10-1:100)	1:20 (1:10-1:50)
Pressure Cooker (EDTA buffer)	NS (1:10-1:100)	1:20 (1:10-1:50)

Table 3.16 Antibody optimisation for HPV L1 (K1H8) antibody

* = HPV-16 positive normal cervical tissue tested by HPV-16 type specific PCR, ** = HeLa cells were transfected with two plasmids to induce the expression of HPV-16 L1, these cells were formalin-fixed and paraffin-embedded as a positive control, cit buffer = citrate buffer, NS = no nuclear staining present.

(xvi) HPV-16 L1 (CamVir 1)

This mouse monoclonal had been previously tested by the manufacturer for use in IHC on paraffin sections but no information on antigen retrieval or antibody dilution was given. Only background staining was identified at any antibody dilution using both microwave and pressure cooker with citrate and EDTA (Table 3.17). Therefore this antibody could not be optimised for use on formalin-fixed paraffin embedded cervical samples.

	Positive Control Tissue	
	Cervix*	HeLa Cells expressing HPV-16 L1**
	Dilution (Range)	Dilution (Range)
Microwave (cit buffer)	NS (1:10-1:100)	1:25 (1:10-1:50)
Pressure Cooker (EDTA buffer)	NS (1:10-1:100)	1:30 (1:10-1:50)

Table 3.17 Antibody optimisation for HPV-16 L1 (CamVir 1) antibody

* = HPV-16 positive normal cervical tissue tested by HPV-16 type specific PCR, ** = HeLa cells were transfected with two plasmids to induce the expression of HPV-16 L1, these cells were formalin-fixed and paraffin-embedded as a positive control, cit buffer = citrate buffer, NS = no nuclear staining present.

(xvii) HPV L1 (ab11777)

This rabbit polyclonal had been previously tested by the manufacturer for use in IHC using paraffin embedded sections. An antibody dilution of 1:40 was suggested, antigen retrieval was also suggested but no details were given. Only background staining was identified at any antibody dilution using both microwave and pressure cooker with citrate and EDTA (Table 3.18). Therefore this antibody could not be optimised for use on formalin-fixed paraffin embedded cervical samples.

	Positive Control Tissue	
	Cervix* Dilution (Range)	HeLa Cells expressing HPV-16 L1** Dilution (Range)
Microwave (cit buffer)	NS (1:10-1:100)	1:20 (1:10-1:100)
Pressure Cooker (EDTA buffer)	NS (1:10-1:100)	1:40 (1:10-1:100)

Table 3.18 Antibody optimisation for HPV L1 (ab11777) antibody

* = HPV-16 positive normal cervical tissue tested by HPV-16 type specific PCR, ** = HeLa cells were transfected with two plasmids to induce the expression of HPV-16 L1, these cells were formalin-fixed and paraffin-embedded as a positive control, cit buffer = citrate buffer, NS = no nuclear staining present.

(xviii) Cytokeratin-5

Cytokeratin 5 antibody is a marker of the basal epithelial cells. This mouse monoclonal had been previously tested by the manufacturer for use in IHC using paraffin-embedded sections. An antibody dilution of 1:100 – 1:200 was suggested, high temperature antigen retrieval in 0.01mol/l citrate buffer (pH 6) was also suggested. The recommended positive control was prostate. Microwave treatment with citrate buffer using an antibody dilution of 1:40 was chosen as the optimal conditions for this antibody although at all concentrations this antibody stained all layers of the epithelium (Table 3.19), (Fig. 3.11).

	Positive Control Tissue	
	Prostate	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:25-1:200)	NS (1:10-1:200)
Microwave (cit buffer)	1:150 (1:25-1:200)	**1:40** (1:25-1:200)

Table 3.19 Antibody optimisation for CK-5 antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(xix) Cytokeratin -13

Cytokeratin 13 antibody is a marker of suprabasal epithelium. This mouse monoclonal had been previously tested by the manufacturer for use in IHC using paraffin-embedded sections. An antibody dilution of 1:100 – 1:200 was suggested, high temperature antigen retrieval was also suggested. The recommended positive control was tonsil. Microwave treatment with citrate buffer using an antibody dilution of 1:200 was chosen as the optimal conditions for this antibody (Table 3.20), (Fig. 3.12). Clear epithelial staining with clean background was observed.

	Positive Control Tissue	
	Tonsil	Cervix
	Dilution (Range)	Dilution (Range)
Protease	1:150 (1:50-1:250)	1:200 (1:50-1:250)
Microwave (cit buffer)	1:200 (1:25-1:300)	**1:200** (1:25-1:300)

Table 3.20 Antibody optimisation for CK-13 antibody

Cit buffer = citrate buffer.

(xx) CK-14

Cytokeratin 14 antibody is a marker of basal epithelium. This mouse monoclonal had been previously tested by the manufacturer for use in IHC using paraffin-embedded sections. An antibody dilution of 1:20 was suggested, high temperature antigen retrieval in 0.01mol/l citrate buffer (pH 6) was also suggested. The recommended positive control was skin. Microwave treatment with citrate buffer using an antibody dilution of 1:50 was chosen as the optimal conditions for this antibody (Table 3.21), (Fig. 3.13). Clear epithelial staining with clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Protease	NS (1:10-1:200)	NS (1:10-1:200)
Microwave (cit buffer)	1:200 (1:10-1:200)	**1:50** (1:10-1:200)

Table 3.21 Antibody optimisation for CK-14 antibody

Cit buffer = citrate buffer, NS = no nuclear staining present.

(xxi) Syndecan-1

This mouse monoclonal had been previously tested by the manufacturer for use in IHC using paraffin-embedded sections. An antibody dilution of 1:100 – 1:500 was suggested, heat-mediated antigen retrieval using sodium citrate buffer (pH 6) was also suggested. Pressure cooker treatment with citrate buffer using an antibody dilution of 1:200 was chosen as the optimal conditions for this antibody (Table 3.22), (Fig. 3.14). Clear epithelial staining with clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Microwave (cit buffer)	1:100 (1:50-1:500)	1:50 (1:50-1:500)
Pressure Cooker (cit buffer)	1:200 (1:50-1:500)	**1:200** (1:50-1:500)

Table 3.22 Antibody optimisation for Syndecan-1 antibody

Cit buffer = citrate buffer.

(xxii) PCNA

This mouse monoclonal had been previously tested by the manufacturer for use in IHC using paraffin-embedded sections. An antibody concentration of 0.5-1μg/ml of was suggested. Also heat-meditated antigen retrieval was suggested (boiling in citrate buffer, 10mmol/l, pH 6 for 10 minutes followed by cooling at room temperature for 20 minutes). Pressure cooker treatment with citrate buffer using an antibody dilution of 1:700 was chosen as the optimal conditions for this antibody (Table 3.23), (Fig. 3.15). Clear epithelial nuclear staining with clean background was observed.

	Positive Control Tissue	
	Skin	Cervix
	Dilution (Range)	Dilution (Range)
Microwave (cit buffer)	1:600 (1:400-1:1000)	1:500 (1:400-1:1000)
Pressure Cooker (cit buffer)	1:700 (1:400-1:1000)	**1:700** (1:400-1:1000)

Table 3.23 Antibody optimisation for PCNA antibody

Cit buffer = citrate buffer.

(xxiii) p16INK4a

This mouse monoclonal had been previously tested by the manufacturer for use in IHC using paraffin-embedded sections. An antibody dilution of 1:25 – 1:50 was suggested. Also heat-meditated antigen retrieval was suggested (boiling in citrate buffer, 10mmol/l, pH 6 for 10 minutes followed by cooling at room temperature for 20 minutes). Pressure cooker treatment with citrate buffer using an antibody dilution of 1:50 was chosen as the optimal conditions for this antibody (Table 3.24), (Fig. 3.16). Clear epithelial staining with clean background was observed.

	Positive Control Tissue Cervix Dilution (Range)
Microwave (cit buffer)	1:50 (1:10-1:200)
Pressure Cooker (cit buffer)	**1:50** (1:10-1:200)

Table 3.24 Antibody optimisation for p16INK4a antibody

Cit buffer = citrate buffer.

The optimal antibody dilutions and antigen retrieval methods for each antibody are summarised in Table 3.25.

Protein	Optimised Antibody Dilution	Optimised Antigen Retrieval
SR	1:80	MW citrate buffer
9G8	unable to optimise	unable to optimise
ASF/SF2	1:50	PC EDTA buffer
SRp20*	1:25	PC EDTA buffer
SC35	unable to optimise	unable to optimise
hnRNP A1	1:25	PC Trilogy
hnRNP C1/C2	1:500	MW citrate buffer
hnRNP E1	unable to optimise	unable to optimise
hnRNP E2	unable to optimise	unable to optimise
hnRNP H	1:400	PC citrate buffer
hnRNP I	1:400	MW citrate buffer
hnRNP K	unable to optimise	unable to optimise
HuR	1:10000	MW citrate buffer
$U2AF^{65}$	1:25	PC Trilogy
HPV L1 (K1H8)	unable to optimise	unable to optimise
HPV-16 L1 (CamVir1)	unable to optimise	unable to optimise
HPV L1 (ab11777)	unable to optimise	unable to optimise
CK-5	1:40	MW citrate buffer
CK-13	1:200	MW citrate buffer
CK-14	1:50	MW citrate buffer
Syndecan-1 (CD-138)	1:200	PC citrate buffer
PCNA	1:700	PC citrate buffer
p16INK4a	1:50	PC citrate buffer

Table 3.25 Optimal antibody dilutions and antigen retrieval methods for immunohistochemical staining of formalin-fixed paraffin embedded cervical epithelium

MW = microwave treatment, PC = pressure cooker treatment. * = antibody was not successfully used with National Maternity Hospital LLETZ biopsy samples.

3.3.2 Expression of hnRNPs and SR proteins in cervical lesions and cancer cells compared to histologically normal epithelia

In order to investigate if expression levels of regulatory RNA-binding proteins that are known to interact with HPV mRNA were altered in cervical premalignant and malignant lesions, 50 archival histological cervical tissue samples that were used for HPV detection in the previous chapter were immunostained with various optimised antibodies against RNA-binding proteins. These 50 samples were divided into 4 groups, 10 patients that were diagnosed with no CIN lesion, 10 with low-grade lesions (CIN 1), 20 with high-grade lesions (CIN 2 and 3) and 10 with cancer of the uterine cervix. We speculated that expression of RNA-binding proteins that have been shown previously to bind to and/or regulate HPV mRNA processing could be altered in CIN lesions. All antibodies that had been successfully optimised were used to stain the 50 samples.

(i) SR proteins

SR proteins (Fig. 3.3) were detected in the intermediate layer of normal epithelium, with some staining seen in both basal and superficial layers. Cervical intraepithelial neoplasia (CIN) lesions showed an increase in staining as the lesions increased in severity with the majority of high-grade CIN nuclei showing expression (Fig. 3.3). In SCC samples nearly all nuclei were positive for SR protein expression.

(ii) ASF/SF2

Normal epithelium displayed ASF/SF2 (Fig. 3.4) staining within the nuclei of the basal and intermediate layers, but no positive nuclei were seen in the upper (differentiated) superficial

layer. From low-grade CIN through to high-grade CIN there was an increase in the number of positive nuclei seen. SCC samples showed nuclear staining in the majority of tumour cells.

(iii) hnRNP A1

hnRNP A1 (Fig. 3.5) staining was distributed mostly in the intermediate layer of histologically normal cervical epithelium, with some scattered staining in the basal layer but no expression in the superficial layer. In low-grade CIN, high-grade CIN and SCC samples approximately half of lesional cells showed nuclear expression of hnRNP A1 (Fig. 3.5).

(iv) hnRNP C1/C2

hnRNP C1/C2 (Fig. 3.6) expression was present in occasional basal cells of normal epithelium, expression was also present in most intermediate cells and occasional superficial cells. Expression levels show a slight increase in lesional cells from low-grade CIN to high-grade CIN. Tumours showed a very high level of hnRNP C1/C2 expression.

(v) hnRNP H

hnRNP H (Fig. 3.7) nuclear staining was detected in the basal and intermediate layers of normal epithelium, while the superficial layers remained negative. All CIN grades and SCC samples had high levels of nuclear expression (Fig. 3.7). In low-grade CIN lesions the superficial layers were negative. The majority of SCC nuclei had positive staining for hnRNP H.

(vi) hnRNP I

In normal epithelium, hnRNP I (Fig. 3.8) was detected in the intermediate layer with occasional basal expression seen. There was no staining in the superficial layers this

epithelium. An increase in the number of positive nuclei was seen as the CIN grade deteriorated from low-grade to high-grade. In SCC cases nearly all nuclei showed hnRNP I expression.

(vii) HuR

Normal epithelium displayed HuR (Fig. 3.9) expression in basal and intermediate layer, scattered staining was seen in the superficial layers but this layer remained mostly negative. In low-grade CIN through to high-grade CIN cases most lesional cell nuclei demonstrated positive staining with the superficial layer in low-grade CIN cases showing no positive staining. SCC nuclei had a high level of expression of HuR.

(viii) U2AF^{65}

U2AF^{65} (Fig. 3.10) expression was seen in the basal and intermediate layers of normal epithelium, expression was occasionally seen in the superficial layers. In low-grade and high-grade CIN the majority of nuclei demonstrated U2AF^{65} expression and nearly all tumour nuclei showed expression.

In general, the RNA-binding proteins that were examined all showed expression in the intermediate layers of the normal epithelium. There was varying expression in the proliferative basal layer of these tissues, for example, hnRNP A1 expression was only seen in a few basal cells when compared to ASF/SF2 which was detected in the majority of basal cells, but overall there was very little protein expressed in the superficial layers. The CIN lesions showed a general increase in RNA-binding protein expression as the CIN grade deteriorated from low-grade to high-grade CIN. Expression of all proteins was seen in SCC samples. Table 3.26 gives a summary of expression in cervical epithelium.

3.3.3 Correlation of RNA-binding protein expression in cervical epithelium to known markers of differentiation

To correlate these results to cell differentiation and proliferation, the sections were all stained with antibodies against proteins with a known expression pattern in cervical epithelia. Further sections from the same 50 patients were stained using immunohistochemistry with these markers, cytokeratin 5 (CK-5), CK-13, CK-14, Syndecan-1, PCNA and p16INK4a in order to compare with the staining of the RNA-binding proteins.

(i) Cytokeratin 5

Cytokeratin 5 (Fig. 3.11) is a marker for the basal layer of epithelium. Staining was seen in all layers of normal and neoplastic epithelium, therefore no differential expression during epithelial differentiation was observed.

(ii) Cytokeratin 13

Cytokeratin 13 (Fig. 3.12) is a marker for the differentiated cells in the upper epithelial layers. Normal epithelium demonstrated positive staining from the intermediate to superficial layers with occasional weak staining seen in basal cells. Low-grade CIN through to high-grade CIN cases had mostly weak staining with some lesions lacking expression, occasional basal layer staining was present. SCC samples showed a focal staining pattern with varying expression from case to case. No relationship to RNA-binding protein expression was observed.

(iii) Cytokeratin 14

Cytokeratin 14 (Fig. 3.13) is a basal cell layer marker. Normal epithelium demonstrated staining in the basal layer and in two to three layers above the basal layer, with the remainder of the epithelium negative for cytokeratin 14 expression. CIN sections showed a mixed pattern of staining ranging from weak to strong staining intensity. In high-grade cases staining was mostly confined to the basal layers. All SCC samples were positive for cytokeratin 14 expression, mostly at the edges of the tumour. This did not correspond to RNA-binding protein expression, which was detected throughout the tumours.

(iv) Syndecan-1

Syndecan-1 (Fig. 3.14), a transmembrane heparin sulphate proteoglycan expressed on epithelial cells, was highly expressed on cell membranes of basal and intermediate cells in normal epithelium and lesional cells in CIN samples. Most lesions have extensive clear membranous staining. Tumours differed from premalignant lesions as a number of tumours had little or reduced staining. This would suggest an inverse relationship to RNA-binding protein expression.

3.3.4 Correlation of RNA-binding protein expression in cervical epithelium to known markers of proliferation

(v) Proliferating Cell Nuclear Antigen (PCNA)

PCNA (Fig. 3.15), a marker for proliferating cells, was expressed by occasional basal cells in epithelium with no CIN, the majority of intermediate cells and very few superficial cells. This is a similar pattern of expression to that of the RNA-binding proteins. Most lesional cells in low-grade CIN to high-grade CIN had strong nuclear PCNA expression. High-

grade lesions also had weak cytoplasmic staining. Most nuclei in the tumour samples expressed PCNA, they also had strong cytoplasmic staining.

(vi) p16INK4a

p16INK4a (Fig. 3.16), which has been reported to be a marker of high-grade cervical lesion, was undetectable in normal cervical epithelium. The majority of low-grade CIN samples lacked p16INK4a, low-expression was detected in a few scattered nuclei (Fig. 3.16). Nuclear and cytoplasmic staining was present in high-grade CIN lesions, while adjacent normal epithelium remained negative (Fig. 3.16). Tumours were all positive, most nuclei exhibited p16INK4a expression, occasional cases had a reduced nuclear staining but extensive cytoplasmic staining.

3.3.5 RNA-binding protein expression and HPV-16

When the RNA-binding protein expression was compared between the 27 cases that had been identified as being HPV-16 positive and the 22 HPV-16 negative cases no differences were noted. The expression of cytokeratin 13, cytokeratin 14, syndecan-1, PCNA and p16INK4a were also compared between the HPV-16 positive and HPV-16 negative cases, again no differences were noted.

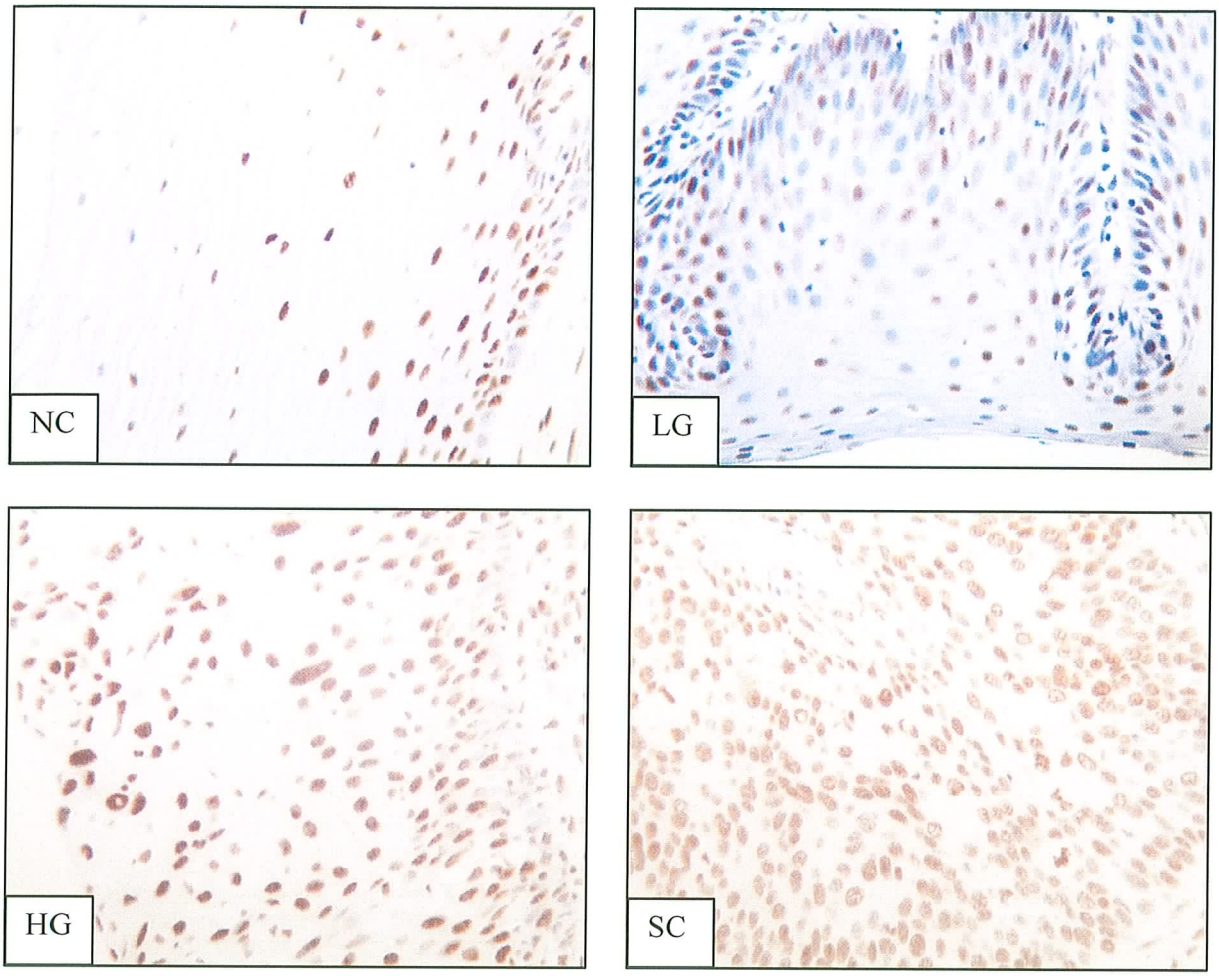


Figure 3.3 SR protein expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Microwave antigen retrieval with citrate buffer, antibody diluted 1:80. Magnification 400x.

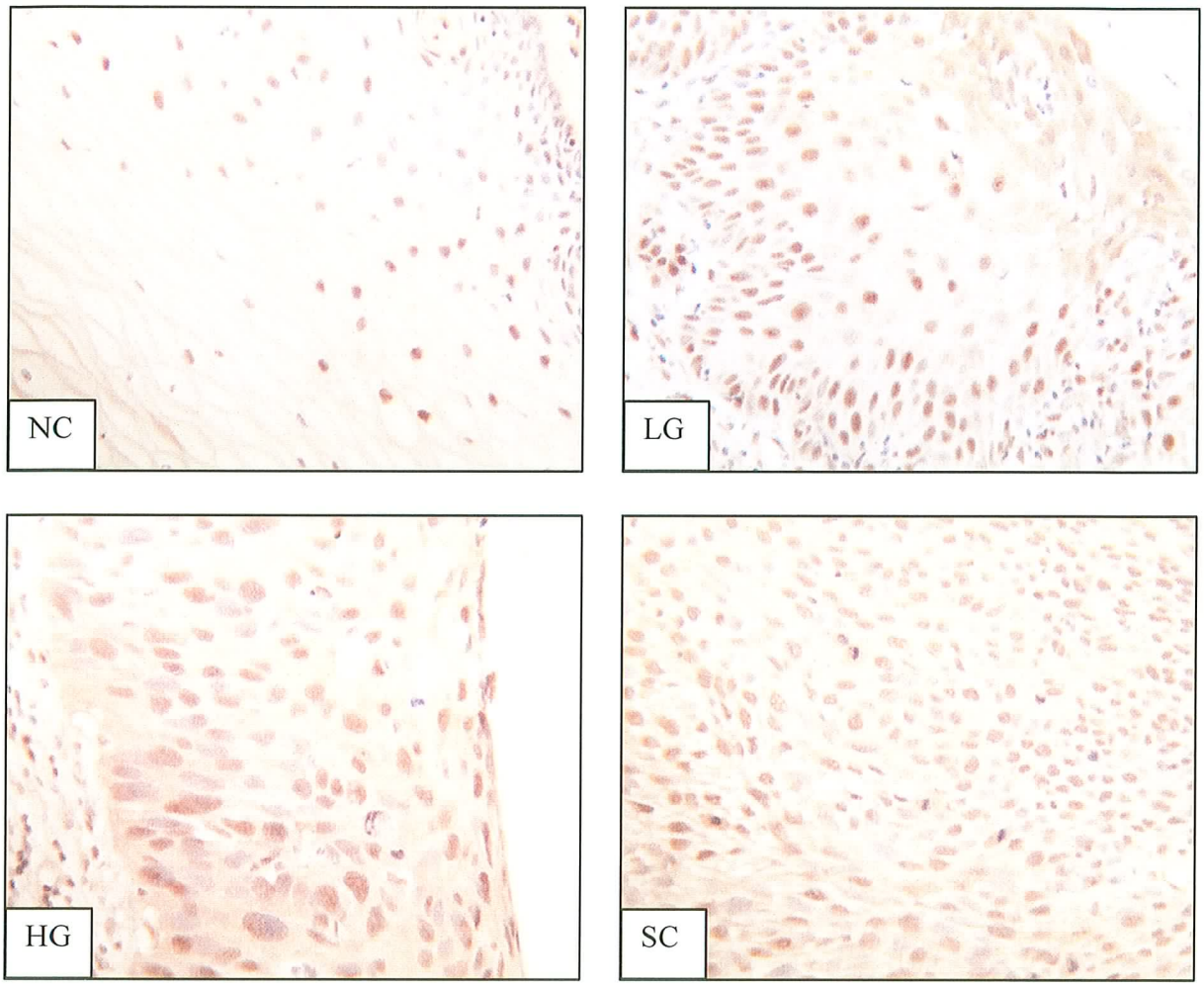


Figure 3.4 ASF/SF2 expression detected by using immunohistochemical staining.

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Pressure cooker antigen retrieval with EDTA buffer, antibody diluted 1:50. Magnification 400x.

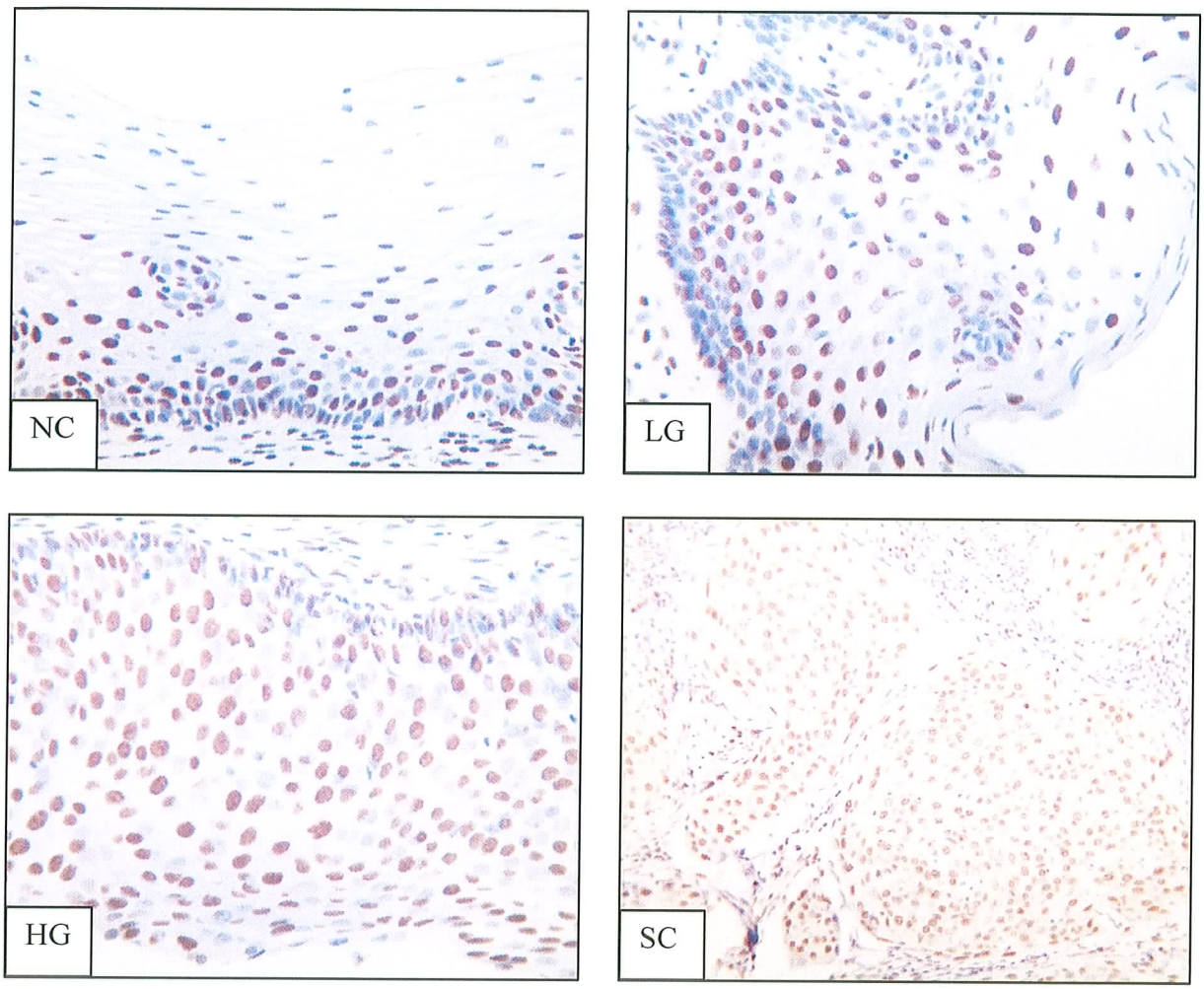


Figure 3.5 hnRNP A1 expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Pressure cooker antigen retrieval with Trilogy buffer, antibody diluted 1:25. Magnification 400x.

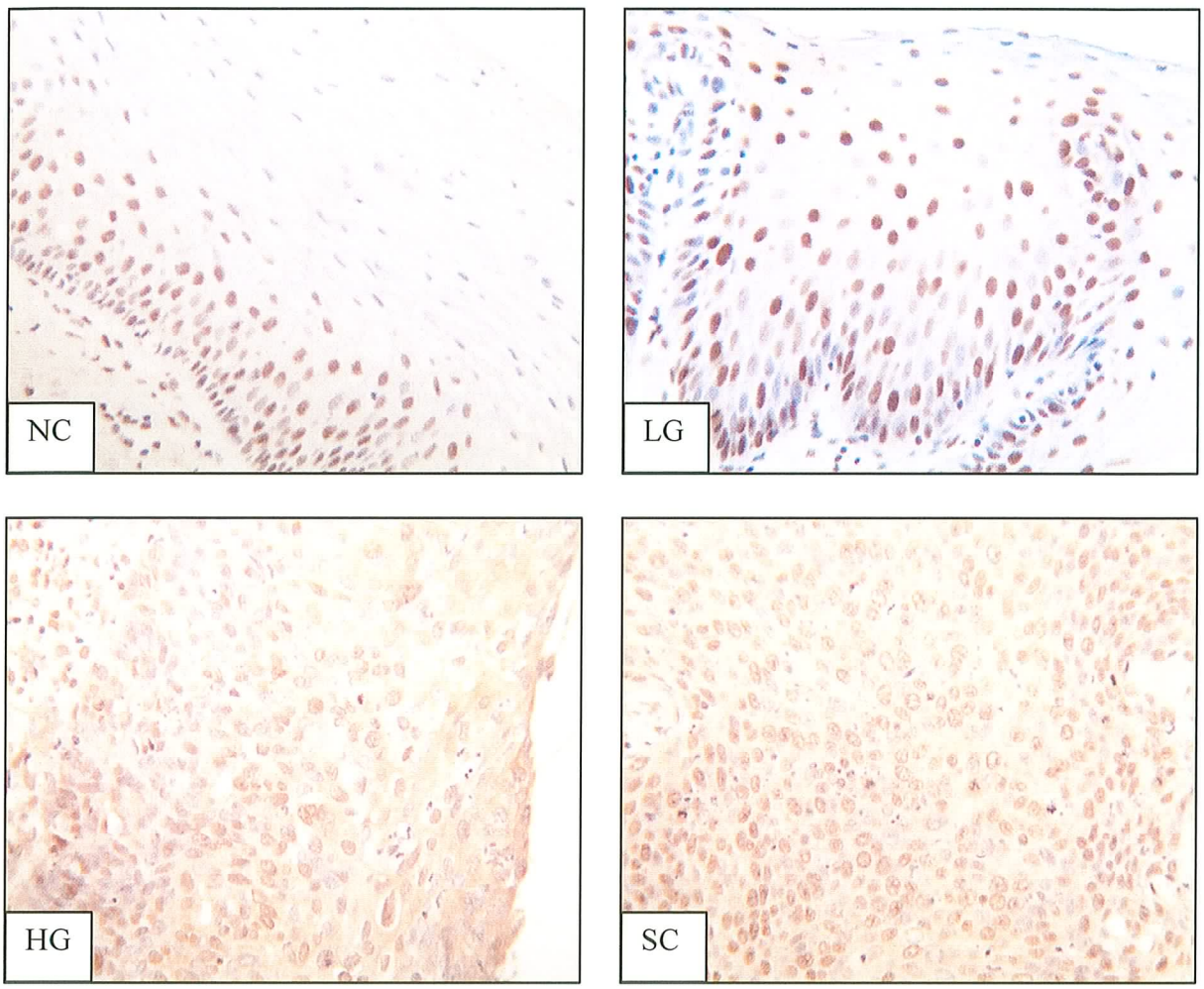


Figure 3.6 hnRNP C1/C2 expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Microwave antigen retrieval with citrate buffer, antibody diluted 1:500. Magnification 400x.

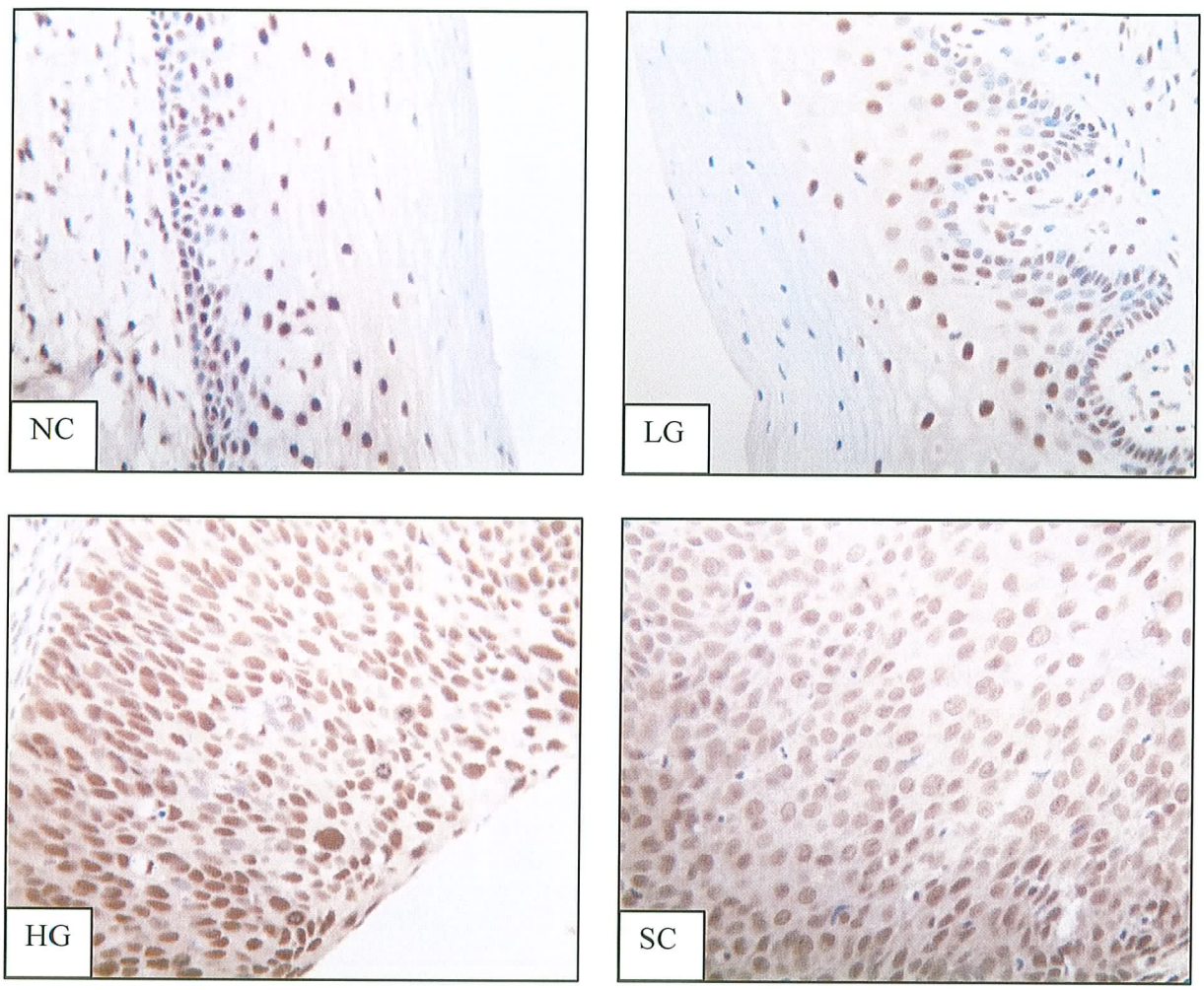


Figure 3.7 hnRNP H expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Pressure cooker antigen retrieval with citrate buffer, antibody diluted 1:400. Magnification 400x.

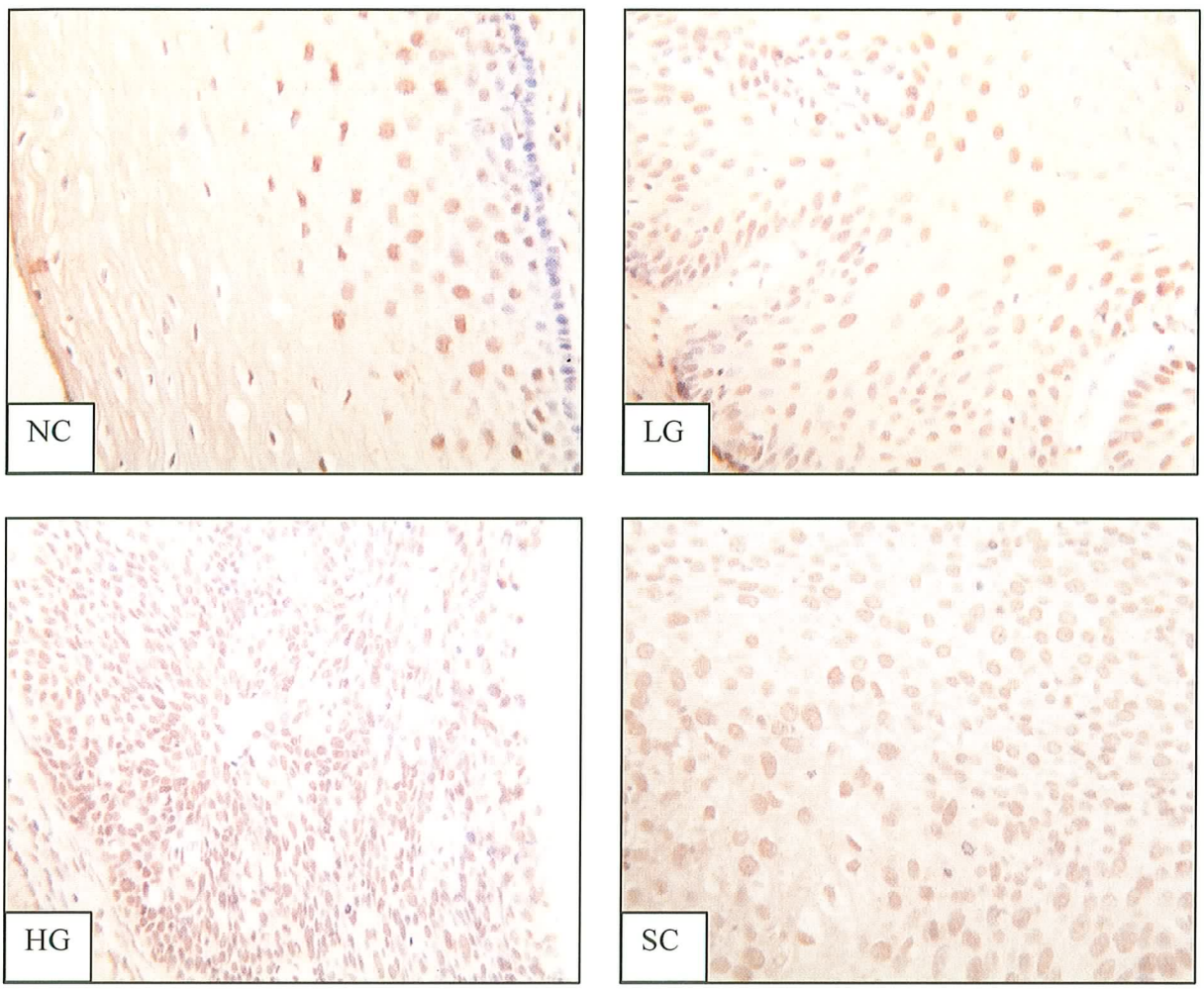


Figure 3.8 hnRNP I expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Microwave antigen retrieval with citrate buffer, antibody diluted 1:400. Magnification 400x.

