
Doctoral

Science

2012

Analysis of HPV-16 Late Gene Expression

Beatrice Orru

Technological University Dublin

Follow this and additional works at: <https://arrow.tudublin.ie/sciendoc>



Part of the [Biology Commons](#)

Recommended Citation

Orru, B. (2012). *Analysis of HPV-16 Late Gene Expression*. Doctoral Thesis. Technological University Dublin. doi:10.21427/D7JG68

This Theses, Ph.D is brought to you for free and open access by the Science at ARROW@TU Dublin. It has been accepted for inclusion in Doctoral by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie, vera.kilshaw@tudublin.ie.



Analysis of HPV-16 late gene expression

A thesis submitted for the degree of Doctor of Philosophy

Beatrice Orrù

Supervisors

Fergus Ryan

Stefan Schwartz

School of Biological Sciences,
Dublin Institute of Technology,
Kevin Street,
Dublin 8

2012

Abstract

Human papillomaviruses (HPVs) are present in 99.7% of all cervical cancers and HPV type 16 (HPV-16) is the major cause of cervical cancer. Expression of the viral capsid gene L1 and L2 can be detected only in the terminally epithelial cells and we speculate that inhibition of HPV-16 late gene expression in the early stage of the life cycle is probably a prerequisite for persistence of infection. The products of the late genes, L1 and L2, are highly immunogenic and expression of these proteins in the lower layers of the cervical epithelium could lead to clearance of the virus. Therefore, it is of interest to understand how HPV late gene expression is regulated. The goal of this thesis was to examine the regulation of late genes in HPV-16. To this end we wished to generate reporter plasmids based on the HPV-16 genome with the L1 gene replaced by an easily measurable reporter gene, such as chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), secreted alkaline phosphatase (SEAP) or luciferase, and to establish reporter stable cell lines useful for large scale screening of small molecules or cellular factors that influence RNA processing events during late gene expression. CAT and GFP proved to be functional surrogate markers of late gene expression and their expression was dependent on the levels of known inducers of HPV-16 late gene expression such as adenovirus E4orf4 protein (E4orf4), polypyrimidine tract binding protein (PTB), arginine/serine-rich SRp30c protein (SRp30c) or alternative splicing factor/splicing factor 2 (ASF/SF2). Functional stable cell lines with CAT reporter plasmids, separately integrated into the HeLa cellular genome, were also generated allowing the identification of a number of small molecules capable of modulating CAT expression. Phorbol 12-myristate 13-acetate (TPA), valproic acid and tannic acid were identified as inducers of HPV-16 late gene expression. Further experiments identified the TPA-

inducible, hnRNP A2/B1 protein as a novel regulator of HPV-16 late gene expression. Immunohistochemical analysis of this protein in cervical epithelium at the different stages of the development of cervical cancer demonstrated that hnRNP A2/B1 is highly expressed in normal cervical epithelium and low-grade squamous intraepithelial lesion (LSIL) and decreased in high-grade squamous intraepithelial lesion (HSIL) and squamous cell carcinoma (SCC). In conclusion, the HPV-16 reporter plasmids and reporter cell lines described herein are functional and can be used for the investigations of HPV-16 late gene expression.

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 another award in any other third level institution.

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

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

Signature _____ Date _____

To my parents

Acknowledgements

I would like to extend my sincere gratitude to all the people who have made the completion of this thesis possible

I am deeply indebted to my supervisors, Dr. Fergus Ryan and Prof. Stefan Schwartz for their constant support and invaluable guidance throughout the project. I could not have imagined having better supervisors for my PhD study.

I would like to thank all staff in the School of Biological Sciences in particular Brid Ann Ryan, Louis Armstrong, Mary Hunt, Jan Guerin for continuous support and Derek Neylan, Joseph Vaughan, Helen Lambkin, Patrick McHale, Brenda Brankin, Fabian McGrath, Greg Byrne for being always available for lab meetings and assistance. Thanks to all the technicians, Stephanie, Tricia, Ted, Kathleen and especially Daireen for all their help and advice.

I would like to thank all the post-grad lab staff, in particular Rali, Sangeetha, Ciaran, Paul, Dave and Swarna for making the lab a nice environment where to work. Thanks also to the group in Sweden, Monika, Cecilia and Xiaoze for providing plasmids and Annelie the Erasmus student.

To all my friends especially Ilaria, Sara, Claudia, Silvia, Tatiana, Elisa, Roberta, Valentina, Daniela, Giusi, Riccardo, Federico, thank you for always being there. Special thanks to Giuseppe who helped me get through the difficult times and enriched my life in Dublin all these years.

Last but not least, I wish to thank my beloved family, Antonio, Fatma and Gianfranco for supporting and encouraging me through the duration of my studies. This thesis would certainly not have existed without them!

Abbreviations

Abbreviation	Full term
Ad E4orf4	adenoviral E4orf4
ASF/SF2	alternative splicing factor/splicing factor 2
BP	branch point
BPS	branch point sequence
CBC	cap binding complex
CDK	cyclin-dependent kinase
CFI	cleavage factors I
CFII	cleavage factors II
CIN	cervical intraepithelial neoplasia
CKI	cyclin-dependent kinase inhibitor
CMV	cytomegalovirus immediate early
CPSF	cleavage and polyadenylation specificity factor
CstF	cleavage stimulation factor
DNA	deoxyribonucleic acid
E6AP	E6-associate protein
ECM	extracellular matrix
EGFR	epidermal growth factor receptor
ESS	exonic splicing silencers
GA	golgi apparatus
GAGs	glycosaminoglycans
GR	glucocorticoid receptor

GMP	guanosine mono-phosphate
hnRNPs	heterogeneous nuclear ribonucleoproteins
hnRNP A2/B1	heterogeneous nuclear ribonucleoprotein A2/B1
HPV	human papillomavirus
hTERT	human telomerase reverse transcriptase
IRES	internal ribosomal entry site
ISS	intronic splicing silencer
kDa	kilodalton
KH	K homology domain
LCR	long control region
mRNA	messenger ribonucleic acid
MOI	multiplicity of infectivity
NLS	nuclear localization signals
nt	nucleotide
ORF	open reading frame
pAE	early polyadenylation signal
pAL	late polyadenylation signal
PAB II	poly(A) binding II
PAP	poly(A) polymerase
Pap	papanicolaou
PDZ domain	PSD-95/discs large/zonaoccludens 1
PP2A	protein phosphatase 2
pRb	retinoblastoma protein
PPT	polypyrimidine tract

PTB	polypyrimidin tract binding protein
PV	papillomavirus
RISC	RNA-induced Silencing Complex
RNAi	RNA interference
RNA pol II	RNA polymerase II
RRM	RNA recognition domain
SA	splice acceptor
SD	splice donor
siRNA	short interfering RNA
snRNA	small nuclear RNA
shRNA	short-hairpin RNA
snRNPs	small nuclear nucleoproteins
SR protein	serine-arginine rich protein
SRp30c	serine-arginine rich protein 30c
U2AF	U2 snRNP auxiliary factor
VLP	virus-like particle
VPA	valproic acid
TPA	12-O-tetradecanoylphorbol-13-acetate
TMA	tissue microarrays
TNF	tumor necrosis factor
YB-1	Y -box protein-1
3'ss	3' splice site
5'ss	5' splice site

Table of Contents

Abstract.....	i
Acknowledgment.....	v
Abbreviation.....	vii
1. Chapter 1 - Introduction.....	11
1.1 Viruses.....	11
1.2 Papillomavirus.....	12
1.3 HPVs and Cervical cancer.....	15
1.4 The cervix.....	19
1.5 HPV life cycle.....	21
1.6 The HPV genome.....	23
1.6.1 E1 and E2 proteins.....	25
1.6.2 E4 and E5 proteins.....	26
1.6.3 E6 and E7 proteins.....	27
1.6.4 L1 and L2 proteins.....	30
1.7 DNA integration.....	30
1.8 Regulation of gene expression.....	31
1.8.1 Capping.....	32
1.8.2 Splicing.....	33
1.8.3 Polyadenylation.....	36
1.8.4 Alternative splicing.....	38
1.9 Regulation of HPV gene expression.....	40
1.10 HPV-16 late gene expression.....	43

1.11 SR proteins	44
(i) ASF/SF2	46
(ii) SRp30c	47
(iii) SC35	48
1.12 hnRNP family.....	49
(i) hnRNP A2/B1	51
(ii) PTB.....	53
1.13 E4orf4.....	54
1.14 Transient and stable transfection.....	55
1.15 Small molecules used in this study	57
(i) TPA.....	57
(ii) Emetine.....	58
(iii) Cycloheximide	59
(iv) Ceramide.....	60
(v) Amiloride.....	61
(vi) Scriptaid.....	61
(vii) Valproic acid	62
(viii) Tannic acid	63
1.16 RNAi pathway.....	65
1.17 Immunohistochemistry.....	69
1.17.1 Avidin-Biotin Conjugate (ABC) Method.....	70
1.17.2 Tissue Microarray	71
1.17.3 p16INK4a.....	72
1.18 Aims	74

2. Chapter 2 - Material and Methods	75
2.1 Generation of subgenomic HPV-16 reporter plasmids	75
2.1.1 Plasmid construction	75
2.1.2 TOPO TA Cloning	83
2.1.3 Transformation into TOP10 Chemically competent <i>E. coli</i>	83
2.1.4 Purification of PCR product and vector reactions.....	83
2.1.5 Ligation Reaction	84
2.1.6. Analysis of recombinant plasmids	85
2.1.7 Restriction of mini preparation plasmid DNA	85
2.1.8 Isolation of high quality plasmid.....	86
2.2 Analysis of the reporter genes.....	88
2.2.1 Cell culture and transfections	88
2.2.2 CAT ELISA assay	88
2.2.3 SEAP assay	89
2.2.4 GFP assay	89
2.2.5 β -galactosidase assay	89
2.3 Establishment of stable cell lines	90
2.3.1 Stable transfection and clones selection.....	90
2.3.2 RNA extraction and reverse transcription (RT)-PCR	91
(i) RNA extraction.....	91
(ii) RT-PCR.....	92
2.4 Treatment with small molecules drug	94
2.4.1 Stable transfection with small molecules drugs	94
2.4.2 Transient transfection with small molecules drugs	94
2.5 Protein analysis	94

2.5.1 Protein extraction	94
2.5.2 Western blot	95
2.6 Apoptosis-program.....	96
2.6.1 Serum deprivation.....	96
2.6.2 Heat shock.....	97
2.7 Infection of cells with Lentiviruses expressing SiRNA-GFP constructs	97
2.7.1 Transduction.....	97
2.7.2 Detection of GFP fluorescence	98
2.8 Tissue microarray preparation and analysis	98
2.8.1 Samples collection and tissue microarray preparation.....	98
2.8.2 Immunohistochemistry (Avidin-Biotin Complex Immuno-peroxidase method)99	
3. Chapter 3 - Results - Generation and validation of HPV-16 late gene reporter constructs	
.....	101
3.1 Generation of reporter plasmids.....	101
3.2 Determination of CAT expression in pBELCAT and pBELMCAT.....	117
3.3 Adenovirus E4orf4, polypyrimidine tract binding protein, ASF/SF2 and SRp30c induce CAT expression from pBELCAT and pBELMCAT reporter plasmids.....	121
3.4 Determination of GFP levels in pBELMGFP	123
3.5 Determination of SEAP levels in pBELSEAP and pBELMSEAP	128
3.6 Identification and characterization of cis-acting regulatory RNA elements in HPV-16 CAT and SEAP reporter plasmids	132
4. Chapter 4 - Results - Generation and characterization of stable cell lines.....	137
4.1 Establishment of two novel HPV-16 reporter cell lines that can serve as tools to investigate regulation of HPV-16 gene expression	137
4.2 Analysis of overexpression of E4orf4, PTB and SRp30c in pBELCAT and pBELMCAT stable cell lines	140
4.3 RT-PCR analysis.....	148

4.4 Treatment of CAT reporter cell lines with small molecules	150
(i) TPA.....	150
(ii) Emetine, ceramide and cycloheximide.....	151
(iii) Valproic acid.....	156
(iv) Scriptaid and amiloride.....	158
(v) Tannic acid	159
4.5 Treatment of pBELCAT or pBELMCAT plasmids transiently transfected in HeLa cells with TPA, valproic acid or tannic acid.....	161
4.6 Treatment of pBELCAT-67 cell line with TPA, valproic acid or tannic acid combined	162
4.7 Apoptosis-program.....	164
4.8 Establishment of pBELMGFP reporter cell lines	165
4.9 Treatment of pBELMGFP with TPA	168
4.10 Gene Knockdown.....	170
5. Chapter 5 - Results – Identification and localization of a RNA-binding protein involved in the regulation of splicing in normal and neoplastic cervical epithelium	177
5.1 Identification of a novel splicing factor involved in HPV-16 late gene regulation	177
5.2 Detection of hnRNP A2/B1 protein on cervical tissue arrays.....	184
5.3 Correlation of RNA-binding protein expression in cervical epithelium to p16INK4a, a known marker of differentiation	189
6. Chapter 6 - Discussion	193
7. References	208
Appendix I	239
Publications	251

Table of Figures

Figure 1.1 Human papillomavirus family tree	13
Figure 1.2 Cervical cancer incidence worldwide	15
Figure 1.3 The major steps in cervical carcinogenesis.....	16
Figure 1.4 Representation of cervix epithelium pre- and post-infection.....	18
Figure 1.5 The uterus	20
Figure 1.6 Structure of the ectocervix.....	20
Figure 1.7 Representation of the different layers in the epithelium and HPV gene expression in infected cells.....	22
Figure 1.8 Genomic organisation of HPV-16	24
Figure 1.9 p53 and Rb pathway	29
Figure 1.10 RNA processing.....	32
Figure 1.11 Splicing	35
Figure 1.12 The structure of the pre-mRNA and RNA processing.....	37
Figure 1.13 Schematic representation of alternative splicing	39
Figure 1.14 HPV-16 genome and transcription map.....	42
Figure 1.15 Human SR proteins family	45
Figure 1.16 Schematic representation of the hnRNP family structure.....	50
Figure 1.17 Structure of hnRNPs A2/B1 and transcript in mouse	52
Figure 1.18 RNAi mechanisms	67
Figure 1.19 shRNA Lentiviral particle delivery.....	68
Figure 1.20 IHC Methods.....	70
Figure 1.21 Representation of tissue microarrays.....	71
Figure 1.22 p16INK4a pathway	73

Figure 2.1 pBEL and pBELM plasmids.....	76
Figure 2.2 Representation of the reporter plasmids generated in this study	77
Figure 2.3 Representation of the reporter plasmids carrying only 2 splice sites SD3632 and SA5639	78
Figure 3.1 Schematic representation of the subgenomic HPV-16 expression plasmids, pBEL and pBELM, and the major potential mRNAs that can be produced.....	102
Figure 3.2 Schematic representation of cloning steps for insertion of IRES sequence and reporter gene into pBEL plasmid	103
Figure 3.3 Generation of reporter plasmids	104
Figure 3.4 Restriction enzymes in pBEL plasmid.....	105
Figure 3.5 PCR amplification of IRES sequence with IRES set primers	106
Figure 3.6 Analysis of recombinant plasmids TOPO + IRES	107
Figure 3.7 Analysis of recombinant plasmids pBEL + IRES	109
Figure 3.8 Analysis of recombinant plasmids pBELM + IRES.....	109
Figure 3.9 Schematic representation of the reporter plasmids.....	110
Figure 3.10 PCR amplification of secreted alkaline phosphatase (SEAP), β -galactosidase (LacZ) and chloramphenicol acetyltransferase (CAT) reporter genes.....	111
Figure 3.11 Analysis of recombinant plasmids TOPO + CAT	112
Figure 3.12 Analysis of recombinant plasmids pBEL-IRES + CAT	113
Figure 3.13 Analysis of recombinant plasmids pBELM-IRES + CAT.....	114
Figure 3.14 Analysis of recombinant plasmids TOPO + SEAP	114
Figure 3.15 Analysis of recombinant plasmids pBEL-IRES + SEAP and pBELM-IRES + SEAP	115
Figure 3.16 Analysis of recombinant plasmids pBELM-IRES + GFP	116
Figure 3.17 Analysis of recombinant plasmids TOPO + LacZ.....	117
Figure 3.18 Northern blot analysis on pBEL and pBELM	118