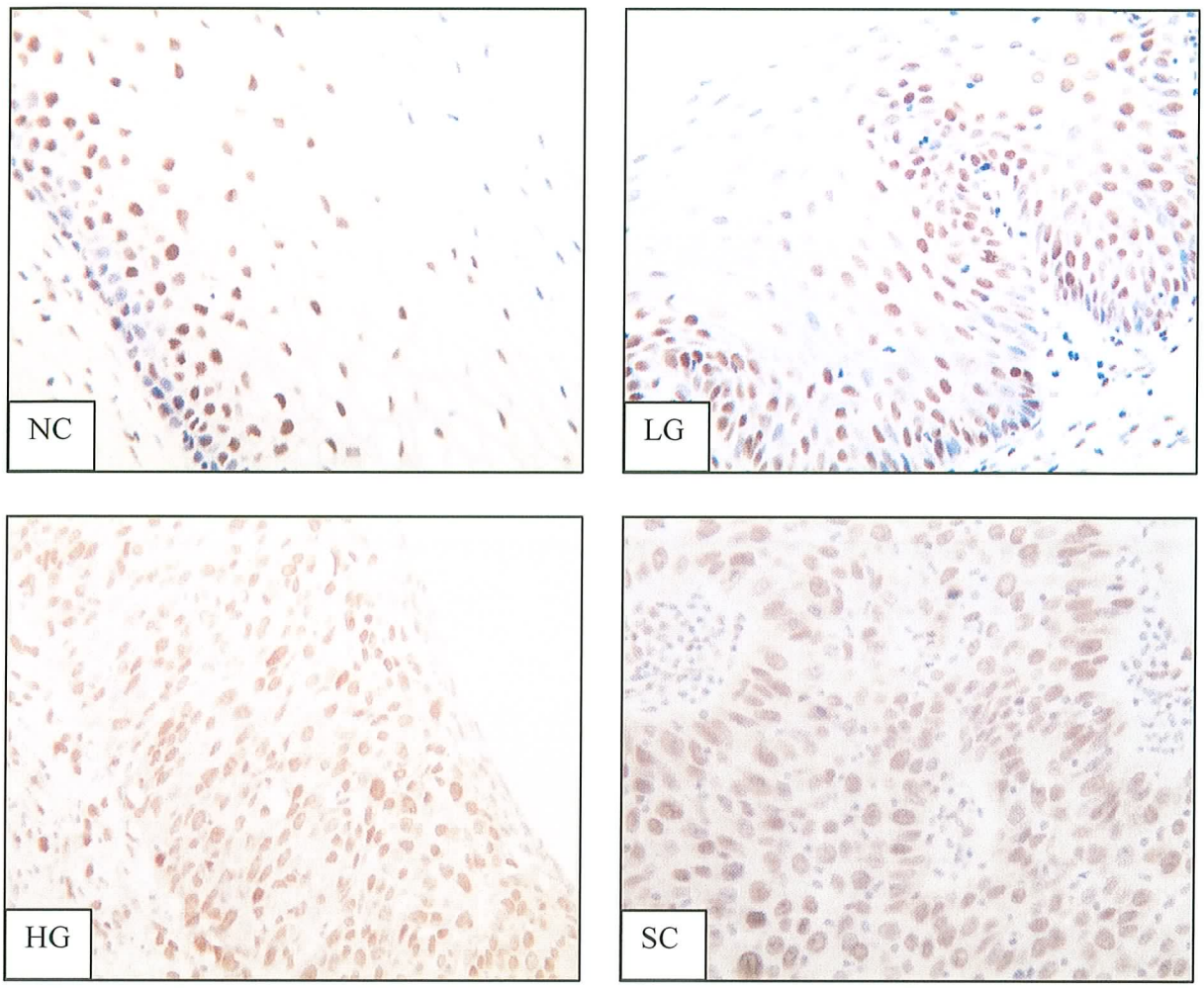


Figure 3.9 HuR expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Microwave antigen retrieval with citrate buffer, antibody diluted 1:10000. Magnification 400x.

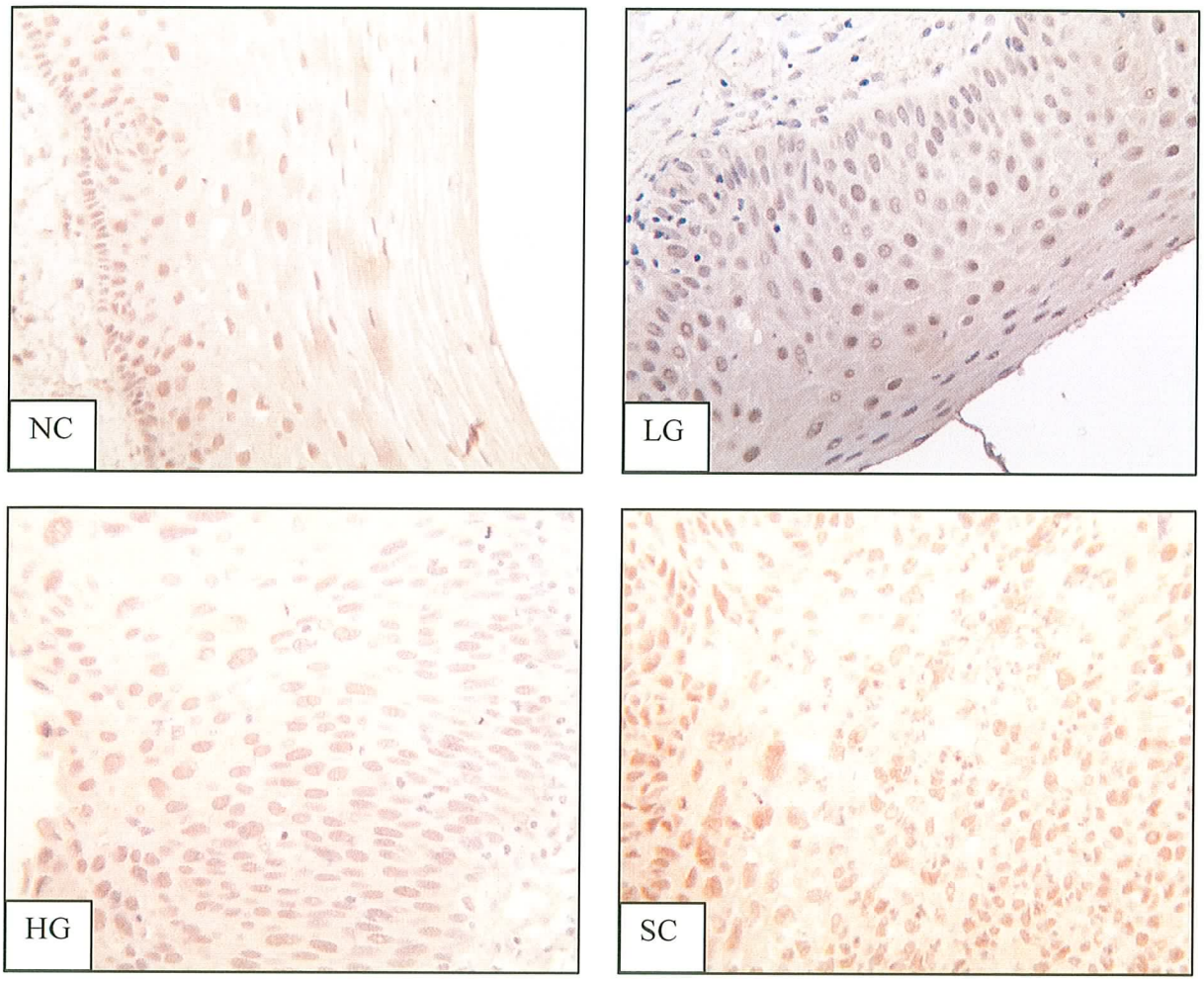


Figure 3.10 U2AF^{65} expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Pressure cooker antigen retrieval with Trilogy buffer, antibody diluted 1:25. Magnification 400x.

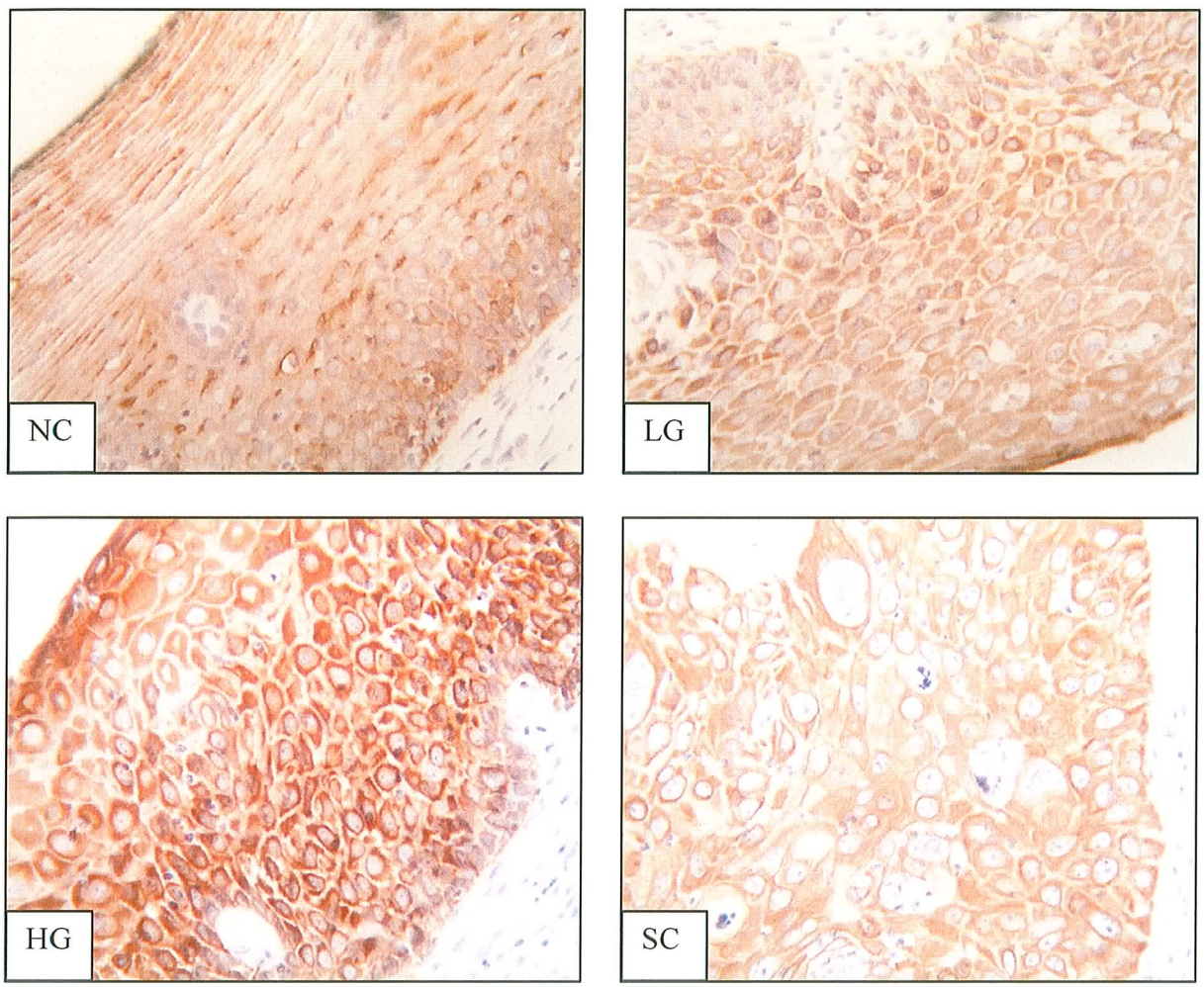


Figure 3.11 Cytokeratin 5 expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Microwave antigen retrieval with citrate buffer, antibody diluted 1:40. Magnification 400x.

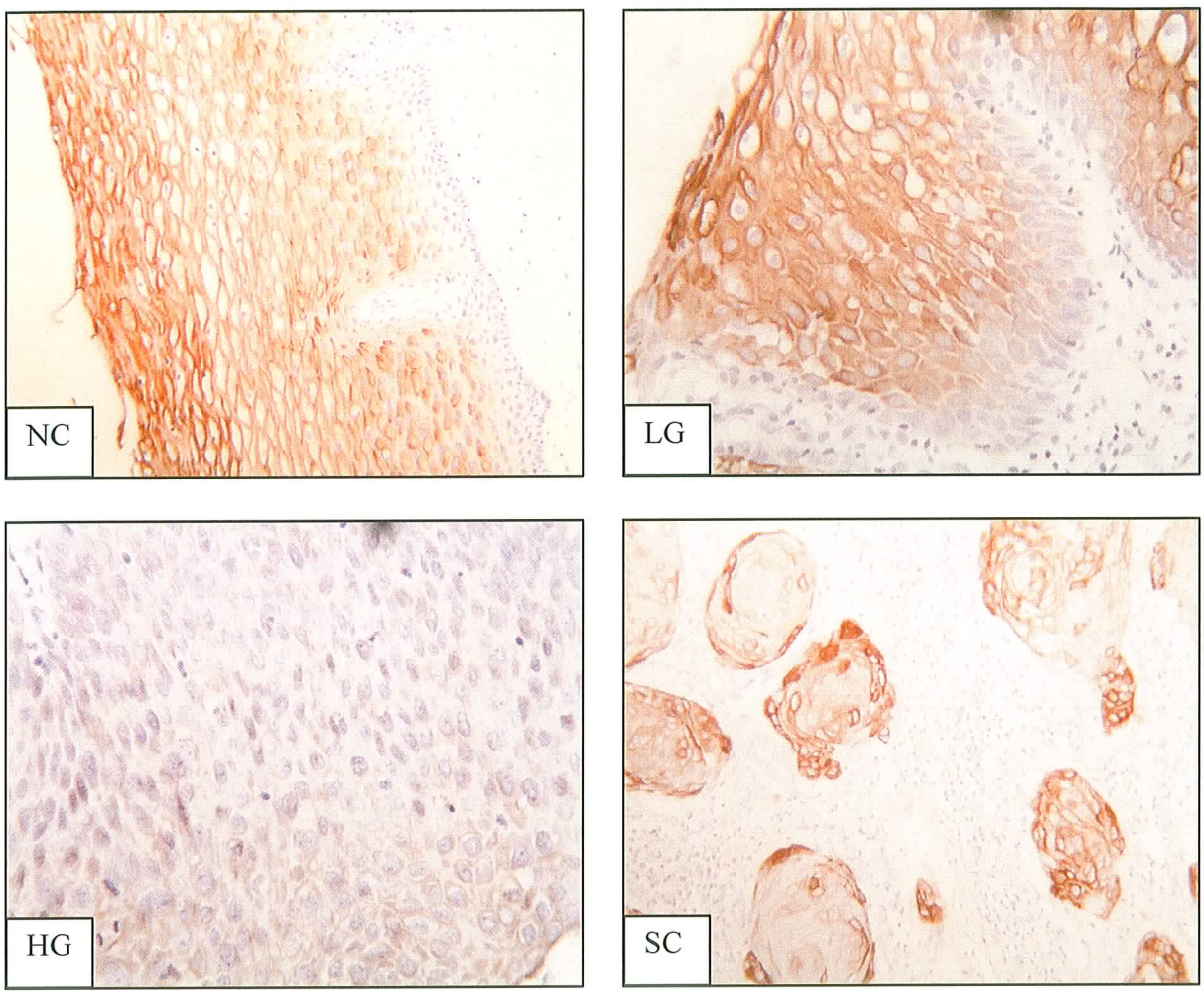


Figure 3.12 Cytokeratin 13 expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Microwave antigen retrieval with citrate buffer, antibody diluted 1:200. Magnification 400x.

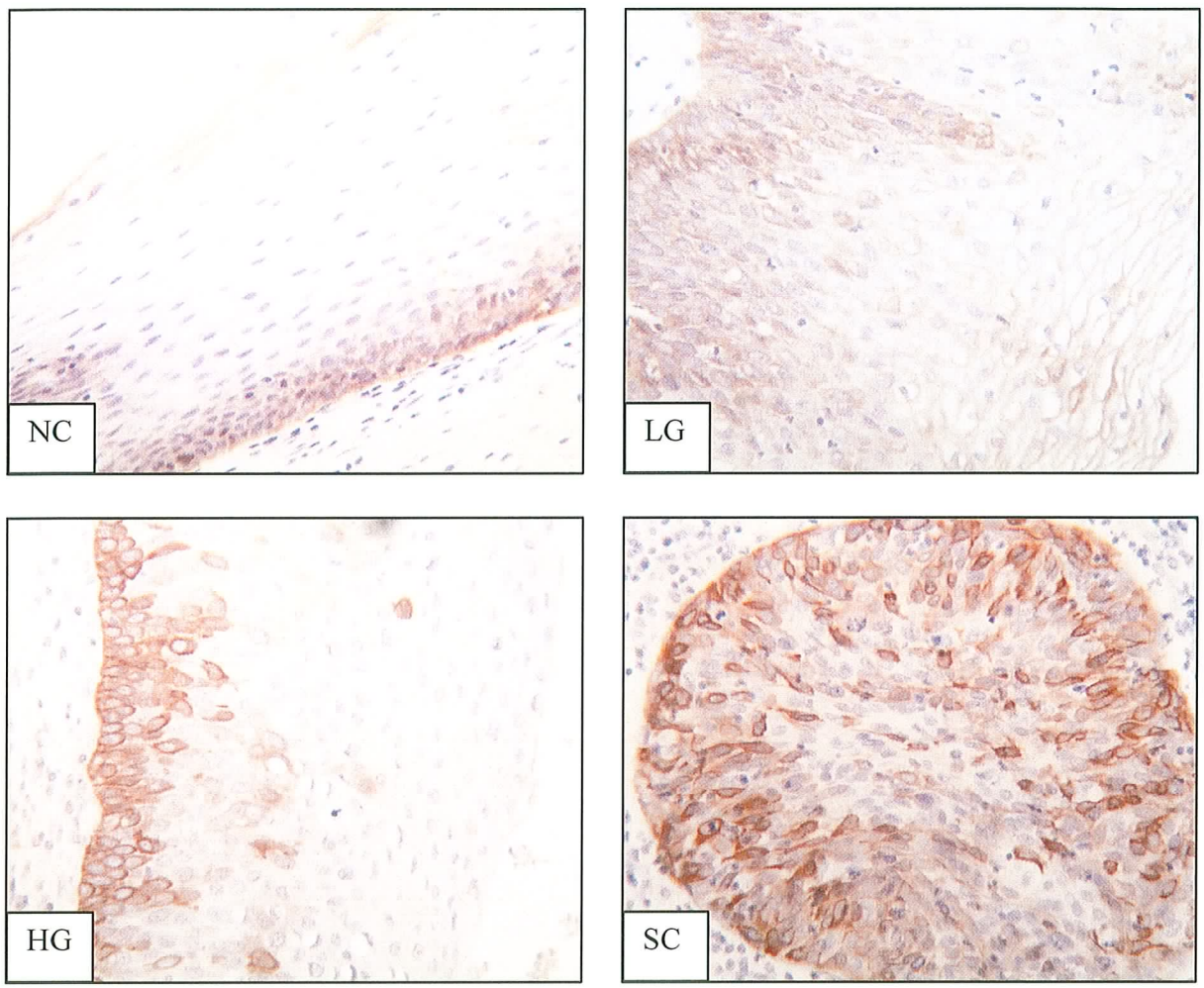


Figure 3.13 Cytokeratin 14 expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Microwave antigen retrieval with citrate buffer, antibody diluted 1:50. Magnification 400x.

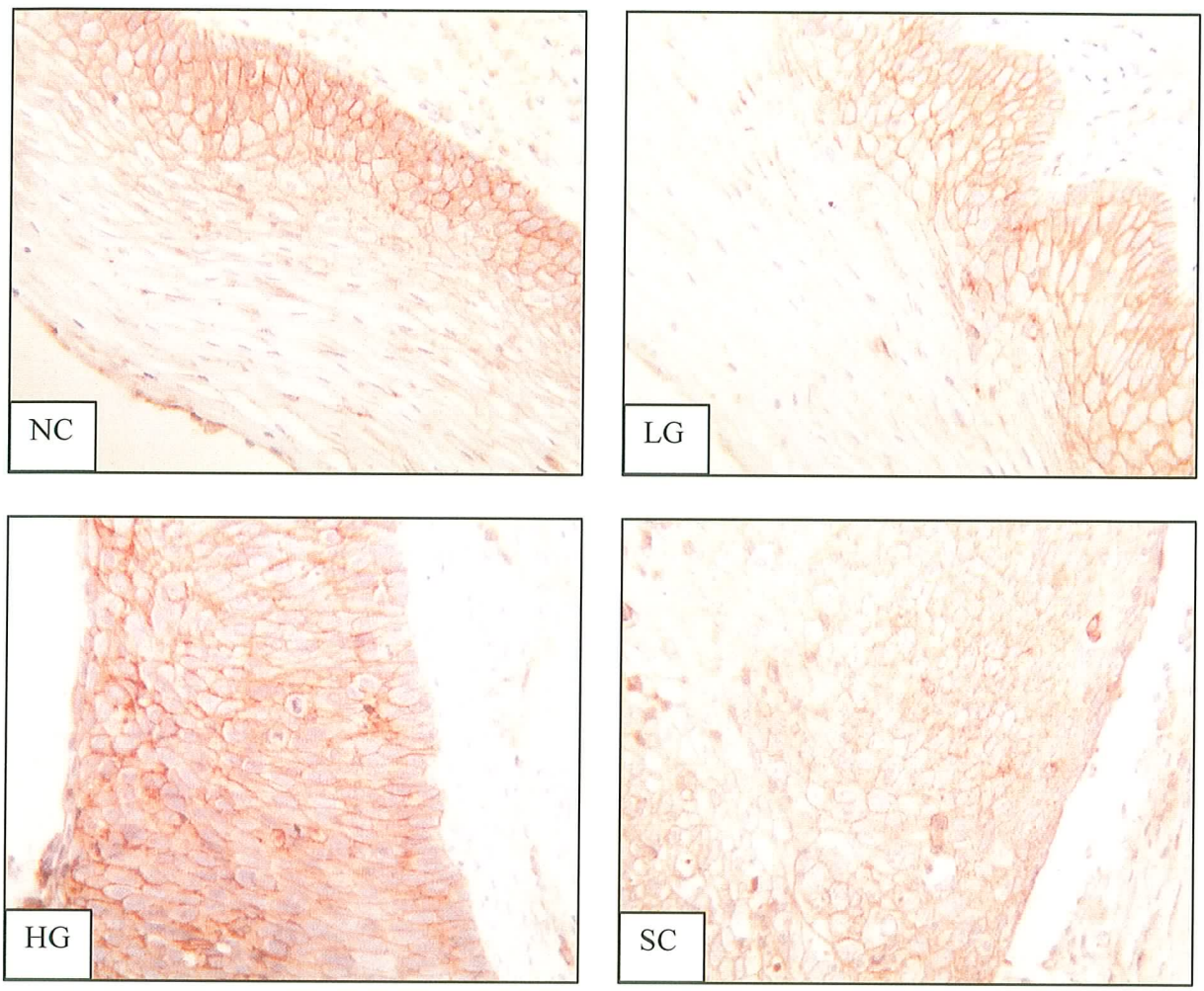


Figure 3.14 Syndecan-1 expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Pressure cooker antigen retrieval with citrate buffer, antibody diluted 1:200. Magnification 400x.

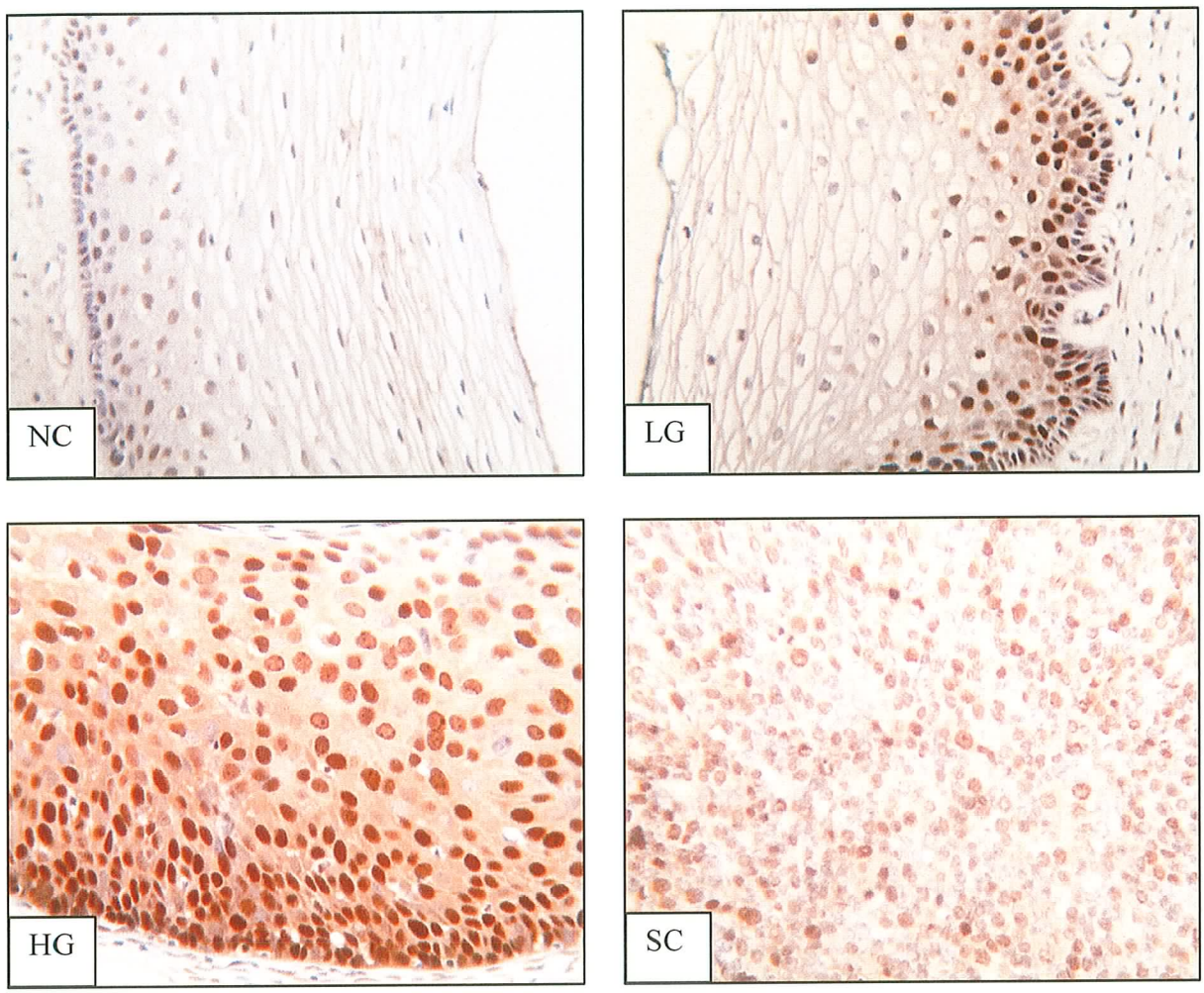


Figure 3.15 PCNA expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Pressure cooker antigen retrieval with citrate buffer, antibody diluted 1:700. Magnification 400x.

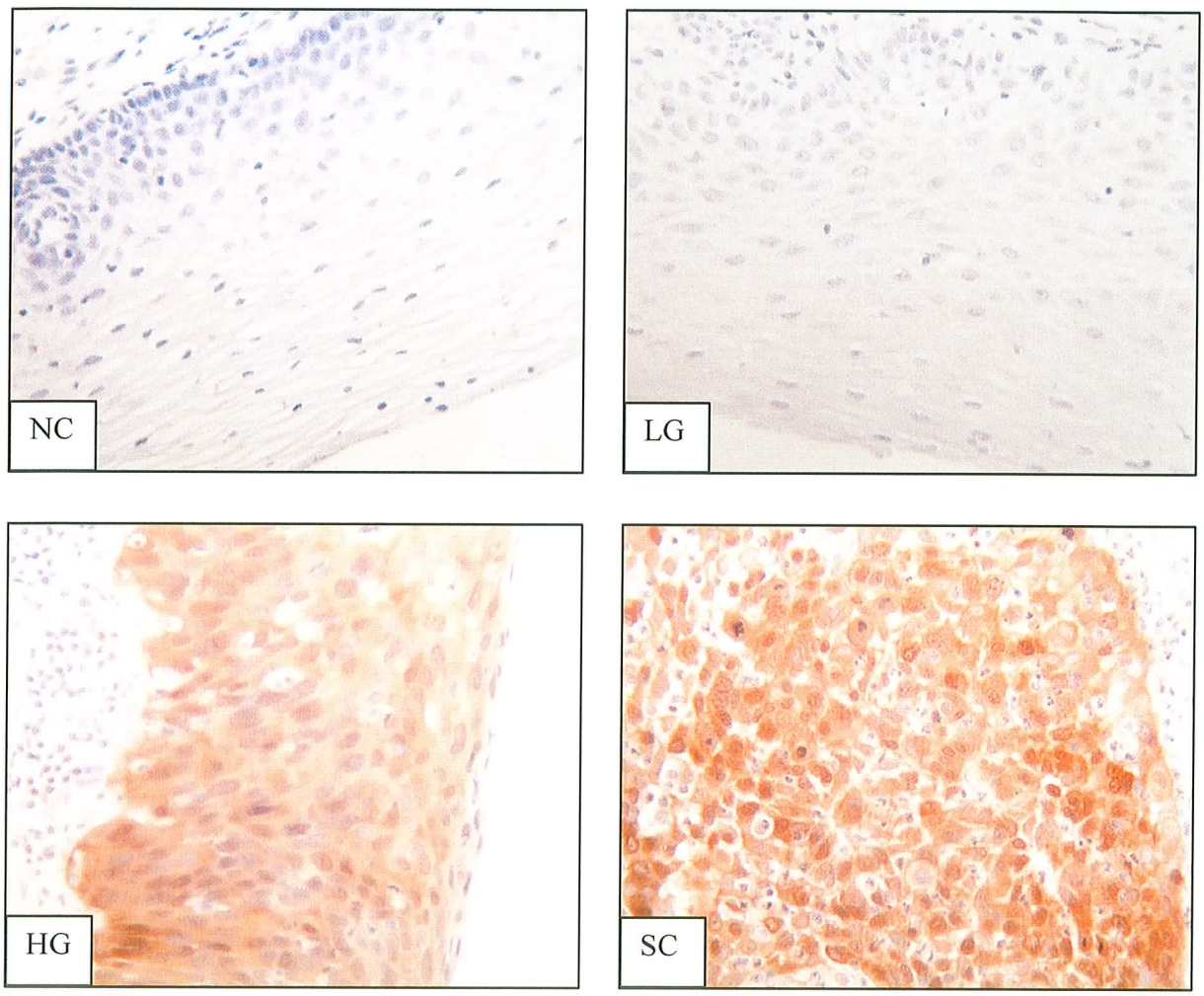


Figure 3.16 p16INK4a expression detected by using immunohistochemical staining

(NC) normal cervical epithelium, (LG) low grade cervical lesion, (HG) high grade cervical lesion, (SC) squamous cell carcinoma of the cervix. Pressure cooker antigen retrieval with citrate buffer, antibody diluted 1:50. Magnification 400x.

	No CIN	Low Grade CIN	High Grade CIN	SCC
hnRNP A1	+	++	+++	+++
hnRNP C1/C2	+	++	+++	+++
hnRNP H	++	+++	+++	+++
hnRNP I	+	++	+++	+++
ASF/SF2	+	++	+++	+++
SR		++	+++	+++
HuR	++	+++	+++	+++
$U2AF^{65}$	++	+++	+++	+++
CK-5	+++	+++	+++	+++
CK-13	+++	+/-	+/-	+
CK-14	+	+	+	++
Sydecan-1	++	+++	+++	+/-
p16INK4a	-	-	+++	+++
PCNA	+	++	+++	+++

Table 3.26 Overall expression of proteins in sections of cervical epithelium.

+/- not all samples in this CIN grade expressed this protein.

3.3.6 Key Findings

- All RNA-binding proteins were highly expressed in all neoplastic lesions
- Most squamous cell carcinoma cells were positive for the expression of the RNA-binding proteins
- Some RNA-binding proteins were not very commonly expressed in the basal layer of normal epithelium where HPV establishes a reservoir of infection
- Cytokeratin distribution was not related to RNA-binding protein expression
- Syndecan-1 was reduced in invasive cancers
- PCNA has a very similar expression pattern to the RNA-binding proteins
- p16INK4a expression is confined to high grade lesions and localised in both the nucleus and cytoplasm

3.4 Discussion

3.4.1 Antibody optimisation

23 antibodies directed against RNA-binding proteins, differentiation markers and proliferation markers were identified for use in this study. 15 antibodies were successfully optimised, all using heat mediated antigen retrieval with varying buffers. One optimised antibody SRp20 did not work when used on samples from the National Maternity hospital and re-optimisation of this antibody using NMH samples was not successful.

Eight antibodies could not be optimised. Of these, two, 9G8 and SC35, had not been tested for use in immunohistochemistry, they had only been used for Western blotting. Therefore these two antibodies are probably not suitable for use in immunohistochemistry.

Three antibodies, hnRNP E1, hnRNP E2 and hnRNP K, had been previously used in immunohistochemistry but no details on what type of tissue section used was given. As these three antibodies could not be optimised using formalin-fixed tissue, the process of formalin fixation may adversely affect the antigens that they are directed against. These antibodies may only be useful for immunohistochemistry using non-formalin-fixed samples, for example fresh-frozen tissue (cryostat sections).

The three different HPV L1 antibodies, K1H8, CamVir 1 and ab11777, had previously been tested for use in immunohistochemistry with paraffin-embedded samples. None of these three antibodies could be optimised for this study. Samples from NMH that had been previously confirmed as being HPV positive were used for their optimisation. The degree of masking that occurs as a result of formalin-fixation depends on the length of time that the tissue is exposed to the formalin [184], for this reason for tissues to be used in

immunohistochemistry fixation in formalin should be as short as possible while still allowing complete fixation of the tissue. As the samples used here were between 10 and 16 years old no details were available as to how long they had remained in formalin. If severe cross-linking of the antigen had occurred due to prolonged exposure to formalin this may be the reason these antibodies could not be optimised with these samples.

3.4.2 Expression of RNA-binding proteins in cervical epithelium

HPV genes are transcribed as polycistronic mRNAs. In order to differentially express the viral proteins that are required, alternative splicing and polyadenylation must take place to create each individual mRNA. These processes must be tightly regulated to ensure that the viral proteins are only expressed at an appropriate time and level. RNA-binding proteins are major factors that help to regulate both alternative splicing and polyadenylation in HPV [38]. This is demonstrated in HPV-16 by the tight control over the expression of the late genes. Inappropriate early expression of L1 and L2 could alert the immune response to the viral infection and result in clearance of the infection.

hnRNP A1 binds to multiple splicing silencers located in the L1 gene, binding inhibits the production of the late mRNAs [140]. Therefore, it would be expected that in the lower layers of HPV infected epithelium the levels of hnRNP A1 would be high, inhibiting late gene expression. But as differentiation occurs it would be expected that a drop in hnRNP A1 expression would occur allowing the capsid proteins to be expressed. In this study in epithelium that was infected with HPV but was histologically normal, there was expression of hnRNP A1 in the intermediate layers, but as these cells differentiated no expression was present. There was also low expression in the basal cells in these samples. In low-grade

lesions where productive infection can still occur while most lesional cells expressed high levels of hnRNP A1 the superficial layers did not, indicating that the late gene could be produced. In high-grade lesions and tumour samples expression of hnRNP A1 remained high, this is supported by the fact that late gene expression does not occur in these samples.

hnRNP H is similar to hnRNP A1 in the fact that it inhibits the expression of the late genes in HPV-16 by binding to GGG-motifs in L2 resulting in up-regulation of the use of the early polyadenylation signal [177]. hnRNP H demonstrated a similar pattern to the expression of hnRNP A1, supporting its role as an activator of early gene expression thereby inhibiting late gene expression.

hnRNP I and hnRNP C1/C2 bind to an element in the HPV-16 early UTR that stimulates early mRNA polyadenylation [142]. This would suggest that these proteins are involved in the regulation of HPV-16 early mRNA polyadenylation. Both proteins were expressed in the lower layers of normal epithelium but as differentiation occurred and production of the late genes would be required, there was only occasional expression in the nuclei of the superficial layers. High levels of both proteins were present in lesional cells, these results support the idea that these proteins stimulate HPV-16 early mRNA expression.

ASF/SF2 binds to a negative element in the late UTR of HPV-16 [136]. Again expression was present in the lower epithelial layers in normal cervical samples, decreasing as differentiation and late gene expression occurred. Expression levels remained high in low-grade, high-grade and SCC samples, with a decrease in expression only occurring in the superficial layers of the low-grade CIN samples. McPhilips et al [136] reported that ASF/SF2 was up-regulated upon differentiation of W12 cells (derived from a HPV-16 CIN 1 lesion) in a raft culture. This was not demonstrated in this study, as differentiation

occurred in the CIN 1 lesion ASF/SF2 was down-regulated and this was not dependant on HPV-16 infection as there was no overall difference between HPV-16 positive cases and HPV-16 negative cases. The discrepancy between these two results may be due to the differences between the tissue types used, raft culture versus patient tissue.

HuR and U2AF^{65} both bind to the late UTR of HPV-31 [152], HuR also binds to the AU-rich instability element in HPV-1 [150] and U2AF^{65} to a negative regulatory element in HPV-16 [148]. The expression profile for these proteins would support a role in inhibiting late gene expression. In histologically normal epithelium expression was present in the lower layers but no expression was present in the superficial layers. HuR and U2AF^{65} expression was present in the majority of lesional cells in low-grade, high-grade and SCC samples. HuR has been shown to protect unstable mRNAs from degradation and also, under stress conditions, appears to be involved in nuclear export of mRNAs that contain AU-rich RNA elements [149]. However, the expression of HuR in the cervical epithelium does not support it having a positive role in HPV late gene expression.

No RNA-binding protein was found that was induced or activated by terminal differentiation of cervical epithelial cells. In contrast, all RNA-binding proteins analysed here were down-regulated in response to terminal cell differentiation, suggesting that HPV late gene expression was dependant on the down-regulation of negative factors rather than differentiation-dependant activation of positive factors. Interestingly some of the RNA-binding proteins showed an increased expression in the intermediate layer when compared to the basal layer. This expression pattern was similar to PCNA expression which has been used as a surrogate marker for E6/E7 expression in HPV infected epithelia [186]. This result suggested and supported a role for the RNA-binding proteins in HPV early gene

expression. The up-regulation of PCNA has also been suggested to be linked to the presence of high-risk HPV and as such as an aid to the diagnosis of cervical abnormalities [187-189], the results in this study would support this, demonstrating an up-regulation in PCNA expression as CIN grade progressed from CIN 1 to CIN 3.

Some of the RNA-binding proteins were expressed in only a few basal cells, further work using double immunostaining with these RNA-binding proteins and markers of proliferation or antibodies against early HPV gene products to localise the virus would be of value to assess why only some of these basal cells demonstrated RNA-binding protein expression. The expression of the RNA-binding proteins also correlated with the expression of Syndecan-1 but only up to the high-grade lesions. SCC samples demonstrated a reduction in the expression of Syndecan-1 while RNA-binding protein expression remained high.

Increasingly, detection of protein expression by immunohistochemical analysis is being looked at as a method of marking tumours or pre-malignant lesions that may progress to cancer; for example, p16INK4a has been identified as a marker of high-grade CIN lesions and cervical SCC [166, 190]. Increased expression of certain hnRNP members has also been seen in some cancers and they may be useful as tumour markers. It has been shown that hnRNP A1 expression is increased in both lung and colon cancer [191, 192]. hnRNP B1 has been demonstrated to be over-expressed in lung cancer, over-expression can be detected at a very early stage and may be a marker of early lung cancer [193, 194]. hnRNP K is over-expressed in colorectal cancer and is associated with poor prognosis [195]. hnRNP H expression is increased in head and neck cancer, as well as in liver, pancreas and stomach cancer [143, 196], and hnRNP I is increased in brain tumours [197]. Here we

report that hnRNP A1, hnRNP H and hnRNP I are all over-expressed in cervical cancers compared to histologically normal cervical epithelium. In addition we have identified a number of other RNA-binding proteins including SR proteins, HuR, hnRNP C1/C2 and $U2AF^{65}$ that are over-expressed in cervical cancers. To our knowledge only one RNA-binding protein, poly(rC) binding protein-1 (PCBP-1, also named hnRNP E, hnRNP E1 and PBP), has previously been analysed in cervical epithelium and cervical lesions [198]. Similar to our results these investigators also found a differentiation dependant down-regulation of the RNA-binding protein hnRNP E1 (PCBP-1) as well as an over-expression in high-grade lesions and cervical cancer. We also determined the expression of known markers of differentiation (CK-5, CK-13 and CK-14), proliferation (PCNA) and CIN (p16INK4a) and compared these to the expression of the RNA-binding proteins to establish if they could act as markers for HPV infection, CIN lesions or SCC. While the RNA-binding proteins generally exhibited an expression pattern similar to that of PCNA, no one protein was seen to be more highly expressed than PCNA. p16INK4a is very important in current literature due to its strong correlation with high-grade cervical lesions [199-203], so results of its expression along with RNA-binding protein expression are of value. When compared to p16INK4a expression, which gives a clear indication of high-grade lesions or the presence of cervical cancer, the RNA-binding proteins did not have such a strong correlation with aberrant cell cycle regulation. This is due to the relatively high expression of the RNA-binding proteins seen in histologically normal samples and in low-grade lesions. While RNA-binding protein expression is clearly linked to differentiation, they do not match the present markers for indication of either proliferation or high-grade neoplasia. These results also suggest that the up-regulation of p16INK4a is a later event in the

development of cervical neoplasia than the up-regulation of the RNA-binding protein expression.