Figure 3.19 Schematic representation of the subgenomic HPV-16 expression plasmids, pBELCAT and pBELMCAT, and the majors potential mRNAs that can be produced	119
Figure 3.20 Determination of CAT levels in transient transfection	120
Figure 3.21 Cotransfection of CAT reporter plasmids with E4orf4, PTB, SRp30c or ASF/SF2 expression vector.....	122
Figure 3.22 Titration of E4orf4, PTB, SRp30c or ASF/SF2 expression vector in pBELMCAT reporter plasmid.....	122
Figure 3.23 Schematic representation of subgenomic expression plasmid pBELMGFP ..	123
Figure 3.24 Examination of GFP expression after 24 hours posttransfection	124
Figure 3.25 Determination of GFP levels in transient transfection.....	125
Figure 3.26 Determination of GFP levels in pBELMGFP transient transfection	126
Figure 3.27 Examination of GFP expression after 48 hours posttransfection	127
Figure 3.28 Schematic representation of subgenomic expression plasmids pBELSEAP and pBELMSEAP	128
Figure 3.29 Determination of SEAP levels in SEAP reporter plasmids transient transfection	129
Figure 3.30 Optimization of SEAP assay.....	130
Figure 3.31 SEAP levels of SEAP reporter plasmids transiently transfected in HeLa cells	131
Figure 3.32 Schematic representation of subgenomic expression plasmids pBspliceCAT, pBspliceMCAT and pMt1sdCAT	133
Figure 3.33 Determination of CAT levels in CAT reporter plasmids.....	134
Figure 3.34 Schematic representation of subgenomic expression plasmid pMt1sdSEAP .	135
Figure 3.35 Determination of SEAP levels in SEAP reporter plasmid.....	136
Figure 4.1 PCR amplification of DNA extracted from pBELCAT and pBELMCAT stable cell lines using CAT or IRES set primers	138
Figure 4.2 Determination of CAT levels in stable transfection	139
Figure 4.3 Induction of CAT expression in pBELCAT and pBELMCAT derived-cell lines	141
Figure 4.4 Titration of PTB expression plasmid in pBELCAT stable cell lines.....	142

Figure 4.5 Titration of E4orf4 expression plasmid in pBELCAT stable cell lines.....	143
Figure 4.6 Titration of inducers of late gene expression in pBELMCAT stable cell lines	145
Figure 4.7 Induction of CAT expression in pBELCAT-67 and pBELMCAT-31 stable cell lines	147
Figure 4.8 Characterization of pBELCAT-67 and pBELMCAT-31 stable cell lines.....	149
Figure 4.9 Treatment of pBELCAT- and pBELMCAT-derived cell lines with TPA.....	150
Figure 4.10 Treatment of CAT reporter cell lines with emetine, ceramide or cycloheximide	152
Figure 4.11 Serial dilution of emetine and ceramide	153
Figure 4.12 CAT induction in stable cell lines treated with TPA, emetine, cycloheximide or ceramide and Western blot analysis.....	155
Figure 4.13 Serial dilution of valproic acid.....	156
Figure 4.14 Treatment of CAT stable cell lines with valproic acid	157
Figure 4.15 Western blot analysis on stable cell lines treated with valproic acid.....	157
Figure 4.16 Treatment of CAT stable cell lines with scriptaid or amiloride	158
Figure 4.17 Treatment of CAT stable cell lines with tannic acid	160
Figure 4.18 Western blot analysis on stable cell lines treated with tannic acid	160
Figure 4.19 Effect of TPA, valproic acid or tannic acid in cells transiently transfected with pBELCAT or pBELMCAT plasmids.....	162
Figure 4.20 Treatment of pBELCAT-67 stable cell line with TPA, valproic acid and tannic acid combined.....	163
Figure 4.21 Apoptosis program.....	164
Figure 4.22 PCR amplification of DNA extracted from pBELMGFP-derived cell lines using GFP or IRES set primers	166
Figure 4.23 Fluorescence of pBELMGFP stable cell lines	167
Figure 4.24 Treatment of pBELMGFP with TPA.....	169
Figure 4.25 Detection of infected cells at MOI 2 with confocal microscope	172

Figure 4.26 Detection of infected cells at MOI 5 with confocal microscope	173
Figure 4.27 Detection of infected cells at MOI 10 with confocal microscope	174
Figure 4.28 CAT expression in pBELMCAT-31 stable cell line transduced with lentiviruses carrying shRNA ASF/SF2, PTB or scramble	175
Figure 4.29 Western blot analysis on pBELMCAT-31 stable cell line uninfected/infected with lentiviruses	176
Figure 5.1 Co-transfection of CAT reporter plasmids with SC35 or hnRNP A2/B1 expression vector.....	178
Figure 5.2 Western blot analysis on pBELCAT-67 and pBELMCAT-31 stable cell lines	179
Figure 5.3 Induction of CAT by hnRNP A2/B1	180
Figure 5.4 Induction of CAT in pBspliceMCAT and pMt1sdCAT by hnRNP A2/B1	182
Figure 5.5 Induction of CAT in stable pBELCAT-67 and pBELMCAT-31 cell lines by hnRNP A2/B1	183
Figure 5.6 hnRNP A2/B1 protein expression in NC	185
Figure 5.7 hnRNP A2/B1 protein expression in LSIL	186
Figure 5.8 hnRNP A2/B1 protein expression in HSIL.....	187
Figure 5.9 hnRNP A2/B1 protein expression in SCC.....	188
Figure 5.10 p16INK4a protein expression	190
Figure 5.11 Comparison of immunostaining with hnRNP A2/B1 or p16INK4a antibodies on the same tissue microarray slide.....	192

1. Chapter 1 - Introduction

1.1 Viruses

A virus is a microscopic particle consisting of a core of nucleic acid, DNA or RNA, surrounded by a protein shell called a capsid; some viruses have an envelope of lipids that surrounds the protein coat when they are outside a cell. Viruses are typically not considered living organisms because they are unable to replicate without a host cell, but inside of a living cell are capable of taking over the cells metabolic machinery for the production of new virus particles called virions. Viruses are the most abundant biological entities that exist in nature and generally are classified by the organisms they infect (animals, plants, or bacteria) (Suttle, 2007). They are responsible for many diseases with different symptoms.

The first viral agent to be discovered was the tobacco mosaic virus in 1892 by Dmitry Iosifovich Ivanovsky while he was investigating a disease affecting tobacco (Levine, 2001). He demonstrated that this virus is capable of permeating filters that restrain bacteria. These particles were only classified definitively as viruses by Martinus Willem Beijerinck in 1898. He demonstrated that the tobacco mosaic was caused by an infective agent much smaller than bacteria (Levine, 2001). Most viruses are about one hundred times smaller than bacteria.

1.2 Papillomavirus

Papillomaviruses are small, non-enveloped DNA viruses that belong to the Papillomaviridae family (Bernard, 2005). A papillomavirus (PV) was first described as the responsible agent for papillomas in the cotton-tail rabbit by Richard Scope in 1933 (Lowy & Howley, 2001). Despite the lack of an appropriate cell-culture system, papillomaviruses continued to be studied intensively.

They are ubiquitous and can infect many mammals, including humans, and are absolutely species-specific. Papillomaviruses are epitheliotropic viruses with a predilection for infection of either cutaneous or internal squamous mucosal surfaces. Approximately 130 papillomavirus types have been completely described so far (Bernard, 2005) and among all the papillomaviruses the most studied is the Human Papillomavirus (HPV). The 130 HPV types are subdivided in 5 groups distinguished by a letter of the Greek alphabet: Alpha, Beta, Gamma, Mu and Nu (Fig. 1.1) (Bernard, 2005). The largest group of HPVs, containing approximately 60 members, comprises the Alpha papillomavirus, which infect the genital/mucosal area. HPV types from the Beta, Gamma, Mu and Nu groups primarily infect cutaneous surfaces (Doorbar, 2006). The Alpha papillomaviruses can be classified as high-risk, low-risk and cutaneous types depending on the frequency they cause cervical cancer (Table 1.1). The cutaneous viruses, such as HPV-2, cause common warts and very rarely cause cancer (Tate, 2004). The low-risk types can cause genital HPV infections, which are believed to be the most common sexually transmitted viral infection in many countries. Low-risk HPVs, such as HPV-11, are very seldom associated with cancer (de Villiers *et al.*, 2004). The high-risk type can lead to cancers of the cervix, vulva, vagina, anus, oesophagus (Collins *et al.*, 2005) and larynx (Tsutsumi *et al.*, 1993) in women or less common cancers of the anus and penis in men. HPV infections are widely spread and in fact

more than the 70% of women are infected by the virus during their lifetime. HPV-16 is the most common high-risk type and is present in approximately 50% of all cervical cancers (Munoz *et al.*, 2003).

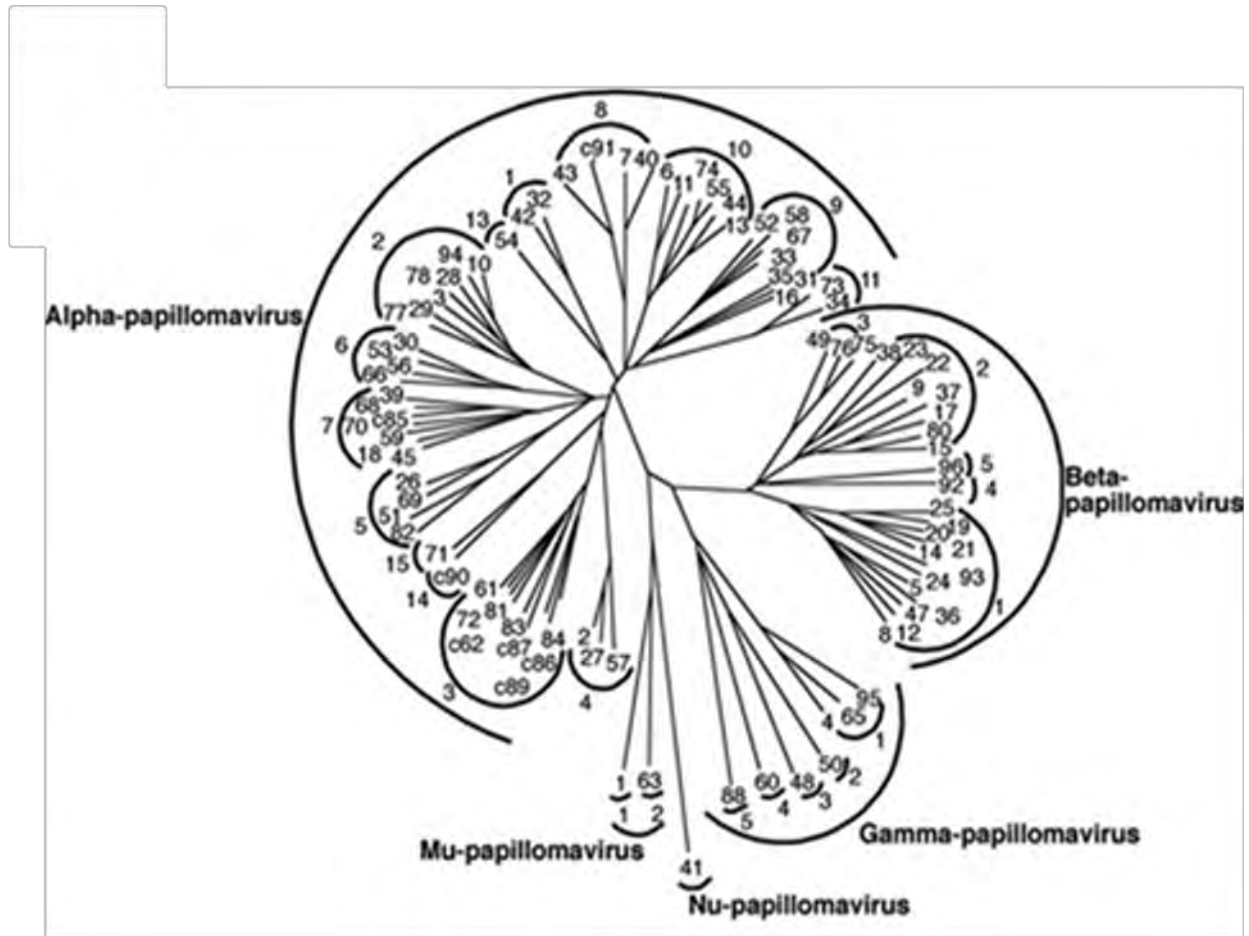


Figure 1.1 Human papillomavirus family tree

HPVs are subdivided into five evolutionary groups. The Alpha group infects the cervix. HPV types from the Beta, Gamma, Mu and Nu groups primarily infect cutaneous sites (Doorbar, 2006).

Risk group	HPV types
High-risk	16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, 73 and 82
Low-risk	6, 11, 42, 43, 44, 55, 81, 83
Cutaneous	3, 7, 10, 28, 29, 40, 43, 77, 78, 91, 94

Table 1.1 Risk groups of α HPV types (risk relates to the risk of developing cervical cancer).

1.3 HPVs and Cervical cancer

Cervical cancer remains the second most common malignancy and second most common cause of cancer-related death in women worldwide (Fig. 1.2) (WHO, 2006).

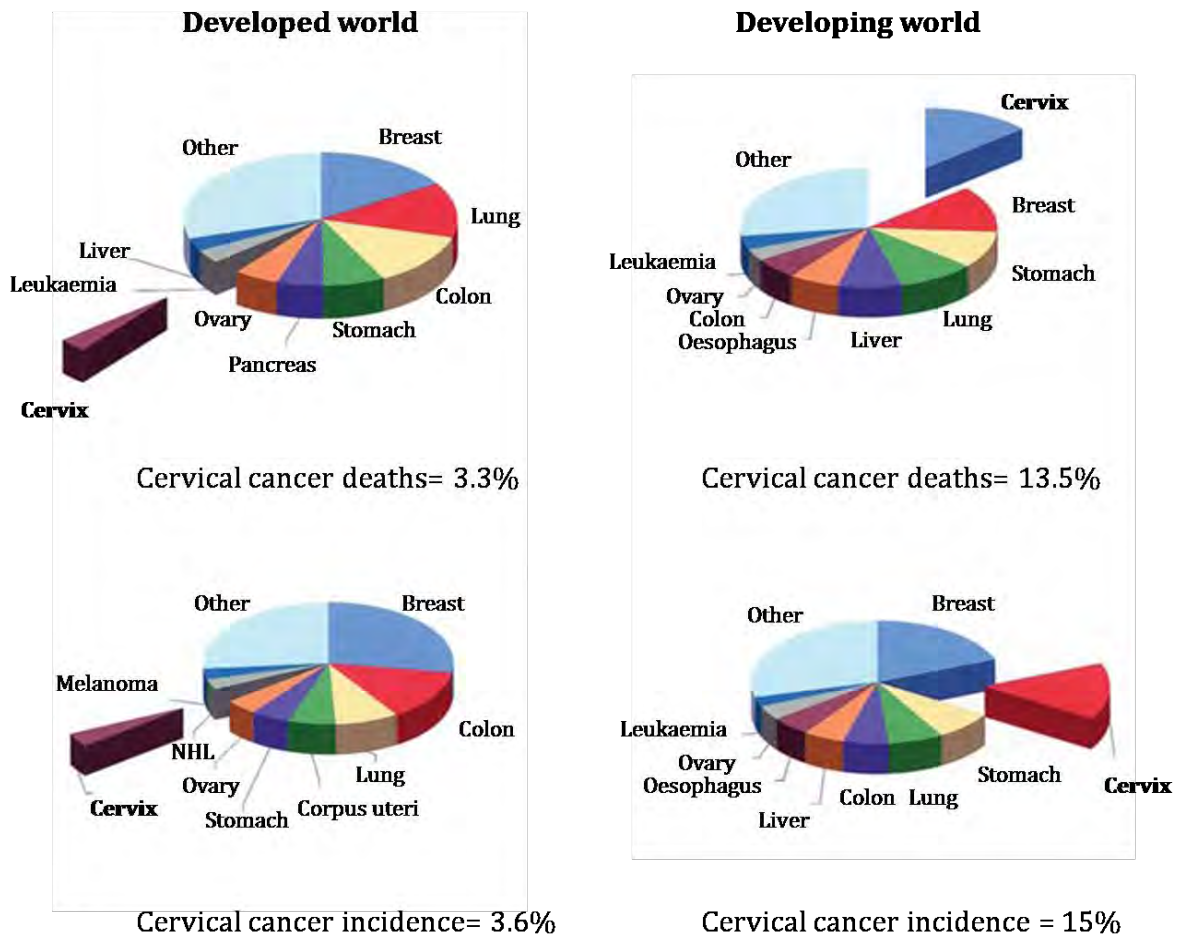


Figure 1.2 Cervical cancer incidence worldwide

Incidence of cervical cancer and deaths related to cervical cancer in the developed and developing world. The pie charts are colour-coded according to relative prevalence. (N.J. Maitland, modified from IARC website (www.iarc.fr)).