When the expression of all proteins were compared between HPV-16 negative cases and HPV-16 positive cases no overall difference in expression profile was noted, although in this study the low number of samples in each group does not allow a comprehensive review of this. This is especially true in the group diagnosed with no CIN and the SCC group as they only contain one HPV-16 negative and one HPV-16 positive patient respectively. To evaluate any differences in expression due to the presence of different HPV types a larger cohort of samples would be required.

In summary, the expression profile of a panel of RNA-binding proteins (hnRNP A1, hnRNP C1/C2, hnRNP H, hnRNP I, ASF/SF2, SR proteins, HuR and $U2AF^{65}$), differentiation markers (CK-13, CK-14 and Syndecan-1) and proliferation markers (PCNA, and p16INK4a) was analysed in histologically normal cervical epithelium and in HPV induced low-grade and high-grade pre-malignant lesions as well as cervical cancer. These results demonstrate different expression levels of the various proteins in basal epithelial cells in histologically normal epithelium followed by an increase in expression in the intermediate layers, whereas the superficial levels remain negative for all tested RNA-binding proteins. Expression of all RNA-binding proteins increases in lesions and highest expression is seen in cervical cancers. The expression profile of the RNA-binding proteins is similar to PCNA expression in histologically normal epithelium as well as in lesions (low-grade and high-grade) and cervical cancers. As PCNA expression has been shown to mimic HPV E6/E7 expression in cervical epithelium [186] the results suggest the RNA-binding proteins analysed here regulate HPV early gene expression directly and late gene

expression indirectly as negative factors that allow late gene expression to occur as they themselves decrease in abundance.

4. Chapter 4 - Investigation and characterisation of the effect of RNA-binding proteins on HPV late gene expression

4.1 Introduction

4.1.1 HPV Late Gene Expression

The HPV late genes, L1 and L2, code for the viral capsid proteins [35]. They are immunogenic proteins that are produced only in the superficial layers of the epithelium [24, 31]. The production of both the L1 and L2 proteins is highly regulated to prevent production of these proteins in the lower epithelial layers, possibly to avoid exposure to the immune system [39]. Both alternative splicing and alternative polyadenylation are used to regulate the expression of these genes. In both HPV-16 and HPV-31, regulatory elements located around the early polyadenylation signal promote its use preventing read-though of this site and production of L2 [75, 177, 204]. In HPV-16 there are two known splice sites that are involved in the expression of the L1 mRNA, SD3632 (located between the E4 and E5 coding regions) and SA5639 (located immediately upstream of the L1 start codon) [39]. Both of these splice sites are actively suppressed by *cis*-acting elements in proliferating cells and inactivation of this suppression must occur in order to produce L1 mRNA [39, 92, 140].

4.1.2 Regulation of Alternative Splicing

Due to the small size of their genome, viruses make extensive use of alternative processing in the regulation of their gene expression [39]. In alternative splicing, the choice of splice site is flexible to allow changes in protein expression during viral infection or the stage of the cell cycle that the infected cell is in [39]. So the question arises as to how it is determined which alternative splice site is selected? Selective regulation of a splice site requires an extra set of signals over and above what is needed for constitutive splicing in assembly of the spliceosome. The choice of a splice site by exon definition (defining the exon by the donor and acceptor sites that flank it [205]) can be regulated by signals located within the exon or intron called splicing enhancers or splicing silencers [206]. These signals can promote or suppress the use of sub-optimal splice sites and so regulate the expression of protein [92, 140].

Exonic splicing enhancers (ESEs) can determine if the exon on which it is located should be included in the transcribed RNA, while exonic splicing silencers (ESSs) cause the skipping of the exon; they can also suppress the splicing of an adjacent intron [207] (Fig. 4.1).

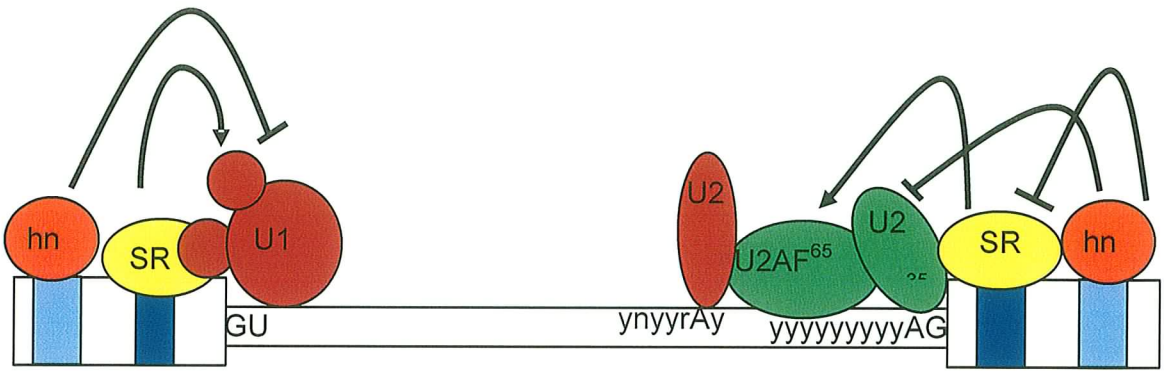


Figure 4.1 Regulation of alternative splicing by SR proteins and hnRNPs

SR proteins (SR) can bind to ESEs (dark blue boxes) and enhance the binding of U1 and U2AF. hnRNPs (hn) can bind to ESSs (light blue boxes) and repress the binding of the spliceosome factors U1 and U2AF and also repress the binding of SR proteins to ESEs.

ESEs can be both purine- or non-purine-rich depending on the consensus sequence of which they are comprised [208]. Purine-rich ESEs usually contain a core sequence of alternating As and Gs and most bind members of the serine-arginine (SR) family of proteins [207]. They can regulate the use of either 3' splice sites, for example in the human fibronectin ED1 gene [209], or 5' splice acceptors, in exon 5 of the caldesmon gene [210]. ESEs can also antagonise the function of ESSs, such as in exon M2 of the IgM gene [211]. Non-purine rich ESEs include the AC-rich enhancers (ACEs) which have been shown to enhance splicing at an upstream 3' splice site in an NS2-specific exon of parvovirus Minute virus of mice [212] and also at a differentiation-specific 3' splice donor in bovine papillomavirus type 1 [213].

ESSs are not as well studied as ESEs, although they may be more common than previously thought and play a crucial role in exon exclusion [214]. ESSs inhibit the use of upstream 3' splice sites and sometimes may antagonise the use of ESEs [215]. ESSs are not as

conserved as the ESEs but they are commonly bound by members of the hnRNP family, including hnRNP A1 and hnRNP H [216, 217].

Although enhancers and silencers located within introns (ISEs and ISSs) are not as well documented as those in exons, there are some examples, such as the ISE in the *Drosophila* Msl2 intron that binds T1A-1 protein stimulating the binding of U1 snRNP to the 5' splice acceptor [218]. ISSs that block the use of downstream 3' splice donors have also been reported, for example a purine-rich intronic sequence in the adenovirus L1 mRNA that blocks splicing when bound by SR proteins during early viral infection [219, 220].

The above *cis*-acting elements (ESEs, ESSs, ISEs and ISSs) act as binding sites for a number of proteins that are involved in regulation of splice-site choice, including the SR proteins and hnRNP proteins. SR proteins are a family of 10 highly conserved proteins involved in both constitutive and alternative splicing [221]. SR proteins contain either one or two RNA recognition motifs (RRMs) at their N-terminus and a serine-arginine-rich domain at their C-terminus. The RRMs are required for sequence-specific binding of the SR proteins to RNA. They are believed to recruit splicing factors to from the initial complexes during spliceosome assembly.

hnRNPs are a large and structurally diverse family of proteins. Some hnRNPs are abundant in the nucleus of vertebrate cells [137] and many have been demonstrated to shuttle between the nucleus and the cytoplasm. A number of different functions have been associated with hnRNPs, including regulation of RNA splicing (hnRNP A1) [222], mRNA polyadenylation (PTB or hnRNP I) [223] and mRNA export (hnRNP L) [224]. hnRNPs generally inhibit splicing by binding to the exonic or intronic regulatory sequences for

example the binding of hnRNP A/B, PTB and hnRNP H to the sequences UAGG, UCUU and polyG motifs respectively [225-227].

4.1.3 Regulation of HPV Gene Expression

The regulation of gene expression in HPV is a complex process involving many *cis*-acting sequences and trans-acting factors [39]. A number of proteins have previously been identified as playing a role in the regulation of gene expression, in this study hnRNP A1, hnRNP H, PTB, ASF/SF2 and HuR were analysed to asses their effect on HPV late gene expression. A number of other proteins that are known to be involved in the regulation of splicing were also examined. These included Adenoviral E4orf4, SR protein SRp30c and HPV-16 E2 [220, 228, 229].

(i) hnRNP A1

hnRNP A1 generally acts as a splicing silencer [139]. In HIV-1 the binding of hnRNP A1 to high affinity sites on the pre-mRNA has been shown to nucleate additional hnRNP A1 molecules [215]. This leads to hnRNP A1 binding low-affinity sites in a co-operative manner. With this co-operative binding, SR proteins are unable to block the binding of hnRNP A1 to the RNA leading to the silencing of splicing [215]. hnRNP A1 represses the use of the late splice site SA5639 by binding to multiple sites within the L1 coding region [140].

(ii) hnRNP H

hnRNP H expression has been reported in some tumours in tissues that usually have low cytoplasmic expression levels [143]. It has also been shown to bind to GGG motifs within

the HPV-16 L2 coding region possibly increasing the efficiency of the early polyadenylation signal [177].

(iii) Polypyrimidine Tract Binding Protein (PTB or hnRNP I)

PTB acts as a splicing silencer by binding to the RNA through its RNA-binding domain thereby blocking the binding of $U2AF^{65}$ to the polypyrimidine tract [145]. PTB contains 4 RRMs and optimal silencer binding sites for the protein contain motifs such as UCUUC and CUCUCU in a pyrimidine-rich region [145]. In HPV-16 PTB binds to the early untranslated region regulating it use [142].

(iv) ASF/SF2

ASF/SF2 is a member of the SR protein family. SR proteins can bind to ESEs and stimulate splicing [80]. ASF/SF2 can regulate HPV-16 gene expression by indirectly interacting with a negative regulatory element, this interaction may be involved in negative regulation of the HPV-16 late genes [136].

(v) HuR

HuR is an RNA-binding protein that binds to RNAs containing AU-rich elements, helping to regulate the lifespan of the RNA. HuR is believed to inhibit late gene expression in HPV-1, HPV-16 and HPV-31 through its interaction with the late untranslated region [148, 150-152].

(vi) Ad E4orf4

The adenovirus E4orf4 protein is a 14kDa protein that has been shown to interact with the cellular protein phosphatase P2 (PP2A), a serine/threonine phosphatase that consists of 3 subunits: core A subunit, catalytic C subunit and regulatory B subunit to which Ad E4orf4 binds [230]. The interaction of Ad E4orf4 and PP2A can act to regulate the alternative

splicing of the adenovirus L1 gene. This gene contains a common 5' splice site and two competing 3' splice sites giving rise to two separate transcripts 52,55K and IIIA. Splicing of the IIIA transcript is under the control of two elements, a repressor element, 3RE, and a virus-infection dependent enhancer, 3VDE. In the early stages of infection SR proteins bind to 3RE repressing IIIA splicing resulting in the 52,55K transcripts being produced exclusively. In the later stages of infection Ad E4orf4 induces the dephosphorylation of the SR proteins. The dephosphoryalted SR proteins no longer bind to the 3RE therefore lifting the repression on the IIIA transcript [220]. Ad E4orf4 has been shown to interact efficiently with 2 SR proteins, ASF/SF2 and SRp30c. Interaction with both PP2A and ASF/SF2 is required to relieve the repression on the splicing of IIIA [231].

(vii) SRp30c

SRp30c is another member of the SR protein family. It is a 25kDa protein [228] that has been shown to help regulate splicing and may possibly have a role in shuttling of proteins into the nucleus [232]. SRp30c is the predominant SR protein in neutrophils and in these cells it is required for the alternative splicing of the glucocorticoid receptor (GCR) pre-mRNA creating mRNA that encodes the GCRβ isoform [233]. SRp30c also has a role in the alternative splicing of hnRNP A1 by repressing the use of a 3' splice site [228]. Over-expression of SRp30c has also been reported to be involved in the shuttling of the Y-box protein-1 (YB-1) into the nucleus, and interaction with this protein has also been demonstrated to have an effect on the selection of alternative splice sites [232].

(viii) HPV-16 E2

The HPV E2 protein plays a role in both the replication and the transcription of the viral genome. It interacts with activators of the anaphase promoting complex (APC) resulting in

the blocking of mitosis and the build-up of genetic instability in the infected cell [234]. There have also been reports that E2 proteins from certain HPV types can help to regulate splicing through interaction with or regulation of SR proteins. HPV-5 E2 has been shown to bind to several members of the SR protein family through the RS-rich hinge region of the protein [229]. The interaction with SR proteins suggest that E2 may have a role in the splicing of HPV-5 mRNA [229]. It has been demonstrated that HPV-16 E2 may regulate splicing of its own late genes by controlling the expression of the SR protein ASF/SF2. HPV-16 E2 can bind a number of cellular transcription factors and through this action may regulate ASF/SF2. ASF/SF2 interacts indirectly with a negative regulatory element in the late region resulting in splicing into this region, therefore regulation of ASF/SF2 expression by HPV-16 E2 would help in regulating late gene expression [136].

4.1.4 mRNA Analysis

While analysis of the DNA of an organism provides much useful information, analysis of the mRNA will give more detail about which genes are actively transcribed in certain cells and if those genes are spliced, either constitutively or alternatively. The HPV-16 genome contains many alternative splice sites and in order to assess which particular protein sequence is being coded for it is necessary to look at which mRNA species are being produced. In this study two RNA analysis techniques are used – Northern Blot and Real-time PCR.

(i) Northern Blot

Northern blot is a type of nucleic acid hybridisation assay which allows the binding of a single-stranded labelled probe to a single-stranded target RNA molecule. After RNA

extraction RNA is run on an agarose gel to separate the molecules by their size [235]. The RNA molecules are denatured (e.g. by exposure to formaldehyde) so the molecules have an unfolded linear conformation [236]. Once separated by their size on the gel, the RNA is transferred by blotting onto a nitrocellulose membrane to which single stranded nucleic acids bind readily [235]. The RNA molecules are immobilised on the membrane and so their position on the membrane is a reflection of their size. The membrane is exposed to a complementary labelled probe that will bind to its target RNA molecule giving off a signal at its location [236]. In these experiments, a ^{32}P radio-labelled probe had been designed to bind to a region in the L1 coding sequence. This probe bound to L1 containing mRNAs on the membrane and when the membrane was exposed to X-ray film the resulting autoradiograph showed their location. The ability to create specific probes for a known RNA sequence allows information to be gathered on the varied mRNA expression in cells and tissues.

(ii) Real-time PCR

The basic principal of real-time PCR is much the same as conventional PCR – a target sequence is amplified and is then subjected to gel electrophoresis to visualise results. In real-time PCR there are some differences, the accumulation of amplified material is detected and quantified using a fluorescent reporter and the accumulation of amplicon is measured during each cycle [237]. Here SYBR Green I is used as the fluorescent reporter. This is a fluorescent dye that will bind only to the minor grove of double stranded DNA [237, 238]. Thus, the amount of fluorescence that is emitted is proportional to the amount of double stranded DNA present in the reaction [238]. SYBR Green is excited at 498nm

and emits light at a wavelength of 522nm [237]. Due to the fact that SYBR Green will bind to any double stranded DNA is it important to perform a melt curve analysis at the end of each run to ensure that a single product has been amplified and also to ensure that primer-dimers have not interfered with the reaction [238]. A melt curve is produced by heating the sample in small increments resulting in melting of the PCR product and releasing of the SYBR Green into solution. Once SYBR Green is in solution its fluorescence decreases. The real-time PCR software will then generate a graph of fluorescence intensity over temperature indicating the melting temperature (Tm) of the PCR product where fluorescence decreased [238]. The real-time PCR has amplified the specific product if the Tm is close to its expected Tm (analysed by gel electrophesis). By designing specific primers, individual mRNA species can be analysed for changes in expression and/or splice sites used due to experimental conditions.

4.1.5 Aims

- To evaluate the role of selected splice regulatory proteins in HPV-16 late gene expression

Objectives

- To co-transfect sub-genomic HPV-16 expression plasmids with selected protein expression plasmids in HeLa cells
- To asses, using RT-PCR, if any protein used previously has an effect on late gene expression in HPV-16
- To examine, using Northern blot analysis, the effect of splicing regulatory proteins and other viral proteins on the regulation of late gene expression in HPV-16
- To optimise a real-time RT-PCR to quantify changes in late gene expression
- To use the real-time RT-PCR system to quantify changes in late gene expression found by RT-PCR and Northern blot

4.2 Methods

In order to determine if any of the proteins that were examined previously had an effect on the HPV-16 late gene expression, transfection studies were carried out in HPV-18 positive HeLa cells and HPV negative A549 cells (a lung adenocarcinoma cell line) using protein expression plasmids. The protein expression plasmids were co-transfected individually into HeLa or A549 cells with one of three HPV-16 sub-genomic expression plasmids, pC97EL, pBEL or pBELM. Total RNA was extracted and mRNA was examined using either Northern blot, RT-PCR or real-time PCR.

4.2.1 Cell Culture

HeLa and A549 cells were grown in complete RPMI medium (Sigma) (including 2mmol/l L-glutamine (Gibco), 10% FCS (Gibco) and 40U penicillin /40mg streptomycin (Gibco)) at 37°C and 5% CO_2 in a humidified atmosphere. Cells were grown to confluence before being trypsinised with 2% trypsin (Gibco) solution for 10 min at 37°C. 10ml complete RPMI medium (Sigma) was added to the culture dish to inactivate trypsin. 5 x 10^4 cells were added to a 6mm cell culture dish containing 2ml complete RPMI and were grown overnight at 37°C and 5% CO_2 in a humidified atmosphere overnight. 5 x 10^4 cells were added to a 10mm cell culture dish to continue the culture and cells were grown to confluence. HeLa cells were used for all transfections unless otherwise stated.

4.2.2 Transfection

100µl of serum free medium (SFM) (Sigma) was added to a sterile eppendorf. 3µl Genejuice (Novagen) per 1µg DNA was added drop-wise to the SFM and was mixed thoroughly by vortexing. The mixture was incubated at room temperature for 5 min. 1µg of each plasmid to be transfected was added to the SFM/Genejuice mixture, mixed by gentle pipetting and incubated at room temperature for 15 min. The entire volume was added drop-wise to the cell culture in complete RPMI, and the dish was rocked to distribute the transfection mixture. The cells were incubated at 37°C and 5% CO_2 in a humidified atmosphere. Twenty-four hours post-transfection the culture medium was poured off and the plate was washed with ice-old PBS (appendix I) before total RNA was extracted using the Qiagen RNeasy Mini kit. Three sub-genomic HPV-16 expression plasmids were used in this study (Fig. 4.2 shows the 2 main plasmids). The first, pC97EL contains the full HPV-16 genome but with the weak viral promoter replaced by the strong human cytomegalovirus (CMV) immediate-early promoter (Fig.4.2). All early and late splice sites are present in the plasmid allowing the examination of the full range of splicing events that can occur in HPV-16. The second plasmid used was pBEL, this plasmid contains all HPV-16 genes except E6 and E7, this plasmid also has a strong CMV promoter replacing the weak viral promoter (Fig. 4.2). pC97EL and pBEL when transfected alone will only express the early genes, no late genes will be produced. Late gene expression can only be induced from these plasmids by co-transfection with protein expression plasmids that can either stimulate read-through of the early polyadenylation signal (pAE), or splicing to the late splice acceptor, SA5639. A third plasmid was used in this study, pBELM. This plasmid contained a slight modification of the pBEL plasmid within the L1 coding region. The late splice acceptor,

SA5639, is negatively regulated by hnRNP A1 binding to multiple sites within L1, in this plasmid these sites have been mutated to reduce the silencing of this splice acceptor [140]. This allows the pBELM plasmid to express the L1 gene when transfected alone.

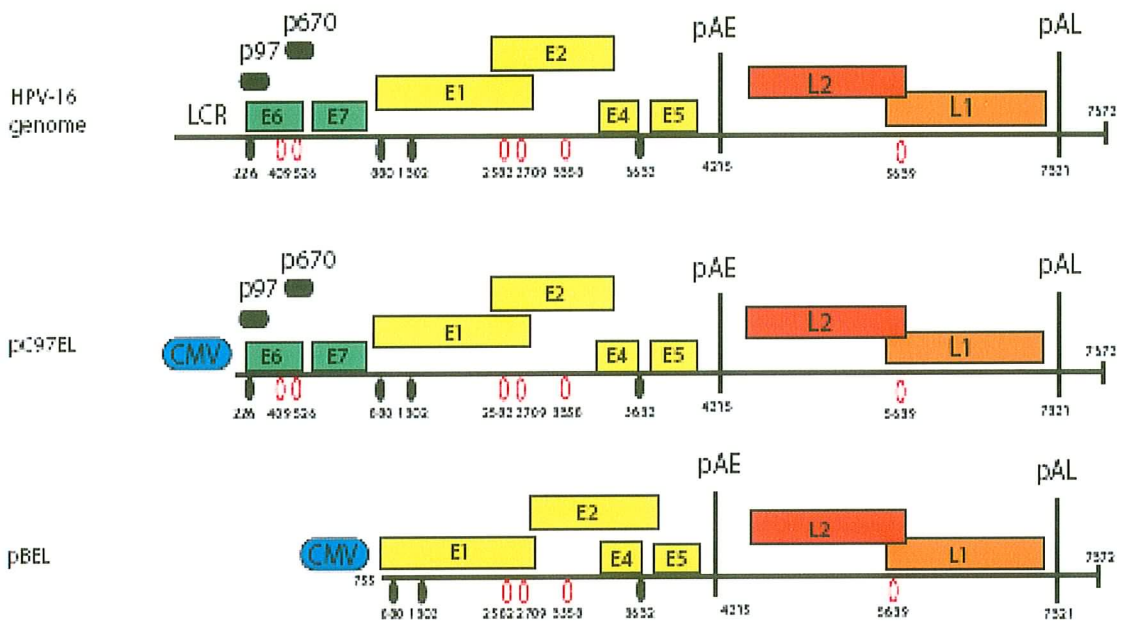


Figure 4.2 The HPV-16 complete genome and sub-genomic expression plasmids pC97EL, pBEL

Schematic representation of the HPV-16 genome. Boxes indicate protein-coding regions. The major promoter, p97, and the differentiation dependent promoter, p670, are shown. Positions for 5´splice sites or splice donors (SD) and 3´splice sites or splice acceptors (SA) are identified; numbers refer to nucleotide positions in the HPV-16 sequence. The early and late polyadenylation signals, pAE and pAL respectively, are also indicated. CMV – human cytomegalovirus immediate-early promoter, black circle – splice donor, red circle – splice acceptor.

4.2.3 RNA Extraction

Total RNA extraction was carried out with the Qiagen RNeasy Mini Kit using the following protocol:

Direct lysis of the transfected cells was carried out by adding 600µl Buffer RTL (Qiagen) to the cell culture dish. The lysate was collected with a cell scraper (Sarstedt) and was pipetted into a sterile 1.5ml eppendorf before being vortexed to mix. The lysate was homogenised by applying the mixture directly into a QIAshredder (Qiagen) and was spun for 2 min at full speed. 1 volume of 70% ethanol (appendix I) was added to the homogenised lysate and was mixed by pipetting. 700µl of the sample was transferred to an RNeasy spin column (Qiagen) and was spun at 10000rpm for 15 seconds. The flow-through was discarded and the remainder of the sample was added to the same spin column and spun as before. 700µl Buffer RW1 (Qiagen) was added to the column and was spun at 10000rpm for 15 seconds. The flow-though was discarded. 500µl Buffer RPE (Qiagen) was added to the column and spun at 10000rpm for 15 seconds. The flow-though was again discarded. 500µl Buffer RPE was added to the column and spun at 10000rpm for 2 min. Again the flow-though was discarded. The RNeasy column was placed in a new collection tube and spun again at full speed for 1 min. The column was placed in a 1.5 ml eppendorf tube and 20µl RNase-free water (Qiagen) was added to the column. The column was spun at 10000rpm for 1 min. The eluted 20µl was reapplied to the same column and was spun again at 10000rpm for 1 min. The sample was then either DNase treated directly, or was stored at -80°C before being DNase treated.

4.2.4 DNase Treatment

To the 20µl RNA sample, 1µl 10x DNase I buffer (Invitrogen) and 1µl DNase I (Invitrogen) was added. The sample was mixed gently by pipetting and was incubated at room temperature for 15 min. The DNase I was inactivated by the addition of 1µl EDTA (25mM) (Invitrogen) to the mixture and heating to 65°C for 10 min. 25µl phenol/chloroform (BDH) was added to the solution and was mixed by vortexing. The sample was spun at full speed for 2 min at 4°C. The upper layer was transferred to a new tube. 1/10 volume of 3mol/l sodium acetate (appendix I) and 2.5 volumes 100% ethanol (Merck) was added and mixed by vortexing. The mixture was placed at -80°C for at least 2 hours. The mixture was spun at full speed for 20 min at 4°C. The supernatant was removed by pipetting and the pellet was air-dried. The pellet was resuspended in 15µl RNase-free water (Qiagen) and stored at -80°C.

4.2.5 Northern Blot

(i) Agarose-Formaldehyde Gel

A 1% agarose-formaldehyde gel (appendix I) was prepared and after it had set it was pre-run in 1x MOPS buffer (appendix I) for 5 min at 80V. 5µg RNA was added to a sterile eppendorf, 2.5µl 10x MOPS buffer (appendix I), 4.5µl 37% formaldehyde (BDH) and 12.5µl formamide (BDH) were also added to the tube. The mixture was incubated at 60°C for 15 min. The sample was spun briefly and placed on ice. 5µl RNA loading dye (appendix I) was added to the sample before loading onto the gel. The gel was run at 80V for 3 to 4 hours.

(ii) Blotting the Gel

Nitrocellulose membrane (Sigma) was cut to the size of the gel and placed in 20x SCC (appendix I) for 5 min. The blotting sandwich was set up (Fig. 4.3), and the gel was blotted overnight. After blotting the membrane was dried by heating it under pressure for 2 hours. Dried membranes were stored at 4°C for subsequent probing.

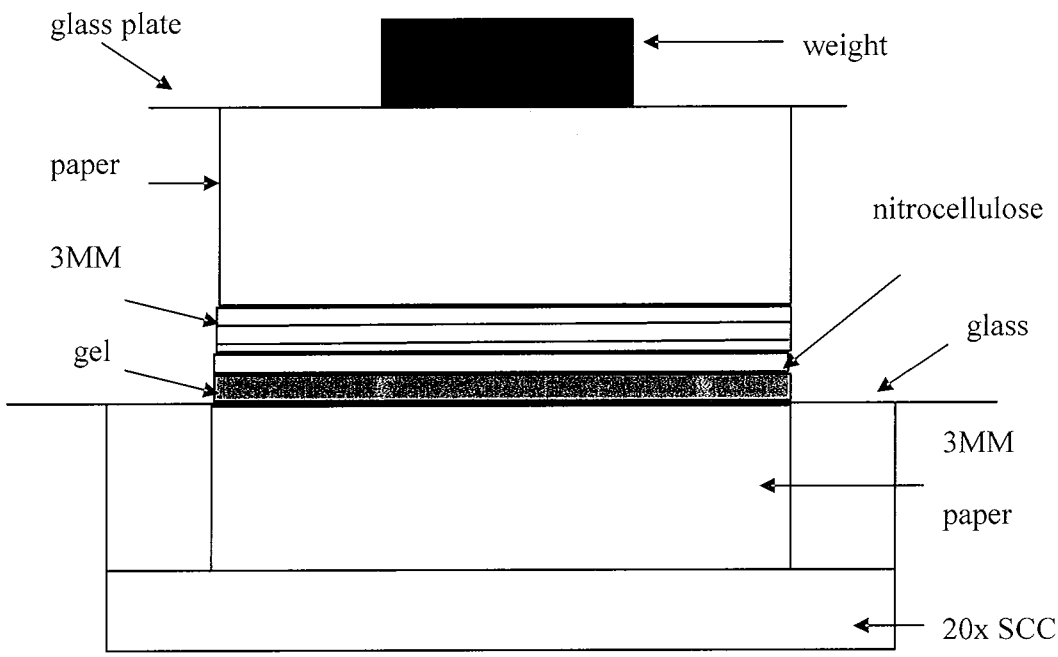


Figure 4.3 Composition of the Northern blot sandwich

(iii) Probing the Membrane for the L1 Gene

Radio-labeled probe, specific for the HPV L1 coding region, was made in the following way using the DECAprime™ II kit (Ambion):

To produce the template DNA a HPV-16 sub-genomic expression plasmid was digested using the restriction enzymes BamH1 (New England BioLabs) and Xho1 (New England

BioLabs), generating a 1005bp DNA sequence that binds to the HPV-16 L1 sequence between base-pair 6152 and base-pair 7157. The template DNA was diluted to approximately 1–2.5 ng/µl. 2.5µl of 10x decamer solution (Ambion) was added and the mixture was incubated at 95–100°C for 3–5 min. The mixture was snap-frozen in liquid nitrogen, thawed, briefly centrifuged and placed on ice. The Klenow reaction was assembled on ice by adding the following components to the DNA/decamer mixture: 5µl dCTP 5x reaction buffer (Ambion), 5µl [α-32P]dCTP (Ambion), water to bring volume to 24µl and 1µl Exo-Klenow (Ambion). The reaction was mixed by gently flicking or pipetting. The mixture was incubated at 37°C for 5–10 min. 1µl of 0.5mol/l EDTA (appendix I) was added to stop the reaction. The probe was either used immediately or it was stored at –20°C.

Once the probe was ready, the membranes to be labeled were wetted with 6x SCC (appendix I). The membranes were placed (RNA side facing in) in hybridisation tubes. 1ml of pre-hybridisation solution (appendix I) was added to the tubes and was left at 42°C for at least 1 hour. The pre-hybridisation solution was poured off and approximately 50µl of radio-labeled probe was added. 1ml hybridisation solution (appendix I) was also added and the tube was incubated at 42°C overnight. After overnight incubation the membrane was washed in the following way: two 5 min washes with 2x SCC, 0.1% sodium dodecylsulfate (SDS) (appendix I) at room temperature, two 5 min washes with 0.2x SCC, 0.1% SDS (appendix I) at room temperature, two 15 min washes with 0.2x SCC, 0.1% SDS (appendix I) at 42°C and two 15 min washes with 0.1x SCC, 0.1% SDS (appendix I) at 42°C. The radioactivity of the membrane was monitored at the end of each wash and the washing steps were stopped when a difference between the background radiation and the probed

areas was noted. The membrane was then exposed to autoradiography film for varying lengths of time (between 12 and 96 hours).

4.2.6 Reverse Transcription of extracted RNA to cDNA

1µl Random Primers (100ng) (Invitrogen) and 1µl dNTP mix (10mmol/l) (Invitrogen) were added to a sterile eppendorf. 200ng RNA was added to the tube and the mixture was made up to 12µl using RNase-free water (Qiagen). The mixture was heated to 65°C for 5 min and then quickly chilled on ice. The tube was spun quickly to collect all its contents. 4µl 5x first strand buffer (Invitrogen) and 2µl dithiothreitol (DTT) (0.1mol/l) (Invitrogen) were added to the tube. The contents were mixed gently and incubated at 37°C for 2 min. 1µl Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT) (Invitrogen) was added to the tube and was incubated at 25°C for 10 min. The mixture was incubated at 37°C for 50 min and the reaction was inactivated by heating to 70°C for 15 min. The cDNA was stored at -20°C until PCR could be carried out.

4.2.7 PCR

Presence of cDNA was confirmed by PCR analysis using GAPDH primers (Table 4.1 and 4.2) and transfection products were analysed using primer sets detailed in Table 4.2.

Primer	Sequence (5'-3')
GAPDHF	AGGTCGGAGTCAACGGATTTGG
GAPDHR	ACAGTCTTCTGGGTGGCAGTGATG
p97s	GCGCGCGTCGACCTGCAATGTTTCAGGACCCACAGG
757s	TCGGTTGTGCGTACAAAGCACACACG
3515s	GCGCGCTCTAGACGACTATCCAGCGACCAAC
M10a	CCGAACGTCGACTCACACTTGCAACAAAAGG
E4a	GCGCGCCTGCCTAATAATTTCAGG
3515a	CTTGGTCGCTGGATAGTCG
L1a wild-type (wt)	CCGTGCTTACAACCTTAGATACTGGGACAG
L1a mutant (mut)	AAGCTTGGATCCCGCTGGGCAGCCACAG

Table 4.1 Sequences of primers used in the experiments for HPV-16 late gene analysis.

Primer Pair	Annealing Temperature	Gene(s) Amplified	Size of Product
GAPDHF – GAPDHR	55°C	GAPDH	550bp
p97s – M10a	55°C	E6 / E7	400bp / 500bp
p97s – E4a	55°C	E4	180bp / 700bp
p97s – 3515a	55°C	E4	260bp / 800bp
p97s – L1a (wt)	55°C	E4 / L1	450bp / 900bp
757s – E4a	55°C	E4	190bp
757s – L1a (wt)	55°C	E4 / L1	200bp / 470bp
757s – L1a (mut)	55°C	E4 / L1	160bp / 430bp
3515s – L1a (wt)	55°C	E4 / L1	250bp

Table 4.2 Primer pairs, annealing temperature, HPV-16 genes amplified and expected product size

All PCR reactions were carried out in a final volume of 50µl containing 1x reaction buffer (200mmol/l Tris-HCl (pH8.4), 500mmol/l KCl) (Invitrogen), 2mmol/l dNTPs (Invitrogen), 3.75mmol/l $MgCl_2$ (Invitrogen), 0.5mmol/l forward primer, 0.5mmol/l reverse primer, 1U Platinum Taq polymerase (Invitrogen) and 1µl template DNA (Table 4.3). The PCR protocol was as follows: an initial 5 min denaturation step at 94°C, followed by a cycle of 30 sec at 94°C, 30 sec at 55°C, 1 min at 72°C and a final elongation step of 5 min at 72°C (Table 4.4). 40 cycles were used for all PCRs.

Solution	Concentration	1x
Buffer (Invitrogen)	10x	5µl
dNTP mix (Invitrogen)	10mmol/l	5µl
$MgCl_2$ (Invitrogen)	50mmol/l	3.75µl
Forward primer	10µmol/l	2.5µl
Reverse primer	10µmol/l	2.5µl
Water	-	29.75µl
Taq polymerase (Invitrogen)	5U/µl	0.5µl
cDNA	-	1µl
Total		50µl

Table 4.3 PCR reaction mix for GAPDH, early HPV-16 region and late HPV-16 region PCR

Temperature	Time
94°C	5 min
94°C	30 sec
55°C	30 sec
72°C	1 min
72°C	5 min

Table 4.4 PCR cycling conditions for GAPDH, early HPV-16 region and late HPV-16 region PCR

4.2.8 Gel Electrophoresis

All PCR products were analysed on 5% acrylamide gels (appendix I) or 1% agarose gels (appendix I). All gels were run at 80V, acrylamide gels were run until the 1st dye band reached the end of the gel (approx 40 min) and agarose gels were run for 1 hour and 15 min. Bands were visualised using ethidium bromide staining (0.5μg/ml ethidium bromide dissolved in gel).

4.2.9 Real-time PCR

Real-time PCR was carried out on reverse transcribed cDNA products using the GAPDH and 757s-L1a (wild-type (wt) and mutant (mut)) primer pairs listed in Table 4.2. For all real-time PCR experiments transfections were carried out on 3 separate occasions and all real-time PCRs were carried out 3 times. All data on CT values (threshold cycle – the point at which the fluorescence crosses the threshold) was normalised to an internal house-keeping gene, GAPDH, in order to control for differences in starting cDNA concentration.

All PCRs were carried out using LightCycler FastStart DNA Master SYBR Green I Kit (Roche). A 9μl PCR reaction mix (without template cDNA) was prepared as detailed in Table 4.5 and added to pre-cooled LightCycler capillaries (Roche). 1μl of template DNA was added to the capillaries. The real-time PCR protocol was as follows (Table 4.6): an initial 10 min denaturation step at 95°C, followed by a cycle of 5 sec at 95°C, 5 sec at 55°C, 15 sec at 72°C. The number of cycles varied for each primer set used, 35 cycles for GAPDH and 757s-L1a(mut) primers, and 40 cycles for 757s-L1a(wt). A melt curve analysis was preformed on all reactions, this comprised of denaturation for 5 sec at 95°C, annealing for 10 sec at 75°C for GAPDH and 70°C for both 757s-L1a(wt) and 757s-

L1a(mut), and a melt curve (with an increase in temperature of 0.1°C/sec) up to 97°C for GAPDH, 95°C for 757s-L1a(wt) and 98°C for 757s-L1a(mut). The LightCycler was cooled at 30°C for 1 min.

Solution	Concentration	1x
Water	-	6.6µl for 757s-L1a(wt) 6.2µl for GAPDH and 757s-L1a(mut)
$MgCl_2$ (Roche)	25mmol/l	0.4µl for 757s-L1a(wt) 0.8µl for GAPDH and 757s-L1a(mut)
Forward Primer	10mmol/l	0.5µl
Reverse Primer	10mmol/l	0.5µl
FastStart DNA Master SYBR Green 1 mix (Roche)	10x	1µl
cDNA	-	1µl
Total		10µl

Table 4.5 PCR reaction mix for GAPDH, 757s-L1a(wt) and 757s-L1a(mut) real-time PCRs

Analysis Mode	Cycles	Segment	Target Temperature	Hold Time	Acquisition Mode
Pre-Incubation					
None	1		95°C	10 min	none
PCR Amplification					
		Denaturation	95°C	5 seconds	none
Quantification	35-40*	Annealing	55°C	5 seconds	none
		Extension	72°C	15 seconds	single
Melting Curve					
		Denaturation	95°C	5 seconds	none
Melting Curve	1	Annealing	70-75°C**	10 seconds	none
		Melting	95-98°C*** (slope 0.1°C/sec)	0 seconds	continuous
Cooling					
None	1		30°C	1 min	none

Table 4.6 Real-time PCR cycling conditions for GAPDH, 757s-L1a(wt) and 757s-L1a(mut) primer pairs

* indicates number of cycles varied for each real-time PCR, GAPDH and 757s-L1a(mut) – 35 cycles, 757s-L1a(wt) – 40 cycles. ** annealing temperature varied for each real-time PCR, GAPDH - 75°C, 757s-L1a(mut) and 757s-L1a(wt) - 70°C. *** melting temperature varied for each real-tine PCR, GAPDH - 97°C, 757s-L1a(mut) - 98°C, 757s-L1a(wt) - 95°C.

4.2.10 Real-time PCR Analysis

All real-time PCRs were analysed for specificity by melt curve analysis to ensure only specific peaks for expected products were present. CT values for all GAPDH and 757s-L1a(wt and mut) were inputted into MS Excel. GAPDH from each individual transfection was used to normalise the L1a values for that transfection. Peak area data was calculated using the LightCycler 3 software and this data was also inputted into MS Excel. MS Excel was used to calculate L1a averages and peak area percentages for graphs and standard error for error bars. T-tests were carried out on L1a CT values to determine significance, the 'T-Test: Two-Sample Assuming Equal Variances' in MS Excel was used for this calculation.

4.2.11 Transfection Efficiency Assay

The transfection efficiency of both HeLa and A549 cell lines used was established using a β-galactosidase liquid assay. 2μg of pCMVSport plasmid (Gibco), which expresses β-galactosidase, was transfected into 12 separate plates of the cell line using the Genejuice method described in section 4.2.2. Two plates were left untransfected as a negative control. Twenty-four hours post-transfection the medium was aspirated off the cells and the plates were washed twice with 2ml PBS (appendix I). After all the PBS was removed, 200μl of lysis buffer (Promega) was added to the plates and they were incubated at room temperature for 15 min. The plates were periodically rocked to ensure full coverage of the plate. The cells were scraped from the plate and the lysate was collected at the bottom of the plate. The cells were transferred to 1.5ml eppendorf tubes, vortexed and spun at maximum speed for 2 min. 150μl of the supernatant was transferred to a new 1.5ml eppendorf tube and 150μl of 2x assay buffer (appendix I) was added. The tube was

vortexed to mix. The tube was incubated at 37°C for 5-30 min until a faint yellow colour developed. 500µl of 1mol/l sodium carbonate (appendix I) was added to the tube to stop the reaction. The absorbance of the samples at 420nm was recorded.

4.3 Results

4.3.1 Sub-genomic Expression Plasmids

Three plasmids were used in transfection studies for the analysis of HPV late gene expression, pC97EL, pBEL and pBELM. All three plasmids have a strong CMV promoter instead of the weak viral promoter. pC97EL contains the full HPV-16 genome and will constitutively express the genes within the early region (Fig. 4.4). This plasmid transfected alone will not express detectable levels of the late genes L1 and L2: these can only be detected when co-transfected with proteins that can induce HPV late gene expression.

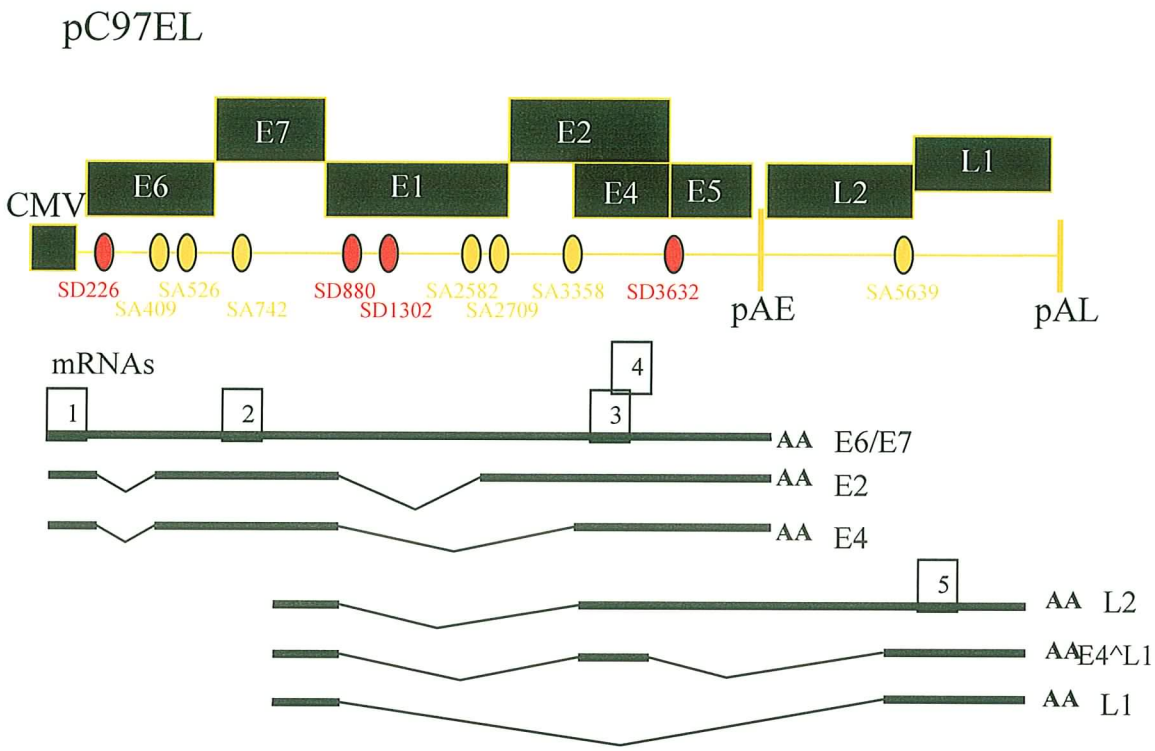


Figure 4.4 pC97EL plasmid and the major mRNA species it can produce

CMV – cytomegalovirus promoter, pAE – early polyadenylation signal, pAL – late polyadenylation signal. White boxes on mRNAs indicate approximate position of primers, 1 – p97s, 2 – M10a, 3 – E4a, 4 – 3515a, 5 – L1a(wt). Thick black lines indicate sections of the mRNA that remain after RNA splicing and will be used for translation, thin black lines indicate sections of the mRNA removed by splicing.

pBEL and pBELM do not contain the very early part of the HPV-16 genome, these plasmids therefore do not code for the E6 or E7 proteins (Fig. 4.5). When transfected alone the expression of E1, E2, E4 and E5 can be detected from both plasmids. Like pC97EL, pBEL does not produce any detectable levels of the late mRNAs. pBELM can express detectable levels of the late genes when transfected alone as it contains a mutation in the L1 coding region that destroys splicing silencers that bind hnRNP A1, decreasing the negative regulation on the SA5639 splice site (Fig. 4.5). The two main sets of primers that are used to detect RNAs expressed from these plasmids are the 757s-E4a and 757s-L1a primers.

These primers use a common forward primer, 757s that binds just after the CMV promoter in both pBEL and pBELM. The E4a reverse primer binds within the E4 coding region, approximately 20bp down-stream from the SA3358 splice site, and the L1a reverse primer binds within the L1 coding region approximately 40bp downstream from the SA5639 splice site. As this area is mutated to destroy the splicing silencers that are located here two L1a primers were required, L1a(wt) that binds to the un-mutated pBEL and L1a(mut) that binds to the mutated pBELM. As the L1a primers (both wt and mut) are designed to specifically detect the mRNAs that contain the L1 coding region they cannot detect the mRNAs that contain the L2 coding region. This means that a negative PCR result using these primers cannot exclude induction of the late genes if only L2 is induced.

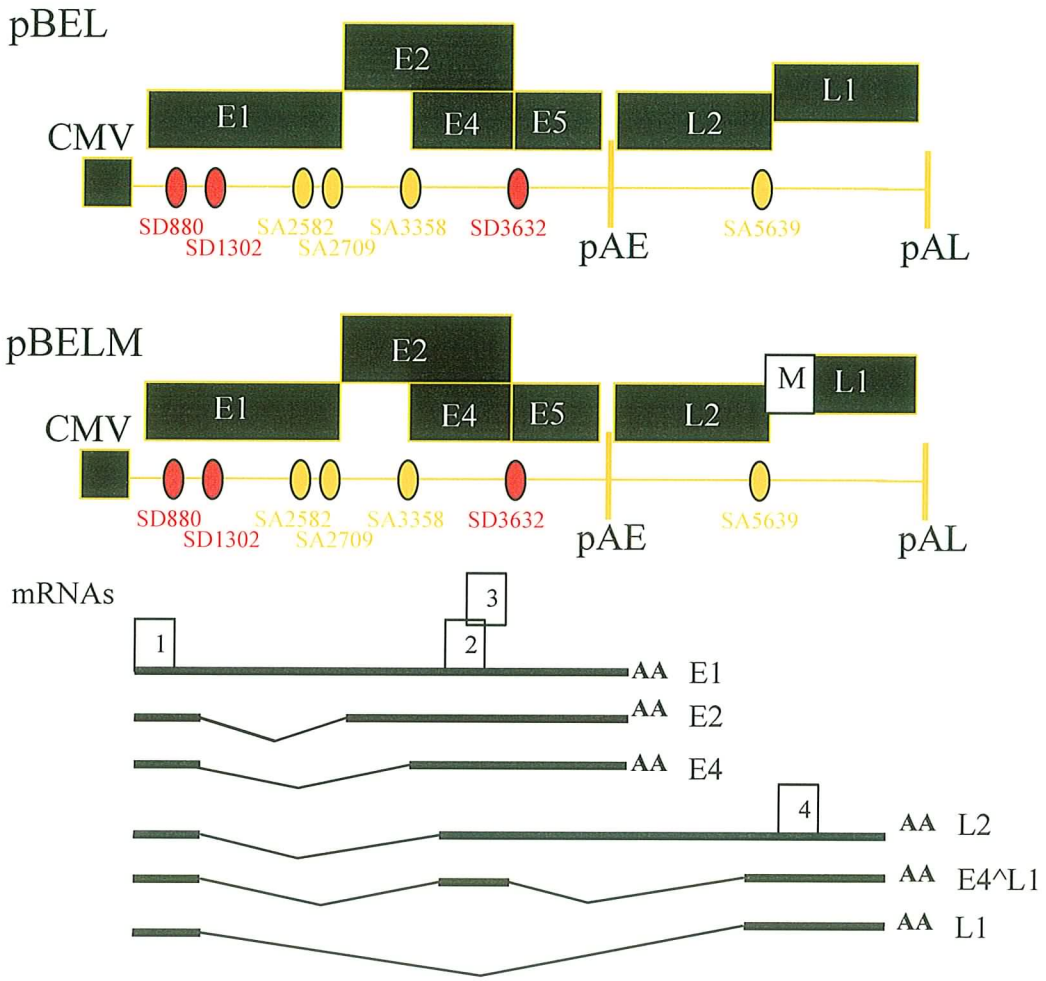


Figure 4.5 pBEL and pBELM plasmids and the major mRNA species they can produce

CMV – cytomegalovirus promoter, pAE – early polyadenylation signal, pAL – late polyadenylation signal, M – mutation located within the L1 coding region. White boxes on mRNAs indicate approximate position of primers, 1 – 757s, 2 – E4a, 3 – 3515s and 3515a, 4 – L1a(wt) and L1a mut. Thick black lines indicate sections of the mRNA that remain after RNA splicing and will be used for translation, thin black lines indicate sections of the mRNA removed by splicing.

4.3.2 The Effect of Viral and RNA-Binding Proteins on HPV Late Gene Expression

A number of proteins that have been previously identified as being involved in splice regulation were examined to identify if they could induce the late genes in the pBEL and the pC97EL plasmids. Two viral proteins, HPV-16 E2 and Adenoviral E4for4, and one protein from the SR protein family, SRp30c, were analysed. Protein expression plasmids for these three proteins were co-transfected individually into HeLa cells with either pBEL or pC97EL to assess the effect on late gene expression. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) PCR was performed on all samples to ensure the presence and quality of cDNA for PCR amplification. GAPDH primers amplify a 550bp region of the GAPDH gene. All samples were positive for GAPDH cDNA (Fig. 4.6).

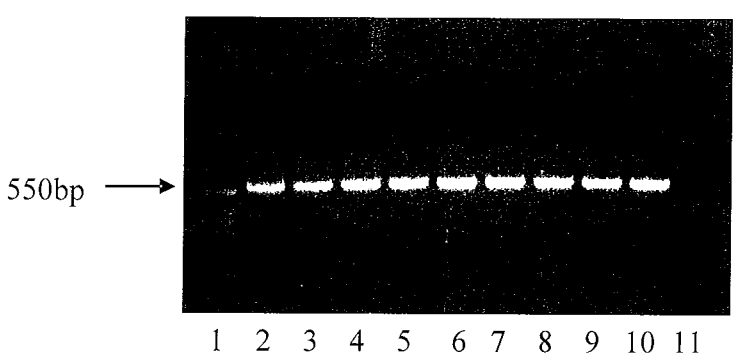


Figure 4.6 PCR amplification of the GAPDH cDNA using the GAPDHF/GAPDHR primers

Lane 1 1kb DNA ladder, lane 2 positive control, lane 3 pBEL, lane 4 pBEL + Ad E4orf4, lane 5 pBEL, lane 6 pBEL + HPV-16 E2, lane 7 pC97EL, lane 8 pC97EL + Ad E4orf4, lane 9 pC97EL, lane 10 positive control, lane 11 negative control. Positive control, RNA extracted from untransfected HeLa cells. Products were run on a 5% acrylamide gel.

HPV-16 E2 was found to be able to induce the expression of late RNA products by RT-PCR using two separate primer pairs (Fig. 4.7). The late RNA products were only observed when the HPV-16 E2 protein was expressed alongside the pBEL HPV expression plasmid using both the 3515s-L1a(wt) and the 757s-L1a(wt) primer pairs. These results were similar to the results for the Adenoviral E4orf4 protein, the late gene products were only detected when Ad E4rof4 is expressed (Fig. 4.8).

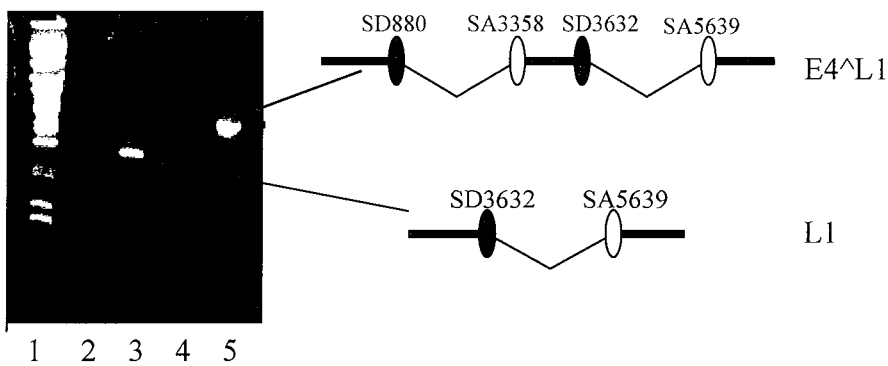


Figure 4.7 PCR amplification of the late HPV-16 region using the 3515s-L1a(wt) and 757s-L1a(wt) primer pairs after co-transfection with pBEL and HPV-16 E2

Lane 1 DNA ladder, lane 2 pBEL, lane 3 pBEL + HPV-16 E2, lane 4 pBEL, lane 5 pBEL + HPV-16 E2. Lanes 2 and 3, 3515s-L1a(wt) primers used, lanes 4 and 5, 757s-L1a(wt) primers used. Products run on a 5% acrylamide gel.

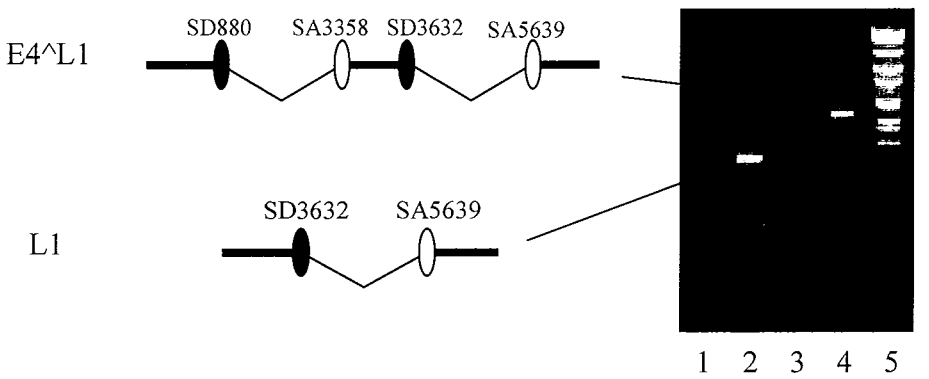


Figure 4.8 PCR amplification of the late HPV-16 region using the 3515s-L1a(wt) and 757s-L1a(wt) primer pairs after co-transfection with pBEL and Adenoviral E4orf4

Lane 1 pBEL, lane 2 pBEL + Ad E4orf4, lane 3 pBEL, lane 4 pBEL + Ad E4orf4 Lane 5 DNA ladder. Lanes 1 and 2, 3515s-L1a(wt) primers used, lanes 3 and 4, 757s-L1a(wt) primers used. Products run on a 5% acrylamide gel.

In order to determine if Ad E4orf4 and SRp30c had an effect on the early gene products as well as the late gene products, the proteins were co-transfected using the pC97EL HPV-16 expression plasmid that contained both the early and late regions. RT-PCR was carried out using 3 primer pairs for the early region and one primer pair for the late region to examine the mRNA species produced. No difference in splicing pattern or intensity of any band in the absence or presence of the Ad E4orf4 protein was detected in the early region using the primer pairs p97s-E4a, p97s-3515s and p97s-M10a (Fig. 4.9). When Ad E4orf4 was expressed the late genes were induced (Fig. 4.9). This study also determined that SRp30c expression did not affect early gene expression but it induced a stronger expression of the late gene products (Fig. 4.10).

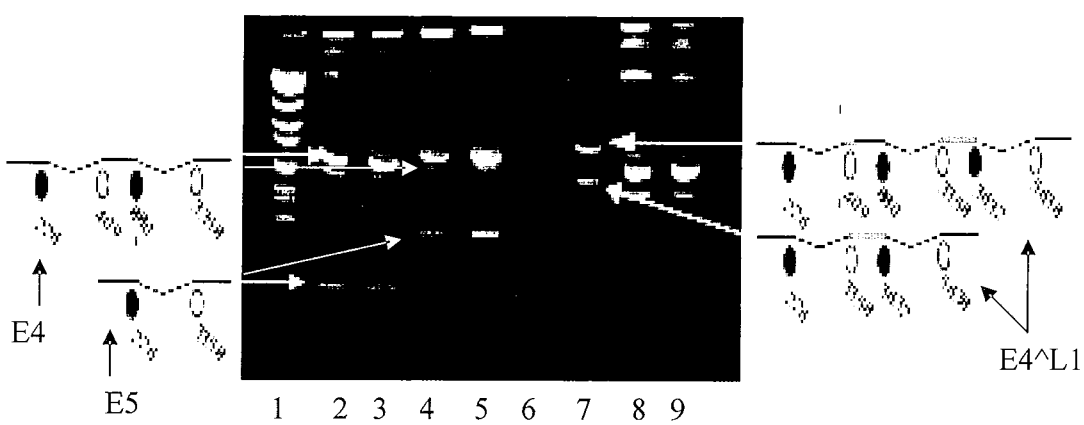


Figure 4.9 PCR amplification of the early and late HPV-16 region using the p97s-E4a, p97s-3515a, p97s-L1a(wt) and p97s-M10a primer pairs after co-transfection with pC97EL and Adenoviral E4orf4

Lane 1 DNA ladder, lane 2 pC97EL, lane 3 pC97EL + E4orf4, lane 4 pC97EL, lane 5 pC97EL + E4orf4, lane 6 pC97EL, lane 7 pC97EL + E4orf4, lane 8 pC97EL, lane 9 pC97EL + E4orf4. Splice products in lanes 8 and 9 are the following: lower band, a splice from SD226 to SA526, upper band, a splice from SD226 to SA409. The following primers were used: lane 2 and 3 - p97s-E4a primers, lane 4 and 5 - p97s-3515a primers, lane 6 and 7 - p97s-L1a(wt) primers and lane 8 and 9 - p97s-M10a primers. Products run on a 5% acrylamide gel.

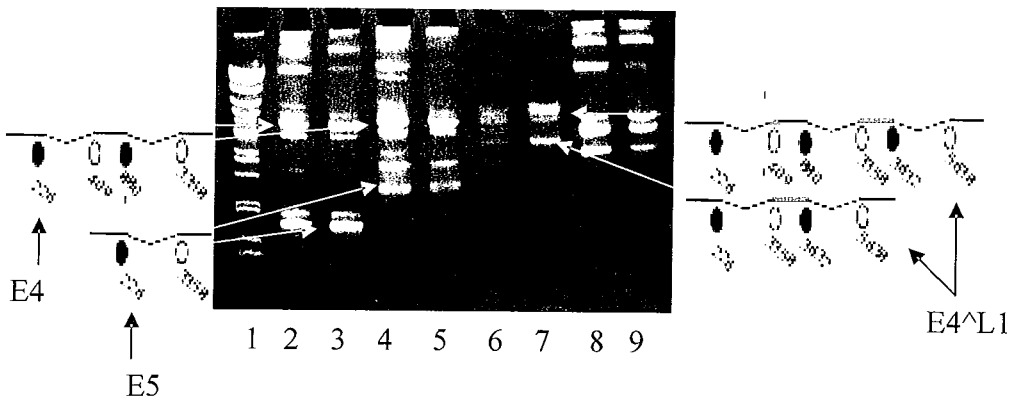


Figure 4.10 PCR amplification of the early and late HPV-16 region using the p97s-E4a, p97s-3515a, p97s-L1a(wt) and p97s-M10a primer pairs after co-transfection with pC97EL and SRp30c

Lane 1 DNA ladder, lane 2 pC97EL, lane 3 pC97EL + SRp30c, lane 4 pC97EL, lane 5 pC97EL + SRp30c, lane 6 pC97EL, lane 7 pC97EL + SRp30c, lane 8 pC97EL, lane 9 pC97EL + SRp30c. Splice products in lanes 8 and 9 are the following: lower band, a splice from SD226 to SA526, upper band, a splice from SD226 to SA409. The following primers were used: lane 2 and 3 - p97s-E4a primers, lane 4 and 5 - p97s-3515a primers, lane 6 and 7 - p97s-L1a(wt) primers and lane 8 and 9 - p97s-M10a primers. Products run on a 5% acrylamide gel.

4.3.3 Effect of RNA-Binding Proteins on HPV Late Gene Expression

In order to determine if any of the proteins that were examined previously had an effect on the HPV-16 late gene expression, transfection studies were carried out in HeLa cells using HPV-16 sub-genomic expression plasmids and plasmids expressing the various proteins. Five plasmids that expressed proteins that had been previously studied using immunohistochemistry were available for the transfection studies. These proteins were ASF/SF2, hnRNP A1, hnRNP H, HuR and PTB. These 5 proteins have all been reported to be involved in the regulation of HPV gene expression. ASF/SF2 possibly helps to regulate late mRNA transcripts by its interaction with the HPV-16 negative regulatory element

[136]. hnRNP A1 binds to multiple silencers in the L1 coding region, inhibiting the use of the late 3' splice site [140]. hnRNP H binds to GGG motifs within the L2 coding region, this binding correlates with efficient use of the early polyadenylation signal [177]. HuR binds to the late untranslated region in HPV-16 possibly inhibiting late gene expression [148]. PTB binds to the early untranslated region of HPV-16 possibly affecting the use of the early polyadenylation signal [142]. Analysis of late gene expression after over-expression of each of these proteins allowed their effect on this process to be evaluated.

The 5 protein expression plasmids were co-transfected individually into HeLa cells with either of two HPV-16 sub-genomic expression plasmids, pBEL or pBELM. Total RNA was extracted and RT-PCR was performed for early and late HPV-16 mRNA as well as GAPDH as a control. GAPDH PCR products were detected in all samples (Fig. 4.11).

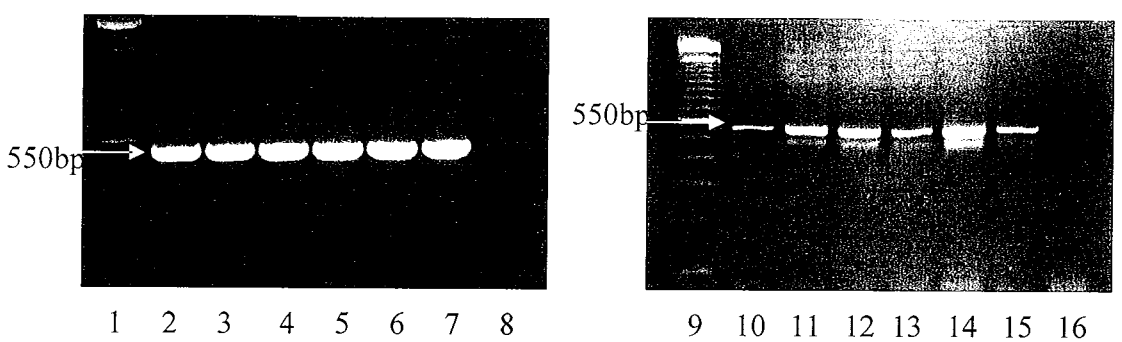


Figure 4.11 PCR amplification of the GAPDH cDNA using the GAPDHF/GAPDHR primers

Lane 1 100bp DNA ladder, lane 2 pBEL, lane 3 pBEL + ASF/SF2, lane 4 pBEL + hnRNP A1, lane 5 pBEL + hnRNP H, lane 6 pBEL + HuR, lane 7 pBEL + PTB, lane 8 negative control, lane 9 100bp DNA ladder, lane 10 pBELM, lane 11 pBELM + ASF/SF2, lane 12 pBELM + hnRNP A1, lane 13 pBELM + hnRNP H, lane 14 pBELM + HuR, lane 15 pBELM + PTB, lane 16 negative control. Run on a 1% agarose gel.

PCR was carried out using the 757s-E4a primers (which amplify a 190bp region of the HPV-16 genome) on the early RNA products from both the pBEL and pBELM plasmids in order to establish if any of the proteins had altered early gene expression. A single splice product was identified: this comprised of a splice from SD880 to SA3358. No change in expression of this splice product was identified when the early region PCR products were analysed on a gel (Fig. 4.12).

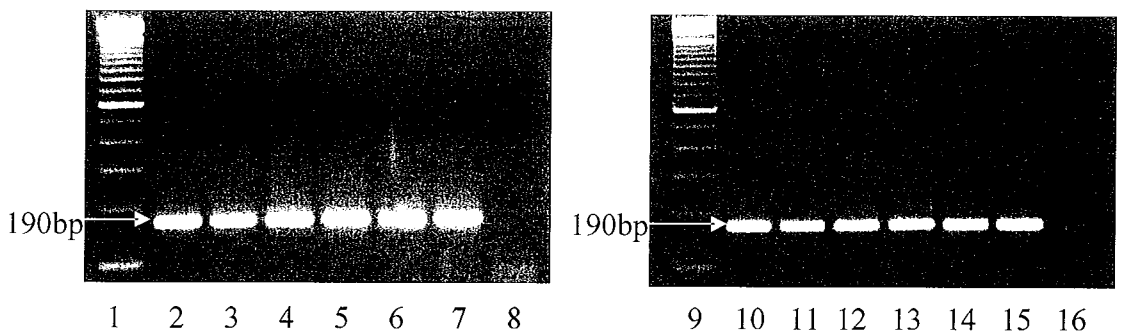


Figure 4.12 PCR amplification of the early HPV-16 region using the 757s-E4a primers

Lane 1 100bp DNA ladder, lane 2 pBEL, lane 3 pBEL + ASF/SF2, lane 4 pBEL + hnRNP A1, lane 5 pBEL + hnRNP H, lane 6 pBEL + HuR, lane 7 pBEL + PTB, lane 8 negative control, lane 9 100bp DNA ladder, lane 10 pBELM, lane 11 pBELM + ASF/SF2, lane 12 pBELM + hnRNP A1, lane 13 pBELM + hnRNP H, lane 14 pBELM + HuR, lane 15 pBELM + PTB, lane 16 negative control. Run on a 1% agarose gel.

When the late region RNA products from the pBEL plasmid were analysed by PCR amplification of the late HPV-16 region using the 757s-L1a(wt) primers (Fig. 4.13) it was found that none of the proteins evaluated induced the production of the late RNA products.

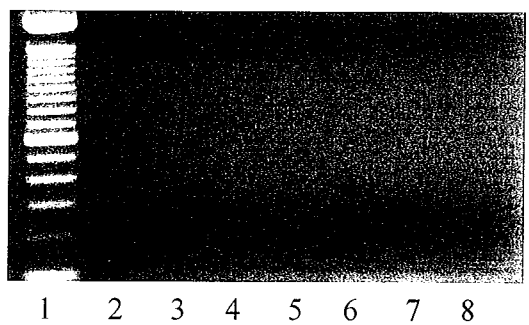


Figure 4.13 PCR amplification of the late HPV-16 region using the 757s-L1a (wt) primers

Lane 1 100bp DNA ladder, lane 2 pBEL, lane 3 pBEL + ASF/SF2, lane 4 pBEL + hnRNP A1, lane 5 pBEL + hnRNP H, lane 6 pBEL + HuR, lane 7 pBEL + PTB, lane 8 negative control. Run on a 1% agarose gel.

When the PCR products (amplified using the 757s-L1a(mut) primers, amplifying two splice variants, a 430bp splice product and a 160bp splice product) from the pBELM plasmid were analysed it was discovered that this plasmid on its own produced two individual RNA species (Fig. 4.14). When the sizes of the two spliced products were compared to the HPV-16 DNA sequence the 430bp band was identified as a splice from the E1 (SD880) coding region to the E4 (SA3358) coding region and a second splice from E4 (SD3632) to the L1 (SA5639) coding region (E4^L1, Fig. 4.15). The second band (160bp) was identified as a direct splice from the E1 (SD880) coding region into L1 (SA5639) (L1, Fig. 4.15). It was also found that the co-transfected protein plasmids altered the relative levels of these two RNA species (Fig. 4.14).

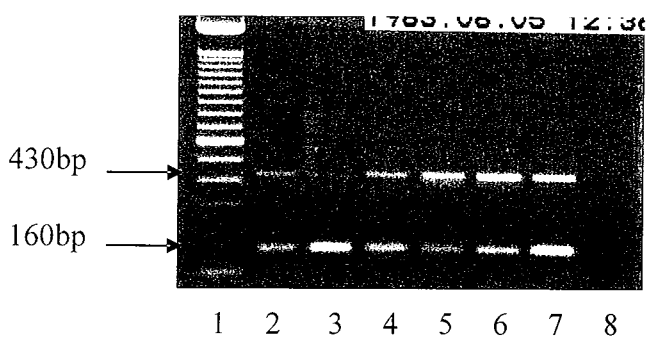


Figure 4.14 PCR amplification of the late HPV-16 region using the 757s-L1a (mut) primers

Lane 1 100bp DNA ladder, lane 2 pBELM, lane 3 pBELM + ASF/SF2, lane 4 pBELM + hnRNP A1, lane 5 pBELM + hnRNP H, lane 6 pBELM + HuR, lane 7 pBELM + PTB, lane 8 negative control. Run on a 1% agarose gel.

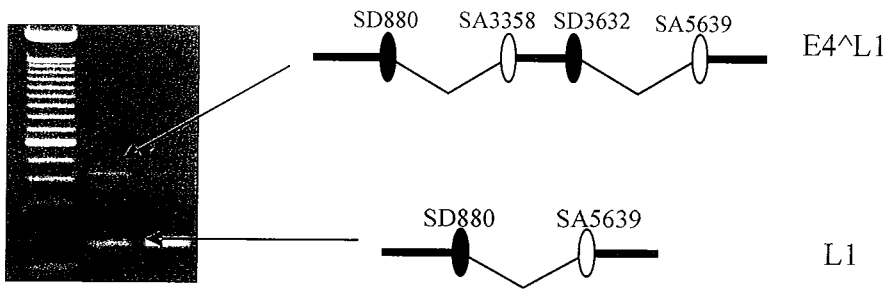


Figure 4.15 PCR amplification of the late HPV-16 region using the 757s-L1a(mut) primers

Lane 1 100bp DNA ladder, lane 2 pBELM, lane 3 pBELM + ASF/SF2. Two splice products were identified by size analysis as being the E4^L1 and L1 splice products. Products were run on a 1% agarose gel.

4.3.4 Optimisation of Real-time PCR for the Analysis of HPV-16 Late Gene Expression

In order to measure changes in late gene expression caused by the protein expression plasmids, a real-time PCR protocol was established. This method used SYBR Green fluorescence to measure double stranded DNA in the PCR reaction. GAPDH real-time PCR was used as an internal control for the late gene PCR. Both the GAPDH PCR, using GAPDHF-GAPDHR primers, and the late gene PCR, using the 757s-L1a(wt and mut) primers, were optimised for the Roche LightCycler system (Figs. 4.16 – 4.21). 757s-L1a(wt and mut) both amplify two splice variants, for the wild-type primers, a 470bp splice product and a 200bp splice product are amplified (Figs. 4.18 and 4.19) and for the mutant primers a 430bp splice product and a 160bp splice product are amplified (Figs. 4.20 and 4.21). The size difference between the products amplified by these primers is due to the deletion in the

pBELM plasmid. Optimisation involved a magnesium titration and DNA dilution series to establish the optimal PCR conditions in this system.

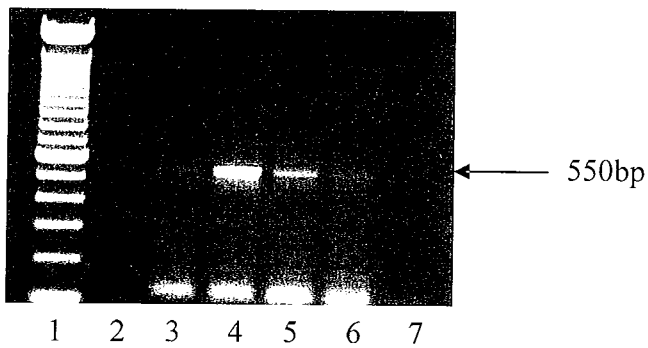


Figure 4.16 Real-time PCR magnesium titration for GAPDHF/GAPDHR primers

Lane 1 100bp DNA ladder, lane 2 1mmol/l magnesium, lane 3 2mmol/l magnesium, lane 4 3mmol/l magnesium, lane 5 4mmol/l magnesium, lane 6 5mmol/l magnesium, lane 7 negative control. Products run on a 1% agarose gel.

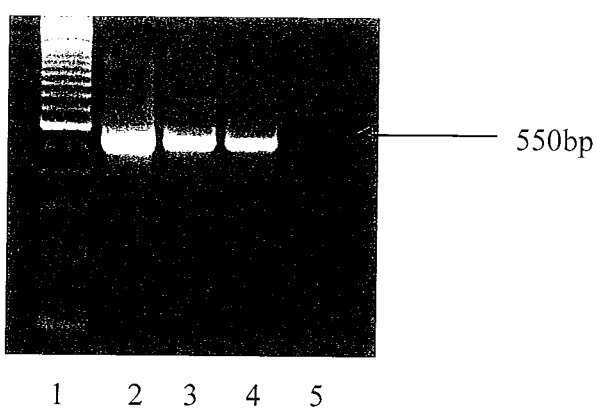


Figure 4.17 Real-time PCR cDNA dilution for GAPDHF/GAPDHR wild-type primers

Lane 1 100bp DNA ladder, lane 2 neat, lane 3 1/5 dilution, lane 4 1/10 dilution, lane 5 negative control. Products run on a 1% agarose gel.

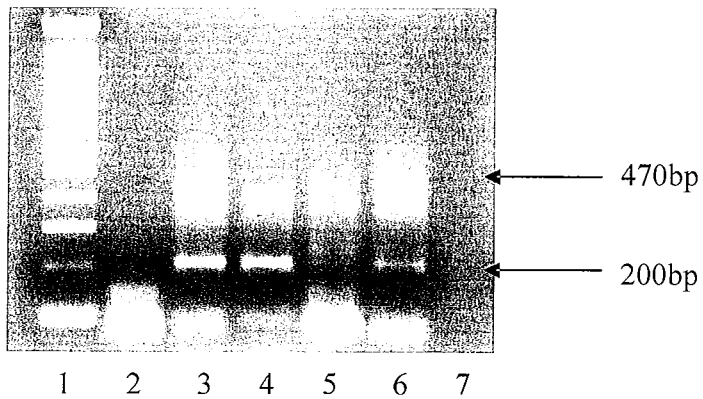


Figure 4.18 Real-time PCR magnesium titration for 757s-L1a wild-type primers

Lane 1 100bp DNA ladder, lane 2 1mmol/l magnesium, lane 3 2mmol/l magnesium, lane 4 3mmol/l magnesium, lane 5 4mmol/l magnesium, lane 6 5mmol/l magnesium, lane 7 negative control. Products run on a 1% agarose gel.

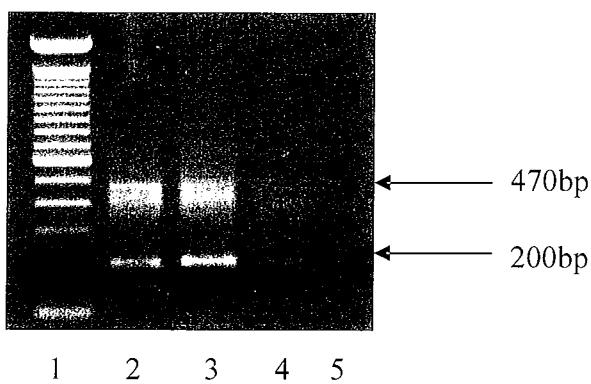


Figure 4.19 Real-time PCR cDNA dilution for 757s-L1a wild-type primers

Lane 1 100bp DNA ladder, lane 2 neat, lane 3 1/5 dilution, lane 4 1/10 dilution, lane 5 negative control. Products run on a 1% agarose gel.

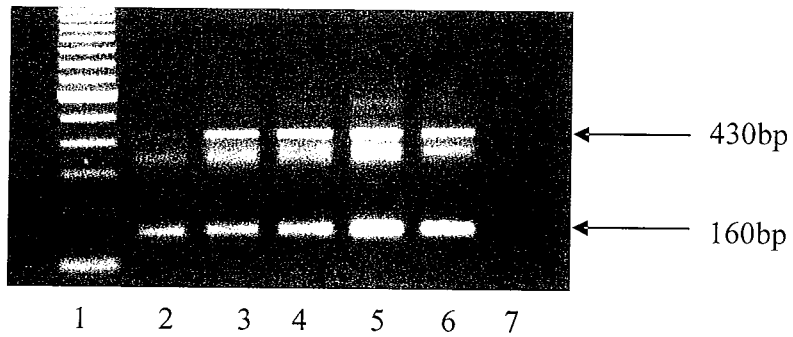


Figure 4.20 Real-time PCR magnesium titration for 757s-L1a(mut) primers

Lane 1 100bp DNA ladder, lane 2 1mmol/l magnesium, lane 3 2mmol/l magnesium, lane 4 3mmol/l magnesium, lane 5 4mmol/l magnesium, lane 6 5mmol/l magnesium, lane 7 negative control. Products run on a 1% agarose gel.

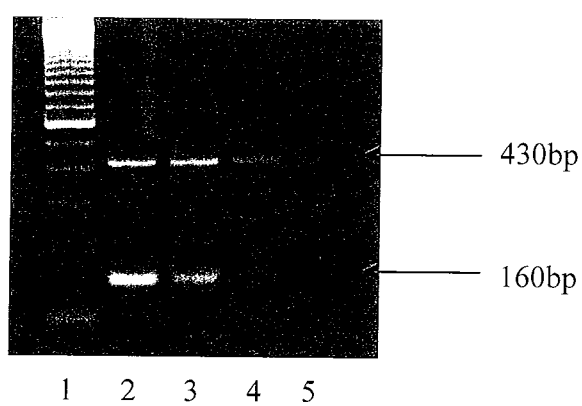


Figure 4.21 Real-time PCR cDNA dilution for 757s-L1a mutant primers run on a 1% agarose gel

Lane 1 100bp DNA ladder, lane 2 neat, lane 3 1/5 dilution, lane 4 1/10 dilution, lane 5 negative control. Products run on a 1% agarose gel.

The optimised real-time PCR conditions were established as 3mmol/l magnesium for GAPDH (Fig. 4.16) and 757s-L1a (mut) (Fig. 4.20) PCRs and 2mmol/l for 757s-L1a (wt) (Fig. 4.18) PCR. All samples were diluted 1/5 as PCR reactions were more specific with

the reduced cDNA levels (Figs. 4.17, 4.19 and 4.21). Melt curve analysis carried out at the end of each PCR allowed the identification of the specific peaks for each band identified on the gel (Figs. 4.22 - 4.24).

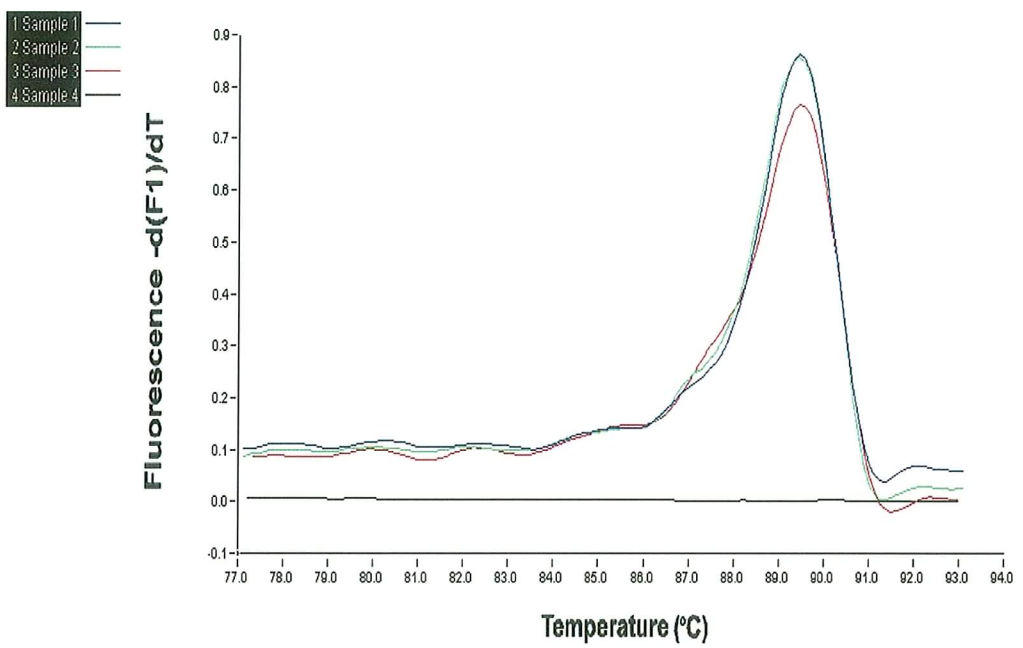


Figure 4.22 Melt curve analysis of GAPDH expression using GAPDHF/GAPDHR primers showing one specific peak for the PCR product

Sample 1 – pBEL, sample 2 – pBEL + Ad E4orf4, sample 3 – pBEL + SRp30c, sample 4 – negative control. The melt curve peak was equivalent to the 550bp band in Figs 4.16 and 4.17.