Most recent estimates suggest that each year there are more than 500,000 new cases of cervical cancer and over 250,000 deaths from the disease (Wheeler, 2007).

The most important risk factor in the development of cervical cancer is a persistent infection with a high-risk strain of human papillomavirus, in fact high-risk HPV strains are responsible for more than 99.7% of cervical cancer worldwide (Walboomers *et al.*, 1999). Most cases of cervical cancer are attributable to sexually acquired HPV infection (Bosch *et al.*, 2002).

HPVs infections are normally cleared within 12-18 months after infection, but high-risk types may persist in the infected host and lead to cancer (Baseman & Koutsky, 2005) (Fig. 1.3). The interval between the acquisition of HPV infection and malignant progression usually takes at least 10 years and is frequently longer (Snijders *et al.*, 2006).

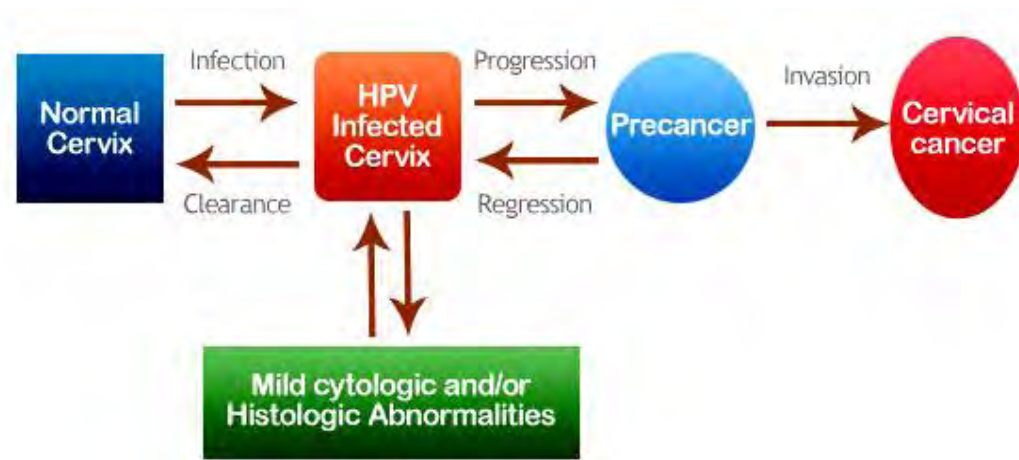


Figure 1.3 The major steps in cervical carcinogenesis

HPV infection followed by clearance by the immune system or progression to precancer and invasion. (Schiffman & Kjaer, 2003).

Cervical cancer arises from non invasive pre-malignant lesions known as cervical intraepithelial neoplasias (CINs). These lesions are divided into three groups that represent the different stages: CIN 1 for mild dysplasia; CIN 2 for moderate dysplasia and CIN 3 for both severe dysplasia and carcinoma in situ (Fig. 1.4). CIN lesions are also graded as low-grade squamous intraepithelial lesion (LSIL) or high-grade squamous intraepithelial lesion (HSIL) (Fig. 1.4). LSIL is the indication of mild cervical dysplasia seen on Papanicolaou (Pap) cytology testing and generally corresponds to the histological classification of CIN 1, whereas HSIL indicates moderately to severely abnormal-appearing cells on a Pap smear and corresponds to the histological classification of CIN 2 or 3 (Fig. 1.4). CIN lesions can be detected by screening programs before cancer develops. In developed countries, regular cervical cancer screening with Pap and HPV tests have been shown to decrease the incidence of cervical cancer by 70% (Kitchner *et al.*, 2006). In less-developed countries, where sufficient resources for high-quality cervical cancer screening are not available, the risk of developing cervical cancer from CIN 3 is estimated to be 80% (Wright *et al.*, 2002). The recent introduction of a vaccine against HPVs is an important development in the fight against cervical cancer. Early studies show that HPV vaccination, recommended for young women and men between 9-26 years old, not exposed to HPV prior to immunization, has the potential to significantly reduce the rate of cervical cancer worldwide and the burden of precancerous cervical lesions in women (Paavonen, 2009). Gardasil (CDC, 2010; Lowy & Schiller, 2006; NCI, 2011) and Cervarix (Hitti, 2006; Kresge, 2006) are the vaccines currently available, which prevent cervical dysplasia from the HPV types 16, 18, 6 and 11 and types 16 and 18, respectively. This effect has lasted 6 years after vaccination for Gardasil and more than 8 years for Cervarix (Harper *et al.*, 2008). To date, it seems that booster vaccines will not be necessary (Committee opinion no. 467, 2010). Both vaccines are subunit virus-like

particle (VLP) vaccines composed of the viral L1 protein, which is the major structural protein of the virus and includes the immunodominant neutralization epitopes of the virus. When expressed in eukaryotic cells, L1 is capable of self-assembly into VLPs (Kirnbauer *et al.*, 1993) that can induce high levels of neutralizing antibodies (Roden, 1996; Rose *et al.*, 1994).

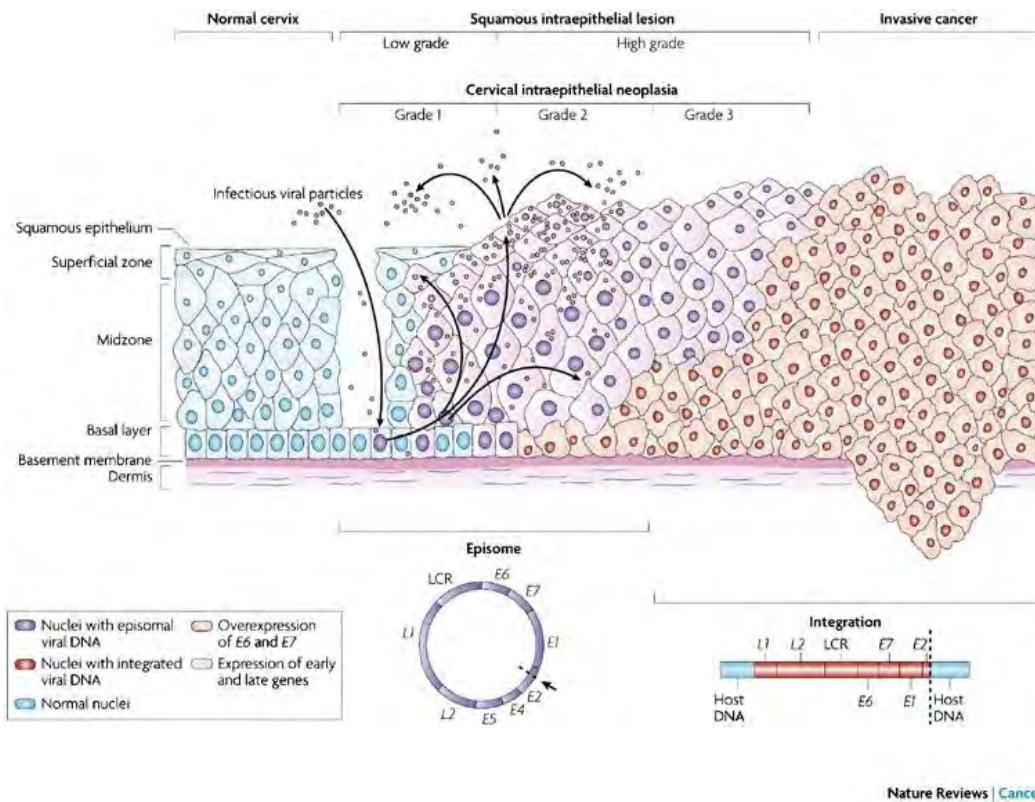


Figure 1.4 Representation of cervix epithelium pre- and post-infection

Normal cervix and different grade of squamous intraepithelial lesions: low-grade squamous intraepithelial lesion (LSIL) and high-grade squamous intraepithelial lesion (HSIL); a small percentage of cases progresses to cervical cancer that, if left untreated, will invade surrounding tissues. The histology grade of cervical intraepithelial neoplasia is also shown: CIN 1 for mild dysplasia; CIN 2 for moderate dysplasia and CIN 3 for both severe dysplasia and carcinoma in situ (Ciaran *et al.*, 2007).

1.4 The cervix

The cervix is the cylindrical lower portion of the uterus that protrudes into the vagina (Junqueira & Carneiro, 2005) (Fig. 1.5). Two narrow openings are present in the cervix called the internal and external *os*. The internal *os* is located on the topmost portion of the cervix and opens into the uterus (Stevens & Lowe, 2005). The external *os* is located at the minor portion of the cervix and opens into the vagina (Ross & Pawlina, 2006). The passageway between the external *os* and the uterine cavity is called endocervical canal or canal of the cervix which can change in width and length. The cervix is divided in endocervix, the portion proximal to the uterus and ectocervix proximal to the vagina. The endocervix has a fusiform shape and is composed of a single layer of mucous-secreting simple columnar epithelium. The ectocervix has a convex, elliptical surface and is composed of nonkeratinized stratified squamous epithelium (Arends *et al.*, 1998) which comprises diverse layers: basal, parabasal, intermediate and superficial. The basal layer is composed of a single row of cylindrical cells where mitosis takes place. Above the basal cell layer lie the parabasal and intermediate layers which together constitute the prickle cell layer. The superficial layer is made up of the most mature cells and varies in thickness, depending on the degree of estrogen stimulation (Fig. 1.6). The portion adjacent to the edge of the endocervix and ectocervix is known as the transformation zone (TZ) (Ross & Pawlina, 2006), or squamo-columnar junction (Stevens & Lowe, 2005), in which columnar epithelium is converted to squamous epithelium by a process called squamous metaplasia. The TZ is the area where most abnormal change occurs and is susceptible to carcinogens and diseases.

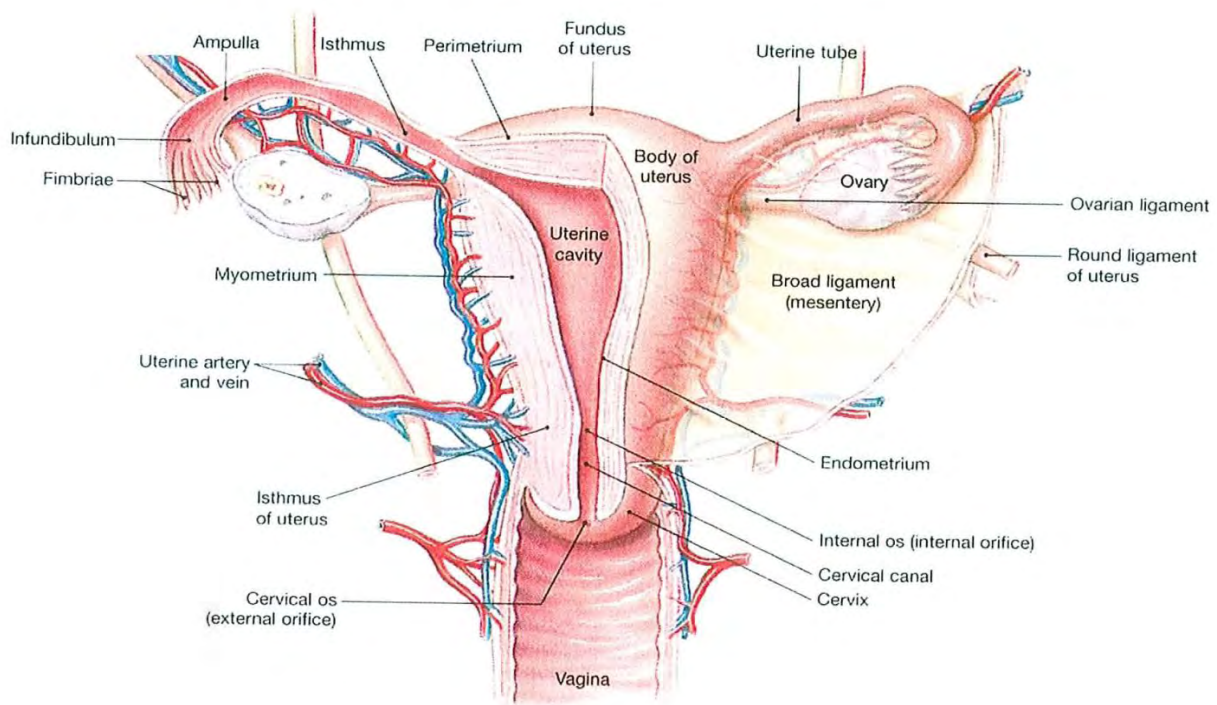


Figure 1.5 The uterus

Diagram of the uterus indicating location of the cervix, the internal and external *os* and the cervical canal (from Martini & Bartholomew, 2007).

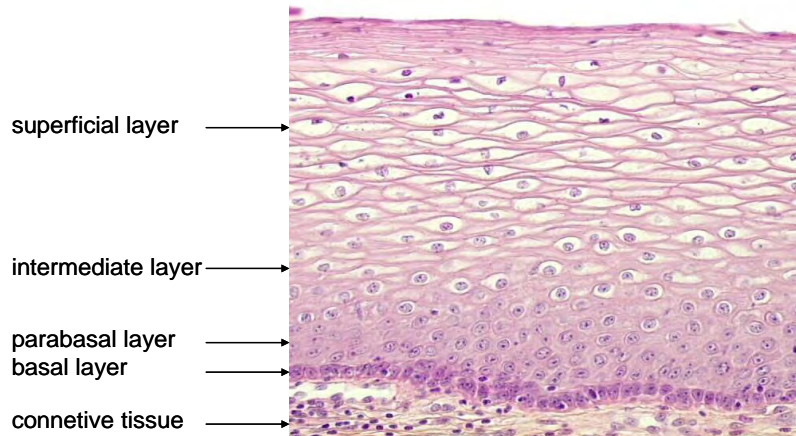


Figure 1.6 Structure of the ectocervix

1.5 HPV life cycle

HPVs infect primitive basal keratinocytes (Lowy & Howley, 2001) through micro-wounds that expose the cells to the incoming viruses (Doorbar, 2005). The HPV-16 viral life cycle is strictly linked to the differentiation of these cells (Fig. 1.7). The virus enters the cell by binding to different receptors on the surface of the cell (Evander *et al.*, 1997; Giroglou *et al.*, 2001; Joyce *et al.*, 1999). Glycosaminoglycans (GAGs), especially heparan sulphate, have been suggested as initial attachments receptors for HPV VLPs (Drobni *et al.*, 2003; Joyce *et al.*, 1999). Another candidate receptor for HPV, laminin-5 in the extracellular matrix (ECM) has been suggested (Culp *et al.*, 2006a; Selinka *et al.*, 2007), with an affinity even higher than heparan sulfate; however infectious transfer from the ECM seems to also require heparan sulfate binding (Culp *et al.*, 2006b; Giroglou *et al.*, 2001; Selinka *et al.*, 2007). A second receptor, $\alpha 6$ -integrin, has been shown to be involved in HPV infection, but its role is rather controversial (Culp *et al.*, 2006b; Sapp & Day, 2009). Once the virus has entered the basal epithelial cells and migrated to the nucleus, the viral genome starts its replication producing approximately 50-100 episomes of viral DNA per cell (Stunberauch & Laimins, 1999). At this stage only low levels of the early E1 and E2 genes are expressed, along with the early oncogenes E6 and E7, whereas expression of the late genes is suppressed.