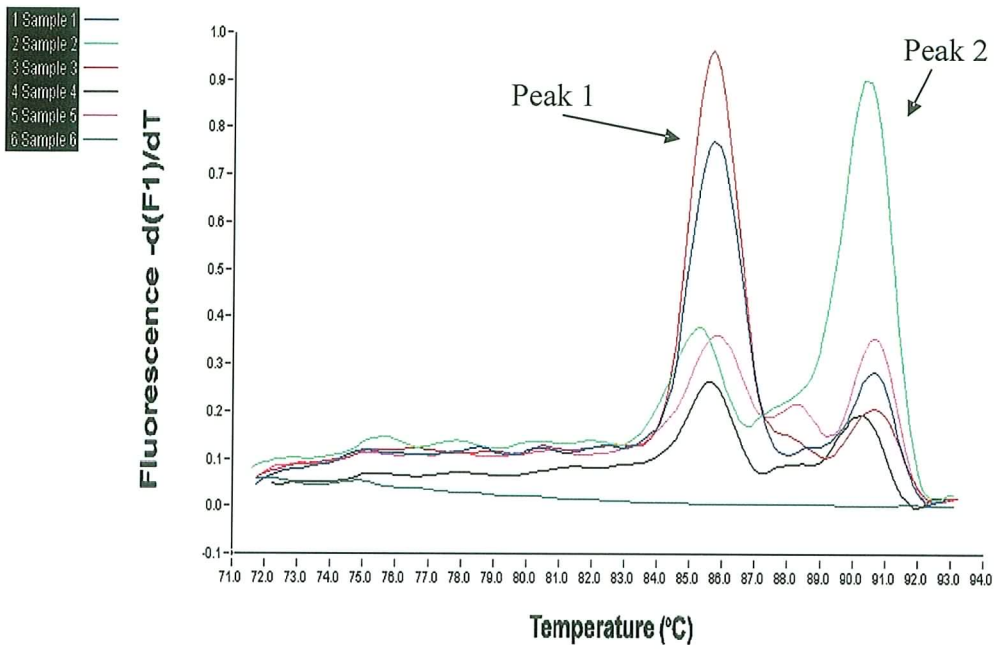


Figure 4.23 Melt curve analysis of HPV-16 late gene expression using 757s-L1a(wt) primers showing specific peaks for the two splice products identified by PCR

Sample 1 - pBEL, sample 2 – pBEL + Ad E4orf4, sample 3 – pBEL + SRp30c, sample 4 – pBEL + PTB, sample 5 – pBEL + HPV-16 E2, sample 6 – negative control. The two peaks represent the two splice products that can be induced in this plasmid (Fig. 4.5), peak 1 = E4^L1 and peak 2 = L1. The melt curve peaks were equivalent to the 470bp (peak 1) and 200bp (peak 2) bands in Figs. 4.18 and 4.19.

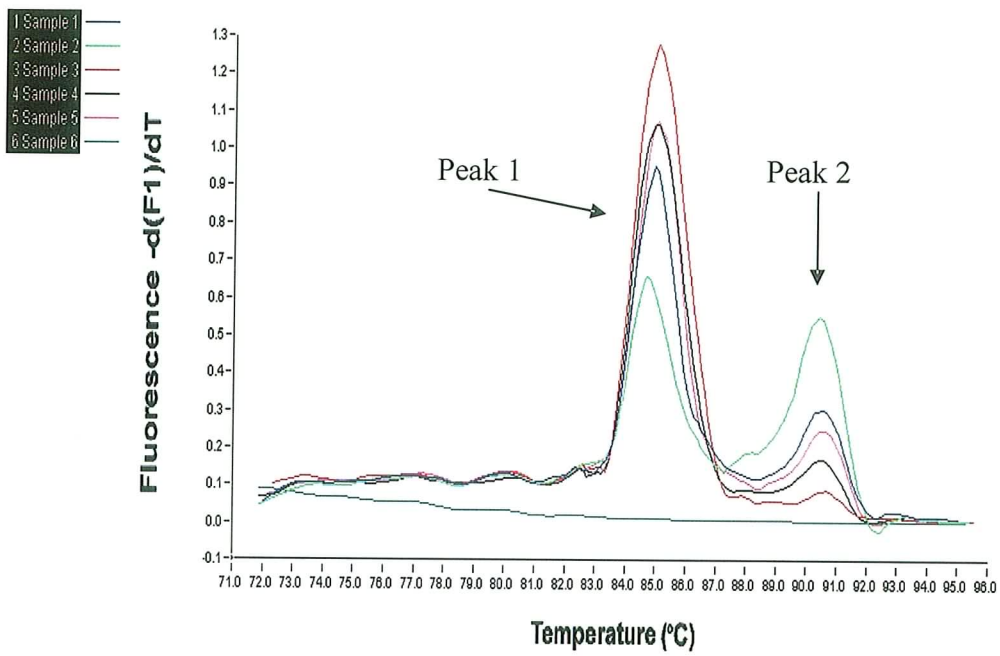


Figure 4.24 Melt curve analysis of HPV-16 late gene expression using 757s-L1a(mut) primers showing specific peaks for the two splice products identified by PCR

Sample 1 pBELM, sample 2 pBELM + ASF/SF2, sample 3 pBELM + hnRNP A1, sample 4 pBELM + hnRNP H, sample 5 pBELM + HuR, sample 6 pBELM + PTB, sample 7 negative control. The two peaks represent the two splice products that can be induced in this plasmid (Figs. 4.5 and 4.15), peak 1 = E4^L1 and peak 2 = L1. The melt curve peaks were equivalent to the 430bp (peak 1) and 160bp (peak 2) bands in Figs. 4.20 and 4.21.

4.3.5 Real-time PCR Analysis of HPV-16 Late Gene Expression

As previously shown when expression plasmids for RNA-binding proteins were transfected with the pBELM HPV-16 expression plasmid late gene expression was induced (Fig. 4.14). To examine the changes that each protein induced, the samples were analysed using the real-time PCR that had been optimised for both GAPDH (as a control) and late gene products.

Melt curves for all PCRs showed that all reactions produced specific peaks for the expected products (Figs. 4.25 and 4.26). This allowed analysis of the CT values and peak areas for the exanimation of late gene expression. The CT value can gave a measure of the amount of template that was added to the reaction. All of the real-time PCR experiments contained a transfection of the HPV-16 expression plasmid on its own as a control. CT values of the HPV-16 expression plasmid co-transfected with the various protein expression plasmids could then be related to this to measure the effect on the expression of the late genes.

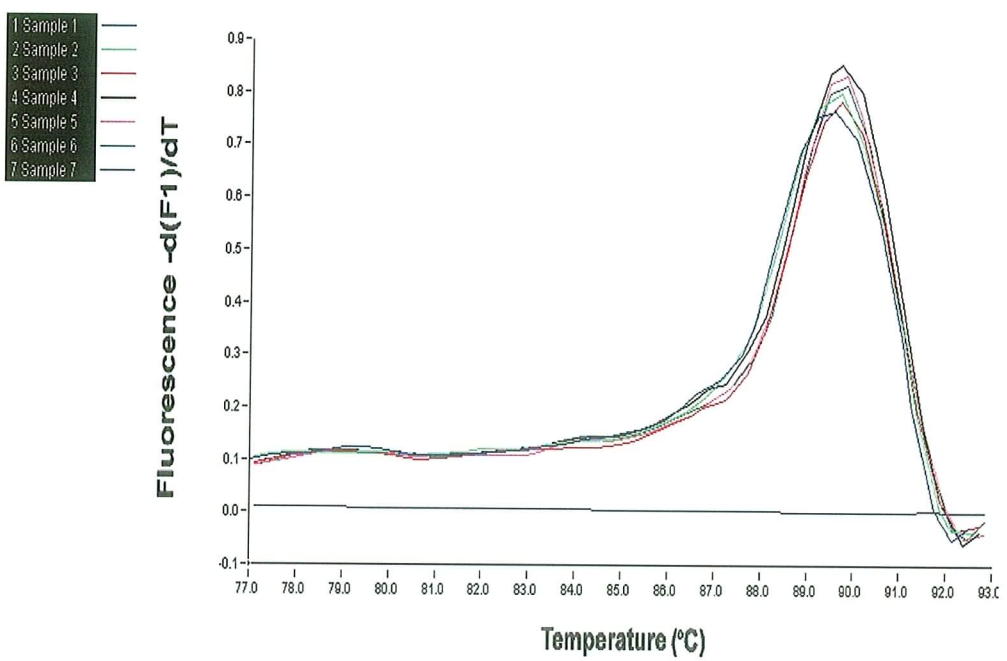


Figure 4.25 Melt curve analysis of GAPDH expression for the experiment pBELM co-transfected with splicing regulatory proteins using GAPDHF/GAPDHR primers

Sample 1 pBELM, sample 2 pBELM + ASF/SF2, sample 3 pBELM + hnRNP A1, sample 4 pBELM + hnRNP H, sample 5 pBELM + HuR, sample 6 pBELM + PTB, sample 7 negative control.

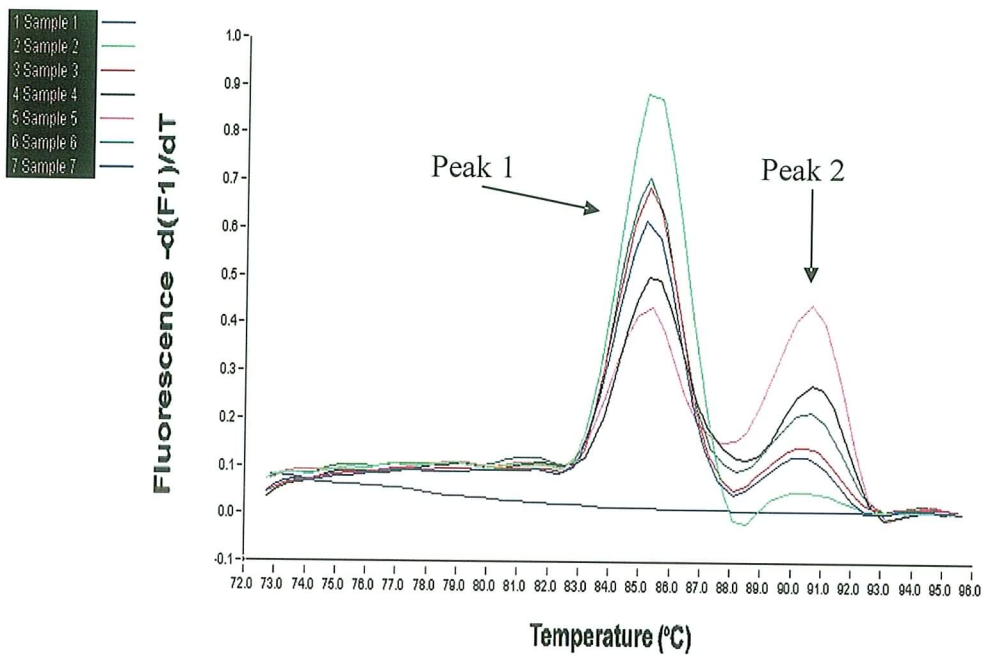


Figure 4.26 Melt curve analysis of late gene expression of pBELM co-transfected with splicing regulatory proteins using 757s – L1a(mut) primers

Sample 1 pBELM, sample 2 pBELM + ASF/SF2, sample 3 pBELM + hnRNP A1, sample 4 pBELM + hnRNP H, sample 5 pBELM + HuR, sample 6 pBELM + PTB, sample 7 negative control. The two peaks represent the two splice products that can be induced in this plasmid (Figs. 4.5 and 4.15), peak 1 = E4^L1 and peak 2 = L1.

When RNA-binding proteins were co-expressed with a sub-genomic HPV-16 expression plasmid (pBELM) it was determined that ASF/SF2, hnRNP A1 and hnRNP H did not alter the expression of the late genes (Fig. 4.28). HuR and PTB were found to increase the level of late gene expression (indicated by a decrease in CT value). HuR was the only protein to indicate a significant change when compared to the HPV-16 plasmid transfected alone.

Using melt curve analysis to examine the expression of the two splice products that could be expressed (Fig. 4.27) it was determined that the ratio of the splice products was altered by ASF/SF2, hnRNP H and HuR (Fig. 4.29). Peak 1 represents the E4^L1 splice product and peak 2 represents the L1 splice product. ASF/SF2 had no effect on the overall level of late gene expression but its expression results in no direct splicing into the late region, only the E4^L1 product was formed. hnRNP H reduces the expression of E4^L1 and increases the expression of L1 whereas HuR reverses the expression ratio of E4^L1 to L1, leading to the production of more L1 compared to E4^L1.

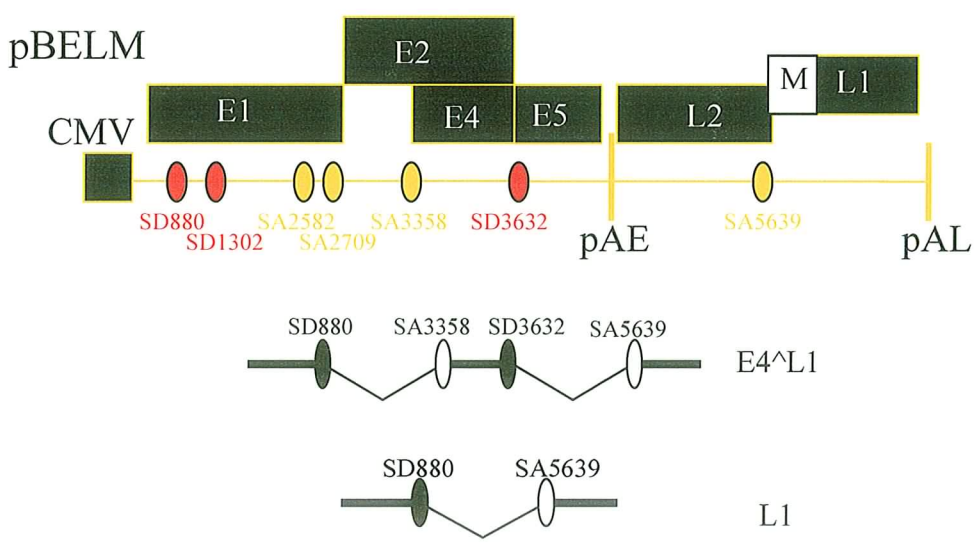
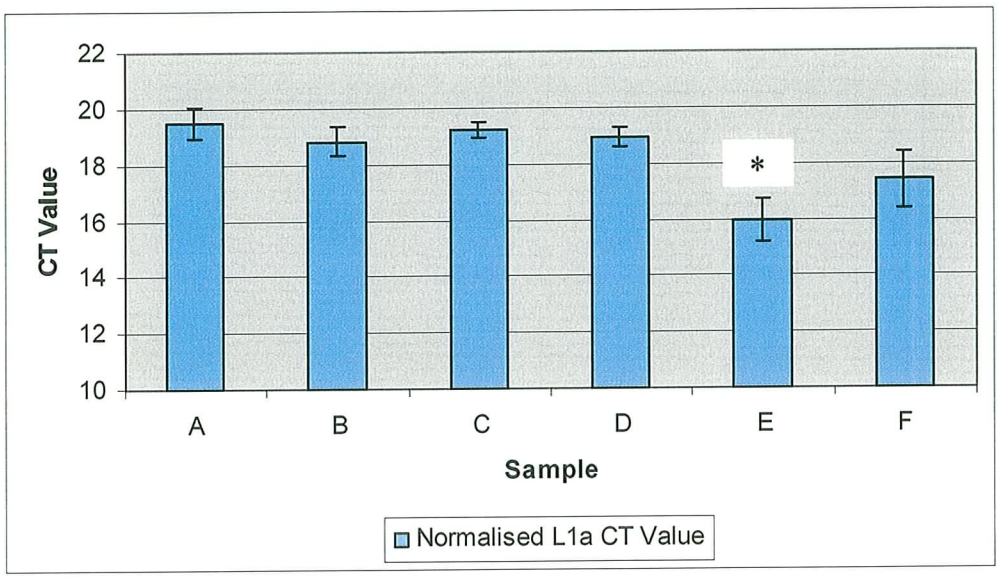


Figure 4.27 pBELM plasmid and the two late splice products E4^L1 and L1

Two late splice products can be induced in this plasmid, E4^L1, which includes the E4 coding region along with the L1 coding region, and L1 which is a direct splice from SD880 to the beginning of the L1 coding region.

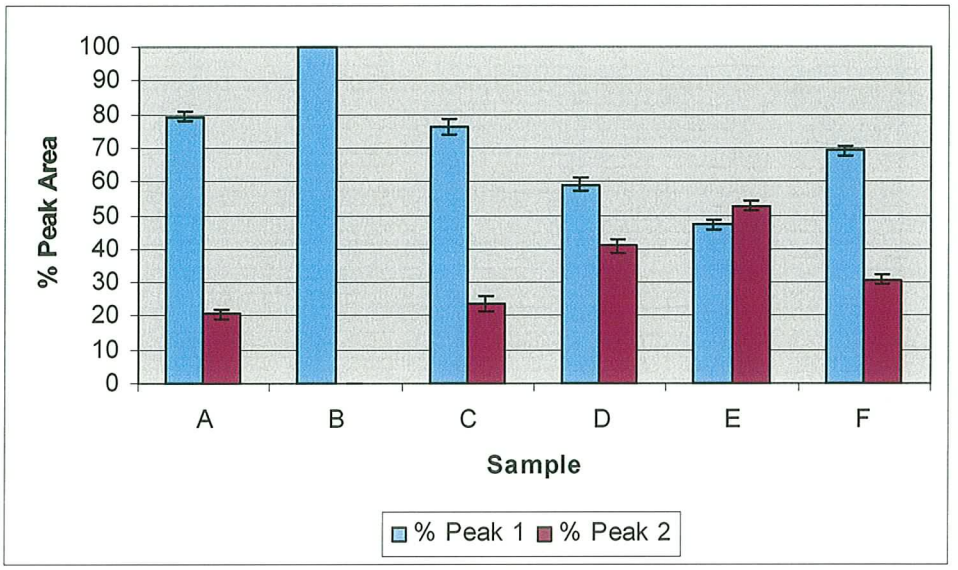
In summary, hnRNP A1 had no effect on late gene expression as expected [140], ASF/SF2 and hnRNP H did not increase the level of late gene production but altered the ratio of the production of E4^L1 to L1. PTB increased the level of late gene production but did not change the ratio of the two splice products and HuR not only increases the level of late gene expression and also alters the ratio of the splice products.



	Normalised L1a CT Value	Standard Error	p Value (*)
A pBELM	19.496	0.553	
B pBELM+ASF/SF2	18.829	0.537	0.4001
C pBELM+hnRNP A1	19.240	0.295	0.6876
D pBELM+hnRNP H	18.930	0.333	0.3938
E pBELM+HuR	15.982	0.753	<0.001
F pBELM+PTB	17.392	0.990	0.0820

Figure 4.28 Graph of Normalised CT values for 757s-L1a(mut) real-time PCR for pBELM co-transfected with splicing regulatory proteins

A pBELM (plasmid control), B pBELM + ASF/SF2, C pBELM + hnRNP A1, D pBELM + hnRNP H, E pBELM + HuR, F pBELM + PTB. Error bars indicate standard error for each sample. * indicates significance in relation to pBELM alone.



	% Peak 1	% Peak 2	Standard Error
A pBELM	79.30	20.69	1.374
B pBELM+ASF/SF2	100	0	0
C pBELM+hnRNP A1	76.18	23.81	2.325
D pBELM+hnRNP H	59.13	40.86	2.140
E pBELM+HuR	47.24	52.75	1.512
F pBELM+PTB	69.25	30.74	1.339

Figure 4.29 Percentage peak areas for 757s-L1a(mut) real-time PCR for pBELM co-transfected with splicing regulatory proteins

A pBELM (plasmid control), B pBELM + ASF/SF2, C pBELM + hnRNP A1, D pBELM + hnRNP H, E pBELM + HuR, F pBELM + PTB. Peak 1: splice from SD880 to SA5639 (E4^L1 splice product), Peak 2: splice from SD 880 to SA3358 and SD3632 to SA5639 (L1 splice product). Error bars indicate standard error for each sample.

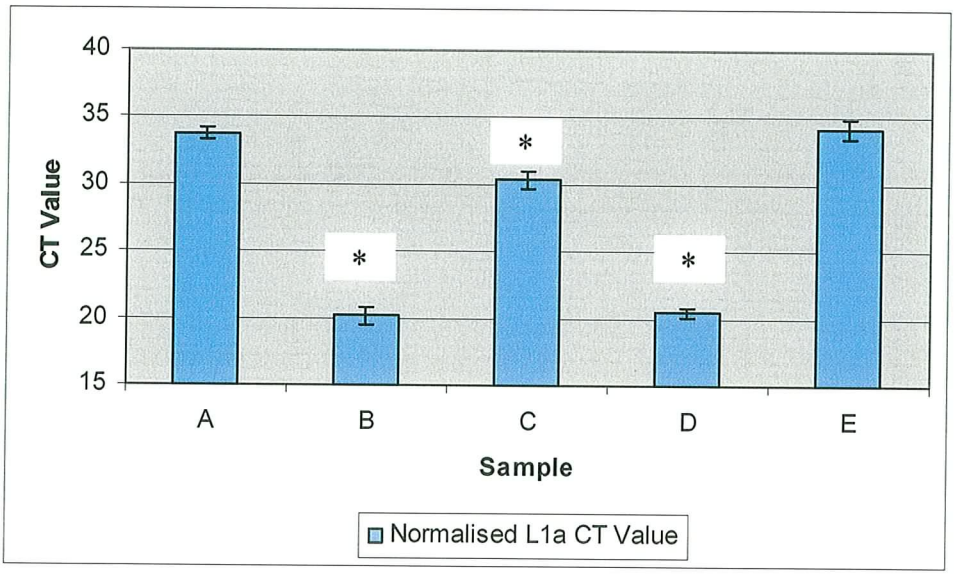
4.3.6 Late Gene Expression Can Be Altered By Viral and Splice Regulatory Proteins

Using immunohistochemistry (by means of a general antibody directed against SR proteins) it was demonstrated that SR proteins were down-regulated in response to cell differentiation. Only one SR protein, ASF/SF2, had an antibody available that allowed its expression to be examined individually. This protein was also down-regulated as cells underwent differentiation. In order to mimic this situation in HeLa cells, Ad E4orf4, which binds to and inactivates ASF/SF2 [231], was over-expressed to identify if the inactivation of this SR protein affected the expression of the late genes. The study by Estmer Nilsson *et al.* also identified a second SR protein, SRp30c, that interacted with Ad E4orf4 [231], this protein was also over-expressed in HeLa cells to asses its effect in HPV late gene expression. Ad E4orf4 can induce HPV-16 late gene expression (Fig. 4.9) and has been demonstrated to induce apoptosis by inhibiting the anaphase-promoting complex/cyclosome (APC/C) [239]. High-risk HPV E2 proteins can also bind to the APC/C [234] so it was investigated if over-expression of HPV-16 E2 could also have an effect on the expression of the late genes. Using real-time PCR the effect of the expression of these four proteins on the amount of late gene expression and the two splice products was examined. The effect of expression was looked at using both the pBEL plasmid and the pBELM plasmid.

When the CT values for the experiment using the pBEL plasmid were evaluated it was found that HPV-16 E2 did not alter the level of the production of the late genes, whereas Ad E4orf4, SRp30c and PTB increased the level of late gene expression (Fig. 4.30). When the values for these proteins were statistically analysed it was found that all 3 induced a

significant change in late gene expression when compared to the expression of pBEL plasmid alone (Fig. 4.30).

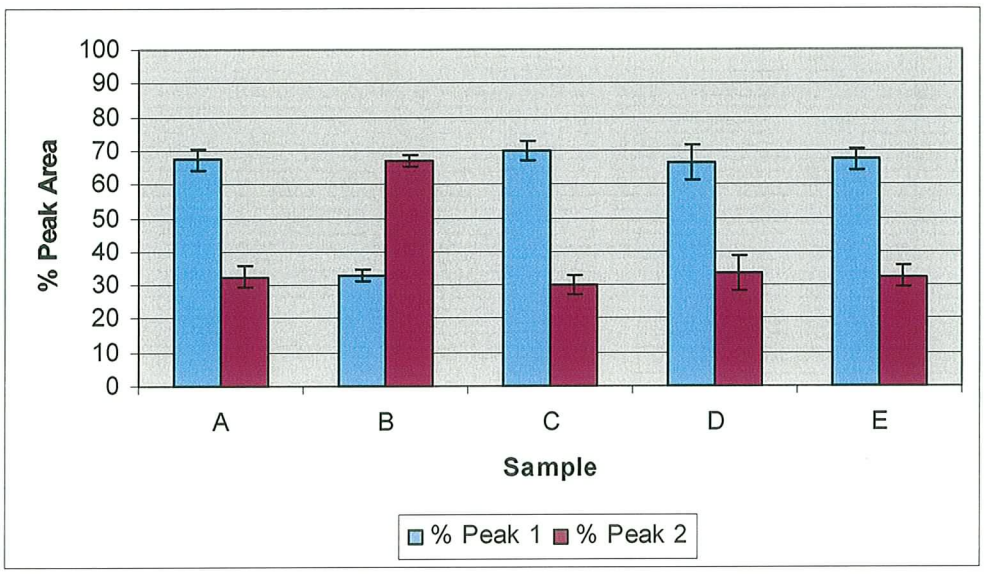
When the expression of the two splice products was evaluated by real-time PCR it was determined that the expression of SRp30c, PTB and HPV-16 E2 had no effect on the ratio of production of E4^L1 to L1 when compared to that of the pBEL plasmid alone (Fig. 4.31). The expression of Ad E4orf4 induced a reversal of the E4^L1 to L1 ratio. When pBEL was transfected alone, 67.4% of the late gene production was the E4^L1 product and 32.6% was the L1 splice product. When pBEL was co-transfected with Ad E4orf4 this became 32.7% E4^L1 and 63.3% L1.



	Normalised L1a CT Value	Standard Error	p value (*)
A pBEL	33.782	0.458	
B pBEL+Ad E4orf4	20.186	0.658	<0.001
C pBEL+SRp30c	30.371	0.667	<0.001
D pBEL+PTB	20.482	0.330	<0.001
E pBEL+HPV-16 E2	34.228	0.750	0.6188

Figure 4.30 Graph of normalised CT values for 757s-L1a(wt) real-time PCR for pBEL co-transfected with viral and splice regulatory proteins

A pBEL (plasmid control), B pBEL + Ad E4orf4, C pBEL + SRp30c, D pBEL + PTB, E pBEL + HPV-16 E2. Error bars indicate standard error for each sample. * indicates significance in relation to pBEL alone.



	% Peak 1	% Peak 2	Standard Error
A pBEL	67.43	32.56	3.053
B pBEL+Ad E4orf4	32.67	67.32	1.740
C pBEL+SRp30c	70.11	29.88	2.816
D pBEL+PTB	66.70	33.29	5.219
E pBEL+HPV-16 E2	67.41	32.58	3.046

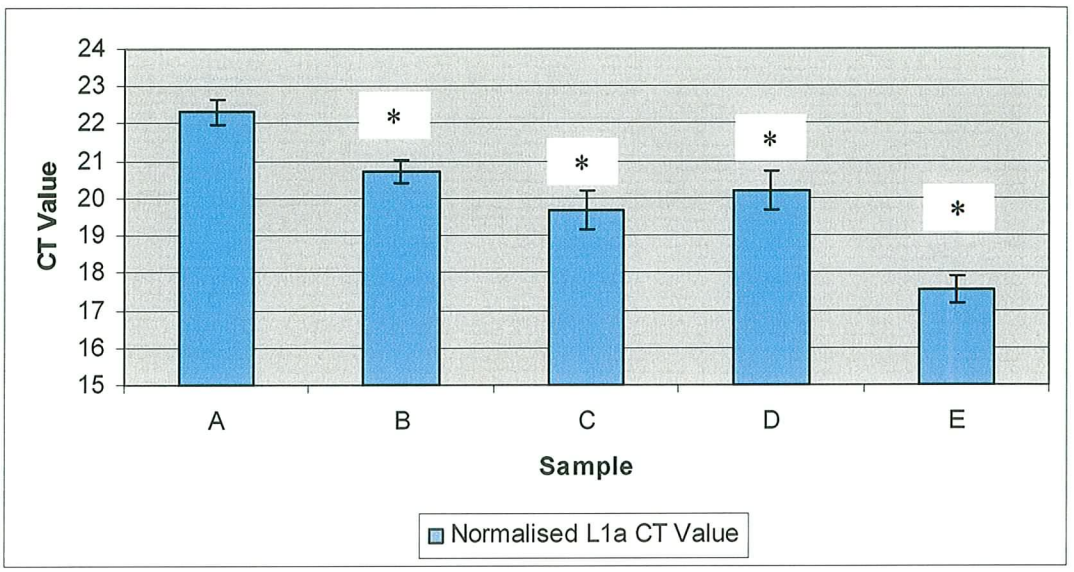
Figure 4.31 Percentage peak areas for 757s-L1a(wt) real-time PCR for pBEL co-transfected with viral and splice regulatory proteins

A pBEL (plasmid control), B pBEL + Ad E4orf4, C pBEL + SRp30c, D pBEL + PTB, E pBEL + HPV-16 E2. Peak 1: splice from SD880 to SA5639 (E4^L1 splice product), Peak 2: splice from SD 880 to SA3358 and SD3632 to SA5639 (L1 splice product). Error bars indicate standard error for each sample.

When the pBELM plasmid was co-transfected with Ad E4orf4, or SRp30c, or PTB or HPV-16 E2 it was observed that these four proteins had an effect on the expression of the late genes (Fig. 4.32). The expression of HPV-16 E2 induced the largest change in the expression of the late genes, but all four proteins increased in the expression of the late genes significantly.

Although HPV-16 E2 and PTB induced a significant increase in the overall late gene expression there was no major change in the ratio of the expression of the individual splice products when compared to the expression of pBELM alone (Fig. 4.33). SRp30c altered the ratio of the splice products when compared to pBELM alone, changing from 75.4% E4^L1 production for pBELM to 92.2% for pBELM + SRp30c and changing L1 production from 24.6% for pBELM to 7.8% when co-transfected with SRp30c. Ad E4orf4 induced the greatest change in the production of the two splice products, its expression reduces E4^L1 production to 54.9% and increases L1 production to 45.1%.

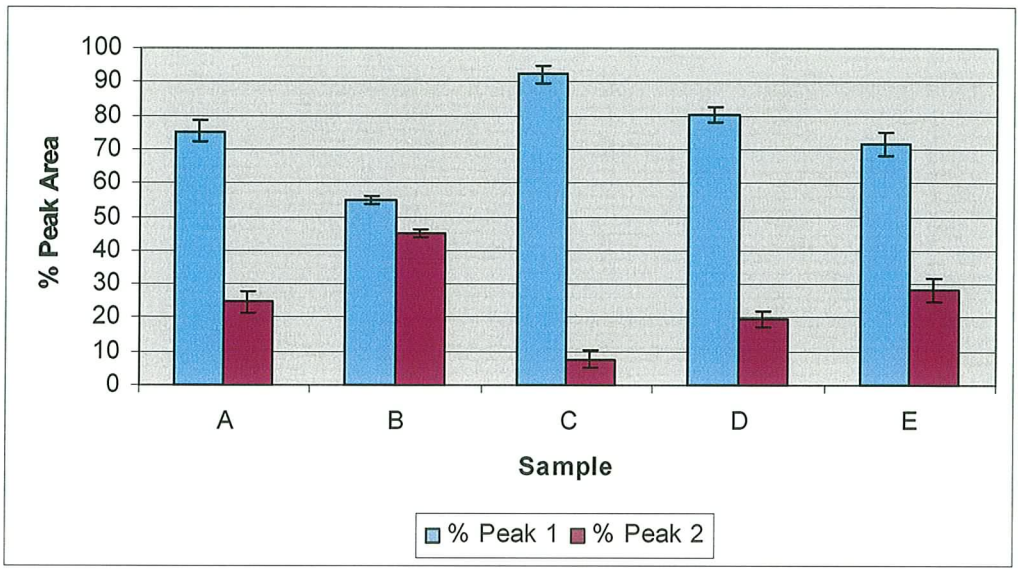
The results of these experiments indicated that Ad E4orf4, SRp30c and PTB induced late gene expression in the pBEL plasmid and the pBELM plasmid, while HPV-16 E2 could only induce late gene expression in pBELM. Also it was demonstrated that Ad E4orf4 could alter the ratio of the two splice products in both pBEL and pBELM plasmids, while SRp30c could induce a slight change in the ratio of the splice products in the pBELM plasmid only. PTB and HPV-16 E2 did not cause any change in the ratio of splice products in either plasmid.



	Normalised L1a CT Value	Standard Error	p value (*)
A pBELM	22.312	0.333	
B pBELM+Ad E4orf4	20.719	0.328	0.003
C pBELM+SRp30c	19.687	0.534	<0.001
D pBELM+PTB	20.198	0.505	0.003
E pBELM+HPV-16 E2	17.557	0.366	<0.001

Figure 4.32 Graph of normalised CT values for 757s-L1a(mut) real-time PCR for pBELM co-transfected with viral and splice regulatory proteins

A pBELM (plasmid control), B pBELM + Ad E4orf4, C pBELM + SRp30c, D pBELM + PTB, E pBELM + HPV-16 E2. Error bars indicate standard error for each sample. * indicates significance in relation to pBELM alone.



	% Peak 1	% Peak 2	Standard Error
A pBELM	75.36	24.63	3.154
B pBELM+Ad E4orf4	54.90	45.09	0.912
C pBELM+SRp30c	92.23	7.76	2.575
D pBELM+PTB	80.33	19.66	2.034
E pBELM+HPV-16 E2	71.68	28.31	3.716

Figure 4.33 Percentage peak areas for 757s-L1a(mut) real-time PCR for pBELM co-transfected with viral and splice regulatory proteins

A pBELM (plasmid control), B pBELM + Ad E4orf4, C pBELM + SRp30c, D pBELM + PTB, E pBELM + HPV-16 E2. Peak 1: splice from SD880 to SA5639 (E4^L1 splice product), Peak 2: splice from SD 880 to SA3358 and SD3632 to SA5639 (L1 splice product). Error bars indicate standard error for each sample.

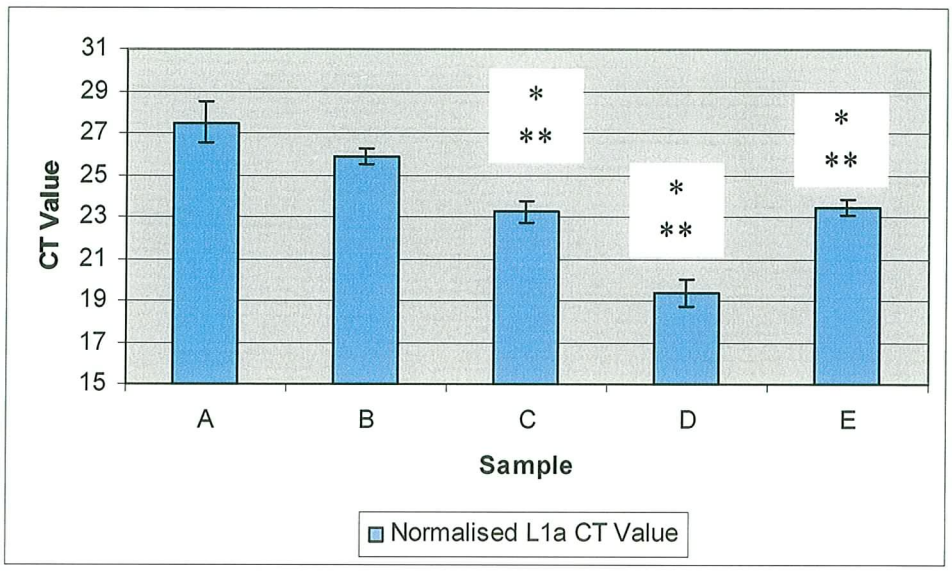
4.3.7 Determination of Effect on Late Gene Induction by Ad E4orf4 with hnRNP A1

In order to examine if the induction of the late genes by Ad E4orf4 could be reversed by the expression of the splicing silencer hnRNP A1, the pBEL plasmid was co-transfected with Ad E4orf4 and varying amounts of hnRNP A1 (4μg, 2μg and 1μg).

Interestingly the results indicated that when hnRNP A1 was transfected along with Ad E4orf4, a greater induction of the late genes was detected when compared to transfection with only Ad E4orf4 (Fig. 4.34). The three concentrations of hnRNP A1 induced a significant increase in late gene expression when compared to both pBEL alone and pBEL + Ad E4orf4. 4μg and 1μg caused a similar increase in the induction of the late genes, while 2μg of hnRNP A1 caused the largest increase.

When the expression of the two splice products was examined (Fig. 4.35) it was discovered that expression from pBEL alone results in 70.0% of the late gene expression being in the E4^L1 form and 30.0% in the L1 form, however when Ad E4orf4 was co-transfected with pBEL this was altered to 39.0% E4^L1 and 61.0% L1. When pBEL was co-transfected with both Ad E4orf4 and hnRNP A1 this ratio of approximately 40% to 60% E4^L1 to L1 was not altered at any of the three concentrations of hnRNP A1.

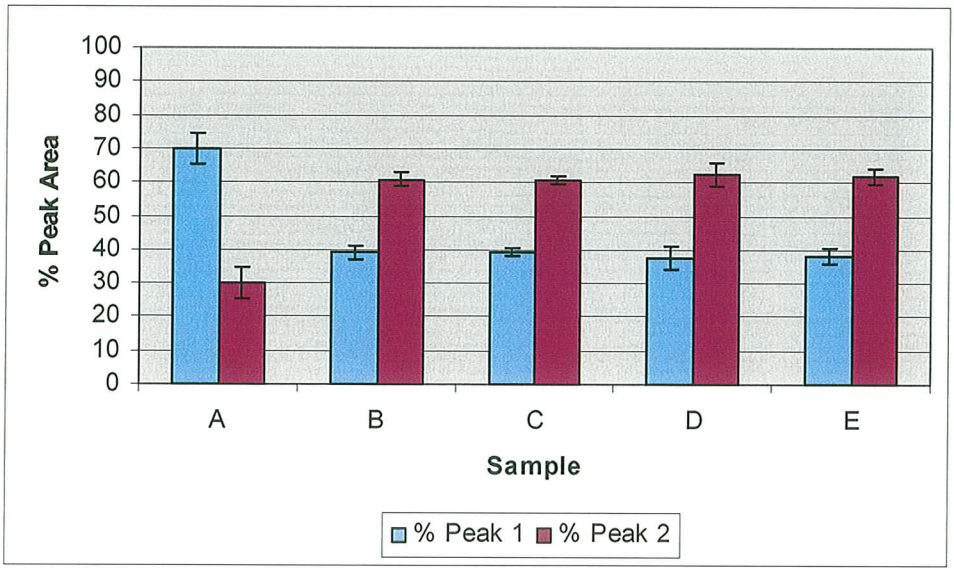
Overall this experiment demonstrated that when hnRNP A1 was co-transfected with Ad E4orf4 there was a stronger induction of the late genes when compared to the induction caused by Ad E4orf4 alone, and although hnRNP A1 increases the overall level of late gene expression it had no effect on the ratio of the two splice products induced by Ad E4orf4.



	Normalised L1a CT Value	Standard Error	p value (*)	p value (**)
A pBEL	27.502	0.965		
B pBEL+Ad E4orf4	25.878	0.392	0.1385	
C pBEL+Ad E4orf4+4μg hnRNP A1	23.266	0.509	<0.001	<0.001
D pBEL+Ad E4orf4+2μg hnRNP A1	19.376	0.687	<0.001	<0.001
E pBEL+Ad E4orf4+1μg hnRNP A1	23.456	0.391	<0.001	<0.001

Figure 4.34 Graph of normalised CT values for 757s-L1a(wt) real-time PCR competition analysis between Ad E4orf 4 and hnRNP A1

A pBEL (plasmid control), B pBEL + Ad E4orf4, C pBEL + Ad E4orf4 + 4μg hnRNP A1, D pBEL + Ad E4orf4 + 2μg hnRNP A1, E pBEL + Ad E4orf4 + 1μg hnRNP A1. Error bars indicate standard error for each sample. * indicates significance in relation to pBEL alone, ** indicates significance in relation to pBEL + E4orf4.



	% Peak 1	% Peak 2	Standard Error
A pBEL	70.09	29.90	4.507
B pBEL+Ad E4orf4	39.02	60.97	1.931
C pBEL+Ad E4orf4+4μg hnRNP A1	39.30	60.69	1.132
D pBEL+Ad E4orf4+2μg hnRNP A1	37.60	62.39	3.320
E pBEL+Ad E4orf4+1μg hnRNP A1	38.07	61.92	2.175

Figure 4.35 Percentage peak areas for 757s-L1a(wt) real-time PCR competition analysis between Ad E4orf 4 and hnRNP A1

A pBEL (plasmid control), B pBEL + Ad E4orf4, C pBEL + Ad E4orf4 + 4μg hnRNP A1, D pBEL + Ad E4orf4 + 2μg hnRNP A1, E pBEL + Ad E4orf4 + 1μg hnRNP A1. Peak 1: splice from SD880 to SA5639 (E4^L1 splice product), Peak 2: splice from SD 880 to SA3358 and SD3632 to SA5639 (L1 splice product). Error bars indicate standard error for each sample.

4.3.8 Determination of the Effect on Late Gene Induction by SRp30c With hnRNP A1

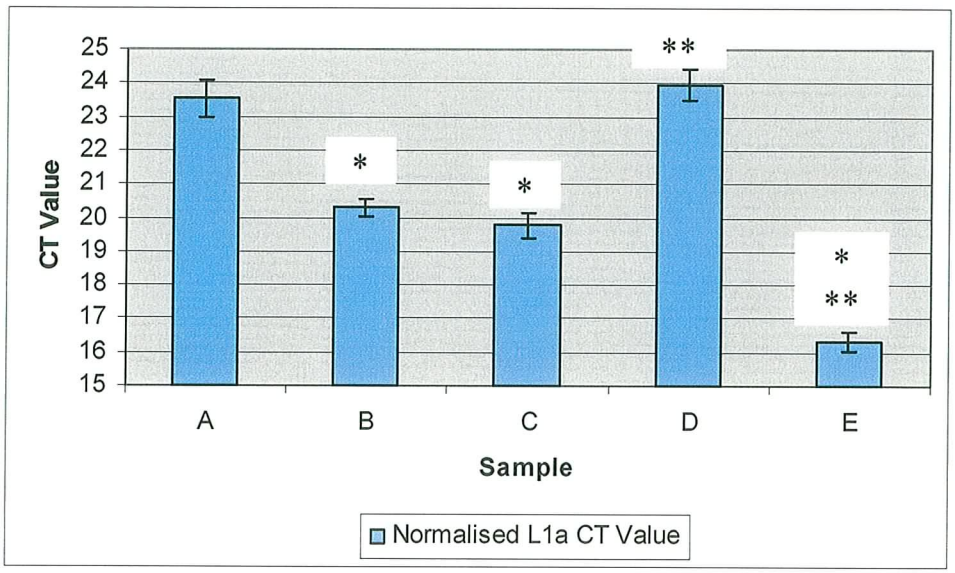
The effect of co-transfecting SRp30c with the splicing silencer hnRNP A1 was examined by transfecting the pBEL plasmid with SRp30c and varying amounts of hnRNP A1 (4μg, 2μg and 1μg). This was to examine if the induction of the late genes by SRp30c could be repressed by hnRNP A1.

When SRp30c was transfected with pBEL in the absence of hnRNP A1 a significant increase in late gene production was detected (Fig. 4.36). When 4μg and 1μg hnRNP A1 were included in the transfection the induction of the late genes compared to pBEL alone was again statistically significant. It was also demonstrated that 1μg hnRNP A1 surprisingly induced a further increase in late gene expression even when compared to pBEL + SRp30c, this change was also statistically significant. When 2μg hnRNP A1 was transfected with pBEL and SRp30c there was a significant decrease in late gene expression when compared to pBEL + SRp30c. The level of late gene expression when 2μg hnRNP A1 was present was similar to that of pBEL alone indicating that this concentration of hnRNP A1 may inhibit the induction of the late genes caused by SRp30c.

The expression of SRp30c changes the ratio of the two splice products from 70% E4^L1 and 30% L1 when pBEL was transfected alone to 50% E4^L1 and 50% L1 when SRp30c was present (Fig. 4.37). Expression of hnRNP A1 with pBEL and SRp30c altered this ratio slightly, to approximately 55% E4^L1 and 45% L1. The concentration of hnRNP A1 had no effect on the ratio of the splice products.

Overall these results suggest that at a concentration of 2μg hnRNP A1 may inhibit the action of SRp30c in the induction of the late genes, although this action is concentration

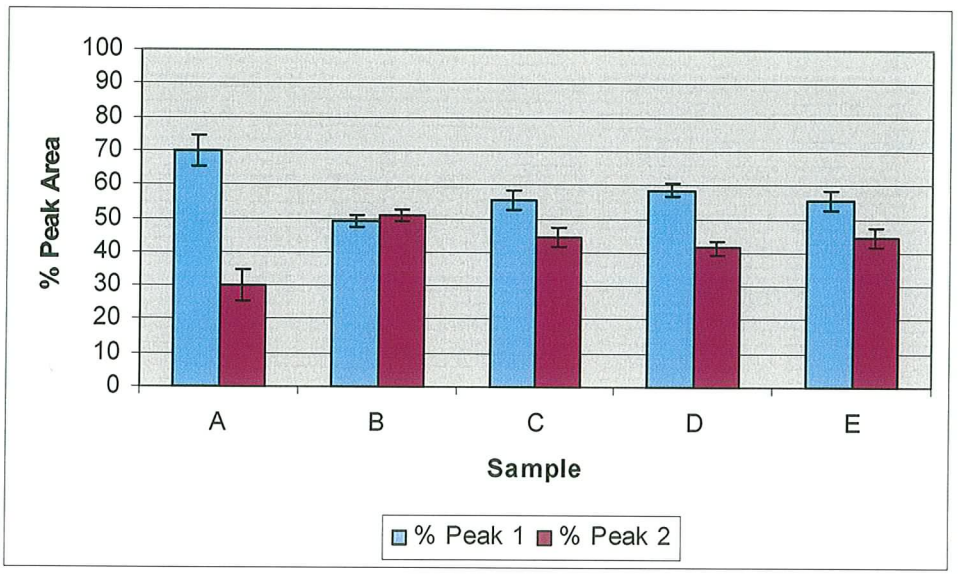
dependant as 1μg hnRNP A1 seems to induce the late genes to a greater degree than SRp30c alone. hnRNP A1 has only a very small effect on the ratio of E4^L1 and L1 produced and this effect is not concentration dependant.



	Normalised L1a CT Value	Standard Error	p value (*)	p value (**)
A pBEL	23.537	0.545		
B pBEL+SRp30c	20.306	0.270	<0.001	
C pBEL+SRp30c+4μg hnRNP A1	19.784	0.376	<0.001	0.276
D pBEL+SRp30c+2μg hnRNP A1	23.953	0.465	0.5698	<0.001
E pBEL+SRp30c+1μg hnRNP A1	16.354	0.287	<0.001	<0.001

Figure 4.36 Graph of normalised CT values for 757s-L1a(wt) real-time PCR competition analysis between SRp30c and hnRNP A1

A pBEL (plasmid control), B pBEL + SRp30c, C pBEL + SRp30c + 4μg hnRNP A1, D pBEL + SRp30c + 2μg hnRNP A1, E pBEL + SRp30c + 1μg hnRNP A1. Error bars indicate standard error for each sample. * indicates significance in relation to pBEL alone, ** indicates significance in relation to pBEL + SRp30c.



	% Peak 1	% Peak 2	Standard Error
A pBEL	69.69	30.30	4.594
B pBEL+SRp30c	49.13	50.86	1.823
C pBEL+SRp30c+4μg hnRNP A1	55.65	44.34	2.819
D pBEL+SRp30c+2μg hnRNP A1	58.44	41.55	1.965
E pBEL+SRp30c+1μg hnRNP A1	55.25	44.74	2.852

Figure 4.37 Percentage peak areas for 757s-L1a(wt) real-time PCR competition analysis between SRp30c and hnRNP A1

A pBEL (plasmid control), B pBEL + SRp30c, C pBEL + SRp30c + 4μg hnRNP A1, D pBEL + SRp30c + 2μg hnRNP A1, E pBEL + SRp30c + 1μg hnRNP A1. Peak 1: splice from SD880 to SA5639 (E4^L1 splice product), Peak 2: splice from SD 880 to SA3358 and SD3632 to SA5639 (L1 splice product). Error bars indicate standard error for each sample.

4.3.9 HPV-16 Late Gene Expression in A549 Cell Line

In order to establish if using a HPV positive cell line (HeLa cell line) had an effect on the results of experiments that used that cell line one experiment was repeated in a HPV negative cell line. The cell line used was the A549 cell line, these cells originated from a patient with lung adenocarcinoma. DNA was extracted from these cells and a PCR for β-globin gene (110bp PCR product) was run on the extracted DNA to ensure the presence of amplifiable DNA (Fig. 4.38). Once the presence of good quality DNA was confirmed a PCR was then set up for the HPV late region using the DNA extracted from the A549 cells. DNA from the SiHa cell line was used as a positive control. This was to ensure that the A549 cells that were being used were truly HPV negative. No HPV DNA was amplified from the A549 DNA using the MY09/MY11 primers (450bp PCR product) indicating the absence of HPV DNA in the A549 cells (Fig. 4.39). DNA extraction, β-globin and MY09/MY11 PCRs were carried out as detailed in chapter 2 of this thesis.

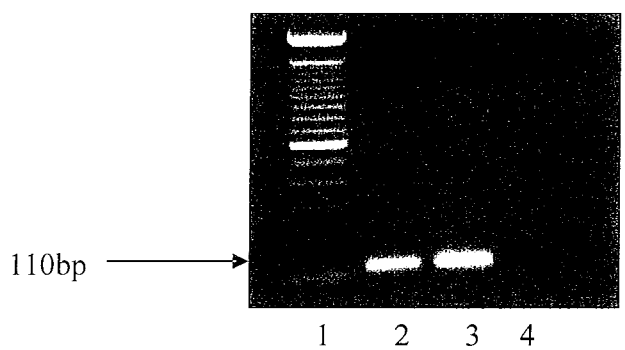


Figure 4.38 PCR amplification of the β-globin gene using the PC03/PC04 primers

Lane 1 100bp DNA ladder, lane 2 positive control (SiHa cell line DNA), lane 3 A549 cell line DNA, lane 4 negative control. Products run on a 1% agarose gel.

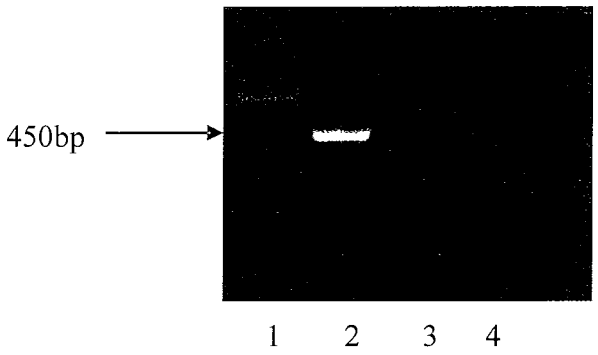


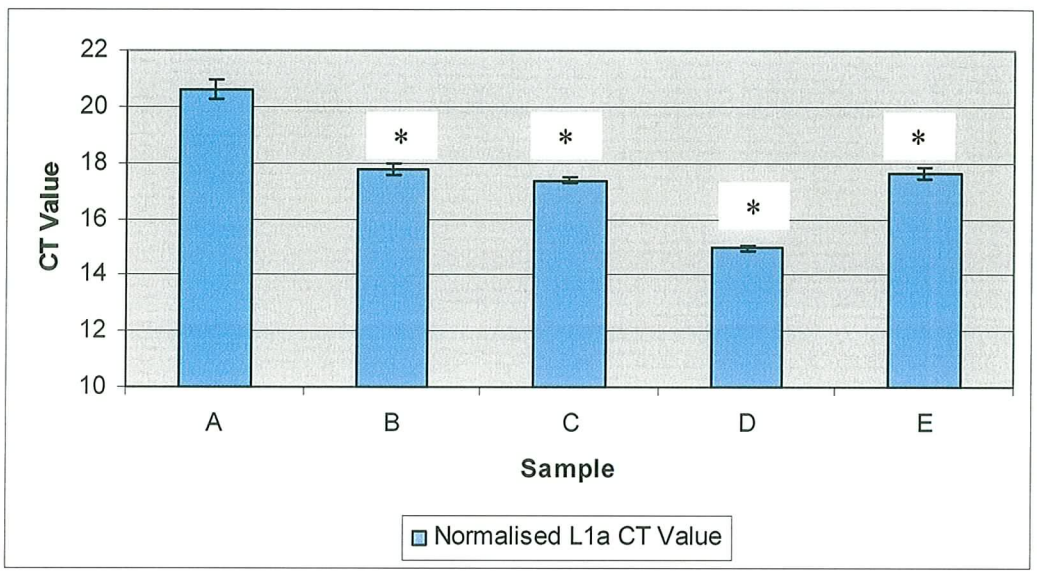
Figure 4.39 PCR amplification of the HPV late region using the MY09/MY11 primers

Lane 1 100bp DNA ladder, lane 2 positive control (SiHa cell line DNA), lane 3 A549 cell line DNA, lane 4 negative control. Products run on a 1% agarose gel.

The A549 cells were transfected as before, pBELM alone, pBELM + Ad E4ofr4, pBELM + SRp30c, pBELM + PTB and pBELM + HPV-16 E2. All real-time PCRs were carried out in triplicate. Analysis of the CT values indicated that all four co-transfected proteins increased the levels of late gene expression from the pBELM plasmid (Fig. 4.40). Also, the induction of the late genes was significant for each protein. When compared to the results for the same experiment carried out in HeLa cells (Fig. 4.32) it can be seen that in both cell lines all four proteins induced an increase in late gene expression and that this change was significant for each of the proteins when compared to pBELM transfected alone.

In A549 cells Ad E4orf4, SRp30c and HPV-16 changed the ratio of E4^L1 to L1 when compared the pBELM alone (Fig. 4.41). SRp30c induced the largest change with the expression of only the E4^L1 product. PTB did not cause a change in the expression of the two splice products. In comparison to the HeLa cell line it can be assessed that the same changes in E4^L1 to L1 ratio was occurring in A549 cells, although HPV-16 E2 had a slighter greater effect in A549 cells.

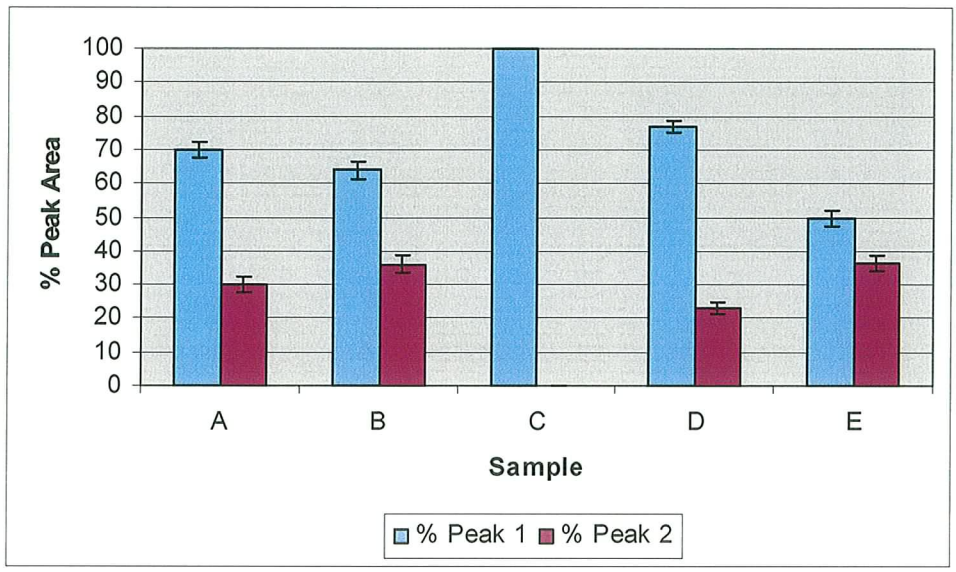
Overall this experiment demonstrated that presence of HPV in the HeLa cell line had very little effect on the results of the experiment, as results in general were similar when a HPV negative cell line was used.



	Normalised L1a CT Value	Standard Error	p value (*)
A pBELM	20.606	0.372	
B pBELM+Ad E4orf4	17.777	0.213	<0.001
C pBELM+SRp30c	17.373	0.114	<0.001
D pBELM+PTB	14.990	0.106	<0.001
E pBELM+HPV-16 E2	17.660	0.207	<0.001

Figure 4.40 Graph of normalised CT values for 757s-L1a(mut) real-time PCR for pBELM co-transfected with viral and splice regulatory proteins in A549 cells

A pBELM (plasmid control), B pBELM + Ad E4orf4, C pBELM + SRp30c, D pBELM + PTB, E pBELM + HPV-16 E2. Error bars indicate standard error for each sample. * indicates significance in relation to pBEL alone.



	% Peak 1	% Peak 2	Standard Error
A pBELM	70.11	29.88	2.223
B pBELM+Ad E4orf4	63.88	36.11	2.399
C pBELM+SRp30c	100	0	0
D pBELM+PTB	76.86	23.13	1.939
E pBELM+HPV-16 E2	49.60	36.57	2.223

Figure 4.41 Percentage peak areas for 757s-L1a(mut) real-time PCR for pBELM co-transfected with viral and splice regulatory proteins in A549 cells

A pBELM (plasmid control), B pBELM + Ad E4orf4, C pBELM + SRp30c, D pBELM + PTB, E pBELM + HPV-16 E2. Peak 1: splice from SD880 to SA5639 (E4^L1 splice product), Peak 2: splice from SD 880 to SA3358 and SD3632 to SA5639 (L1 splice product). Error bars indicate standard error for each sample.

4.3.10 Determination of the Effect of Altering Concentration of Protein Expression Plasmid on Late Gene Expression by Northern Blot

The previous experiments had demonstrated that HPV-16 E2, Adenoviral E4orf4 and SRp30c could induce the expression of the late gene products from sub-genomic HPV expression plasmids. As the mRNAs produced from these plasmids were analysed using L1 specific primers, the induction of L2 could not be ruled out. A radio-labelled probe designed to bind within the L1 region was used to detect RNA species that contained the late coding region: this probe can detect RNAs that contain both the L1 and also the L2 coding regions. Using Northern blot analysis the effect of varying the concentration of the protein expression plasmids was tested on the pBEL expression plasmid. Concentrations of 0.25μg, 0.5μg, 0.75μg and 1μg of each protein expression plasmid were tested.

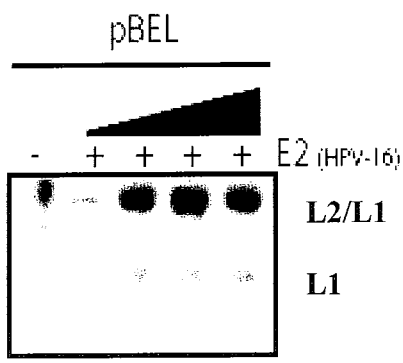


Figure 4.42 Effect of HPV-16 E2 on late gene expression

Transfections were carried out with pBEL only, or pBEL co-transfected with increasing concentrations of HPV-16 E2 starting at 0.25μg and increasing to 0.5μg, 0.75μg and 1μg. The unspliced late mRNA (L2/L1) and spliced late mRNA (L1) are indicated. Membrane radio-labeled with a HPV L1 probe.

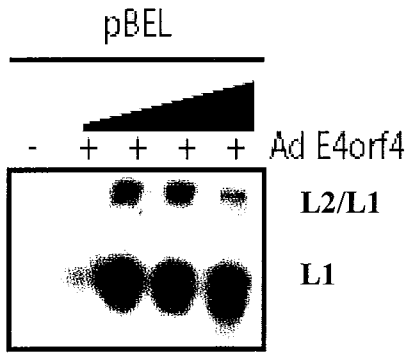


Figure 4.43 Effect of Ad E4orf4 on late gene expression

Transfections were carried out with pBEL only, or pBEL co-transfected with increasing concentrations of Adenoviral E4orf4 starting at 0.25μg and increasing to 0.5μg, 0.75μg and 1μg. The unspliced late mRNA (L2/L1) and spliced late mRNA (L1) are indicated. Membrane radio-labeled with a HPV L1 probe.

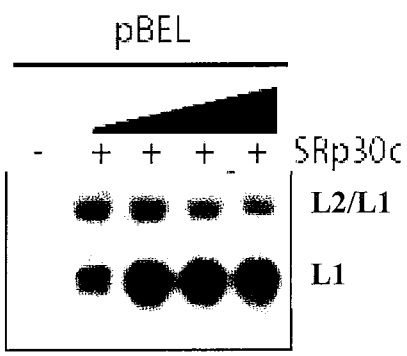


Figure 4.44 Effect of SRp30c on late gene expression

Transfections were carried out with pBEL only, or pBEL co-transfected with increasing concentrations of SRp30c starting at 0.25μg and increasing to 0.5μg, 0.75μg and 1μg. The unspliced late mRNA (L2/L1) and spliced late mRNA (L1) are indicated. Membrane radio-labeled with a HPV L1 probe.

For both HPV-16 E2 (Fig. 4.42) and Ad E4orf4 (Fig. 4.43) plasmids transfection with 0.25μg caused only a small induction of late RNA but for both plasmids 0.5μg demonstrated strong induction. For HPV-16 E2 there was no change in the late RNAs produced as the concentration of the protein increased to 1μg, the concentration of late

RNAs and the ratio of L2/L1 to L1 remained constant. For Ad E4orf4 0.5μg and 0.75μg of plasmid induced the same concentration and ratio of L2/L1 to L1 but when 1μg of Ad E4orf4 expression plasmid was transfected L2/L1 expression was reduced. When 0.25μg of the SRp30c plasmid was transfected a strong induction of both L2/L1 and L1 was detected (Fig. 4.44). When 0.5μg of the Ad E4orf4 expression plasmid is transfected, the L2/L1 expression remains the same while the expression of L1 is increased. There is no change in the expression levels from 0.5μg up to 1μg of expression plasmid.

4.3.11 Transfection Efficiency in HeLa and A549 Cells

In order to determine that transfection efficiency did not change significantly between cell plates a liquid assay for β-galactosidase was carried out on 10 cell plates, the experiment was carried out separately for each cell type. Absorbance readings for each plate were measured 3 times. The percentage variance between the 10 plates was 14.2% for HeLa cells and 4.5% for A549 cells (Tables 4.7 and 4.8). Variance should be below 20%; above this level it is considered that inter-plate transfection variability is too great.

Plate No.	Absorbance
1	0.273
2	0.226
3	0.316
4	0.266
5	0.241
6	0.234
7	0.310
8	0.338
9	0.310
10	0.317
Overall Average	0.283
Std Dev	0.040
% Variance	14.19

Table 4.7 Average absorbance readings for β-galactosidase liquid assay for 10 plates of HeLa cells

Plate No.	Absorbance
1	0.859
2	0.815
3	0.812
4	0.786
5	0.844
6	0.775
7	0.780
8	0.828
9	0.773
10	0.741
Overall Average	0.801
Std Dev	0.036
% Variance	4.547

Table 4.8 Average absorbance readings for β-galactosidase liquid assay for 10 plates of A549 cells

4.3.12 Key Findings

- ASF/SF2, hnRNP A1 and hnRNP H cannot induce HPV late gene expression
- HuR can induce late gene expression
- ASF.SF2, hnRNP H and HuR can induce a change in the ratio of the L1 mRNA splice products produced using a plasmid in which splicing silencers in the L1 coding region are destroyed
- Ad E4orf4, SRp30c, PTB and HPV-16 E2 can induce late gene expression in HPV sub-genomic expression plasmids that cannot on their own produce the HPV late genes
- Ad E4orf4, SRp30c, PTB and HPV-16 E2 can cause a greater induction of the late genes in a plasmid which, on its own, can produce the HPV late genes, as the splicing silencers in the L1 coding region are destroyed
- E4orf4 and SRp30c can induce a change in the ratio of the L1 mRNA splice products produced
- When E4orf4 is co-expressed with hnRNP A1 an increase in late gene expression is induced
- hnRNP A1 may compete with SRp30c during late gene expression

4.4 Discussion

4.4.1 Splicing Proteins

In this study it was investigated whether the proteins that had been evaluated in HPV infected cervical epithelium using immunohistochemistry had an effect at the molecular level on the splicing of the HPV genome. It was determined that ASF/SF2, hnRNP A1, hnRNP H, HuR and PTB had no effect on the splicing in the early region contained in the pBEL plasmid, and these proteins could not induce the expression of the late genes in this plasmid. But when a mutation is introduced into the plasmid (pBELM plasmid) reducing the negative regulation on the splice site in the late region this study showed that there is a regulatory effect by some of these proteins. The pBELM plasmid produced the late genes when transfected alone, but the expression of some of the splice regulatory proteins appeared to cause a change in the ratio of the two splice products. When real-time PCR was carried out on these samples it was found that ASF/SF2, hnRNP A1 and hnRNP H caused no change in overall expression of the late genes when compared to pBELM alone. This would correlate with the previously identified role of hnRNP A1 and hnRNP H as silencers of HPV-16 late gene expression [140, 177]. PTB induced a slight increase in late gene expression but only HuR caused a significant increase in the induction of the late genes. As HuR could not induce late gene expression in the pBEL plasmid, the loss of the negative regulation in pBELM allows HuR to positively influence splicing into the late region.

These 5 proteins have been shown previously to be involved in splicing [80, 136, 140, 142, 143, 148, 150, 177, 215]. Although only HuR and PTB had an effect on the overall amount of late gene expression from the pBELM plasmid, ASF/SF2, hnRNP H and HuR did have an effect on the ratio of the two splice products. hnRNP H and HuR reduced the amount of

E4^L1 and increased the amount of direct splicing into L1. Without the presence of the negative regulation on SA5639 splice site (pBELM) both of these proteins were able to have a positive influence on the direct splicing into the L1 coding region. This means that both proteins have a weak influence on this splice site that can only be seen when the negative regulation is removed. As hnRNP H causes no increase in the expression of the late genes and only a small alteration in the ratio of E4^L1 to L1 production, it can be concluded that this protein has only a very small effect on late gene expression. HuR causes a significant increase in late gene expression and also a large change in the splice product ratio; this would indicate that HuR can have a big influence on late gene expression but only when hnRNP A1 negative regulation is removed (by the mutation in the pBELM plasmid). ASF/SF2 has a major influence on the E4^L1 to L1 ratio. It causes 100% inclusion of the E4 coding region when hnRNP A1 repression is removed. The E4^L1 splice product contains the coding regions for both the E4 and L1 protein: by causing the incorporation of the E4 coding region in 100% of the splice products this shows that ASF/SF2 may be a positive regulator of the SA3358 splice site. This positive effect on a 5' splice site has been demonstrated previously in the alternative splicing of the Adenoviral E1A gene [240].

4.4.2 Ad E4orf4

The adenoviral E4rof4 protein has been shown to interact with the protein phosphatase PP2A, this interaction can regulate the alternative splicing of the adenovirus late genes [220, 230].

Using RT-PCR it has been demonstrated that Ad E4orf4 can induce late gene production from both the pBEL and p97CEL plasmids but that it had no effect on splicing in the early region in the pC97EL plasmid (Rush M, Schwartz S. Unpublished). The Northern blot analysis in this study demonstrated that Ad E4orf4 induced expression of the L1 mRNA more than the L2/L1 mRNA.

Real-time PCR demonstrated Ad E4orf4 had a much greater influence on late gene production from the pBEL plasmid when compared to the pBELM plasmid, although this difference may be due to the high late gene expression from the pBELM alone. Ad E4orf4 also has different effects in the two plasmids when the ratio of the two splice products are analysed. In pBEL Ad E4orf4 can completely reverse the ratio of E4^L1 to L1, while in the pBELM plasmid it can only change the ratio to approximately 50:50 production. This ratio in the pBELM plasmid indicates that there is not as much direct splicing into the late region when compared with the pBEL plasmid. This may indicate that the mutation in pBELM results in decreased direct splicing, possibly due to deletion or mutation of some of the Ad E4orf4 target sites, therefore reducing the positive influence on the SA5639 splice site. As Ad E4orf4 still changes the E4^L1 to L1 ratio in the pBELM plasmid it would indicate that not all the target sites are deleted, some must remain in order for it to influence the change in ratio. The effect could possibly also be that the E4orf4 target sites are weaker due to the mutation and so the strong influence seen in pBEL cannot be reproduced in pBELM.

When it was investigated if the introduction of hnRNP A1 could reverse the effect that Ad E4orf4 has on late gene expression, it was found that the splicing silencer hnRNP A1 actually increased the level of late gene expression. When the splice products were examined, results showed that hnRNP A1 had no effect on the choice of splice site and that

it cannot inhibit the use of the late splice site. This indicates that E4orf4 remains able to positively regulate the use of the SA5639 splice site even in the presence of hnRNP A1, but that hnRNP A1 and E4orf4 together can induce an increase in overall late gene expression.

4.4.3 SRp30c

SRp30c is a member of the SR protein family and is involved in the alternative splicing of the hnRNP A1 gene [228]. SRp30c interacts with Y-box protein 1 leading to the selection of splice sites to be used in alternative splicing [232].

RT-PCR using the pC97EL plasmid demonstrated that SRp30c could induce the expression of the late genes but that it had no effect on the splicing of the early genes (this study and Rush M, Schwartz S. Unpublished). From the Northern blots it was observed that SRp30c mostly influences the expression of the L1 message only. Real-time PCR showed that SRp30c could induce strong expression of the late genes from both the pBEL and pBELM plasmids. In both of the plasmids in this study the increase in late gene expression was significant, and the induction was approximately the same (decrease in CT by approximately 3 cycles for both plasmids). This would indicate that the mutation in the pBELM plasmid has no influence on the action of SRp30c.

Although the mutation has no effect on the ability of SRp30c to induce late gene expression, it does have an effect on the ratio of the splice products. When SRp30c is expressed with pBEL the ratio of E4^L1 to L1 is the same as when pBEL is expressed alone. But when SRp30c is expressed with pBELM there is a dramatic reduction in the amount of direct splicing into the L1 coding region. This indicates that when hnRNP A1 binds to sites within the L1 coding region, negatively regulating the use of the SA5639

splice site, SRp30c can only increase the overall amount of late gene expression. When hnRNP A1 can no longer regulate the SA5639 splice site (in the pBELM plasmid) SRp30c can not only increase the amount of late gene production, but can also cause an increase in the formation of the E4^L1 splice product. This would indicate a role of SRp30c in the positive regulation of the SA3358 splice site, but why only in the pBELM plasmid? The mutation that is present in the pBELM plasmid may disrupt some of the SRp30c binding sites, and with fewer binding sites available for the same amount of SRp30c, the protein may bind to weaker sites located near the SA3358 splice site increasing its use.

When SRp30c is expressed with hnRNP A1 it was found that the effect of hnRNP A1 was concentration-dependant. At the highest concentration used (4μg) no effect on late gene expression was seen, at the lowest hnRNP A1 concentration (1μg), like Ad E4orf4 there was a dramatic increase in late gene expression. At 2μg of hnRNP A1 the level of late gene expression was the same as for pBEL alone. This would indicate that at a concentration of 2μg hnRNP A1 is able to compete with SRp30c and overcome the positive effect that occurs. Although hnRNP A1 is able to compete with SRp30c and return late gene levels to normal, it is unable to completely return the ratio of the splice products to normal values. This indicates that SRp30c still has a positive influence on the SA3358 splice site even in the presence of hnRNP A1.

4.4.4 Polypyrimidine Tract Binding Protein (PTB)

In mammals PTB acts as a key splicing suppressor [90, 145], it generally binds to intronic sequences that mediate splicing repression [80].

Real-time PCR demonstrated that PTB had a significant influence on the induction of late gene expression in both the pBEL and pBELM plasmids. As with Ad E4orf4, PTB had a greater influence on the pBEL plasmid compared to the pBELM. This again is probably due to the strong expression of the late genes form the pBELM plasmid when transfected alone. When PTB is expressed with pBEL it has no effect on the ratio of E4^L1 to L1 production when compare to pBEL alone but when expressed with pBELM there is a slight increase in the E4^L1 splice product. This would indicate that the mutation contained in the pBELM plasmid may have a minor effect on PTB binding close to the SA5639 splice site, decreasing the positive influence on this site and therefore reducing the amount of splicing directly into the L1 coding region.

4.4.5 HPV-16 E2

The HPV-16 E2 protein is involved in both replication and transcription during the viral life-cycle and may be involved in the regulation of splicing of its own late genes by controlling the expression of the SR protein ASF/SF2 [136].

RT-PCR demonstrated that HPV-16 E2 can induce the expression of the late genes from the pBEL plasmid. Using Northern blot technology to look at this induction it can be seen that HPV-16 E2 influences the expression of the L2/L1 message but has very little effect on the production of the L1 message. With real-time PCR it was found that HPV-16 E2 had no effect on the late gene expression when compared to pBEL alone. When HPV-16 E2 was

co-expressed with pBELM a significant increase in late gene expression was observed. From the Northern blot analysis it was shown that HPV-16 E2 had a greater influence on the expression of L2/L1. This increase could not be detected as this product was too large to be amplified using PCR. When the mutation was introduced into the HPV expression plasmid and the inhibition of the SA5639 splice site by hnRNP A1 was removed, HPV-16 E2 significantly increased the late gene production. This may be due to hnRNP A1 having a stronger control over the SA5639 splice site and only when this negative inhibition is removed can HPV-16 E2 have a positive influence on this splice site. The HPV-16 E2 protein has no effect on the ratio of the two splice products, it only affects the amount of late gene production. This again could be because hnRNP A1 inhibition cannot be overcome by HPV-16 E2. Only when the negative regulation by hnRNP A1 is removed by mutation can HPV-16 E2 have any influence, as increasing the use of just the late splice site would have an equal effect on both splice products therefore maintaining the same E4^L1 to L1 ratio.

4.4.6 Summary and Conclusions

HPV-16 late gene expression is a complex process involving many regulatory steps. This study has established the role of a number of proteins in the regulation of late gene expression and mRNA splicing. Further work exploring the relationship of these proteins to the neoplastic lesions would be of interest.

5. Chapter 5 - General Discussion

Cervical cancer is the second most common cancer in women worldwide. As a HPV infection is necessary for the development of this cancer much research has been carried out into this virus and its relationship with cervical cancer. Approximately 80% of women will contract a HPV infection at some point in their lives [27] but only a small percentage of these will ever go on to develop cervical cancer or a pre-malignant lesion that could progress to cancer. Persistence of a HPV infection is known to be necessary for cancer to develop [4], but why some women can clear the infection and others cannot is not known. In all populations, the HPV types present in the healthy population and in women with cervical abnormalities must be determined in order to develop an overall understanding of this virus, its relationship with cervical cancer and appropriate screening methods. With the recent development of a vaccine for cervical cancer further research into HPV will need to be reviewed. Due to the type specificity of current vaccines (only protecting against HPV-16 and -18 and with the Merck quadrivalent vaccine also HPV-6 and -11), they will only be able to prevent approximately 65-75% of cervical cancers [29]. The vaccines are based on the ability of the L1 protein to self-assemble into virus-like particles (VLPs) that are morphologically indistinguishable from the authentic virions [241]. As these VLPs do not contain any HPV DNA they are non-infectious [242]. Upon vaccination with these VLPs, high titres of virion-neutralising antibodies are produced inducing a strong and sustained but HPV type-specific immune response [241, 243]. These vaccines are prophylactic and have unknown therapeutic value for those with established HPV infections [244]. 83% of cervical cancer cases occur in developing countries [245] and so the high cost of

vaccination will mean that the areas in the world that are most affected by cervical cancer will not be able to afford to take full advantage of the availability of the vaccines. Therefore cervical smear screening will continue to play an important part in the control of cervical cancer as it is a more affordable option when compared to mass vaccination. Continued research into biomarkers of HPV related disease will be important in order to improve the sensitivity of diagnosis of cervical abnormalities during screening. Also research into how the virus alters the infected cell during infection, what cellular proteins are used for the life-cycle and how these proteins are utilised by the virus will possibly provide clues into new lower cost treatments or vaccines for HPV infections in order to prevent cervical cancer.

An important part of the viral life-cycle is the regulation of production of the viral capsid proteins (the late genes) L1 and L2. These two proteins are the most immunogenic of the HPV proteins, a fact that has been used in the development of the cervical cancer vaccine [24]. Since these proteins can induce an immune response these genes are tightly regulated, their production is inhibited in the lower layers of the epithelium where detection by the host immune system can occur [16]. This facilitates persistence of HPV infection by evasion of the immune response. Once the infected cell reaches the upper epithelial layers, where immune surveillance is minimal, expression of these proteins can occur, requiring a down-regulation of the negative inhibition of these genes. A key hypothesis of this study is that the expression of L1 and L2 could be promoted using various regulatory proteins in HPV positive neoplastic cells, this could trigger an immune response towards the HPV infected cells, and prevent progression to cancer.

Overall the aims of this study were to investigate the expression profile of RNA-binding proteins that have been shown to be involved in HPV gene regulation and to investigate if

the expression profile of these proteins could be used as biomarkers of high-grade lesions. This study had three main sections, firstly the selection of both a cohort of patients undergoing treatment for cervical abnormalities and a smaller group of patients for nucleic acid detection of HPV and HPV-16. Secondly immunohistochemistry was utilised on tissue samples from this small group of patients to determine the presence and location of proteins involved in HPV gene regulation. These proteins had previously been shown to alter HPV gene expression in transfected cells [136, 140, 142, 148, 152, 177]. Following on from this work, molecular investigations (RT-PCR, Northern blot and real-time PCR) were performed to provide evidence that proteins binding to HPV RNA also function as regulators of HPV gene expression in transfected cells.

In the first part of this study 961 patients undergoing treatment for abnormal smears in the National Maternity Hospital were selected. In this group of patients a large number (21.85%) were diagnosed as having no neoplasia present in the LLETZ sample, this figure was similar to a previous study [106], which also reported that patients with negative LLETZ pathology still had a risk of recurrence of cervical disease. Nested PCR was used to detect HPV DNA in these patients, 98% were found to be HPV positive. The one HPV negative patient was diagnosed with no lesion present, this could be an indicator of reduced risk for recurrence. A nested PCR for detection of HPV-16 infection was developed for this study. Of the 49 HPV positive patients, 55.1% were HPV-16 positive, the number of HPV-16 positive patients increased as the patients deteriorated from no neoplasia (10% HPV-16 positive) to invasive SCC (80% HPV-16 positive). HPV-16 has a known high oncogenic potential, it is linked to 50% of all cervical carcinomas [244] and is also the most common HPV type detected in cervical neoplastic lesions [22]. The high prevalence if HPV-16 is

seen throughout the world including, Asia, Europe, Africa and North, Central and South America [22]. It is for this the reason that HPV-16, among others, is being targeted by both the current vaccines [24] in order to prevent the majority of cervical disease.

The second part of the study was to identify a panel of antibodies that can be used to detect a number of RNA-binding proteins involved in the regulation of HPV gene expression. The expression of these factors was compared with markers of differentiation, markers of proliferation and markers of neoplasia. In epithelium, with no neoplasia present, staining of basal cells differed between the RNA-binding proteins, but all of these proteins demonstrated strong staining in intermediate layers and no staining in the superficial layers. Overall lesional cells in CIN and SCC samples demonstrated strong staining for the presence of RNA-binding proteins.

The deregulation of the cell cycle through the action of the HPV proteins is still a major driving force behind the development of cervical neoplasia. In HPV infection the inactivation of pRB by HPV E7 results in the accumulation of p16 [166]. Many studies have been carried out in order to evaluate the usefulness of p16INK4a as a marker of neoplasia [115, 190, 200, 202, 246-248]. A number of RNA-binding proteins have also previously been reported to have either an increase in expression or be over-expressed in different cancers and are possibly useful as markers of cancer [143, 191, 195, 249]. An increase in SRp40 expression in breast cancer was reported to be correlated with alternative splicing of CD44 and possibly serving as an early marker of the risk of breast cancer patients developing lymph node metastasis [249]. hnRNP A1 and hnRNP B1 have been demonstrated to have an increased expression in lung cancer [191, 193, 194]. hnRNP A1 has also been demonstrated to be over-expressed in colon cancer [192]. hnRNP K and

hnRNP I are over-expressed in colorectal cancer and brain tumours respectively [195, 197]. hnRNP H is over-expressed in cancer of the liver, pancreas, stomach and head and neck [143, 196].

In this study p16INK4a clearly selected for high-grade lesions as all squamous cell carcinoma samples and 19 out of 20 high-grade CIN lesions (95%) were p16INK4a positive. The one negative p16INK4a high-grade lesion was also HPV-16 negative, possibly indicating a lower risk of progression. Six of ten low grade lesions (60%) were weakly positive for p16INK4a expression, only 2 of these p16INK4a samples were HPV-16 positive. These results are in line with previous studies, with p16INK4a expression demonstrated in 88% to 100% of high-grade lesions [203, 247, 250, 251]. Low-grade lesions demonstrate a wider variance in p16INK4a expression, ranging from 15% to 88% [250, 252], but in studies that examined the HPV type as well as p16INK4a expression, the presence of high-risk HPV types correlated with the over-expression of p16INK4a [247, 250]. When the expression of p16INK4a was compared to the expression of the RNA-binding proteins, the RNA-binding proteins did not have such a high association with increasing grade of CIN as demonstrated by p16INK4a.

As only one patient in the group of 50 was found to be HPV negative no conclusions could be made as to whether the expression profile of the RNA-binding proteins that was identified in the tissue samples was altered by the presence of the virus. In future studies it would be interesting to evaluate a larger group of patients who are HPV negative in order to compare the expression of the RNA-binding proteins to similar patients who are HPV positive. Overall the expression profile of the RNA-binding proteins was similar when compared between those patients who were HPV-16 positive and those who were infected

with another HPV type. This suggests that the type of HPV that the patient is infected with does not play a role in the increase in RNA-binding protein expression as CIN progresses to high grade CIN.

The third aim of this study was to identify proteins that induce the expression of the late HPV genes. An *in vitro* transfection method was used to assess the affect of the proteins when transfected with sub-genomic HPV expression plasmids. The effect on late gene expression was analysed by RT-PCR, Northern blot and a real-time PCR method that was established in this study. It was determined that hnRNP H and hnRNP A1 did not induce late gene expression in the HPV expression plasmid. This is in line with known negative regulation of the late gene expression [140, 177]. These results were also supported by the staining pattern for each of these proteins determined using immunohistochemistry where there was no staining present in the superficial epithelial layers.