(Stunberauch & Laimins, 1999). Once the cell migrates to the upper layers, the viral DNA replicates to several thousand copies and the late promoter is activated inducing expression of the late genes L1 and L2, which encode for the capsid proteins. At this stage the viral genomic DNA is packaged into capsids in the nucleus and the newly formed viruses are released as the superficial layers of the epithelia are sloughed off (Doorbar, 2005).

1.6 The HPV genome

The HPV genome is a double stranded, circular DNA, 7904 base pairs (bp) in size, with all genes located on the same strand (Lowy & Howley, 2001). It can be divided into three domains: a long control region (LCR), an early and a late region (Fig. 1.8 A and B). The early and late gene regions are both protein-encoding whereas the LCR is a non-encoding sequence (Park *et al.*, 1995). The LCR is approximately 800-1000 base pairs (bp) long, located adjacent to the E6 ORF and it contains many cis-regulatory elements necessary for viral transcription and replication (Turek, 1994). The early region is approximately 4 Kb in size and contains six open reading frames (ORFs), E6, E7, E1, E2, E4 and E5; however it has been shown that HPV-31 contains also an E8 ORF (Stubenrauch *et al.*, 2000). The late region is approximately 3 Kb and contains L1 and L2 genes which encode two capsid proteins, the L1 major capsid protein, and L2, the minor capsid protein.

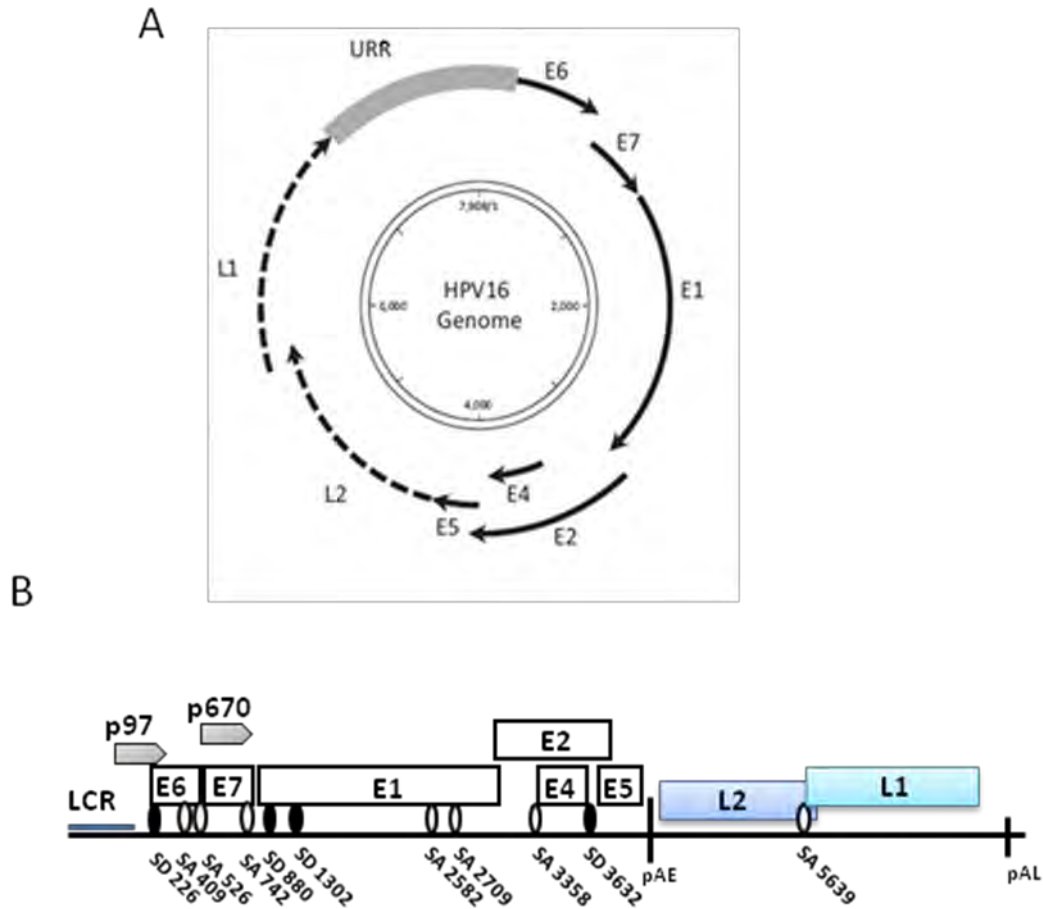


Fig 1.8 Genomic organisation of HPV-16

(A) The HPV-16 genome circular, double stranded (Burk *et al.*, 2009). (B) Linear representation of the HPV-16 genome. The early and late viral promoters, p97 and p670, and the early and late polyadenylation signals pAE and pAL are shown. Boxes indicate protein-coding regions: E represents the early genes and L represents the late genes. The long control region (LCR) is also indicated. Black circles indicate splice donors (5' splice sites) and white circles indicate splice acceptor (3' splice sites). Numbers refer to nucleotide positions in the HPV-16 sequence.

Expression of L1 and L2 is detected only in the upper layers of the epithelium (Zheng & Baker, 2006). Two major promoters are present on the HPV genome; one early promoter (p97) located upstream of the E6 coding region, which drives expression of the early genes and a late promoter (p670) located on the E7 coding region, which drives expression of the late genes (Howley & Lowy, 2001; Longworth & Laiminis, 2004). The late promoter is activated only in the differentiating part of the epithelium (Cumming *et al.*, 2003). The genome also contains two polyadenylation signals: the early polyadenylation signal (pAE) that is located between the early and late region and the late polyadenylate signal (pAL) located in the late region (Zheng & Baker, 2006; Schwartz *et al.*, 2007). All early mRNAs are polyadenylated at the pAE, and the late mRNAs can be polyadenylated at the pAE or pAL (Fig. 1.8).

1.6.1 E1 and E2 proteins

E1 and E2 have a pivotal role in viral replication. The E1 gene product is a 68 kilodalton (kDa) nuclear protein, highly conserved among papillomaviruses. E1 is expressed at very low levels and has been shown to bind to the viral Ori, located in the proximal end of the LCR. The affinity of this binding is very weak and in order to be efficiently increased requires the presence of E2 (Dixon *et al.*, 2000; Frattini & Laimins, 1994). E2 associates with E1 primarily through its N-terminus and binds to DNA as a dimer through its C-terminus (Moscufo *et al.*, 1999; Sarafi & McBride, 1995). The complex E1-E2 binds to DNA with high affinity and can unwind the DNA. It also been shown that E1 forms a specific protein complex with the catalytic p180 subunit of DNA polymerase- α to activate DNA replication (Liu *et al.*, 1995).

The E2 protein is approximately 48 kDa and also has functions in transcriptional regulation (Longworth & Laiminis, 2004). In the undifferentiated cells, the E2 protein acts at low

concentration as a viral transcriptional trans-activator to stimulate transcription whereas at high levels it acts as a repressor by interfering the binding of transcriptional factors such as Sp1 and TFIID to their recognition sequences (Longworth & Laiminis, 2004). In differentiated cells, when the late promoter is activated the E2 repressive function of transcription is arrested, resulting in a rapid increase in the level of E1 and E2 expression (Hebner & Laiminis, 2006). During the integration of the viral genome into chromosomes the E2 gene is broken and has been observed that the loss of E2 expression contributes to the progression from low-grade to high-grade cervical neoplasia (Schneider-Maunoury *et al.*, 1987).

1.6.2 E4 and E5 proteins

The E4 ORF does not contain the AUG start codon and so the E4 protein is formed by a single splice from E1 ORF to E4 ORF, indicated as E1^{E4} protein. The full length of this fusion protein is 17 kDa and contains the first 5 amino acids of the E1 ORF and the last 85 amino acids of the E4 ORF (Doorbar *et al.*, 1986). It is abundantly expressed in all different stages during epithelial cell differentiation and is distributed primarily in the cytoplasm. It has been shown that E4 proteins are associated with the keratin cytoskeleton, causing the collapse of the cytoskeleton. This suggests that E4 protein may allow virion particles to exit the cell.

The E5 proteins of human papillomaviruses are small hydrophobic proteins that are expressed in the early and late stages of the viral life cycle; however, their role in HPV pathogenesis is not clearly understood. It is about 5 kDa in size and is localized predominantly in the endoplasmic reticulum and Golgi apparatus, but is also present in the plasma membrane (McMurray *et al.*, 2001). HPV E5 is known to interact with growth factor receptors and gap junction proteins and is

believed to play a role during the initiation of neoplasia (Leechanachai *et al.*, 1992; Martin *et al.*, 1989).

1.6.3 E6 and E7 proteins

Expression of the viral oncogenes, E6 and E7, has been shown to lead to an increased proliferation of suprabasal epithelial cells (Doorbar, 2006). The E6 oncoprotein is approximately 17 kDa and is present in the nucleus and cytoplasm. It is expressed at the early stage of HPV life cycle and is capable of binding to p53. p53 is a tumor suppressor protein that, when cellular DNA is damaged, induces cell cycle arrest or apoptosis (Fig. 1.9). E6, particularly from high-risk HPV subtypes, binds to p53 through the E6-associate protein (E6AP) ubiquitin ligase. This new E6-E6AP-p53 complex leads to degradation of p53 therefore preventing growth arrest and/or apoptosis (Scheffner *et al.*, 1993). It has been shown that low-risk E6 can also bind to p53 but with lower affinity (Li & Coffino, 1996).

The HPV-16 E6 can activate telomerase via upregulation of human telomerase reverse transcriptase (hTERT) at transcriptional level (Veldman *et al.*, 2001). Telomerase is an enzyme that prevents the loss of important DNA sequences from the end of chromosomes and the shortening of telomere (Fehrmann & Laiminis, 2003). Expression of telomerase prevents senescence of cells. HPV-16 and HPV-18 E6 proteins contain a carboxyl terminal PSD-95/discs large/zonaoccludens 1 (PDZ) domain (Dell & Gaston, 2001). Many PDZ proteins are involved in signal transduction and they are found within cellular protein complexes located in the tight-junctions in epithelial cells (Longworth & Laiminis, 2004). The PDZ domain interacts with proteins such as MUPP-1, hDLG and hSCRIB which can be targeted by E6 protein (Fehrmann & Laiminis, 2003). This binding to E6 protein leads to E6-AP mediated ubiquitination and

proteolysis (Hebner & Laiminis, 2006). Binding to p53, activation of hTERT expression and binding of PDZ proteins are fundamental ways for E6 proteins to immortalize human cells (Hebner & Laiminis, 2006; Veldman *et al.*, 2001).

E7 is a nuclear protein approximately 12 kDa in size. It has been shown to be the major viral protein expressed in cervical cancers and cancer-derived cell lines containing HPV-16 or HPV-18 DNA, and in some cervical cancers only the E7 DNA is retained (Wilczynski *et al.*, 1988).

The HPV-16 E7 protein can form a specific complex with members of the retinoblastoma (Rb) family (Dyson *et al.*, 1989), which are involved in cell cycle regulation. Three proteins belong to this family: Rb, p107 and p130. They prevent excessive cell growth by inhibiting cell cycle progression through G1 into S phase. The binding of E7 with one of the member of the Rb tumor suppressor family leads to an inactivation of Rb protein and a consequent cell cycle progression into S phase (Munger & Phelps, 1993). In addition, it has been shown that inactivation of the cyclin-dependent kinase inhibitors (CKIs) such as p21 and p27, by the E7 oncoprotein is necessary to inhibit cell cycle arrest in human epithelial cells (Demers *et al.*, 1994; Helt *et al.*, 2002) (Fig. 1.9).

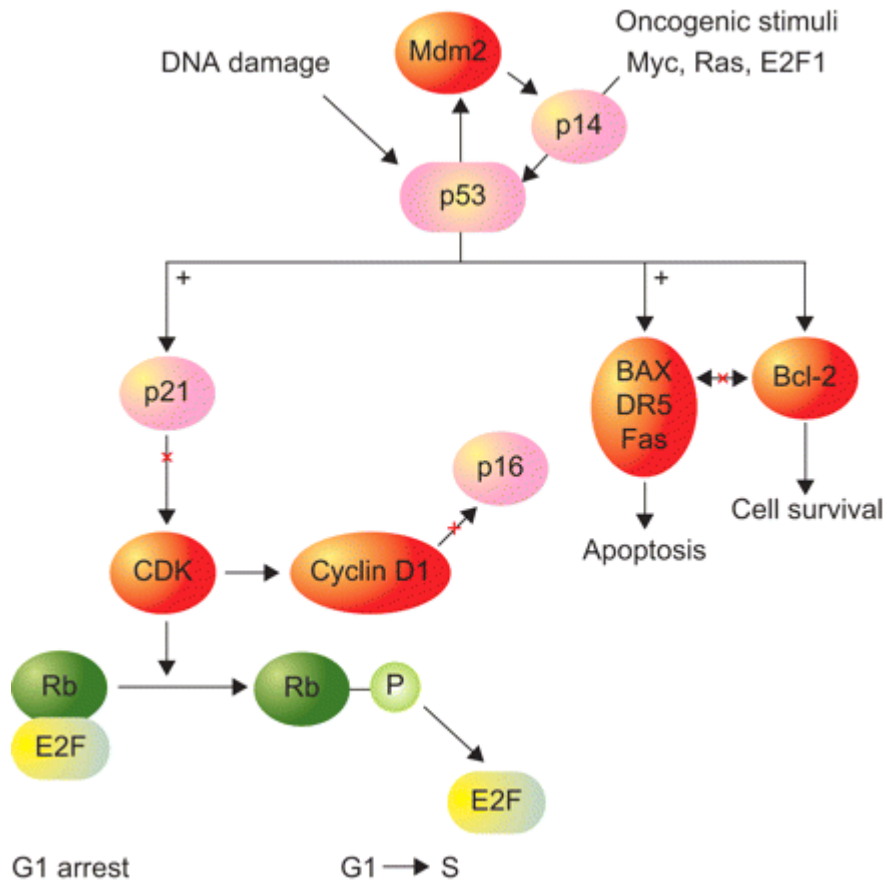


Figure 1.9 p53 and Rb pathway

The tumor suppressor protein, p53, can activate DNA repair and apoptosis. It also regulates cell cycle, through the retinoblastoma (Rb) pathway. The Rb function depends on the level of phosphorylation of Rb which is achieved by cyclin-dependent kinase (CDK) and cyclin D1 complex. When Rb is in its hypophosphorylated state is active and binds to E2F inhibiting cell cycle progression. In its phosphorylated state Rb is inactive and releases E2F that, in turn, activates factors like cyclins which push the cell through the cell cycle (from Brambilla & Gazdar, 2009).

1.6.4 L1 and L2 proteins

L1 and L2 genes are both transcribed from the late promoter (p670) (Stunberauch & Laimins, 1999) and are dependent on epithelial differentiation (Oberg *et al.*, 2003). L1 and L2 are highly immunogenic structural proteins, and their production in the lower layers of the infected epithelium is strongly suppressed, possibly to prevent detection by the host immune system (Schwartz *et al.*, 2007). The L1 protein is around 55 kDa in size and encodes for the major capsid protein. L1 proteins are capable of self-assembly into virus-like particles (VLPs) when expressed as a eukaryotic recombinant protein and the immunisation with these particles can produce an immune response in the infected host (Inglis *et al.*, 2006). The L2 protein is around 50 kDa in size and encodes the minor capsid protein. The function of L2 protein is still elusive but it is thought that it can be important in HPV capsid assembly and DNA binding (Zhou *et al.*, 1994). Both L1 and L2 proteins contain nuclear localization signals (NLS) necessary for transfer into the nucleus after synthesis that occurs in the cytoplasm (Zhou *et al.*, 1991).

1.7 DNA integration

The HPV DNA is usually extrachromosomal or episomal in benign cervical precursor lesions. However, in many cervical cancer cells the HPV DNA is integrated in the cellular genome (Park *et al.*, 1995). In cervical cancer the HPV DNA could be either episomal or integrated at the same time, although integration appears to occur more frequently in HPV-18-associated cervical cancer than in HPV-16-associated cervical cancer (Crusius *et al.*, 1997). Viral integration sites seem to be randomly distributed among all chromosomes with a clear preference for chromosomal breakpoints and genomic fragile sites. In fact viral DNA sequences have been

mapped to different chromosomal loci in independent cervical cancer cell lines (Durst *et al.*, 1987; Popescu *et al.*, 1987).