In total, 5 proteins, 3 cellular, PTB (hnRNP I), HuR and SRp30c, and 2 viral, Ad E4orf4 and HPV-16 E2 were found to induce late gene expression to varying degrees. HuR was found to only induce the L1 gene when a plasmid containing a mutated late splice site was used. The wild-type splice site is negatively regulated by hnRNP A1, these results suggest that HuR cannot overcome the negative inhibition of L1 expression by hnRNP A1. HuR and also PTB appeared not to have a direct role in the expression of the late genes when their expression was examined using immunohistochemistry. Both proteins were only present in the basal and intermediate layers of normal cervical epithelium and both were strongly expressed in all grades of neoplasia when the late genes are not expressed. But when the effect of these proteins was evaluated using transfection studies in cell lines, both seemed to have an effect on the induction of the late genes. One possible explanation for

these results is the presence of an antagonist to both HuR and PTB. This antagonist could out-compete PTB and HuR for binding sites in the lower layers of the epithelium and in doing so inhibit late gene expression. As differentiation continues the expression of the antagonist decreases along with HuR and PTB alleviating the inhibition of the late genes allowing their expression. In the transfection studies both PTB and HuR are over-expressed resulting in increased competition with the antagonist for binding sites, this could lead to increased HUR and PTB binding and the induction of the late genes. Therefore, while the over-expression of both PTB and HuR *in vitro* can lead to the induction of the late genes, their induction *in vivo* may be due to the decrease of the antagonist rather than an increase in expression of PTB or HuR. Further research is needed to establish the mechanism of induction by PTB and HuR and also to identify how the late genes are induced in terminally differentiating epithelial cells. PTB and HuR are RNA-binding proteins that are involved in the repression of splicing and the stabilisation of mRNAs respectively [145, 149]. Both function by binding to specific sequences in the mRNA, PTB mainly to sites within the polypyrimidine tract and HuR to AU-rich elements [90, 149]. Recently both PTB and HuR have both been reported to up-regulate the translation of hypoxia-inducible factor 1alpha (HIF-1alpha) in response to hypoxia in HeLa cells [253]. This up-regulation is achieved by the binding of PTB and HuR to the 3'UTR and 5'UTR, respectively, of the HIF-1alpha mRNA [253].

To further evaluate the function of these RNA-binding proteins it would be necessary to identify where these proteins, PTB (hnRNP I) and HuR (and also the possible antagonist) bind the HPV RNA. Once these binding sites are identified further investigation of the

mechanism of action of these proteins could be determined, providing more information on how the virus differentially expresses the proteins required during its life-cycle.

A general antibody against SR proteins was used to establish the expression profile of these proteins in tissue samples. This indicated that as the SR proteins were not expressed in the superficial epithelial layers and were present in high levels in neoplastic lesions they did not induce late gene expression. The effect of two individual SR proteins, ASF/SF2 and SRp30c was examined using transfection studies to assess their role in late gene expression. The results of the transfection studies for ASF/SF2 supported the immunohistochemistry result by demonstrating that ASF/SF2 did not increase the expression of the late genes above the basal expression levels. SRp30c on the other hand did seem to induce the expression of the late genes. The general antibody (SR(1H4)) used to detect the SR proteins in the cervical tissue was raised against full-length SR of Xenopus origin and is recommended for the detection of SR proteins from human origin. Western blot analysis on cell extracts demonstrates that the antibody detects proteins in the 50 – 100kDa range (from Santa Cruz antibody datasheet which can be downloaded at: http://www.scbt.com/product.php?datasheet=13509). SRp30c is a 25kDa protein [228], therefore this general SR may not be optimal in detecting SRp30c. Unfortunately no specific antibody against SRp30c could be found, it would be useful to produce a specific antibody for SRp30c, and further investigate expression in the superficial epithelial layers in HPV infected tissue.

Adenoviral E4orf4 is involved in the regulation of late gene expression during adenovirus infection [254]. Ad E4orf4 has been shown to bind to both SR proteins and PP2A (a protein phosphatase) leading to the dephosphorylation and inactivation of the SR proteins [220]. In

adenovirus infection the dephosphorylation of SR proteins by Ad E4orf4 induces late gene expression [220]. E4orf4 enhances the use of the distal adenoviral L1 splice acceptor site promoting the expression of the IIIa protein at the expense of the 52,55K protein [254]. E4orf4 has also been shown to specifically induce apoptosis in cancer cells but not in normal cells [255]. The induction of apoptosis in these cells requires the interaction of E4orf4 with PP2A, although the downstream effectors of this interaction are not known [255].

The transfection studies carried out here demonstrate that Ad E4orf4 can also induce the expression of the HPV late genes. The action of Ad E4orf4 in dephosphorylating SR proteins may mimic the fate of these proteins during epithelial cell differentiation, although further experiments would be necessary to confirm this. Its has also been shown that Ad E4orf4 primarily binds one SR protein, ASF/SF2 [231]. Using immunohistochemistry it was demonstrated that ASF/SF2 is down-regulated during epithelial differentiation, and was found in transfection studies that ASF/SF2 did not to induce late gene expression. Since Ad E4orf4 can induce late HPV gene expression and primarily interacts with ASF/SF2, it could be suggested that ASF/SF2 is a negative regulator of late gene expression, and that the dephosphorylation of ASF/SF2 by Ad E4orf4 in transfected cells causes the induction of the late genes. This role would be supported by the down-regulation of ASF/SF2 demonstrated by immunohistochemistry. Again further studies would be needed to examine if the induction of the HPV late genes by Ad E4orf4 is through its interaction with ASF/SF2 or if Ad E4orf4 is using another mechanism of action to induce the expression of these genes.

HPV is the most common sexually transmitted infection worldwide [13, 16] and accounts for 5% of all cancers [11]. It is estimated that if screening for cervical cancer did not occur cervical cancer would occur in 1% of all women who acquire a HPV infection [15]. For these reasons research into screening methods, HPV detection and vaccine development will continue to be important. It will be necessary in the future to continue the clinical trials for the two vaccines to understand their full potential. Once large numbers of women who have been vaccinated reach the age for routine cervical screening, the effect on the screening process will have to be monitored to assess what decrease in HPV related cervical abnormalities results. When the mass vaccination of individuals occurs the overall HPV prevalence will also have to be examined, to determine if less common high risk HPV types, or HPV variants become more prevalent. The cost and effectiveness of developing vaccines for further high risk HPV types will also be studied to gauge whether their introduction will reduce the cervical cancer rate enough to make development and mass vaccination feasible.

There is also potential for continuing research into the mechanisms of the viral life-cycle, to evaluate splicing and gene expression. Alternative splicing and alternative polyadenylation are two ways the virus uses to increase its coding potential but how these mechanisms are being regulated is still being investigated [38, 39]. Many positive and negative regulatory sites and the proteins that bind to them have been identified, however, the existence or location of regulatory sequences have not been identified for some alternative splice sites; also if regulatory sequences have been identified the proteins that bind to them have not [38, 39]. The negative inhibition of the late genes is an important part of the HPV life-cycle, this regulation has been shown to include positive regulation of the early poly(A)

signal and negative regulation of the late splice site [92, 140, 142, 177]. Most studies have been carried out on *in vitro* transfection studies, but whether these results carry over to the *in vivo* situation is not yet known [40, 152]. Continuing research in this area will hopefully lead to a complete understanding of how the regulation of expression of HPV proteins occurs and the translation of this research for clinical application.

Cervical cancer cells never express the HPV late genes, L1 and L2. Inhibition of the HPV late gene expression is probably a prerequisite for establishment of persistence and progression to cervical cancer. This present study investigated the role of RNA-binding proteins in the regulation of late gene expression in HPV infected cells. It would be interesting to identify if the expression of the late genes could be induced in cancer cells and if this then caused an effective immune response against the cancer. If an effective treatment for HPV in high grade lesions and cervical cancer was developed and used in combination with the HPV vaccines a significant reduction in both the development of and deaths from cervical cancer could take place.

6. References

1. Flint, S.J., et al., *Chapter 1: Foundations*, in *Principles of Virology: Molecular Biology, Pathogensis & Control of Animal Viruses*. 2004, ASM Press: Washington D.C. p. 2-25.
2. Levine, A.J., *Chapter 1: The Origins of Vriology*, in *Fundamental Virology*, D.M. Knipe and P.M. Howley, Editors. 2001, Lipincott, Williams & Wikins: Philadelphia. p. 3-18.
3. Turek, L.P., *The structure, function, and regulation of papillomaviral genes in infection and cervical cancer.* Adv Virus Res, 1994. **44**: p. 305-56.
4. Lowy, D.R. and P.M. Howley, *Chapter 66: Papillomaviruses*, in *Fields Virology*, D.M. Knipe and P.M. Howley, Editors. 2001, Lippincott, Williams & Wilkins. p. 2231-2264.
5. zur Hausen, H., *Papillomavirus infections--a major cause of human cancers.* Biochim Biophys Acta, 1996. **1288**(2): p. F55-78.
6. Orth, G., *Papillomaviruses - Human (Papovaviridae): General Features*, in *Encycolpedia Of Virology*, A. Granoff and R. Webster, Editors. 1999, Academic Press. p. 1105-1114.
7. zur Hausen, H., *Papillomaviruses and cancer: from basic studies to clinical application.* Nat Rev Cancer, 2002. **2**(5): p. 342-50.
8. Bernard, H.U., *The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses.* J Clin Virol, 2005. **32 Suppl 1**: p. S1-6.

9. Howley, P.M. and D.R. Lowy, *Chapter 65: Papillomaviruses and Their Replication*, in *Fields Virology*, D.M. Knipe and P.M. Howley, Editors. 2001, Lippincott, Williams & Wilkins. p. 2197-2230.

10. de Villiers, E.M., et al., *Classification of papillomaviruses.* Virology, 2004. **324**(1): p. 17-27.

11. Lowy, D.R. and J.T. Schiller, *Prophylactic human papillomavirus vaccines.* J Clin Invest, 2006. **116**(5): p. 1167-73.

12. WHO, *Comprehensive cervical cancer control : a guide to essential practice*. 2006, Geneva: WHO Press. 1-282.

13. Tjalma, W.A., et al., *Role of human papillomavirus in the carcinogenesis of squamous cell carcinoma and adenocarcinoma of the cervix.* Best Pract Res Clin Obstet Gynaecol, 2005. **19**(4): p. 469-83.

14. Tiltman, A.J., *The pathology of cervical tumours.* Best Pract Res Clin Obstet Gynaecol, 2005. **19**(4): p. 485-500.

15. Kitchener, H.C., P.E. Castle, and J.T. Cox, *Chapter 7: Achievements and limitations of cervical cytology screening.* Vaccine, 2006. **24 Suppl 3**: p. S63-70.

16. Scheurer, M.E., G. Tortolero-Luna, and K. Adler-Storthz, *Human papillomavirus infection: biology, epidemiology, and prevention.* Int J Gynecol Cancer, 2005. **15**(5): p. 727-46.

17. Motoyama, S., et al., *The role of human papilloma virus in the molecular biology of cervical carcinogenesis.* Kobe J Med Sci, 2004. **50**(1-2): p. 9-19.

18. IARC, *IARC Monographs on the evaluation of carcinogenic risks to humans. Human Papillomaviruses.* 1995, International Agency for Research on Cancer: Lyon.

19. Munoz, N., et al., *Chapter 1: HPV in the etiology of human cancer.* Vaccine, 2006. **24S3**: p. S1-S10.

20. IARC, *IARC Monographs on the evaluation of carcinogenic risks to humans. Human Papillomaviruses.* 2005, International Agency for Research on Cancer: Lyon.

21. Baseman, J.G. and L.A. Koutsky, *The epidemiology of human papillomavirus infections.* J Clin Virol, 2005. **32 Suppl 1**: p. S16-24.

22. Clifford, G., et al., *Chapter 3: HPV type-distribution in women with and without cervical neoplastic diseases.* Vaccine, 2006. **24 Suppl 3**: p. S26-34.

23. Moscicki, A.B., et al., *Chapter 5: Updating the natural history of HPV and anogenital cancer.* Vaccine, 2006. **24 Suppl 3**: p. S42-51.

24. Inglis, S., A. Shaw, and S. Koenig, *Chapter 11: HPV vaccines: Commercial Research & Development.* Vaccine, 2006. **24 Suppl 3**: p. S99-S105.

25. Koutsky, L.A. and D.M. Harper, *Chapter 13: Current findings from prophylactic HPV vaccine trials.* Vaccine, 2006. **24 Suppl 3**: p. S114-21.

26. Frazer, I., *HPV vaccines.* Int J Gynaecol Obstet, 2006. **94 Suppl 4**: p. S81-S88.

27. Wright, T.C., et al., *Chapter 30: HPV vaccines and screening in the prevention of cervical cancer; conclusions from a 2006 workshop of international experts.* Vaccine, 2006. **24 Suppl 3**: p. S251-61.

28. Hildesheim, A., et al., *Chapter 27: Research needs following initial licensure of virus-like particle HPV vaccines.* Vaccine, 2006. **24 Suppl 3**: p. S227-32.

29. Batson, A., F. Meheus, and S. Brooke, *Chapter 26: Innovative financing mechanisms to accelerate the introduction of HPV vaccines in developing countries.* Vaccine, 2006. **24 Suppl 3**: p. S219-25.

30. Favre, M., N. Ramoz, and G. Orth, *Human papillomaviruses: general features.* Clin Dermatol, 1997. **15**(2): p. 181-98.

31. Doorbar, J., *The papillomavirus life cycle.* J Clin Virol, 2005. **32 Suppl 1**: p. S7-15.

32. Longworth, M.S. and L.A. Laimins, *Pathogenesis of human papillomaviruses in differentiating epithelia.* Microbiol Mol Biol Rev, 2004. **68**(2): p. 362-72.

33. Stubenrauch, F. and L.A. Laimins, *Human papillomavirus life cycle: active and latent phases.* Semin Cancer Biol, 1999. **9**(6): p. 379-86.

34. Flores, E.R., et al., *Establishment of the human papillomavirus type 16 (HPV-16) life cycle in an immortalized human foreskin keratinocyte cell line.* Virology, 1999. **262**(2): p. 344-54.

35. Hebner, C.M. and L.A. Laimins, *Human papillomaviruses: basic mechanisms of pathogenesis and oncogenicity.* Rev Med Virol, 2006. **16**(2): p. 83-97.

36. Lambert, P., et al., *Papillomaviruses - Human (Papovaviridae): Molecular Biology*, in *Encycolpedia Of Virology*, A. Granoff and R.G. Webster, Editors. 1999, Academic Press. p. 1115-1120.

37. Stubenrauch, F., et al., *The E8E2C protein, a negative regulator of viral transcription and replication, is required for extrachromosomal maintenance of human papillomavirus type 31 in keratinocytes.* J Virol, 2000. **74**(3): p. 1178-86.

38. Zheng, Z.M. and C.C. Baker, *Papillomavirus genome structure, expression, and post-transcriptional regulation.* Front Biosci, 2006. **11**: p. 2286-302.

39. Schwartz, S., et al., *Regulation of HPV Gene Splicing and Expression*, in *Human Papillomavirus Gene Regulation and Transformation*, B. Norrild, Editor. 2007, Transworld Research Network: Kerala. p. 47-67.

40. Cumming, S.A., et al., *Activity of the human papillomavirus type 16 late negative regulatory element is partly due to four weak consensus 5' splice sites that bind a U1 snRNP-like complex.* J Virol, 2003. **77**(9): p. 5167-77.

41. Syrjanen, K. and S. Syrjanen, *Chapter 2: Molecular Biology of papillomaviruses*, in *Papillomavirus Infections in Human Pathology*. 2000, John Wiley & Sons: New York. p. 11-52.

42. IARC, *Chapter 1: Cervical Cancer and Screening*, in *IARC Handbooks of Cancer Prevention: Cervical Cancer Screening*. 2005, IARC Press. p. 1-57.

43. McMurray, H.R., et al., *Biology of human papillomaviruses.* Int J Exp Pathol, 2001. **82**(1): p. 15-33.

44. Clower, R.V., J.C. Fisk, and T. Melendy, *Papillomavirus E1 protein binds to and stimulates human topoisomerase I.* J Virol, 2006. **80**(3): p. 1584-7.

45. Blachon, S. and C. Demeret, *The regulatory E2 proteins of human genital papillomaviruses are pro-apoptotic.* Biochimie, 2003. **85**(8): p. 813-9.

46. Enari, M., et al., *A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD.* Nature, 1998. **391**(6662): p. 43-50.

47. Fang, L., et al., *The human papillomavirus type 11 E1^E4 protein is not essential for viral genome amplification.* Virology, 2006. **351**(2): p. 271-9.

48. Nath, R., et al., *Analyses of variant human papillomavirus type-16 E5 proteins for their ability to induce mitogenesis of murine fibroblasts.* Cancer Cell Int, 2006. **6**: p. 19.

49. Tsai, T.C. and S.L. Chen, *The biochemical and biological functions of human papillomavirus type 16 E5 protein.* Arch Virol, 2003. **148**(8): p. 1445-53.

50. Cho, N.H., Y.T. Kim, and J.W. Kim, *Alteration of cell cycle in cervical tumor associated with human papillomavirus: cyclin-dependent kinase inhibitors.* Yonsei Med J, 2002. **43**(6): p. 722-8.

51. Song, S., H.C. Pitot, and P.F. Lambert, *The human papillomavirus type 16 E6 gene alone is sufficient to induce carcinomas in transgenic animals.* J Virol, 1999. **73**(7): p. 5887-93.

52. Huibregtse, J.M., M. Scheffner, and P.M. Howley, *Localization of the E6-AP regions that direct human papillomavirus E6 binding, association with p53, and ubiquitination of associated proteins.* Mol Cell Biol, 1993. **13**(8): p. 4918-27.

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

54. Fehrmann, F. and L.A. Laimins, *Human papillomaviruses: targeting differentiating epithelial cells for malignant transformation.* Oncogene, 2003. **22**(33): p. 5201-7.

55. Hubbert, N.L., S.A. Sedman, and J.T. Schiller, *Human papillomavirus type 16 E6 increases the degradation rate of p53 in human keratinocytes.* J Virol, 1992. **66**(10): p. 6237-41.

56. McMurray, H.R. and D.J. McCance, *Human papillomavirus type 16 E6 activates TERT gene transcription through induction of c-Myc and release of USF-mediated repression.* J Virol, 2003. **77**(18): p. 9852-61.

57. Liu, J.P., *Studies of the molecular mechanisms in the regulation of telomerase activity.* Faseb J, 1999. **13**(15): p. 2091-104.

58. Munger, K. and P.M. Howley, *Human papillomavirus immortalization and transformation functions.* Virus Res, 2002. **89**(2): p. 213-28.

59. Lee, S.S., R.S. Weiss, and R.T. Javier, *Binding of human virus oncoproteins to hDlg/SAP97, a mammalian homolog of the Drosophila discs large tumor suppressor protein.* Proc Natl Acad Sci U S A, 1997. **94**(13): p. 6670-5.

60. Clements, A., et al., *Oligomerization properties of the viral oncoproteins adenovirus E1A and human papillomavirus E7 and their complexes with the retinoblastoma protein.* Biochemistry, 2000. **39**(51): p. 16033-45.

61. Munger, K., et al., *Biological activities and molecular targets of the human papillomavirus E7 oncoprotein.* Oncogene, 2001. **20**(54): p. 7888-98.

62. Dell, G. and K. Gaston, *Human papillomaviruses and their role in cervical cancer.* Cell Mol Life Sci, 2001. **58**(12-13): p. 1923-1942.

63. Reinstein, E., et al., *Degradation of the E7 human papillomavirus oncoprotein by the ubiquitin-proteasome system: targeting via ubiquitination of the N-terminal residue.* Oncogene, 2000. **19**(51): p. 5944-50.

64. Cheng, S., et al., *Differentiation-dependent up-regulation of the human papillomavirus E7 gene reactivates cellular DNA replication in suprabasal differentiated keratinocytes.* Genes Dev, 1995. **9**(19): p. 2335-49.

65. Flores, E.R., et al., *The human papillomavirus type 16 E7 oncogene is required for the productive stage of the viral life cycle.* J Virol, 2000. **74**(14): p. 6622-31.

66. Munger, K. and W.C. Phelps, *The human papillomavirus E7 protein as a transforming and transactivating factor.* Biochim Biophys Acta, 1993. **1155**(1): p. 111-23.

67. Weinberg, R.A., *The retinoblastoma protein and cell cycle control.* Cell, 1995. **81**(3): p. 323-30.

68. Slebos, R.J., et al., *p53-dependent G1 arrest involves pRB-related proteins and is disrupted by the human papillomavirus 16 E7 oncoprotein.* Proc Natl Acad Sci U S A, 1994. **91**(12): p. 5320-4.

69. Giarre, M., et al., *Induction of pRb degradation by the human papillomavirus type 16 E7 protein is essential to efficiently overcome p16INK4a-imposed G1 cell cycle Arrest.* J Virol, 2001. **75**(10): p. 4705-12.

70. Davies, R., et al., *Human papillomavirus type 16 E7 associates with a histone H1 kinase and with p107 through sequences necessary for transformation.* J Virol, 1993. **67**(5): p. 2521-8.

71. Missero, C., et al., *Involvement of the cell-cycle inhibitor Cip1/WAF1 and the E1A-associated p300 protein in terminal differentiation.* Proc Natl Acad Sci U S A, 1995. **92**(12): p. 5451-5.

72. Pietenpol, J.A., et al., *TGF-beta 1 inhibition of c-myc transcription and growth in keratinocytes is abrogated by viral transforming proteins with pRB binding domains.* Cell, 1990. **61**(5): p. 777-85.

73. Duensing, S., et al., *Centrosome abnormalities and genomic instability by episomal expression of human papillomavirus type 16 in raft cultures of human keratinocytes.* J Virol, 2001. **75**(16): p. 7712-6.

74. Kieback, E. and M. Muller, *Factors influencing subcellular localization of the human papillomavirus L2 minor structural protein.* Virology, 2006. **345**(1): p. 199-208.

75. Oberg, D., et al., *Mutational inactivation of two distinct negative RNA elements in the human papillomavirus type 16 L2 coding region induces production of high levels of L2 in human cells.* J Virol, 2003. **77**(21): p. 11674-84.

76. Cooper, G.M. and R.E. Hausman, *Chapter 7: RNA Synthesis & Processing*, in *The Cell: A Molecular Approach.* 2007, ASM Press. p. 253-308.

77. Proudfoot, N.J., A. Furger, and M.J. Dye, *Integrating mRNA processing with transcription.* Cell, 2002. **108**(4): p. 501-12.

78. Neugebauer, K.M., *On the importance of being co-transcriptional.* J Cell Sci, 2002. **115**(Pt 20): p. 3865-71.

79. Bentley, D.L., *Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors.* Curr Opin Cell Biol, 2005. **17**(3): p. 251-6.

80. Soller, M., *Pre-messenger RNA processing and its regulation: a genomic perspective.* Cell Mol Life Sci, 2006. **63**(7-8): p. 796-819.

81. Gu, M. and C.D. Lima, *Processing the message: structural insights into capping and decapping mRNA.* Curr Opin Struct Biol, 2005. **15**(1): p. 99-106.

82. Strachan, T. and A.P. Read, *Chapter 1: DNA Structure and Gene Expression*, in *Human Molecular Genetics 2*. 1999, BIOS Scientific Publishers: New York. p. 1-26.

83. Brown, T.A., *Chapter 10: Synthesis & Processing of RNA*, in *Genomes 2*. 2002, BIOS Scientific Publishers New York. p. 273-312.

84. Lopez, A.J., *Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation.* Annu Rev Genet, 1998. **32**: p. 279-305.

85. ten Dam, G.B., et al., *Regulation of alternative splicing of CD45 by antagonistic effects of SR protein splicing factors.* J Immunol, 2000. **164**(10): p. 5287-95.

86. Lodish, H., et al., *Chapter 12: Post-transcriptional Gene Control & Nuclear Transport*, in *Molecular Cell Biology*. 2004, W.H. Freeman: New York. p. 493-532.

87. Gilmartin, G.M., *Eukaryotic mRNA 3' processing: a common means to different ends.* Genes Dev, 2005. **19**(21): p. 2517-21.

88. Little, P.F., *Structure and function of the human genome.* Genome Res, 2005. **15**(12): p. 1759-66.

89. Srebrow, A. and A.R. Kornblihtt, *The connection between splicing and cancer.* J Cell Sci, 2006. **119**(Pt 13): p. 2635-41.

90. Black, D.L., *Mechanisms of alternative pre-messenger RNA splicing.* Annu Rev Biochem, 2003. **72**: p. 291-336.

91. Tang, S., et al., *The E7 oncoprotein is translated from spliced E6*I transcripts in high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via translation reinitiation.* J Virol, 2006. **80**(9): p. 4249-63.

92. Rush, M., X. Zhao, and S. Schwartz, *A splicing enhancer in the E4 coding region of human papillomavirus type 16 is required for early mRNA splicing and polyadenylation as well as inhibition of premature late gene expression.* J Virol, 2005. **79**(18): p. 12002-15.

93. Terhune, S.S., C. Milcarek, and L.A. Laimins, *Regulation of human papillomavirus type 31 polyadenylation during the differentiation-dependent life cycle.* J Virol, 1999. **73**(9): p. 7185-92.

94. Junqueira, L.C. and J. Carneiro, *Chapter 22: The Female Reproductive System*, in *Basic Histology*. 2005, The McGraw-Hill Companies. p. 435-455.

95. Stevens, A. and J. Lowe, *Chapter 17: Female Reproductive System*, in *Human Histology*. 2005, Elsevier Mosby: New York. p. 345-372.

96. Ross, M.H. and W. Pawlina, *Chapter 23: Female Reproductive System*, in *Histology: A Text and Atlas*. 2006, Lipincott, Williams & Wikins: Philadelphia. p. 773-833.

97. Gartner, L.P. and J.L. Hiatt, *Chapter 20: Female Reproductive System*, in *Colour Textbeek of Histology*. 2007, Saunders Elsiver: Philadelphia. p. 463-488.

98. Arends, M.J., C.H. Buckley, and M. Wells, *Aetiology, pathogenesis, and pathology of cervical neoplasia.* J Clin Pathol, 1998. **51**(2): p. 96-103.

99. Martini, F.H. and E.F. Bartholomew, *Chapter 19: The Reproductive System*, in *Essentials of Anatomy and Physiology*. 2007, Pearson Education Inc.: San Francisco. p. 612-643.

100. Lacey, C.J., C.M. Lowndes, and K.V. Shah, *Chapter 4: Burden and management of non-cancerous HPV-related conditions: HPV-6/11 disease.* Vaccine, 2006. **24 Suppl 3**: p. S35-41.

101. Wright, T.C., *Pathology of HPV infection at the cytologic and histologic levels: Basis for a 2-tiered morphologic classification system.* Int J Gynaecol Obstet, 2006. **94**(Supp 1): p. S22-S31.

102. Montag, A. and V. Kumar, *Chapter 19: The Femal Genital System and Breast*, in *Robbins Basic Pathology*, V. Kumar, et al., Editors. 2007, Saunders Elsevier: Philadelphia. p. 711-750.

103. Syrjanen, K. and S. Syrjanen, *Chapter 6: Natural History of Cervical HPV Infections and CIN*, in *Papillomavirus Infections in Human Pathology*. 2000, John Wiley & Sons: New York. p. 143-166.

104. Syrjanen, K. and S. Syrjanen, *Chapter 5: Epidemiology of Genital HPV Infections, CIN and Cervical Cancer*, in *Papillomavirus Infections in Human Pathology*. 2000, John Wiley & Sons: New York. p. 117-142.

105. Ngan, H.Y. and C.L. Trimble, *Clinical manifestations of HPV infection Section A: Benign manifestations of HPV infection: A4. Preinvasive lesions of the cervix.* int J Gynaecol Obstet, 2006. **94**(Supp 1): p. S44-S49.

106. Livasy, C.A., D.T. Moore, and L. Van Le, *The clinical significance of a negative loop electrosurgical cone biopsy for high-grade dysplasia.* Obstet Gynecol, 2004. **104**(2): p. 250-4.

107. Flowers, L.C. and M.A. McCall, *Diagnosis and management of cervical intraepithelial neoplasia.* Obstet Gynecol Clin North Am, 2001. **28**(4): p. 667-684.

108. Prendiville, W., *Large loop excision of the transformation zone.* Clin Obstet Gynecol, 1995. **38**(3): p. 622-39.

109. Prendiville, W., J. Cullimore, and S. Norman, *Large loop excision of the transformation zone (LLETZ). A new method of management for women with cervical intraepithelial neoplasia.* Br J Obstet Gynaecol, 1989. **96**(9): p. 1054-60.

110. McIndoe, G.A., et al., *Laser excision rather than vaporization: the treatment of choice for cervical intraepithelial neoplasia.* Obstet Gynecol, 1989. **74**(2): p. 165-8.

111. Villa, L.L. and L. Denny, *Methods for detection of HPV infection and its clinical utility.* int J Gynaecol Obstet, 2006. **94**(Supp 1): p. S71-S80.

112. Syrjanen, K. and S. Syrjanen, *Chapter 4: Diagnostic Techniques in HPV Detection*, in *Papillomavirus Infections In Human Pathology*. 2000, John Wiley & Sons: New York. p. 89-115.

113. Molijn, A., et al., *Molecular diagnosis of human papillomavirus (HPV) infections.* J Clin Virol, 2005. **32 Suppl 1**: p. S43-51.

114. Burd, E.M., *Human papillomavirus and cervical cancer.* Clin Microbiol Rev, 2003. **16**(1): p. 1-17.

115. Dehn, D., K.C. Torkko, and K.R. Shroyer, *Human papillomavirus testing and molecular markers of cervical dysplasia and carcinoma.* Cancer, 2007. **111**(1): p. 1-14.

116. Iftner, T. and L.L. Villa, *Chapter 12: Human papillomavirus technologies.* J Natl Cancer Inst Monogr, 2003(31): p. 80-8.

117. McPherson, M.M., S, *Chapter 1: An Introduction To PCR*, in *PCR*, E. Owen, Editor. 2006, Taylor & Francis. p. 1-8.

118. Shibata, D.K., N. Arnheim, and W.J. Martin, *Detection of human papilloma virus in paraffin-embedded tissue using the polymerase chain reaction.* J Exp Med, 1988. **167**(1): p. 225-30.

119. Chan, P.K., et al., *Biases in human papillomavirus genotype prevalence assessment associated with commonly used consensus primers.* Int J Cancer, 2006. **118**(1): p. 243-5.

120. Remmerbach, T.W., et al., *PCR detection of human papillomavirus of the mucosa: comparison between MY09/11 and GP5+/6+ primer sets.* J Clin Virol, 2004. **30**(4): p. 302-8.

121. Husnjak, K., et al., *Comparison of five different polymerase chain reaction methods for detection of human papillomavirus in cervical cell specimens.* J Virol Methods, 2000. **88**(2): p. 125-34.

122. Moselhi, M., et al., *Follow-up after negative large loop excision of the transformation zone (LLETZ) of the cervix.* J Obstet Gynaecol, 2002. **22**(2): p. 193-6.

123. Howells, R.E., et al., *How can the incidence of negative specimens resulting from large loop excision of the cervical transformation zone (LLETZ) be reduced? An analysis of negative LLETZ specimens and development of a predictive model.* Bjog, 2000. **107**(9): p. 1075-82.

124. Murta, E.F., A.O. Silva, and E.A. Silva, *Clinical significance of a negative loop electrosurgical excision procedure, conization and hysterectomy for cervical intraepithelial neoplasia.* Eur J Gynaecol Oncol, 2006. **27**(1): p. 50-2.

125. Waggoner, S.E., *Cervical cancer.* Lancet, 2003. **361**(9376): p. 2217-25.

126. Arora, R., et al., *Prevalence of high-risk human papillomavirus (HR-HPV) types 16 and 18 in healthy women with cytologically negative Pap smear.* Eur J Obstet Gynecol Reprod Biol, 2005. **121**(1): p. 104-9.

127. Gjooen, K., et al., *Prevalence of human papillomavirus in cervical scrapes, as analyzed by PCR, in a population-based sample of women with and without cervical dysplasia.* Apmis, 1996. **104**(1): p. 68-74.

128. Burchell, A.N., et al., *Chapter 6: Epidemiology and transmission dynamics of genital HPV infection.* Vaccine, 2006. **24 Suppl 3**: p. S52-61.

129. Clifford, G.M., et al., *Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis.* Lancet, 2005. **366**(9490): p. 991-8.

130. Gok, M., et al., *HPV16 and increased risk of recurrence after treatment for CIN.* Gynecol Oncol, 2007. **104**(2): p. 273-5.

131. Mann, C.H., et al., *LLETZ--evidence of its efficacy against HPV infection.* Gynecol Oncol, 2001. **81**(1): p. 125-7.

132. van Ham, M.A., et al., *High-risk HPV presence in cervical specimens after a large loop excision of the cervical transformation zone: significance of newly detected hr-HPV genotypes.* J Med Virol, 2007. **79**(3): p. 314-9.

133. Sarian, L.O., et al., *Factors associated with HPV persistence after treatment for high-grade cervical intra-epithelial neoplasia with large loop excision of the transformation zone (LLETZ).* J Clin Virol, 2004. **31**(4): p. 270-4.

134. Wang, J., S.H. Xiao, and J.L. Manley, *Genetic analysis of the SR protein ASF/SF2: interchangeability of RS domains and negative control of splicing.* Genes Dev, 1998. **12**(14): p. 2222-33.

135. Caceres, J.F. and A.R. Kornblihtt, *Alternative splicing: multiple control mechanisms and involvement in human disease.* Trends Genet, 2002. **18**(4): p. 186-93.

136. McPhillips, M.G., et al., *SF2/ASF binds the human papillomavirus type 16 late RNA control element and is regulated during differentiation of virus-infected epithelial cells.* J Virol, 2004. **78**(19): p. 10598-605.

137. Dreyfuss, G., V.N. Kim, and N. Kataoka, *Messenger-RNA-binding proteins and the messages they carry.* Nat Rev Mol Cell Biol, 2002. **3**(3): p. 195-205.

138. Caporali, R., et al., *Autoantibodies to heterogeneous nuclear ribonucleoproteins.* Autoimmunity, 2005. **38**(1): p. 25-32.

139. Rooke, N., et al., *Roles for SR proteins and hnRNP A1 in the regulation of c-src exon N1.* Mol Cell Biol, 2003. **23**(6): p. 1874-84.

140. Zhao, X., M. Rush, and S. Schwartz, *Identification of an hnRNP A1-dependent splicing silencer in the human papillomavirus type 16 L1 coding region that prevents premature expression of the late L1 gene.* J Virol, 2004. **78**(20): p. 10888-905.

141. Gontarek, R.R., et al., *hnRNP C and polypyrimidine tract-binding protein specifically interact with the pyrimidine-rich region within the 3'NTR of the HCV RNA genome.* Nucleic Acids Res, 1999. **27**(6): p. 1457-63.

142. Zhao, X., et al., *A 57-nucleotide upstream early polyadenylation element in human papillomavirus type 16 interacts with hFip1, CstF-64, hnRNP C1/C2, and polypyrimidine tract binding protein.* J Virol, 2005. **79**(7): p. 4270-88.

143. Honore, B., U. Baandrup, and H. Vorum, *Heterogeneous nuclear ribonucleoproteins F and H/H' show differential expression in normal and selected cancer tissues.* Exp Cell Res, 2004. **294**(1): p. 199-209.

144. Oberg, D., et al., *A Downstream polyadenylation element in human papillomavirus type 16 encoudes multiple GGG-motifs and interacts with hnRNP H.* J Virol, 2005.

145. Spellman, R. and C.W. Smith, *Novel modes of splicing repression by PTB.* Trends Biochem Sci, 2006. **31**(2): p. 73-6.

146. Wagner, E.J. and M.A. Garcia-Blanco, *Polypyrimidine tract binding protein antagonizes exon definition.* Mol Cell Biol, 2001. **21**(10): p. 3281-8.

147. Forch, P., et al., *U2 small nuclear ribonucleoprotein particle (snRNP) auxiliary factor of 65 kDa, U2AF65, can promote U1 snRNP recruitment to 5' splice sites.* Biochem J, 2003. **372**(Pt 1): p. 235-40.

148. Koffa, M.D., et al., *The human papillomavirus type 16 negative regulatory RNA element interacts with three proteins that act at different posttranscriptional levels.* Proc Natl Acad Sci U S A, 2000. **97**(9): p. 4677-82.

149. Gallouzi, I.E., et al., *HuR binding to cytoplasmic mRNA is perturbed by heat shock.* Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3073-8.

150. Sokolowski, M., H. Furneaux, and S. Schwartz, *The inhibitory activity of the AU-rich RNA element in the human papillomavirus type 1 late 3' untranslated region*

correlates with its affinity for the elav-like HuR protein. J Virol, 1999. **73**(2): p. 1080-91.

151. Sokolowski, M., et al., *AU-rich mRNA instability elements on human papillomavirus type 1 late mRNAs and c-fos mRNAs interact with the same cellular factors.* Oncogene, 1997. **15**(19): p. 2303-19.

152. Cumming, S.A., et al., *The human papillomavirus type 31 late 3' untranslated region contains a complex bipartite negative regulatory element.* J Virol, 2002. **76**(12): p. 5993-6003.

153. Koch, P.J. and D.R. Roop, *The role of keratins in epidermal development and homeostasis--going beyond the obvious.* J Invest Dermatol, 2004. **123**(5): p. x-xi.

154. Vaidya, M.M. and D. Kanojia, *Keratins: markers of cell differentiation or regulators of cell differentiation?* J Biosci, 2007. **32**(4): p. 629-34.

155. Gusterson, B.A., et al., *Basal cytokeratins and their relationship to the cellular origin and functional classification of breast cancer.* Breast Cancer Res, 2005. **7**(4): p. 143-8.

156. Lopes, C.C., C.P. Dietrich, and H.B. Nader, *Specific structural features of syndecans and heparan sulfate chains are needed for cell signaling.* Braz J Med Biol Res, 2006. **39**(2): p. 157-67.

157. Bartlett, A.H., K. Hayashida, and P.W. Park, *Molecular and cellular mechanisms of syndecans in tissue injury and inflammation.* Mol Cells, 2007. **24**(2): p. 153-66.

158. Beauvais, D.M. and A.C. Rapraeger, *Syndecans in tumor cell adhesion and signaling.* Reprod Biol Endocrinol, 2004. **2**: p. 3.

159. Tkachenko, E., J.M. Rhodes, and M. Simons, *Syndecans: new kids on the signaling block.* Circ Res, 2005. **96**(5): p. 488-500.

160. Maga, G. and U. Hubscher, *Proliferating cell nuclear antigen (PCNA): a dancer with many partners.* J Cell Sci, 2003. **116**(Pt 15): p. 3051-60.

161. Martin, I.V. and S.A. MacNeill, *ATP-dependent DNA ligases.* Genome Biol, 2002. **3**(4): p. REVIEWS3005.

162. Bellacosa, A., *Functional interactions and signaling properties of mammalian DNA mismatch repair proteins.* Cell Death Differ, 2001. **8**(11): p. 1076-92.

163. Al-Mohanna, M.A., et al., *The p16INK4a tumor suppressor controls p21WAF1 induction in response to ultraviolet light.* Nucleic Acids Res, 2007. **35**(1): p. 223-33.

164. Ohtani, N., et al., *The p16INK4a-RB pathway: molecular link between cellular senescence and tumor suppression.* J Med Invest, 2004. **51**(3-4): p. 146-53.

165. Satyanarayana, A. and K.L. Rudolph, *p16 and ARF: activation of teenage proteins in old age.* J Clin Invest, 2004. **114**(9): p. 1237-40.

166. O'Neill, C.J. and W.G. McCluggage, *p16 expression in the female genital tract and its value in diagnosis.* Adv Anat Pathol, 2006. **13**(1): p. 8-15.

167. Griesser, H., et al., *Correlation of immunochemical detection of HPV L1 capsid protein in pap smears with regression of high-risk HPV positive mild/moderate dysplasia.* Anal Quant Cytol Histol, 2004. **26**(5): p. 241-5.

168. Melsheimer, P., et al., *Immunocytochemical detection of HPV high-risk type L1 capsid proteins in LSIL and HSIL as compared with detection of HPV L1 DNA.* Acta Cytol, 2003. **47**(2): p. 124-8.

169. Egawa, K., et al., *Synthesis of viral DNA and late capsid protein L1 in parabasal spinous cell layers of naturally occurring benign warts infected with human papillomavirus type 1.* Virology, 2000. **268**(2): p. 281-93.

170. Iwasaki, T., et al., *Detection of capsid antigen of human papillomavirus (HPV) in benign lesions of female genital tract using anti-HPV monoclonal antibody.* J Pathol, 1992. **168**(3): p. 293-300.

171. Schwartz, S., Zhao, X, Oberg, D, Rush, M., *Regulation of papillomavirus late gene expression.* Recent Res. Devel. Virol., 2004. **6**: p. 29-45.

172. Sanford, J.R., J. Ellis, and J.F. Caceres, *Multiple roles of arginine/serine-rich splicing factors in RNA processing.* Biochem Soc Trans, 2005. **33**(Pt 3): p. 443-6.

173. Cartegni, L., S.L. Chew, and A.R. Krainer, *Listening to silence and understanding nonsense: exonic mutations that affect splicing.* Nat Rev Genet, 2002. **3**(4): p. 285-98.

174. Sokolowski, M. and S. Schwartz, *Heterogeneous nuclear ribonucleoprotein C binds exclusively to the functionally important UUUUU-motifs in the human papillomavirus type-1 AU-rich inhibitory element.* Virus Res, 2001. **73**(2): p. 163-75.