During HPV DNA integration, the viral genome usually breaks in the E1/E2 region resulting usually in the loss of the E1 and/or E2 regions. The functional inactivation of E2, which acts as a transcriptional repressor of the E6 and E7, leads to an uncontrolled and increased expression of the transforming genes E6 and E7. This continuous expression of the viral gene E6 and E7 is the major cause of malignant transformation of the host cells and tumor formation (Bosch *et al.*, 1992; Doorbar *et al.*, 1991; Phelps *et al.*, 1988). The integration of HPV DNA into the host DNA appears to increase cellular proliferation and the risk of malignancy (Scheurer *et al.*, 2005).

1.8 Regulation of gene expression

RNA processing is the pivot of regulation of gene expression, delivering a number of modifications in mRNA molecular features. In eukaryotic cells, the first transcript from a DNA template is called pre-mRNA. It is made in the nucleus by RNA polymerase II (RNA pol II) and it is not directly exported to the translation machinery in the cytoplasm. First it must undergo three major processing events to become a mature and stable mRNA. These three events are capping, splicing and polyadenylation and they can occur co-transcriptionally (Proudfoot *et al.*, 2002) (Fig. 1.10).

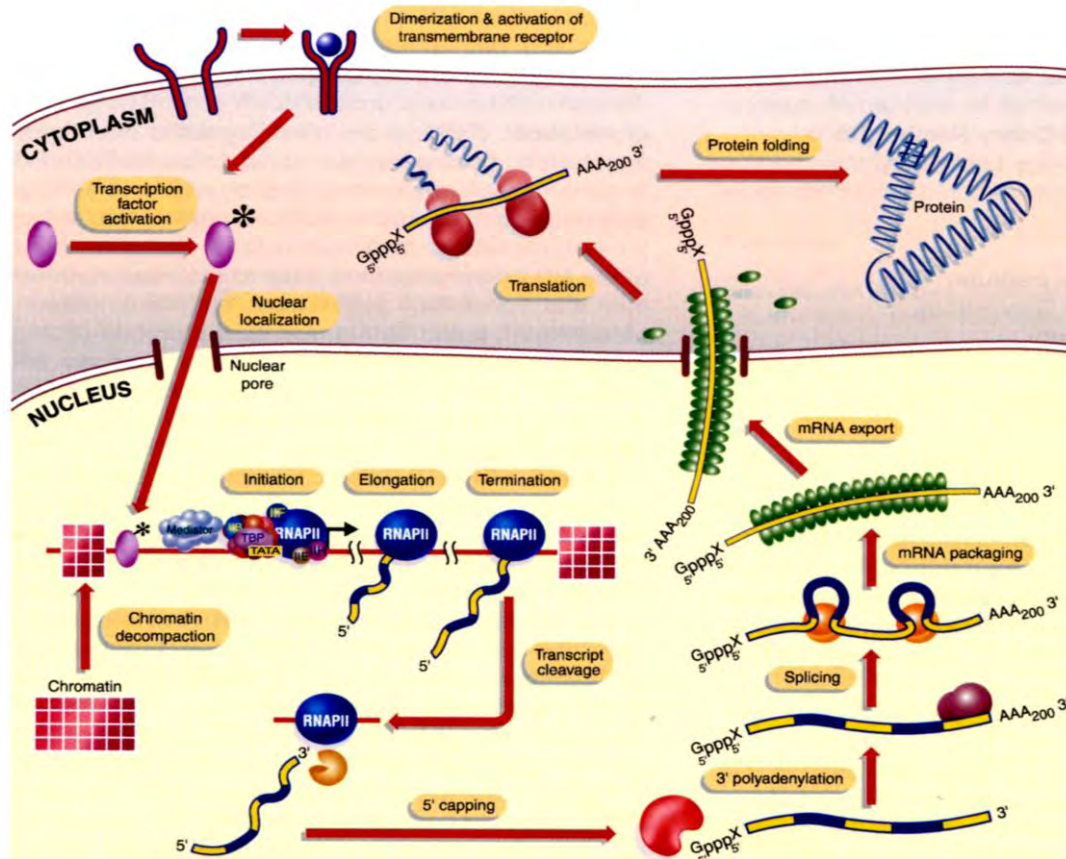


Figure 1.10 RNA processing

In eukaryotes a gene is transcribed from DNA to pre-mRNA. mRNA is then made from pre-mRNA by RNA processing, which includes the capping, splicing and polyadenylation of the transcript. It is then transported from the nucleus to the cytoplasm for translation. (Orpahnides, A unified theory of gene expression. Cell, 2000).

1.8.1 Capping

Capping consists on the addition of 7-methyl guanosine groups (mRNA "cap") to the 5' -ends of the new synthesized pre-mRNA and occurs after about 20-30 nucleotides of the molecule have been transcribed. The 5' -triphosphate is removed from the first nucleotide of the pre-mRNA by RNA triphosphatase, leaving two terminal phosphates. Guanosine mono-phosphate (GMP) is

added to the terminal phosphates by RNA guanylyl transferase. The 7-nitrogen of guanine is then methylated by methyltransferase (Shatkin, 1976). The 5' cap then binds to the cap binding complex (CBC). The new formed complex helps to protect the mRNA from premature degradation and also facilitates binding of ribosomes to the mRNA during translation.

1.8.2 Splicing

Almost all synthesized eukaryotic pre-mRNAs contain coding sequences called exons and non-coding sequences called introns. Pre-mRNA splicing is a process that removes introns and retains exons in the transcript that can be then translated into proteins. Each intron contains all consensus sequences necessary for splicing: a 5' splice site (5' ss), a branch point sequence (BPS) and a 3' splice site (3' ss) containing a polypyrimidine tract (PPT). The 5' ss is identified by the consensus sequence AG | GURAGU (R, purine) where the line | represents the junction between the first exon and the adjacent 5' end of the intron. The consensus sequence of the branch point sequence is YNYURAC (the A is the adenosine conserved in all genes which lies between 18 and 37 nucleotides upstream from the 3' ss). The consensus sequence of the 3' ss is YnAG|RNNN (Y, pyrimidine; N, variable nucleotides) where Yn represents the polypyrimidin tract and the line represents the junction between the 3' end of the intron and the second exon (Fig. 1.11 and 1.12). In the 3' splice site, the AG is normally preceded by a polypyrimidine tract of variable length (Y₁₀₋₂₀) and sequence and if it is long and uninterrupted, the 3' splice site is strong.

The pre-mRNA splicing is carried out by the spliceosome via a 2 step trans-esterification mechanism. In the first step, the 2' OH group of the conserved A residue in the BPS binds to the 5' ss and cleaves the phosphodiester bond at this site. This reaction produces a lariat shaped

intron and releases the 5' exon. In the second step, the 3' OH group at the end of the 5' exon attacks the phosphodiester bond at the 3' splice site releasing the lariat shaped intron and joining the flanking exons (Cooper & Hausman, 2007).

The spliceosome is formed from five ribonucleoprotein subunits, termed uridine-rich small nuclear ribonucleoproteins (snRNPs), transiently associated to non-snRNPs splicing factors (small nuclear RNA (snRNA), SR splicing factors, etc...) (Brown, 2002). Spliceosomes are highly dynamic machines, building anew on each pre-mRNA and their assembly occurs in a stepwise manner. In the first step, U1 snRNP binds to the 5' splice site and U2 snRNP to the branch point sequence via RNA:RNA interactions between snRNA and pre-mRNA. A pre-formed U5-U4-U6 triple-snRNP then binds to this new complex forming an immature spliceosome. This immature spliceosome complex undergoes a series of rearrangements to form a fully functional spliceosome. Several splicing factors are involved in the removal of introns and the process often continues post-transcriptionally (Bentley, 2005).

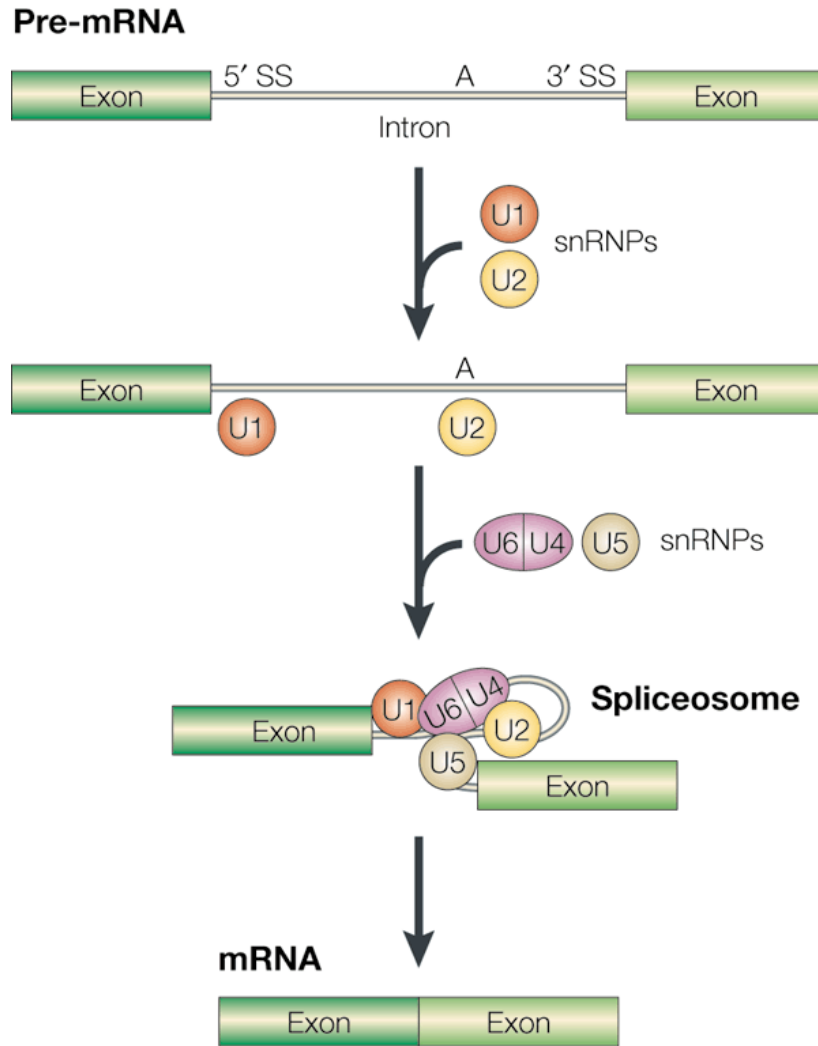


Figure 1.11 Splicing

The structure of the pre-mRNA and spliceosome is shown (from Dredge *et al.*, 2001).

1.8.3 Polyadenylation

Polyadenylation is an important process for the nuclear export, translation and stability of mRNA. It consists on the addition of a poly(A) tail at the RNAs 3' end (Fig. 1.12). First the pre-mRNA is cleaved by the cleavage and polyadenylation specificity factor (CPSF) at the cleavage/polyadenylation site, that usually is the AAUAAA sequence located 10-30 nucleotides upstream of the polyadenylation site. This binding is very weak and is stabilized by binding of the cleavage stimulation factor (CstF) which binds specifically to the GU-rich elements 20-40 nucleotides downstream of the polyadenylation site (Colgan & Manley, 1997). The cleavage factors I and II (CFI and CFII) allow the cleavage of the RNA. Before the cleavage occurs the poly(A) polymerase (PAP) associates with this complex allowing the rapid addition of adenosine at the cleavage site of mRNA. After the first 12 adenosines synthesis, poly(A) binding II (PABII) binds to the poly(A) tail, allowing the synthesis of a full length poly(A) by PAP. When the poly(A) tail is approximately 200 nucleotides long the polyadenylation stops.

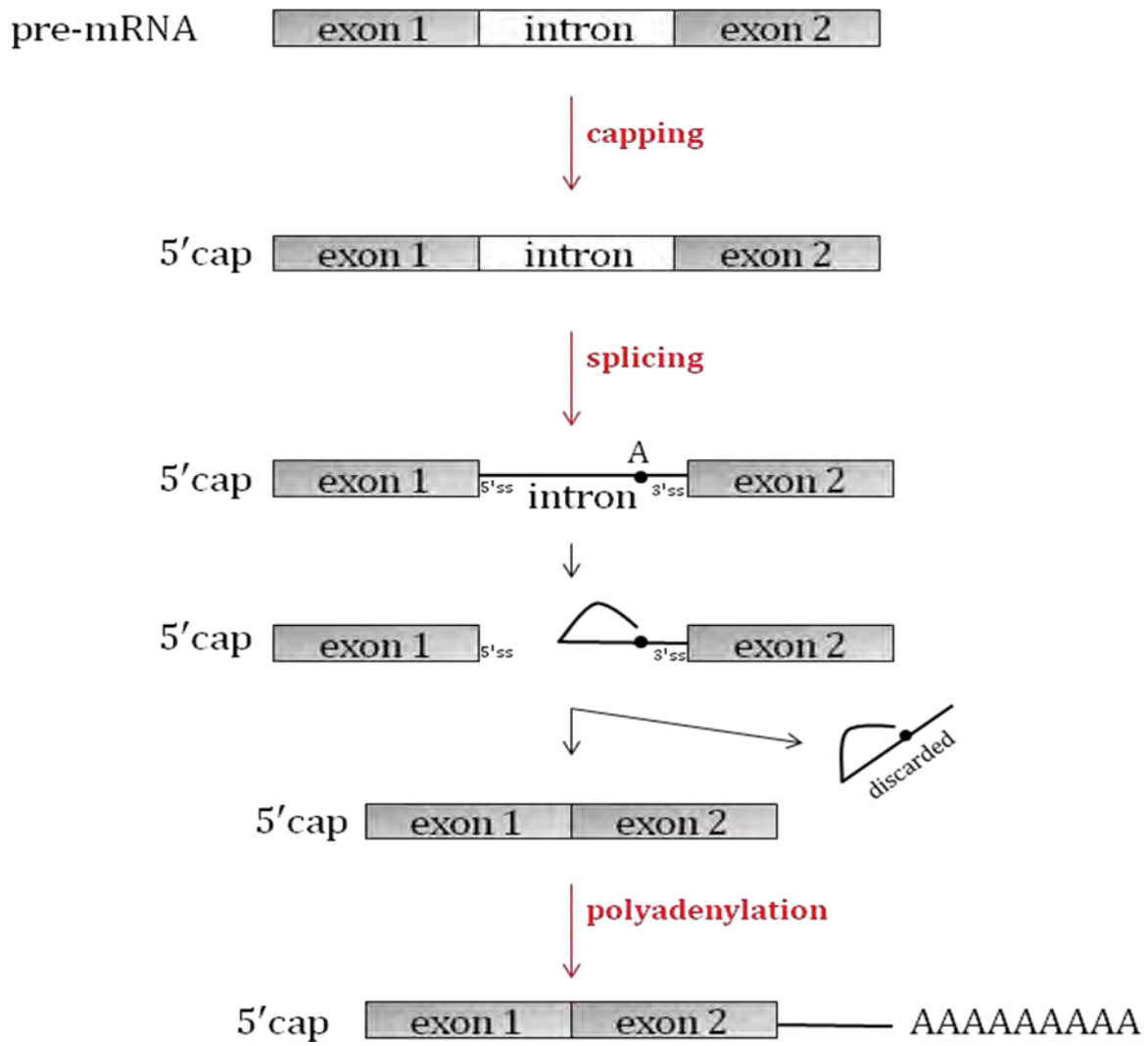


Figure 1.12 The structure of the pre-mRNA and RNA processing

The structure of the pre-mRNA is shown. Schematic representation of the three RNA processing events: capping, splicing and polyadenylation.

1.8.4 Alternative splicing

Alternative splicing is an important process to generate diversity and to regulate gene expression. This is particularly important in HPV gene expression where transcription of all mRNAs is regulated through the use of splicing and polyadenylation signals (discussed in detail below). Six basic modes of alternative splicing have been recognized: constitutive, exon skipping, mutually exclusive exons, intron retention, alternative donor site and alternative acceptor site (Fig. 1.13). The selection of the splice site depends on the strength of each splice site. The presence of splicing enhancer and silencer elements in exon and introns can increase or inhibit respectively the usage of weak splice site (Srebrow & Kornblihtt, 2006). Viruses efficiently exploit alternative splicing to produce many functional mRNAs from small genomes. About 70% of all human genes in the human genome undergo alternative splicing (Venter *et al.*, 2001). Alternative splicing can result in the addition or loss of coding sequences, a frame shift, or the presence of a premature stop codon. It is known that about 80% of alternative splicing can alter the encoded protein due to the addition or removal of many regulatory elements controlling translation, mRNA stability, or localization (Mondrek & Lee, 2002). Even though it is still unclear, different studies have demonstrated a correlation between alternative splicing and cancer (Faustino & Cooper, 2003; Krawczak *et al.*, 1992; Srebrow & Kornblihtt, 2006; Venables, 2004) and many splicing factors are up-regulated in cancer cells, suggesting that it would be opportune to have a better understanding of the diverse mechanisms by which disrupted splicing and splicing regulation contribute to human disease and cancer.

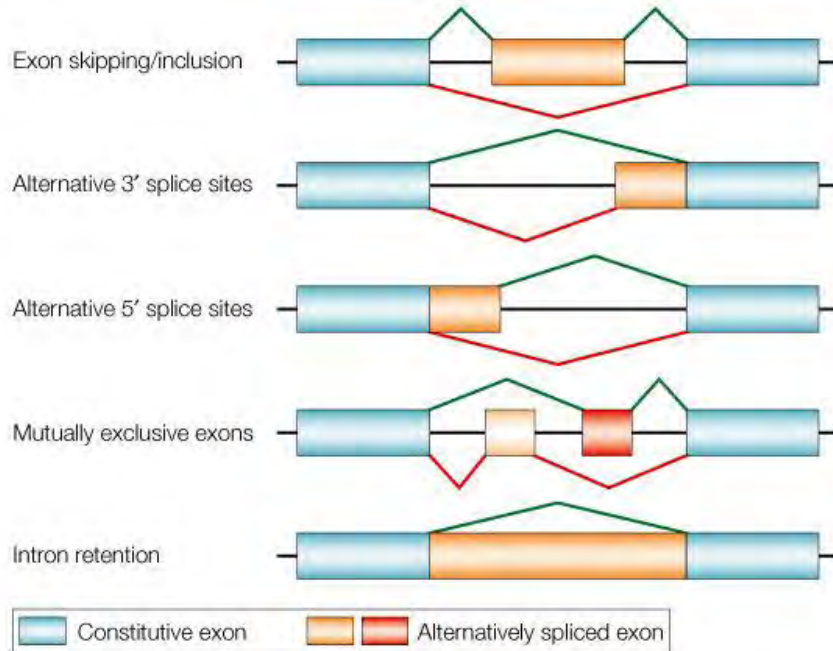


Figure 1.13 Schematic representation of alternative splicing

Different types of alternative splicing are illustrated. For a singular pre-mRNA different alternative exons often show diverse types of alternative-splicing patterns (from Cartegni *et al.*, 2002).

1.9 Regulation of HPV gene expression

HPV gene expression is regulated at the transcriptional level by the use of the two major promoters activated at different stages of cell differentiation (Lambert *et al.*, 1999). Cellular factors, such as TEF-2, Oct1, AP-1 have been shown to have an essential role in HPV transcriptional regulation by regulating promoter activity (Chong *et al.*, 1991). It has been demonstrated that also the viral E2 protein is involved in transcriptional regulation through the interaction with 4 binding sites located in the LCR (Zur Hausen, 1996). Since HPVs have only one strand active, transcription occurs in only one direction (Zur Hausen, 1996). Most HPV mRNAs are polycistronic (Favre *et al.*, 1997), and some of them carry 3 or more messages on one mRNA (Longworth & Laiminis, 2004). During transcription all mRNAs are regulated through the use of splicing and polyadenylation signals. For HPV-16 20 transcripts have been identified (Fig. 1.14), of which 14 are produced from the early promoter to the early polyadenylation signal (Zheng & Baker, 2006). In HPV-16 11 splice sites have been identified, 10 located in the early region and only 1 in the late region (Fig. 1.14) (Doorbar *et al.*, 1990). Splice Donor 226 (SD226) is used exclusively by early mRNAs and is followed by 3 distinct splice acceptors (SAs). This disposition allows the production of 4 different mRNAs, E6*I, E6*II, E6[^]E7 and when no splicing occurs E6E7 (Zheng & Baker, 2006). However in order to reach a correct and proportionate concentration of E6 and E7 proteins, a balance in these splicing events is required. E6 and E7 mRNAs are also spliced directly from SD880 to SA3358 (Fig. 1.14) (Baker & Calef, 1997).

Splicing from SD226 to SD742 in HPV-16 is required also for production of E1 mRNAs (Hubert & Laiminis, 2002). Splicing directly to SA2582 and SA2709 result in production of E2 mRNAs and directly to SA3358 is likely to produce E5 mRNAs (Baker & Calef, 1997).

The E4 mRNA is one of the most abundant HPV-16 mRNAs produced and is generated from SD880 to SA3358. The SA3358 is the most efficient splice site used in HPV-16. In fact it is used for production of E6, E7, E4, L1 and L2 mRNAs and also indirectly blocks HPV-16 late gene expression. It has been shown that an AC-rich splicing enhancer at SA3358 in HPV-16 promotes usage of SA3358 and polyadenylation at pAE, thereby indirectly suppressing HPV-16 late gene expression (Rush *et al.*, 2005). Late mRNAs are transcribed from the late promoter and are presumably spliced from SD880 to SA3358 and from SD3632 to SA5639 or directly from SD880 to SA5639 (Fig. 1.14). SD3632 and SA5639 are used exclusively by late mRNAs and are under control of multiple adjacent splicing silencer elements that suppress these splice sites therefore inhibiting late gene expression (Rush *et al.*, 2005; Zhao *et al.*, 2007). It has been shown that late gene expression is also inhibited in mitotic cells because of a competition between SA3358 and late splice sites (Rush *et al.*, 2005). In fact inactivation of the E4 enhancer and the L1 splicing silencers in a subgenomic HPV-16 plasmid showed production only of L1 mRNA (L1i) but not E4 mRNA production (Rush *et al.*, 2005; Schwartz *et al.*, 2007).

Alternative splicing seems to be necessary to produce L1 since the 3' end of L2 and the 5' end of L1 overlap. Late viral mRNAs are expressed only in differentiating cells and it has been shown that posttranscriptional events are highly involved in late gene regulation (Zheng & Baker, 2006). Viral RNA elements and cellular RNA binding factors have therefore a pivotal role in HPV-16 gene regulation (Mole *et al.*, 2009a; Mole *et al.*, 2006; Schwartz *et al.*, 2007). Thus it is of interest to elucidate the mechanism of HPV-16 gene regulation.

HPV-16 genome

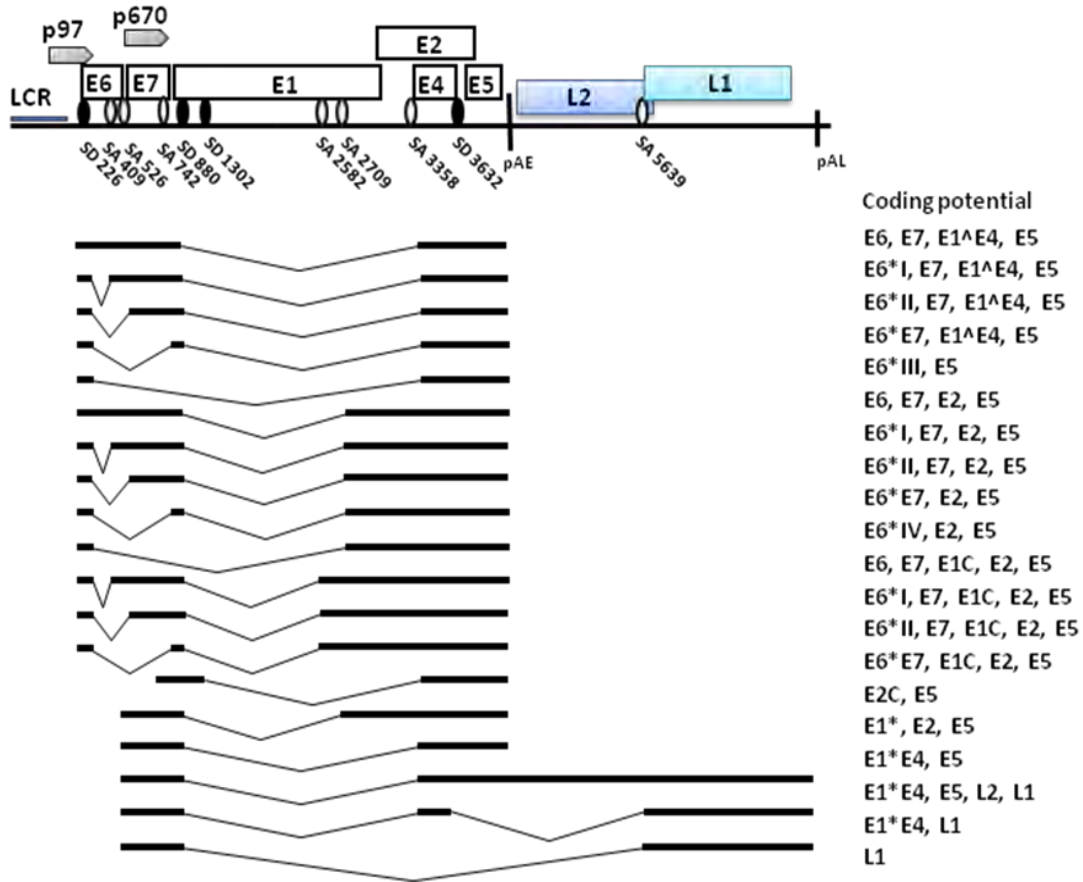


Figure 1.14 HPV-16 genome and transcription map

Schematic representation of the HPV-16 genome and transcription map. The early and late viral promoters, p97 and p670, and the early and late polyadenylation signals pAE and pAL are shown. Boxes indicate protein-coding regions. Black circles indicate splice donors (5' splice sites) and white circles indicate splice acceptor (3' splice sites). Numbers refer to nucleotide positions in the HPV-16 sequence. All 20 different RNA transcripts are shown. The coding potential of each mRNA is shown on the right.

1.10 HPV-16 late gene expression

As already mentioned, the HPV late genes are expressed only in differentiated cells and also it is known that late HPV-16 genes are not expressed in cervical cancer containing HPV-16 DNA (Doorbar, 2005). They encode for the highly immunogenic structural proteins, L1 and L2, and their production in the lower layers is strongly suppressed, possibly to prevent detection from the host immune system (Schwartz *et al.*, 2007). The SD3632 and SA5639 are used exclusively by the late mRNAs and the presence of an adjacent splicing silencer that actively suppresses the use of these splice sites has been shown to inhibit late gene expression (Rush *et al.*, 2005; Zhao *et al.*, 2004; Zhao *et al.*, 2007). Late gene expression is also indirectly inhibited by RNA elements that stimulate early mRNA splicing (Rush *et al.*, 2005) and polyadenylation (Oberberg *et al.*, 2003; Oberberg *et al.*, 2005; Terhune *et al.*, 1999; Zhao *et al.*, 2005). Recent studies show that overexpression of a number of proteins, such as adenovirus E4orf4 (E4orf4), polypyrimidine tract binding protein (PTB) and alternative splicing factor/splicing factor 2 (ASF/SF2), induce late gene expression in subgenomic HPV-16 expression plasmids (Somberg & Schwartz, 2010; Somberg *et al.*, 2009; Somberg *et al.*, 2008). Another member of the serine-arginine rich family (SRp30c) has been shown to regulate splicing (Raffetserder *et al.*, 2003) and it also has a role in late gene expression (Somberg *et al.*, 2011).

1.11 SR proteins

In humans, the SR protein family is encoded by nine genes, named ASF/SF2, SC35, SRp20, SRp40, SRp55, SRp75, SRp30c, 9G8, and SRp54 which have a common structural organization (Fig. 1.15), containing either one or two amino-terminal RNA-binding domains (RNA recognition motifs (RRM)) and a variable-length RS domain rich in arginine (R) and serine (S) amino acid residues at their carboxyl terminus (Shepard & Hertel, 2009). As the name implies the RRM provides RNA-binding specificity through a wide range of interactions, whereas the RS-domain functions as a protein interaction domain that promote recruitment of the spliceosome, splicing and shuttling between the nucleus and cytoplasm (Caceres *et al.*, 1998; Graveley & Maniatis, 1998; Wu & Maniatis, 1993). In addition, the RS domain can function as a nuclear localization signal by mediating the interaction with of the SR protein with the nuclear import receptor, transportin-SR (Cáceres *et al.*, 1997; Lai *et al.*, 2000). The SR proteins play an important role in mRNA export, translation, stability and both activation and repression of constitutive and alternative splicing of pre-mRNA (Shepard & Hertel, 2009; Sanford *et al.*, 2005; Long & Caceres, 2009). They mostly bind exonic splicing enhancer (ESE) therefore enhancing splicing of pre-mRNA into messenger RNA (mRNA).

SR proteins take part in many steps during the splicing reaction (Chew *et al.*, 1999; Fu, 1993; Ge & Manley, 1990; Krainer *et al.*, 1990; Zahler *et al.*, 1992) and require phosphorylation for efficient splice site recognition and dephosphorylation for splicing catalysis (Mermoud *et al.*, 1994; Mermoud *et al.*, 1992).

SF2/ASF and SC35 are two of the best characterized among the nine human SR proteins identified to date. They interact with U1-70K and U2AF by RS domain-mediated protein-protein

interactions and the RRM of these two proteins are responsible for their unique substrate specificities (Chandler *et al.*, 1997; Mayeda *et al.*, 1999).

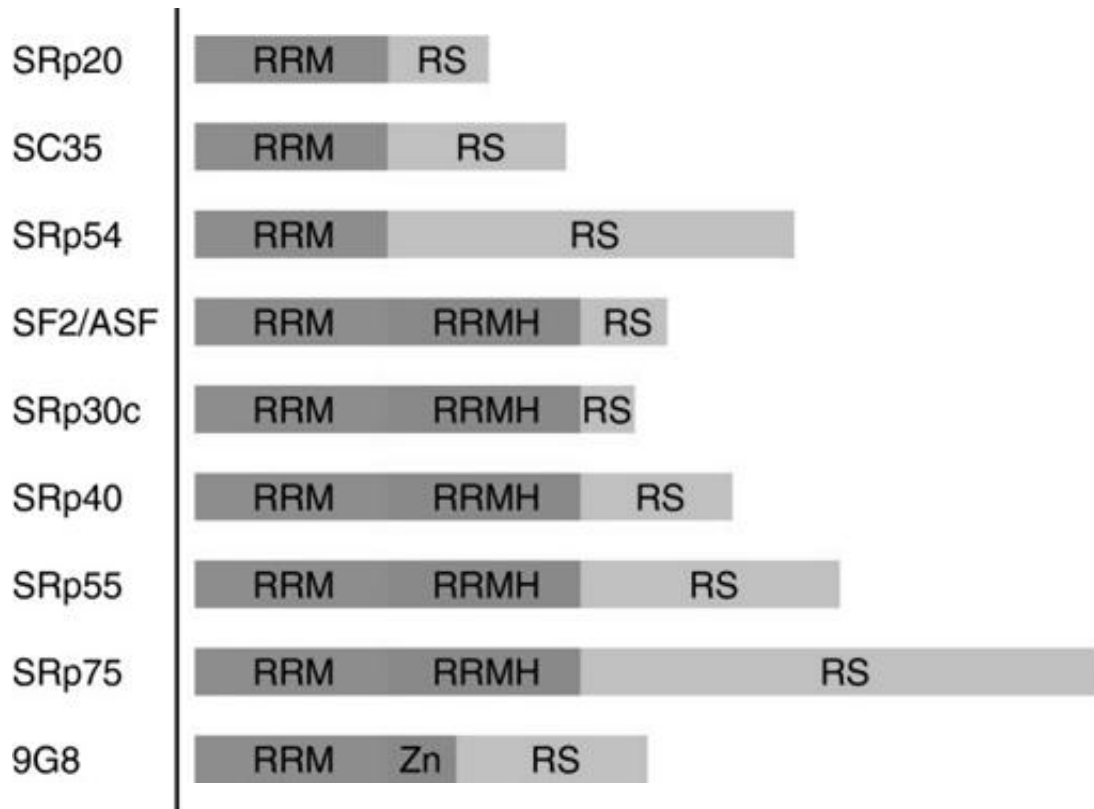


Figure 1.15 Human SR proteins family

The structural organization of the nine human SR proteins is shown. RRM, RNA recognition motif; RRMH, RRM homology; RS, arginine/serine-rich domain; Zn, Zinc knuckle (from Shepard & Hertel, 2009).