175. Dietrich-Goetz, W., et al., *A cellular 65-kDa protein recognizes the negative regulatory element of human papillomavirus late mRNA.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 163-8.

176. Collier, B., et al., *Translational inhibition in vitro of human papillomavirus type 16 L2 mRNA mediated through interaction with heterogenous ribonucleoprotein K and poly(rC)-binding proteins 1 and 2.* J Biol Chem, 1998. **273**(35): p. 22648-56.

177. Oberg, D., et al., *A downstream polyadenylation element in human papillomavirus type 16 L2 encodes multiple GGG motifs and interacts with hnRNP H.* J Virol, 2005. **79**(14): p. 9254-69.

178. Zheng, Z.M., P.J. He, and C.C. Baker, *Structural, functional, and protein binding analyses of bovine papillomavirus type 1 exonic splicing enhancers.* J Virol, 1997. **71**(12): p. 9096-107.

179. Zheng, Z.M., M. Huynen, and C.C. Baker, *A pyrimidine-rich exonic splicing suppressor binds multiple RNA splicing factors and inhibits spliceosome assembly.* Proc Natl Acad Sci U S A, 1998. **95**(24): p. 14088-93.

180. Beesley, J.E., *Chapter 1: Introduction*, in *Immunocytochemistry and In Situ Hybridisation in the Biomedical Sciences*, J.E. Beesley, Editor. 2001, Birkhauser. p. 1-5.

181. Taylor, C.R., et al., *Chapter 1: Techniques of Immunohistochemistry: Principles, Pitfalls and Standardization*, in *Diagnostic Immunohistochemistry*, D. Dabbs, Editor. 2006, Churchill Livingstone Elsevier. p. 1-42.

182. Polack, J.M. and S. Van Noorden, *Chapter 1: Introduction*, in *Introduction to Immunocytochemistry (RMS: Microscopy Handbook Series No. 37)*. 1997, BIOS Scientific Publishers: New York. p. 1-4.

183. Polack, J.M. and S. Van Noorden, *Chapter 3: Requirements for Immunocytochemistry*, in *Introduction to Immunocytochemistry (RMS: Microscopy Handbook Series No. 37)*. 1997, BIOS Scientifis Publishers: New York. p. 11-35.

184. Jackson, P. and D. Blythe, *Chapter 4: Antigen Retrieval*, in *Immunocytochemistry and In Situ Hybridisation in the Biomedical Sciences*, J.E. Beesley, Editor. 2001, Birkhauser. p. 55-72.

185. Polack, J.M. and S. Van Noorden, *Chapter 4: Methods*, in *Introduction to Immunocytochemistry (RMS: Microscopy Handbook Series No. 37)*. 1997, BIOS Scientific Publishers: New York. p. 41-54.

186. Nakahara, T., et al., *Human papillomavirus type 16 E1circumflexE4 contributes to multiple facets of the papillomavirus life cycle.* J Virol, 2005. **79**(20): p. 13150-65.

187. Branca, M., et al., *Up-regulation of proliferating cell nuclear antigen (PCNA) is closely associated with high-risk human papillomavirus (HPV) and progression of cervical intraepithelial neoplasia (CIN), but does not predict disease outcome in cervical cancer.* Eur J Obstet Gynecol Reprod Biol, 2007. **130**(2): p. 223-31.

188. Goel, M.M., et al., *MIB-1 and PCNA immunostaining as a diagnostic adjunct to cervical Pap smear.* Diagn Cytopathol, 2005. **33**(1): p. 15-9.

189. Wang, J.L., et al., *Predictive significance of the alterations of p16INK4A, p14ARF, p53, and proliferating cell nuclear antigen expression in the progression of cervical cancer.* Clin Cancer Res, 2004. **10**(7): p. 2407-14.

190. Murphy, N., et al., *p16INK4A, CDC6, and MCM5: predictive biomarkers in cervical preinvasive neoplasia and cervical cancer.* J Clin Pathol, 2005. **58**(5): p. 525-34.

191. Pino, I., et al., *Altered patterns of expression of members of the heterogeneous nuclear ribonucleoprotein (hnRNP) family in lung cancer.* Lung Cancer, 2003. **41**(2): p. 131-43.

192. Ushigome, M., et al., *Up-regulation of hnRNP A1 gene in sporadic human colorectal cancers.* Int J Oncol, 2005. **26**(3): p. 635-40.

193. Tominaga, M., et al., *Detection and discrimination of preneoplastic and early stages of lung adenocarcinoma using hnRNP B1 combined with the cell cycle-related markers p16, cyclin D1, and Ki-67.* Lung Cancer, 2003. **40**(1): p. 45-53.

194. Wu, S., et al., *hnRNP B1 protein may be a possible prognostic factor in squamous cell carcinoma of the lung.* Lung Cancer, 2003. **41**(2): p. 179-86.

195. Carpenter, B., et al., *Heterogeneous nuclear ribonucleoprotein K is over expressed, aberrantly localised and is associated with poor prognosis in colorectal cancer.* Br J Cancer, 2006. **95**(7): p. 921-7.

196. Rauch, J., et al., *Allogenic antibody-mediated identification of head and neck cancer antigens.* Biochem Biophys Res Commun, 2004. **323**(1): p. 156-62.

197. Jin, W., et al., *Fibroblast growth factor receptor-1 alpha-exon exclusion and polypyrimidine tract-binding protein in glioblastoma multiforme tumors.* Cancer Res, 2000. **60**(5): p. 1221-4.

198. Pillai, M.R., et al., *Expression of folate receptors and heterogeneous nuclear ribonucleoprotein E1 in women with human papillomavirus mediated transformation of cervical tissue to cancer.* J Clin Pathol, 2003. **56**(8): p. 569-74.

199. Monsonego, J., et al., *P16(INK4a) immunocytochemistry in liquid-based cytology samples in equivocal Pap smears: added value in management of women with equivocal Pap smear.* Acta Cytol, 2007. **51**(5): p. 755-66.

200. Iaconis, L., et al., *p16 and Ki-67 immunostaining in atypical immature squamous metaplasia of the uterine cervix: correlation with human papillomavirus detection.* Arch Pathol Lab Med, 2007. **131**(9): p. 1343-9.

201. Feng, W., et al., *Senescence and apoptosis in carcinogenesis of cervical squamous carcinoma.* Mod Pathol, 2007. **20**(9): p. 961-6.

202. Carozzi, F.M., *Combined analysis of HPV DNA and p16INK4a expression to predict prognosis in ASCUS and LSIL pap smears.* Coll Antropol, 2007. **31 Suppl 2**: p. 103-6.

203. Hariri, J. and A. Oster, *The negative predictive value of p16INK4a to assess the outcome of cervical intraepithelial neoplasia 1 in the uterine cervix.* Int J Gynecol Pathol, 2007. **26**(3): p. 223-8.

204. Terhune, S.S., et al., *Early polyadenylation signals of human papillomavirus type 31 negatively regulate capsid gene expression.* J Virol, 2001. **75**(17): p. 8147-57.

205. Buratti, E., M. Baralle, and F.E. Baralle, *Defective splicing, disease and therapy: searching for master checkpoints in exon definition.* Nucleic Acids Res, 2006. **34**(12): p. 3494-510.

206. Ladd, A.N. and T.A. Cooper, *Finding signals that regulate alternative splicing in the post-genomic era.* Genome Biol, 2002. **3**(11): p. reviews0008.

207. Baralle, D. and M. Baralle, *Splicing in action: assessing disease causing sequence changes.* J Med Genet, 2005. **42**(10): p. 737-48.

208. Fairbrother, W.G., et al., *Predictive identification of exonic splicing enhancers in human genes.* Science, 2002. **297**(5583): p. 1007-13.

209. Lavigueur, A., et al., *A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding.* Genes Dev, 1993. **7**(12A): p. 2405-17.

210. Humphrey, M.B., et al., *A 32-nucleotide exon-splicing enhancer regulates usage of competing 5' splice sites in a differential internal exon.* Mol Cell Biol, 1995. **15**(8): p. 3979-88.

211. Kan, J.L. and M.R. Green, *Pre-mRNA splicing of IgM exons M1 and M2 is directed by a juxtaposed splicing enhancer and inhibitor.* Genes Dev, 1999. **13**(4): p. 462-71.

212. Gersappe, A. and D.J. Pintel, *CA- and purine-rich elements form a novel bipartite exon enhancer which governs inclusion of the minute virus of mice NS2-specific exon in both singly and doubly spliced mRNAs.* Mol Cell Biol, 1999. **19**(1): p. 364-75.

213. Zheng, Z.M., E.S. Reid, and C.C. Baker, *Utilization of the bovine papillomavirus type 1 late-stage-specific nucleotide 3605 3' splice site is modulated by a novel exonic bipartite regulator but not by an intronic purine-rich element.* J Virol, 2000. **74**(22): p. 10612-22.

214. Fu, X.D., *Towards a splicing code.* Cell, 2004. **119**(6): p. 736-8.

215. Zhu, J., A. Mayeda, and A.R. Krainer, *Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins.* Mol Cell, 2001. **8**(6): p. 1351-61.

216. Caputi, M., et al., *hnRNP A/B proteins are required for inhibition of HIV-1 pre-mRNA splicing.* Embo J, 1999. **18**(14): p. 4060-7.

217. Jacquenet, S., et al., *A second exon splicing silencer within human immunodeficiency virus type 1 tat exon 2 represses splicing of Tat mRNA and binds protein hnRNP H.* J Biol Chem, 2001. **276**(44): p. 40464-75.

218. Forch, P., et al., *The apoptosis-promoting factor TIA-1 is a regulator of alternative pre-mRNA splicing.* Mol Cell, 2000. **6**(5): p. 1089-98.

219. Kanopka, A., O. Muhlemann, and G. Akusjarvi, *Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA.* Nature, 1996. **381**(6582): p. 535-8.

220. Kanopka, A., et al., *Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins.* Nature, 1998. **393**(6681): p. 185-7.

221. Graveley, B.R., *Sorting out the complexity of SR protein functions.* Rna, 2000. **6**(9): p. 1197-211.

222. Chabot, B., et al., *An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1.* Mol Cell Biol, 1997. **17**(4): p. 1776-86.

223. Moreira, A., et al., *The upstream sequence element of the C2 complement poly(A) signal activates mRNA 3' end formation by two distinct mechanisms.* Genes Dev, 1998. **12**(16): p. 2522-34.

224. Pinol-Roma, S., et al., *A novel heterogeneous nuclear RNP protein with a unique distribution on nascent transcripts.* J Cell Biol, 1989. **109**(6 Pt 1): p. 2575-87.

225. Burd, C.G. and G. Dreyfuss, *RNA binding specificity of hnRNP A1: significance of hnRNP A1 high-affinity binding sites in pre-mRNA splicing.* Embo J, 1994. **13**(5): p. 1197-204.

226. Caputi, M. and A.M. Zahler, *Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family.* J Biol Chem, 2001. **276**(47): p. 43850-9.

227. Chan, R.C. and D.L. Black, *The polypyrimidine tract binding protein binds upstream of neural cell-specific c-src exon N1 to repress the splicing of the intron downstream.* Mol Cell Biol, 1997. **17**(8): p. 4667-76.

228. Simard, M.J. and B. Chabot, *SRp30c is a repressor of 3' splice site utilization.* Mol Cell Biol, 2002. **22**(12): p. 4001-10.

229. Lai, M.C., B.H. Teh, and W.Y. Tarn, *A human papillomavirus E2 transcriptional activator. The interactions with cellular splicing factors and potential function in pre-mRNA processing.* J Biol Chem, 1999. **274**(17): p. 11832-41.

230. Kleinberger, T. and T. Shenk, *Adenovirus E4orf4 protein binds to protein phosphatase 2A, and the complex down regulates E1A-enhanced junB transcription.* J Virol, 1993. **67**(12): p. 7556-60.

231. Estmer Nilsson, C., et al., *The adenovirus E4-ORF4 splicing enhancer protein interacts with a subset of phosphorylated SR proteins.* Embo J, 2001. **20**(4): p. 864-71.

232. Raffetseder, U., et al., *Splicing factor SRp30c interaction with Y-box protein-1 confers nuclear YB-1 shuttling and alternative splice site selection.* J Biol Chem, 2003. **278**(20): p. 18241-8.

233. Xu, Q., D.Y. Leung, and K.O. Kisich, *Serine-arginine-rich protein p30 directs alternative splicing of glucocorticoid receptor pre-mRNA to glucocorticoid receptor beta in neutrophils.* J Biol Chem, 2003. **278**(29): p. 27112-8.

234. Bellanger, S., et al., *High-risk but not low-risk HPV E2 proteins bind to the APC activators Cdh1 and Cdc20 and cause genomic instability.* Cell Cycle, 2005. **4**(11): p. 1608-15.

235. Strachan, T. and A.P. Read, *Chapter 5: Nucleic Acid Hybribization Asseys*, in *Human Molecular Genetics 2.* 1999, BIOS Scientific Publishers New York. p. 95-118.

236. Lodish, H., et al., *Chapter 11: RNA Processing, Nuclear Transport, and Post-Transcriptional Control*, in *Molecular Cell Biology.* 2000, W. H. Freeman and Company: New York.

237. McPherson, M., and Moller, S., *Chapter 9: Real-time RT-PCR*, in *PCR*, E. Owen, Editor. 2006, Taylor & Francis. p. 209-232.

238. Ponchel, F., *Chapter 8: Real-time PCR Using SYBR Green*, in *Real-time PCR*, M.T. Dorak, Editor. 2006, Taylor & Francis. p. 139-154.

239. Sauer, G., et al., *Progression of cervical carcinomas is associated with down-regulation of CD9 but strong local re-expression at sites of transendothelial invasion.* Clin Cancer Res, 2003. **9**(17): p. 6426-31.

240. Dauksaite, V. and G. Akusjarvi, *The second RNA-binding domain of the human splicing factor ASF/SF2 is the critical domain controlling adenovirus E1A alternative 5'-splice site selection.* Biochem J, 2004. **381**(Pt 2): p. 343-50.

241. Saslow, D., et al., *American Cancer Society Guideline for human papillomavirus (HPV) vaccine use to prevent cervical cancer and its precursors.* CA Cancer J Clin, 2007. **57**(1): p. 7-28.

242. Cutts, F.T., et al., *Human papillomavirus and HPV vaccines: a review.* Bull World Health Organ, 2007. **85**(9): p. 719-26.

243. Schiller, J.T. and D.R. Lowy, *Prospects for cervical cancer prevention by human papillomavirus vaccination.* Cancer Res, 2006. **66**(21): p. 10229-32.

244. Ault, K.A., *Epidemiology and natural history of human papillomavirus infections in the female genital tract.* Infect Dis Obstet Gynecol, 2006. **2006 Suppl**: p. 40470.

245. Parkin, D.M. and F. Bray, *Chapter 2: The burden of HPV-related cancers.* Vaccine, 2006. **24 Suppl 3**: p. S11-25.

246. Agoff, S.N., et al., *p16(INK4a) expression correlates with degree of cervical neoplasia: a comparison with Ki-67 expression and detection of high-risk HPV types.* Mod Pathol, 2003. **16**(7): p. 665-73.

247. Klaes, R., et al., *Overexpression of p16(INK4A) as a specific marker for dysplastic and neoplastic epithelial cells of the cervix uteri.* Int J Cancer, 2001. **92**(2): p. 276-84.

248. Guimaraes, M.C., et al., *Immunohistochemical expression of p16INK4a and bcl-2 according to HPV type and to the progression of cervical squamous intraepithelial lesions.* J Histochem Cytochem, 2005. **53**(4): p. 509-16.

249. Huang, C.S., et al., *Increased expression of SRp40 affecting CD44 splicing is associated with the clinical outcome of lymph node metastasis in human breast cancer.* Clin Chim Acta, 2007. **384**(1-2): p. 69-74.

250. Kalof, A.N., et al., *p16INK4A immunoexpression and HPV in situ hybridization signal patterns: potential markers of high-grade cervical intraepithelial neoplasia.* Am J Surg Pathol, 2005. **29**(5): p. 674-9.

251. Hu, L., et al., *Human papillomavirus genotyping and p16INK4a expression in cervical intraepithelial neoplasia of adolescents.* Mod Pathol, 2005. **18**(2): p. 267-73.

252. Eleuterio, J., Jr., et al., *Prognostic markers of high-grade squamous intraepithelial lesions: the role of p16INK4a and high-risk human papillomavirus.* Acta Obstet Gynecol Scand, 2007. **86**(1): p. 94-8.

253. Galban, S., et al., *RNA-binding proteins HuR and PTB promote the translation of hypoxia-inducible factor 1alpha.* Mol Cell Biol, 2008. **28**(1): p. 93-107.

254. Branton, P.E. and D.E. Roopchand, *The role of adenovirus E4orf4 protein in viral replication and cell killing.* Oncogene, 2001. **20**(54): p. 7855-65.

255. Maddika, S., et al., *Cancer-selective therapy of the future: apoptin and its mechanism of action.* Cancer Biol Ther, 2006. **5**(1): p. 10-9.

A. Appendix I – Solutions

Mayer's haematoxylin

2g haematoxylin (Merck)

100g aluminium ammonium sulphate (BDH)

0.4g sodium iodate (EMD Chemicals)

make up to 2 litres with distilled water and leave overnight at room temperature

2g citric acid (BDH)

mix and boil for 5 minutes

allow to cool and filter before use

1% eosin

1g eosin powder (Merck) dissolved in 100ml distilled water

Spirit (96%)

960ml absolute alcohol (Merck)

made up to 1 litre with distilled water

70% Alcohol

700ml absolute alcohol (Merck)

made up to 1 litre with distilled water

1% and 2% agarose gel

1g (1% gel) or 2g (2% gel) agarose powder (Pronadisa)

100ml of TBE (Invitrogen)

boil in microwave until dissolved

cool to 60°C

add 0.5μg/ml ethidium bromide (Fluka)

mix gently and pour into gel template

allow to set

6x DNA/RNA loading dye (50ml)

0.125g bromophenol blue (BDH)

0.125g xylene cyanol (BDH)

15g glycerol (BDH)

0.01mol/l citrate buffer (pH 6)

2.1g citric acid (BDH)

add approximately 800ml distilled water

pH to pH 6 using 2mol/l NaOH (79.98g NaOH (BDH) in 1 litre)

make solution up to 1 litre in a volumetric flask

0.01mol/l EDTA buffer (pH 8)

3.7g EDTA (BDH)

add approximately 800ml distilled water

pH to pH 8 using 2mol/l NaOH (79.98g NaOH (BDH) in 1 litre)

make solution up to 1 litre in a volumetric flask

Phosphate Buffered Saline (PBS)

1 PBS tablet (Oxoid) for every 100ml distilled water, use a magnetic stirrer to mix until tablets have dissolved

0.1% protease

0.1g protease type 24 (Sigma) in 100ml PBS

3% hydrogen peroxide in methanol

1ml 30% hydrogen peroxide (BDH)

9ml methanol (Fluka)

0.06% 2,4-diaminobenzidine (DAB)

16ml PBS injected into 1 container of DAB (Sigma)

1% bovine serum albumin (BSA)

0.2g BSA (Sigma) dissolved in 20ml PBS

0.2% Tween 20

0.2ml Tween 20 (Sigma)

make up to 100ml with PBS

70% Ethanol

700ml ethanol (Merck)

make up to 1 litre with distilled water in a volumetric flask

3mol/l sodium acetate

40.82g sodium acetate (BDH)

make up to 100ml with distilled water in a volumetric flask

1% agarose formaldehyde gel

1g agarose powder (Pronadisa)

80ml of TBE (Invitrogen)

boil in microwave until dissolved and allow to cool slightly

10ml 10x MOPS (Fluka)

18ml 37% formaldehyde (BDH)

pour into template and allow to set

10x MOPS (3-(N-morpholino)propanesulfonic acid) buffer

8.2g sodium acetate (BDH)

83.72g MOPS (Fluka)

20ml 0.5mol/l EDTA (BDH) (146g in 1 litre)

make up to 800ml with distilled water in a volumetric flask

1x MOPS made by a 1/10 dilution of 10x MOPS

10% Sodium dodecyl sulphate (SDS)

10g SDS (BDH)

make up to 100ml with distilled water in a volumetric flask

20x SCC

175.3g NaCl (BDH)

88.2g sodium citrate (BDH)

make up to 1 litre with distilled water in a volumetric flask

6x SCC made by a 1/3.3 dilution of 20x SCC

2x SCC 0.1% SDS made by a 1/10 dilution of 20x SCC and a 1/10 dilution 10% SDS

0.2x SCC 0.1% SDS made by a 1/100 dilution of 20x SCC and a 1/10 dilution 10% SDS

0.1x SCC 0.1% SDS made by a 1/200 dilution of 20x SCC and a 1/10 dilution 10% SDS

0.5mol/l EDTA

146.13g EDTA (BDH)

make up to 1 litre with distilled water in a volumetric flask

Pre-hybridisation solution / Hybridisation solution

3ml 20x SCC (appendix I)

0.6ml Denhart's Solution (appendix I)

1.2g dextransulphate (Fluka)

6ml fromamide (BHD)

0.2ml 10% sodium pyrophosphate (1g sodium pyrophosphate (Fluka) made up to 10ml with distilled water)

1.2ml 10% SDS

100x Denhart's Solution

1g ficoll (Sigma)

1g polyvinylpyrrolidone (Fluka)

1g BSA (Sigma)

make up to 50ml with distilled water in a volumetric flask

5% acrylamide gel

16ml distilled water

2ml 5x TBE (Invitrogen)

2.5ml 40% PAGE (BioRad)

200µl 10% APS (0.1g APS (Sigma) in 10ml distilled water)

20µl TEMED (BDH)

mix gently and pour quickly into gel template

2x assay buffer (50ml)

10ml 1mol/l sodium phosphate buffer pH 7.2 (appendix I)

1ml 1mol/l magnesium chloride (appendix I)

350µl β-mercaptoethanol (Fluka)

66.5mg *ortho*-Nitrophenyl-β-galactoside (ONPG) (Sigma)

make up to 50ml with distilled water in a volumetric flask

1mol/l sodium phosphate buffer pH 7.2 (100ml)

68.4ml 1mol/l disodium hydrogen phosphate (Na_2HPO_4) (141.96g Na_2HPO_4 (BDH) in 1 liter distilled water)

31.6ml 1mol/l monosodium phosphate (NaH_2PO_4) (119.97g NaH_2PO_4 (BDH) in 1 liter distilled water)

1mol/l magnesium chloride

95.2g magnesium chloride (BDH)

make up to 1 litre with distilled water in a volumetric flask

1mol/l sodium carbonate

10.5g sodium carbonate (BDH)

make up to 100ml with distilled water in a volumetric flask

2. Appendix II – β-globin PCR

β-globin PCR for the 50 patient cohort in Chapter 2. PC03/PC04 primers (Table 2.1) used producing a 110bp product.

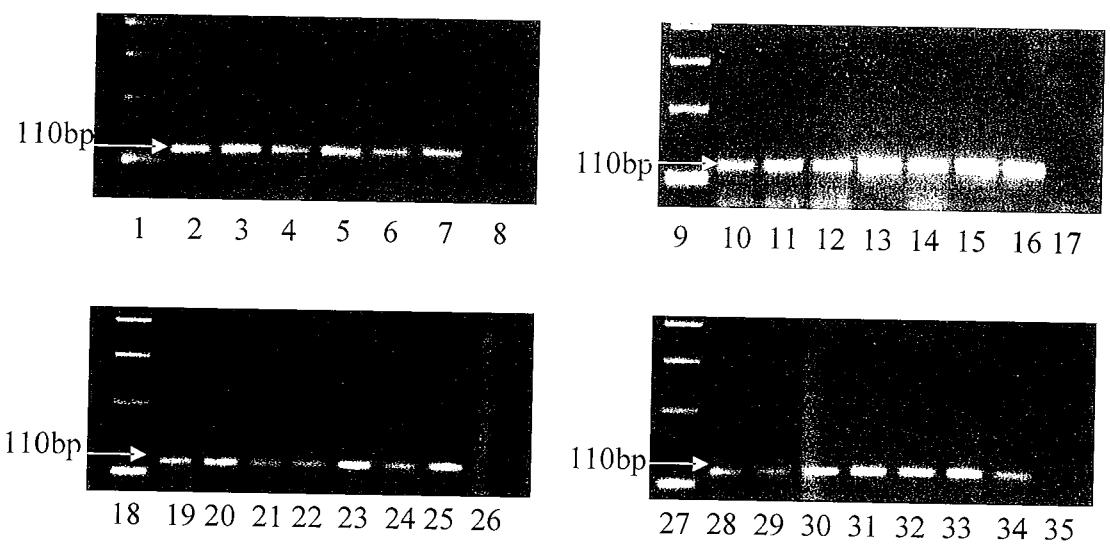


Figure 2.1 PCR amplification of the β-globin gene using the PC03/PC04 primers

Lane 1 100bp DNA ladder, lane 2 positive control, lane 3 – 7 patient 1 – 5, lane 8 negative control, lane 9 100bp DNA ladder, lane 10 positive control, lane 11 – 16 patient 6 – 11, lane 17 negative control, lane 18 100bp DNA ladder, lane 19 positive control, lane 20 – 25 patent 12 – 17, lane 26 negative control, lane 27 100bp DNA ladder, lane 28 – 34 patient 18 – 24, lane 35 negative control. PCR products run on a 1% agarose gel.

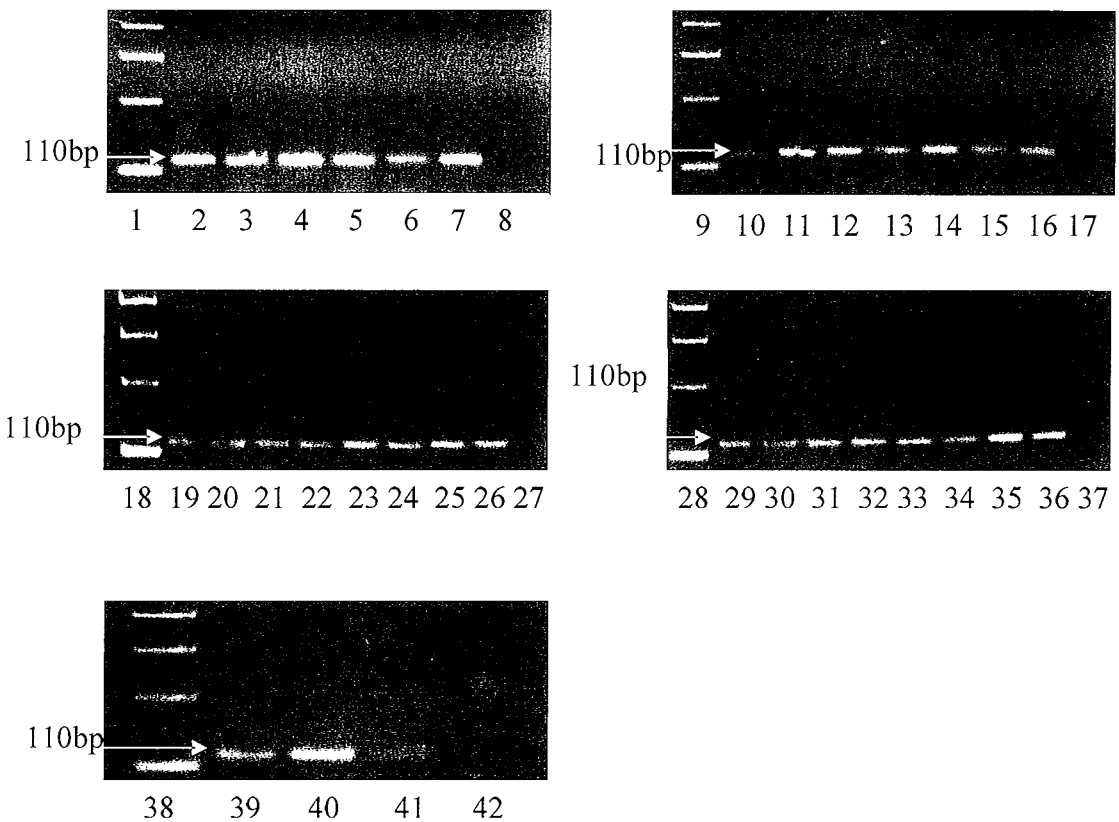


Figure 2.2 PCR amplification of the β-globin gene using the PC03/PC04 primers

Lane 1 100bp DNA ladder, lane 2 positive control, lane 3 – 7 patient 25 – 29, lane 8 negative control, lane 9 100bp DNA ladder, lane 10 positive control, lane 11 - 16 patient 30 – 35, lane 17 negative control, lane 18 100bp DNA ladder, lane 19 positive control, lane 20 – 26 patient 36 – 42, lane 27 negative control, lane 28 100bp DNA ladder, lane 29 positive control, lane 30 – 35 patient 43 - 48, lane 36 positive control, lane 37 negative control, lane 38 100bp DNA ladder, lane 39 positive control, lane 40 – 41 patient 49 – 50 lane 42 negative control. PCR products run on a 1% agarose gel.

3. Appendix III – HPV PCR

HPV PCR for the 50 patient cohort in Chapter 2. MY09/MY11 and GP5+/GP6+ primers (Table 2.1) used in a nested PCR producing a 150bp product

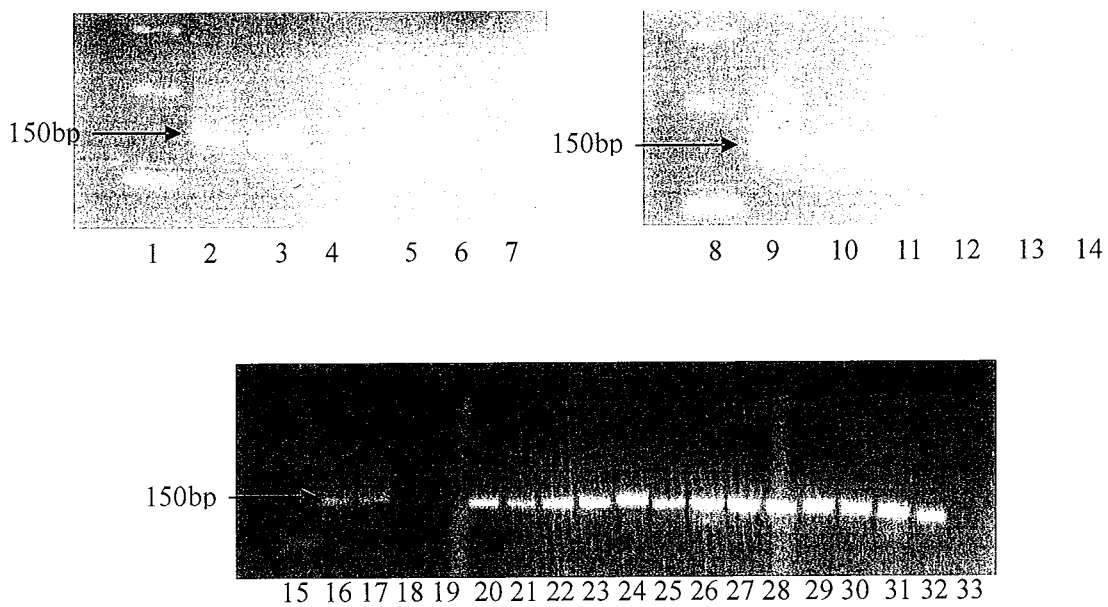


Figure 3.1 Nested PCR amplification of the L1 HPV gene using the MY09/MY11 primers in combination with the GP5+/GP6+ primers

Lane 1 100bp DNA ladder, lane 2 positive control, lane 3 – 6 patient 1 – 4, lane 7 negative control, lane 8 100bp DNA ladder, lane 9 positive control, lane 10 – 13 patient 5 – 8, lane 14 negative control, lane 15 100bp DNA ladder, lane 16 positive control, lane 17 patient 9, lane 18 and 19 lane skipped, lane 20 – 32 patient 10 – 22, lane 33 negative control. PCR products run on a 1% agarose gel.

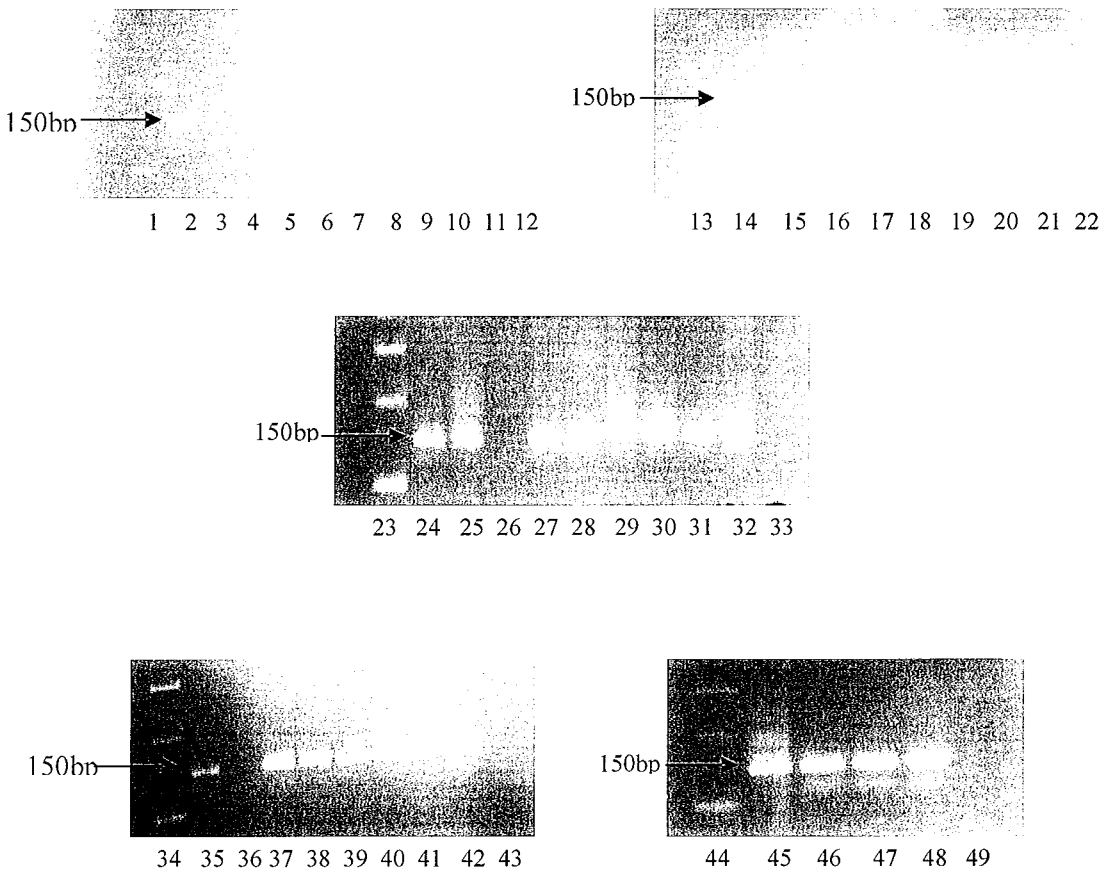


Figure 3.2 Nested PCR amplification of the L1 HPV gene using the MY09/MY11 primers in combination with the GP5+/GP6+ primers

Lane 1 100bp DNA ladder, lane 2 positive control, lane 3 negative control, lane 4 – 10 patient 23 – 29, lane 11 positive control, lane 12 negative control, lane 13 100bp DNA ladder, lane 14 positive control, lane 15 – 20 patient 30 – 35, lane 21 positive control, lane 22 negative control, lane 23 100bp DNA ladder, lane 24 positive control, lane 25 patient 36, lane 26 lane skipped, lane 27 – 31 patient 36 – 41, lane 32 positive control, lane 33 negative control, lane 34 100bp DNA ladder, lane 35 positive control, lane 36 lane skipped, lane 37 – 42 patient 42 - 47, lane 43 negative control, lane 44 100bp DNA ladder, lane 45 positive control, lane 46 – 48 patient 48 – 50, lane 49 negative control. PCR products run on a 1% agarose gel.

4. Appendix IV – HPV-16 PCR

HPV-16 PCR for the 50 patient cohort in Chapter 2. HPV-16O and HPV-16I primers (Table 2.1) used in a nested PCR producing a 119bp product.

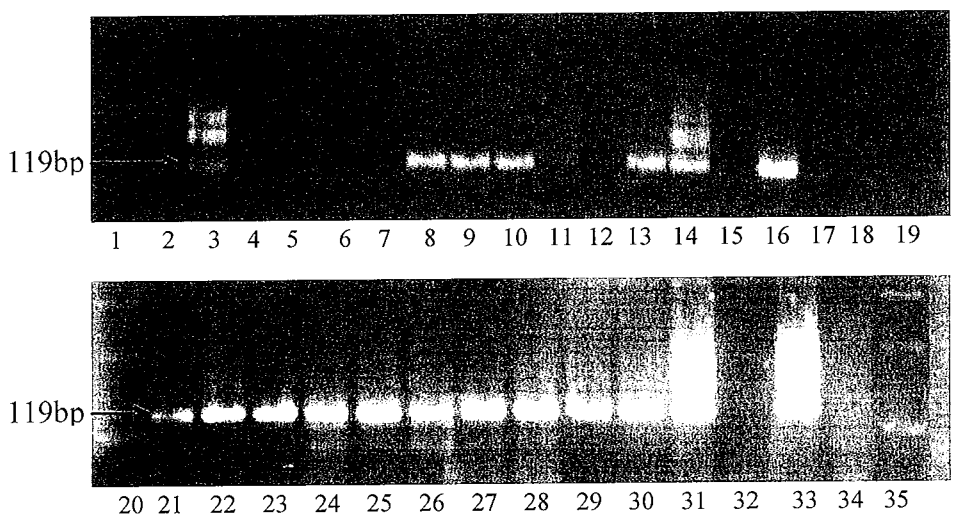


Figure 4.1 Nested HPV-16 PCR using HPV-16O and HPV-16 I primer pairs

Lane 1 DNA ladder, lane 2 round 1 positive control, lane 3 positive control for both rounds, lane 4 – 8 patient 1 - 5, lane 9 – 13 patient 11 - 15, lane 14 – 16 patient 21 – 23, lane 17 round 1 negative control, lane 18 negative control for both rounds, lane 19 100bp DNA ladder, lane 20 100bp DNA ladder, lane 21 round 2 positive control, lane 22 – 23 patient 24 - 25, lane 24 – 28 patient 31 – 35, lane 29 – 33 patient 41 – 45, lane 34 negative control for both rounds, lane 35 100bp DNA ladder. PCR products run on a 1% agarose gel.

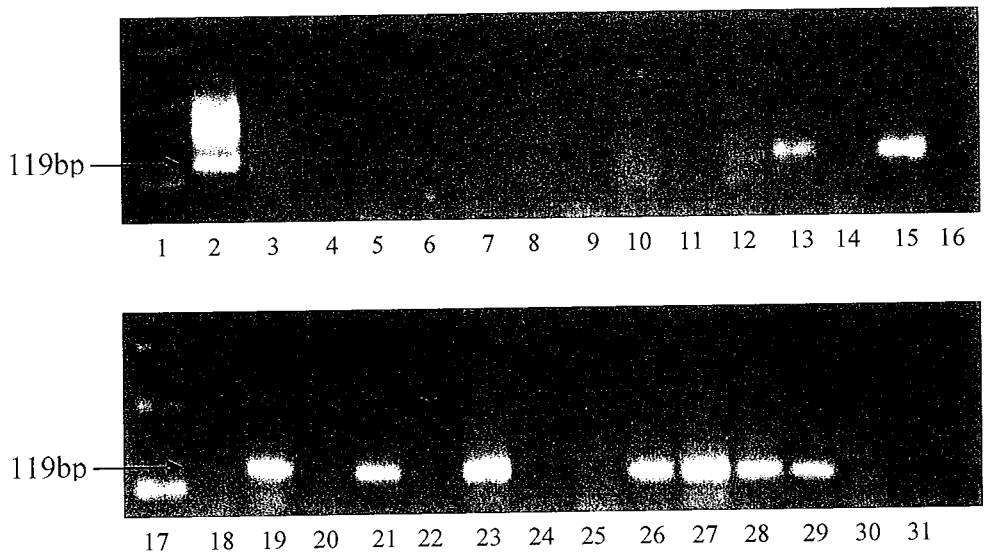


Figure 4.2 Nested HPV-16 PCR using HPV-16O and HPV-16 I primer pairs

Lane 1 100bp DNA ladder, lane 2 positive control for both rounds, lane 3 round 1 positive control, lane 4 – 8 patient 6 – 10, lane 9 – 13 patient 16 – 20, lane 14 - 15 patient 26 - 27, lane 16 negative control for both rounds, lane 17 100bp DNA ladder, lane 18 – 20 patient 28 – 30, lane 21 – 25 patient 36 – 40, lane 26 – 30 patient 46 – 50, lane 31 negative control for both rounds. PCR products run on a 1% agarose gel.

Publications

Oberg D, Fay J, Lambkin H, Schwartz S.

A downstream polyadenylation element in human papillomavirus type 16 L2 encodes multiple GGG motifs and interacts with hnRNP H.

J Virol. 2005 Jul;79(14):9254-69

Zhao X, Fay J, Lambkin H, Schwartz S.

Identification of a 17-nucleotide splicing enhancer in HPV-16 L1 that counteracts the effect of multiple hnRNP A1-binding splicing silencers.

Virology. 2007 Dec 20;369(2):351-63