(i) ASF/SF2

ASF/SF2 (alternative splicing factor/splicing factor 2), also called splicing factor, arginine/serine-rich 1 or SFRS1, is encoded in humans by the *SFRS1* gene. This protein of approximately 30 kDa binds to pre-mRNA transcripts and components of the spliceosome, and can either activate or repress splicing depending on the location of the pre-mRNA binding site.

ASF/SF2 contains an arginine-serine rich region (RS domain), and two RNA recognition motifs (Fig. 1.15). It interacts with RNA and other splicing factors through the RRM (Ngo *et al.*, 2008; Hagopian *et al.*, 2008) and binds directly to RNA at the branchpoint to promote pre-spliceosomal assembly (Shen & Green, 2004; Shen *et al.*, 2004).

It is well-known that ASF/SF2 is a proto-oncogene and it has also been shown that knockdown of this gene inhibits tumor growth *in vitro* and *in vivo* (Karni *et al.*, 2007). ASF/SF2 expression is up-regulated in a number of tumors such as breast, liver, colon, kidney, pancreas, lung (Zerbe *et al.*, 2004) and cervix (Fay *et al.*, 2009). ASF/SF2 and heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1) act antagonistically to modulate splice site selection. Levels of ASF/SF2 along with hnRNP A1 have been shown to be altered in neoplastic lung cells and ASF/SF2 has shown to be localized in cytoplasm and/or nuclei while hnRNP A1 appeared prevalently within tumor nuclei (Zerbe *et al.*, 2004).

It has been shown that ASF/SF2 regulates one of the most used splice sites in HPV-16, the splice acceptor SA3358, specifically binding to exonic sequences located between SA3358 and SD3632 (Somberg & Schwartz, 2010). In particular ASF/SF2 enhances splicing to SA3358 and inhibits usage of splice donor SD3632 (Somberg & Schwartz, 2010).

This SR protein also interacts with U2AF, which binds to the negative regulatory element (NRE) that is located in the late 3' untranslated region of HPV-16 (McPhillips *et al.*, 2004). In addition,

it seems that in epithelia infected by high-risk HPV, ASF/SF2 expression is up-regulated by the virus transcription and replication factor E2 (McPhillips *et al.*, 2004).

(ii) SRp30c

SRp30c is one of several highly conserved serine/arginine-rich (SR) proteins that are involved in both constitutive and alternative splicing in eukaryotic cells. It has been shown to have a role in the alternative splicing of hnRNP A1 by repressing the use of a 3' splice site (Simard & Chabot, 2002). However, SRp30c has been shown to regulate alternative splicing of cellular mRNAs as an inducer and suppressor. SRp30c displays 74% amino acid similarity with ASF/SF2 although no consensus mechanism of action for these two splicing factors has been identified. As shown in figure 1.15 it contains two RRM (RNA recognition motif) domains.

Selection of alternative splice sites can be affected by interaction with SRp30c and also its overexpression is involved in the shuttling of the Y-box protein-1 (YB-1) into the nucleus (Raffetseder *et al.*, 2003). Presence of nuclear YB-1 is a strong adverse survival factor in human breast cancer (Dahl *et al.*, 2009).

It has been shown that SRp30c is required for alternative splicing of the glucocorticoid receptor (GR) pre-mRNA to form mRNA encoding GR β (Xu *et al.*, 2003). Several studies indicate that GR resistance has been associated with increased expression of GR β , therefore upregulation of SRp30c contributes to the GR resistance in several inflammatory cell types (Xu *et al.*, 2003; Jain *et al.*, 2011).

SRp30c has been shown to induce HPV-16 late L1 gene expression in subgenomic HPV-16 plasmids by binding to sequences between SA3358 and SD3632 and specifically interfering with splicing enhancers downstream of SA3358 in an RS-domain independent manner (Somberg *et*

al., 2011). SRp30c also interferes with splicing inhibitory RNA elements in an RS-domain dependent manner activating splicing of late splice site SD3632 leading to production of L1 mRNA (Somberg *et al.*, 2011).

(iii) SC35

Serine/arginine-rich splicing factor 2, also called SC35, belongs to the splicing factor SR family. It contains 1 RRM (RNA recognition motif) domain and 1 RS (arginine/serine-rich domain) (Fig. 1.15).

In vitro it self-associates and binds to ASF/SF2, SNRNP70 and U2AF1 but not U2AF2. SC35, like other members of the SR family of splicing factors, may take part in splice site selection by functioning as a bridge between components bound to the 5' and 3' splice sites (Wu & Maniatis, 1993; Kim *et al.*, 2011; Tronchère *et al.*, 1997). SC35 leads to both an exon inclusion and an intron excision in the 3' untranslated region of its mRNAs. It is required for formation of the earliest ATP-dependent splicing complex and for ATP-dependent interactions of both U1 and U2 snRNPs with pre-mRNA.

Sureau *et al.* reported that endogenous SC35 mRNA levels in HeLa cells are decreased when SC35 is overexpressed. They also demonstrated that SC35 accumulation is correlated to changes in the splicing pattern regulating its expression by promoting alternative splicing events changing the stability of its own mRNAs (Sureau *et al.*, 2001).

SC35 protein plays a critical role in terminal differentiated mature cardiomyocytes in mice and controls cell proliferation during pituitary gland development (Xiao *et al.*, 2007).

Zheng *et al.* demonstrated with DNA microarrays that SC35 expression is induced by the small molecule 12-O-tetradecanoylphorbol-13-acetate (TPA), in leukemia cell lines (Zheng *et al.*, 2002).

1.12 hnRNP family

Heterogeneous nuclear ribonucleoproteins (hnRNPs) are a group of RNA-binding proteins first discovered in 1956 during the observation of lampbrush chromosomes in amphibian oocytes (Gall, 1956). These proteins have molecular masses of 34-120 kDa and play a critical role as trans-factors in regulating gene expression, in processing heterogeneous nuclear RNAs (hnRNAs) into mature mRNAs, in DNA repair and telomere regulation.

These RNA-binding proteins are among the most abundant proteins in the nucleus. The hnRNP proteins generally consist of multiple domains such as RNA recognition domain (RRM), and KH, K homology domain (Fig. 1.16); they are connected by a linker of varying length. Even though the RNA recognition domain is the most common it is not present in all hnRNPs. RRMs have two different structures (Hoffman *et al.*, 1991), which the most prevalent is characterized by a $\beta 1-\alpha 1-\beta 2-\beta 3-\alpha 2-\beta 4$ structure and two RNP consensus sequences called RNP-1 and RNP-2 useful for the interaction with RNA (Birney *et al.*, 1993). The hnRNPs bind to pre-mRNA with sequence specificity to high-affinity binding sites consisting of sequences that code for 5'- and 3'splice sites, polyadenylation elements and the polypyrimidine tract (PPT) (Ghetti *et al.*, 1992).

It is known that there is a correlation between carcinogenesis and some of these proteins such as hnRNP A2/B1 and hnRNP A1 in lung cancer (Tockman *et al.*, 1997; Tockman *et al.*, 1988;

Zerbe *et al.*, 2004; Zhou *et al.*, 1996) hnRNP K in esophageal cancer (Hatakeyama *et al.*, 2006) and in oral squamous cell carcinoma (Roychoudhury & Chaudhuri, 2007).

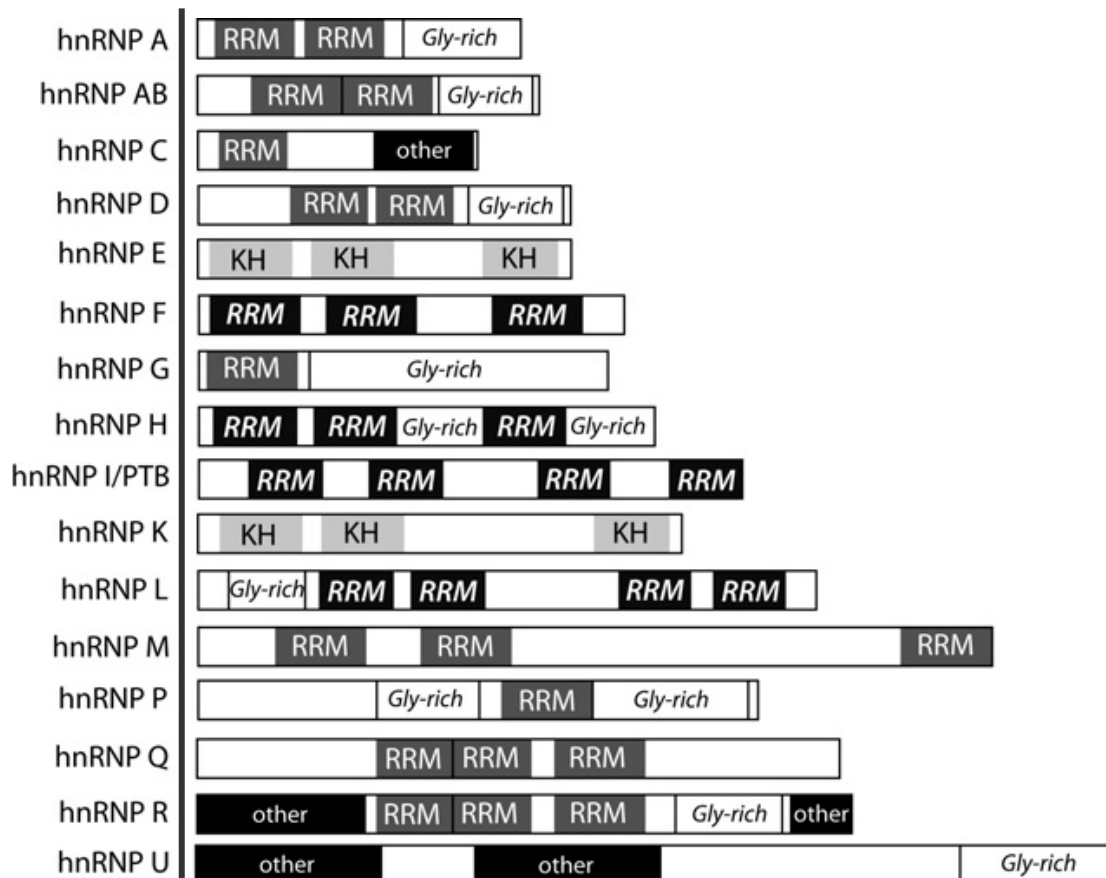


Figure 1.16 Schematic representation of the hnRNP family structure

RRM, RNA recognition domain, KH, K homology domain; RNA binding domain structurally distinct from RRM. (Adapted from Han *et al.*, 2010).

(i) hnRNP A2/B1

Heterogeneous nuclear ribonucleoprotein (hnRNP) A2/B1 belongs to the hnRNPs A/B family that comprises RNA-binding proteins important for alternative splicing. They are also involved in many stages of RNA metabolism and play crucial roles in RNA processing, mRNA trafficking, and telomere maintenance (Krecic & Swanson, 1999; Dreyfuss *et al.*, 2002; Smith, 2004; He & Smith, 2009).

The hnRNP A/B paralogs are structurally similar, but differ especially in their alternative splicing patterns. As shown in figure 1.17 hnRNP A2/B1 is alternatively spliced to produce the transcripts B1, A2, A2b, and B1b (Han *et al.*, 2010). The A2/B1 isoforms have different intracellular localization patterns and expression levels vary between tissue types and across developmental and cell cycle stages (Kamma *et al.*, 2001; Kamma *et al.*, 1999; Matsui *et al.*, 2000; Hatfield *et al.*, 2002; Maggipinto *et al.*, 2004; He *et al.*, 2005). The hnRNP A2/B1 group have been associated with the binding and elongation of single stranded telomere repeats TTAGGG_n (McKay & Cooke, 1992).

It has been shown that the splicing factor hnRNP A2/B1 is overexpressed in glioblastomas and is correlated with poor prognosis (Golan-Gerstl *et al.*, 2011). There have been also some speculations about the role of hnRNP A2/B1 as a presumed proto-oncogene due to the fact that overexpression of hnRNP A2/B1 in immortal cells led to malignant transformation and knockdown of hnRNP A2/B1 in glioblastoma cells inhibited tumor formation in mice (Golan-Gerstl *et al.*, 2011).

Different studies have shown that hnRNP A2/B1 is overexpressed in breast and lung tumors (Tockman *et al.*, 1997; Tockman *et al.*, 1988; Zhou *et al.*, 1996) whereas the hnRNP A2/B1

proteins showed decreased expression or no change in tumors of the colon, thyroid, small intestine and kidney (Karni *et al.*, 2007).

It has been demonstrated that hnRNP A2/B1 levels are increased by the small molecule TPA in leukemia cell lines (Zheng *et al.*, 2002), whereas in human neuroblastoma cell lines its expression level is down-regulated by retinoic acid (Liang *et al.*, 2011).

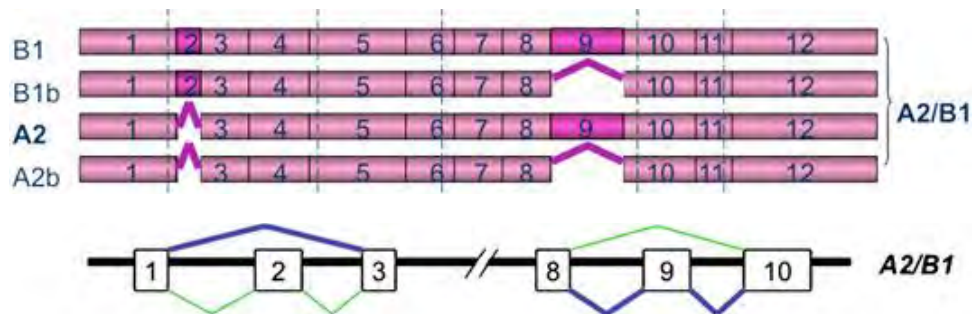


Figure 1.17 Structure of hnRNPs A2/B1 and transcript in the mouse

Transcripts B1, A2, A2b, and B1b produced by alternatively spliced are shown. Each numbered rectangle represents an exon, with lengths adjusted to align homologous exons. Vertical dashed lines highlight the structural similarity in terms of exon lengths, the correspondence of exons to functional protein domains, and the location of alternative exons relative to the entire mRNA. Below are shown the different splicing of alternative exons can occur in hnRNPs A/B in mouse. Exons are represented with boxes and introns with black horizontal lines. Blue lines represent major splicing patterns and green lines represent minor splicing patterns. Note that exon 9 in A2/B1 is generally included whereas exon 2 is excluded. (Adapted from Han *et al.*, 2010).

(ii) PTB

Polypyrimidine tract-binding protein (PTB), called also hnRNP I, is a 58 kDa protein. PTB is a RNA-binding protein and has a regulatory role in alternative splicing. When PTB binds to the RNA it blocks the binding of E2AF to the polypyrimidine tract (Spellman & Smith, 2006). It binds both intronic splicing silencers (ISS) and exonic splicing silencers (ESS) and in particular the ESSs have multiple PTB binding sites regulating the 3' splice site (Garcia-Blanco *et al.*, 1989). The UCUU in a pyrimidine-rich context is an optimal binding site for PTB (Patton *et al.*, 1991).

As it can be seen in figure 1.16 the structure has four weakly conserved RRM s (Oh *et al.*, 1998), of which the most prevalent is characterized by a $\beta\alpha\beta\beta\alpha\beta$ structure and an N-terminal domain involved in nuclear localization and export. It contains also a hinge region between RRM 2 and 3 of different length depending on the alternative splicing of the PTB transcript itself. The RRM s have been associated with the ability of PTB to oligomerize, bind to RNA, and possibly interact with other proteins.

Recent studies showed that PTB is highly upregulated in cervical cancer cells, whereas in the superficial levels of the cervical epithelium is down regulated (Fay *et al.*, 2009). It has been shown that PTB induces HPV-16 late gene expression in subgenomic HPV-16 expression plasmids acting on the inhibitory elements located immediately upstream of SD3632 (Somberg *et al.*, 2008).

1.13 E4orf4

Adenoviral E4orf4 (AdE4orf4) is a protein encoded by adenoviruses which are double stranded DNA viruses that, like HPV, has a life cycle that is characterized by an early and a late phase. E4orf4 protein of human adenovirus plays important roles in the adenoviral life cycle, including the switch from early to late gene expression (Akusjärvi, 2008). It is a 14 kDa protein which it has been shown to induce p53-independent apoptosis, promoting cell death and viral spread. When expressed alone, E4orf4 kills transformed cells but not normal human cells (Roopchand *et al.*, 2001). At the molecular level it has been shown to interact with the cellular protein phosphatase P2 (PP2A), thereby dephosphorylating SR-proteins (Kanopka *et al.*, 1998; Kanopka *et al.*, 1996). In subgenomic HPV-16 expression plasmids, overexpression of E4orf4 protein has been shown to overcome the negative effect of the splicing silencer adjacent to either splice sites, SD3632 and SA5639, inducing late gene expression (Somberg *et al.*, 2009).

1.14 Transient and stable transfection

The artificial uptake of new nucleic acid by mammalian cells, called transfection, was first described in 1965 by Vaheri and Pagano and had a substantial impact on biomedical research. Since then transfections have been widely used in laboratories to analyze gene expression and different methodologies have been developed to enhance the efficiency of delivered nucleic acids directly into the cytoplasm. These include calcium-phosphate transfection technique (Graham & van der Eb, 1973), electroporation (Andreason & Evans, 1988), cationic lipid-mediated transfection (Felgner *et al.*, 1993) and other reagents composed of cellular protein and polyamine. In particular one reagent, based on a nontoxic cellular protein and a small amount of a novel polyamine, is GeneJuice[®] Transfection reagent, which ensures maximal transfection efficiency and minimal cytotoxicity in both stable and transient transfection of eukaryotic cells.

The passage of foreign DNA into the nucleus is primarily mediated by cellular processes (Colosimo *et al.*, 2000; Salman *et al.*, 2001) and once into the nucleus part of the exogenous DNA can be degraded or lost through successive cell divisions (Liang & Jasin, 1996).

Transfected nucleic acids can either be transiently expressed or permanently expressed; unlike transient transfection, in which introduced DNA persists in cells for few days, stable transfection guarantees long term gene expression due to integration of exogenous DNA into the cellular genome and the inheritance of the cells progeny. Foreign DNA integrates in the chromosomal DNA randomly at one or very few sites (Murnane *et al.*, 1990) and site integration influences the transcription rate of the gene of interest (Wurm, 2004). Cellular DNA in the nucleus is combined with proteins to prevent damage, to control gene expression, DNA replication and to reduce the volume to fit in the nucleus. This compact structure is called chromatin and is found in two forms: euchromatin and heterochromatin (Elgin, 1996). Euchromatin is the active form of

chromatin. It is lightly packed allowing RNA polymerase and gene regulatory proteins to bind to DNA sequence therefore delivering gene transcription. The heterochromatin is the compact form that, with his tight structure of packed DNA, impedes the binding of protein factors and RNA polymerase complexes to bind the DNA inhibiting gene expression. Therefore integration of exogenous DNA into active euchromatin generally leads to transgene expression, whereas integration into inactive heterochromatin results in very little or no transgene expression.

Even though integration of foreign DNA into the cellular genome generally does not cause damage to its integrity (McFarlane & Wilson, 1996), it can sporadically have deleterious consequences for the endogenous DNA, leading to different rearrangement at the site of integration (Covarrubias *et al.*, 1987; Kato *et al.*, 1986; Robins *et al.*, 1981; Wilkie & Palmiter, 1987), such as deletions, duplications and translocations (Hamada *et al.*, 1993; Mahon *et al.*, 1988; Mark *et al.*, 1992) or can interfere with coding sequences (Rijkers *et al.*, 1994). Therefore, alteration of chromatin structure and functions leads to gene inactivation and genomic instability. However, it seems that in the endogenous genome there are “integration hotspots”, where DNA integration is more likely to occur (Rijkers *et al.*, 1994).

The development of genome manipulation and the establishment of stable cell lines is one main achievements of molecular biology which allows the analysis of large scale protein production (Wurm, 2004), gene function and regulation (Grimm, 2004), gene therapy and drug discovery (Glover *et al.*, 2005).

1.15 Small molecules used in this study

(i) TPA

TPA (12-O-tetradecanoylphorbol-13-acetate), also called PMA (phorbol-12-myristate-13-acetate), is a small molecule drug, the most commonly used phorbol ester. It was first found in the *Croton tiglium*, commonly called croton, a shrub of the *Euphorbiaceae* family native to Southeast Asia. It is extensively used as a tumor promoter in cancer research (Furstenberger *et al.*, 1981; Scribner & Boutwell, 1972). It binds and activates protein kinase C (PKC), causing numerous effects in cells and tissues (Blumberg, 1988; Blumberg, 1980; Castagnag *et al.*, 1982; Niedel *et al.*, 1983; Rebois & Patel, 1985) although at maximal or submaximal concentration elicits down-regulation of PKC in response to continuous stimulation (Blobe *et al.*, 1996; Liu, 1996). The activation of protein kinase C by TPA is due to its resemblance with diacylglycerol.

Weinberg has shown that in the presence of TPA human leukemia cell lines HL-60 differentiate to nondividing macrophage-like cells. These differentiated cells are cytotoxic for tumor cells *in vitro*, including untreated HL-60 cells (Weinberg, 1981). In 1998 TPA was reported as a therapeutic drug for patients with leukemia (Han *et al.*, 1998; Strair *et al.*, 2002). Intravenous administration of TPA alone or in combination with other drugs in patients with myelocytic leukemia has been demonstrated to reduce the number of myeloblasts in blood and bone marrow (Han *et al.*, 1998). It has been reported that this small molecule can regulate Bcl-x alternative splicing in glioma cells but not in leukemia cells (Li *et al.*, 2004). Bcl-x, a Bcl-2-related gene, is an apoptosis regulator that can function as pro- or anti-apoptotic, depending on the rate of alternative splicing in two distinct Bcl-x mRNAs: the Bcl-xS isoform, which promotes apoptosis, or the Bcl-xL isoform, which is a potent inhibitor of cell death (Boise *et al.*, 1993). BclxS

induces apoptosis in caspase- and BH3-dependent manner (Chang *et al.*, 1999; Lindenboim *et al.*, 2000). In addition, it has been shown that Bcl-xS induces apoptosis via Bak activation and induction of apoptosis is merely dependent on Bak but not on Bax (Lindenboim *et al.*, 2005). Different pathways have been reported for the regulation of apoptosis by Bcl-xL suggesting that there may be involved multiple mechanisms. Recently, it has been demonstrated that Bcl-xL inhibits apoptosis by binding to CED-4-like molecules such as Apaf-1 preventing the activation of downstream caspases (Hu *et al.*, 1998; Pan *et al.*, 1998), or by the heterodimerization with BH3-containing death agonists such as Bax and the ability to form a sustained ion channel (Minn *et al.*, 1999). Bcl-xL is over-expressed in several types of cancers and promotes resistance to chemotherapy in tumors *in vivo* (Liu *et al.*, 1999).

Interestingly, DNA microarray analysis of TPA treated cells have identified several genes, whose expression is altered by this substance; in particular TPA has shown to induce expression of a number of genes that are involved in splicing, such as SC35 (serine/arginine-rich splicing factor 2) and hnRNP A2/B1 (Zheng *et al.*, 2002).

In 1992 Meyers *et al.* demonstrated that TPA induces HPV-31 late gene expression and that production of the L1 proteins occurs simultaneously with the induction of keratin 10 and filaggrin synthesis (Meyers *et al.*, 1992).

(ii) Emetine

Emetine is the principal alkaloid of ipecac root (Wiegrebe *et al.*, 1984). Initially emetine was used as an expectorant and emetic (Grollman, 1966).

In 1968 Grollman demonstrated that emetine inhibits both ribosomal and mitochondrial protein synthesis and interferes with the synthesis and activities of DNA and RNA (Grollman, 1968). This inhibition is caused by the binding of emetine to the 40S subunit of the ribosome (Jimenez *et al.*, 1977). It is also known that it can inhibit viral RNA synthesis in poliovirus-infected HeLa cells (Grollman, 1968) and selectively inhibits mitochondrial protein synthesis in mouse liver (Lietman, 1970).

Treatment with emetine has been shown to be effective in the therapy of leukemia (Jondorf *et al.*, 1970) and to induce apoptosis in cancer cells (Boon-Unge *et al.*, 2007). In fact, it has been shown that emetine modulates the regulation of the apoptotic factors Bcl-x, downregulating Bcl-xL and upregulating Bcl-xS (Boon-Unge *et al.*, 2007). Moreover, it is known that treatment of Jurkat cells with emetine up-regulates a number of genes such as CASP8 (caspase 8), CASP9 (caspase 9), BAK1, DAXX (death-associated protein 6), GZMB (granzyme B) AKT1, MST1, TNFRSF6, TNFRSF11B, and TNFSF13 and downregulates BCL2, EGFR (epidermal growth factor receptor) and TNF (tumor necrosis factor) (Moller *et al.*, 2007).

(iii) Cycloheximide

Cycloheximide is a glutarimide antibiotic which was isolated from the bacterium *Streptomyces griseus*. It was first discovered in 1947 by Leach *et al.*, and it has been defined as an inhibitor of protein biosynthesis in eukaryotic organisms. Grollman has shown that cycloheximide and emetine share certain structural properties around two nitrogen atoms that are essential for their activity and they have a similar site and mode of action for inhibition of protein synthesis

(Grollman, 1966) and just like emetine, cycloheximide regulates alternative splicing of exon 2 in the Bcl-x gene (Boon-Unge *et al.*, 2007).

Due to its heavy toxic side effects, including DNA damage, teratogenesis, and other reproductive effects it is not suitable for human use as a therapeutic compound.

(iv) Ceramide

Ceramides (sphingolipids) are composed of sphingosine and a fatty acid and were first described in 1874 by Thudichum (Thudichum, 1874). Ceramides are found in high concentrations within the cell membrane and they were believed to be merely structural elements but over the years it has been discovered that this compound plays many others important functions in the cell. It has been proposed as a novel lipid second messenger with specific roles in mediating antiproliferative responses including apoptosis and cell cycle arrest. In fact, numerous reports have demonstrated that Bcl-x alternative splicing can be regulated by ceramide in cancer cells (Chalfant *et al.*, 2002; Massiello *et al.*, 2006; Massiello *et al.*, 2004). It has been shown that endogenous ceramide generated through the action of the CoA-dependent ceramide synthase is implicated in regulating the alternative splicing of caspase 9 but not Bax or caspase 2 pre-mRNAs (Chalfant *et al.*, 2002). Other studies reported that ceramide dephosphorylates a certain number of SR-proteins such as SRp70, SRp55, SRp40, and SRp30 in leukemia cell lines (Chalphant *et al.*, 2001).

(v) Amiloride

Amiloride belongs to a class of drugs known as potassium-sparing diuretics (Bull & Laragh, 1968). First used in 1967 for hypertension and congestive heart failure, it is now considered as a novel drug in cancer therapy. In fact it has been shown that amiloride modulates the alternative splicing of a number of cancer genes, such as Bcl-x, HIPK3, and BCR/ABL, affecting the phosphorylation state of serine-arginine-rich (SR) proteins (Chang *et al.*, 2011). Chang *et al.* reported that treatment of cells with amiloride resulted in hypo-phosphorylation of splicing factor ASF/SF2 and levels of SRp20 and two other un-identified SR proteins were decreased (Chang *et al.*, 2011). It is also known that amiloride decreases phosphorylation of AKT, ERK1/2 and PP1, increases phosphorylation of p38 and JNK (Chang *et al.*, 2011), promotes TRAIL-induced apoptotic death through Akt dephosphorylation and activates caspase-3 and -8 in HeLa and LNCaP cells lines (Cho *et al.*, 2005).

It has been demonstrated that prolonged injections of amiloride in rats significantly decrease the incidence of gastric cancers (Tatsuta *et al.*, 1993).

(vi) Scriptaid

Scriptaid belongs to a well-known class of hydroxamic acid-containing histone deacetylase (HDAC). They cause accumulation of hyperacetylated nucleosome core histones in most regions of chromatin affecting the expression of a small subset of genes, leading to transcriptional activation of some genes and repression of others, modulating the regulation of gene transcription. HDAC inhibitors have become a new class of drugs for treatment of a variety of

cancers (Marks & Jiang, 2005). Scriptaid has many similarities with trichostatin A but it has lower toxicity.

Treatment of breast cancer cell lines with scriptaid it has been shown to inhibit cell growth and to increase estrogen receptor α mRNA transcript (Clancy *et al.*, 2003). Scriptaid has been also used *in vivo* on a mouse model bearing MDA-MB-231 tumors showing an inhibition of tumor growth following scriptaid treatment (Clancy *et al.*, 2003).

(vii) Valproic acid

VPA is a histone deacetylase inhibitor. In 1882 Beverly S. Burton was the first to synthesize valproic acid as an analogue of valeric acid, found naturally in valerian and initially was used as an organic solvent (Burton, 1882). It is widely used to treat bipolar disorder, epilepsy, schizophrenia, migraine headaches and also major depression but recent studies have revealed that valproic acid can be used for treatment of HIV (Jennings & Romanelli, 1999; Moog *et al.*, 1996; Albrecht, 2005; Lehrman *et al.*, 2005) and various cancers such as multiple myeloma (Schwartz *et al.*, 2007), glioma (Admirant *et al.*, 2006) melanoma (Valentini *et al.*, 2007) and breast cancer (Munster *et al.*, 2007).

Valproic acid has been demonstrated to affect the expression of a number of splicing factors such as ASF/SF2 and hnRNPA1: in particular it seems to increase the expression of ASF/SF2 and decrease hnRNPA1 levels (Kusuma Harahap *et al.*, 2011). Britcha also demonstrated that VPA up-regulates the SR-proteins ASF/SF2, SRp20 and Htra2-b1 which in particular overexpression of Htra2- β 1 but not ASF/SF2 or SRp20 have been shown to lead to restoration of the splicing pattern of SMN2 transcripts (Britcha *et al.*, 2003; Hofmann *et al.*, 2000; Hofmann & Wirth,

2002). Alternative splicing and a consequent number of SMN2 copies are primarily related to the severity of a common neuromuscular disorder, the spinal muscular atrophy (SMA) disease (Burghes, 1997; Feldkotter *et al.*, 2002), suggesting that VPA is a candidate drug for SMA.

Since January 2011 the salt magnesium valproate is in phase III trials for cervical cancer and ovarian cancer. It has been shown that valproic acid induces apoptosis in HeLa cervical cancer cells by inhibition of gene expression of Akt1 and Akt2 mediated by the caspase-dependent pathways (Chen *et al.*, 2006).

(viii) Tannic acid

Tannic acid is a commercial form of tannins and was first recorded in 1836. Tannins are plant-derived polyphenols and are divided into two groups of hydrolyzable and condensed tannins (proanthocyanidins) (Hagerman, 2002). Hydrolyzable tannins are derivatives of gallic acid (3,4,5-trihydroxyl benzoic acid) in which a variable number of gallic acids are esterified to a core phenol. The simplest hydrolyzable tannins are gallotannins that are polygalloyl esters of glucose. The prototypical gallotannin is tannic acid or 1,2,3,4,6-penta-O-b-D-glucose. Tannins are rich in nuts, red wine, tea and coffee, but hydrolyzable tannins including tannic acid are not rich in tea (Hamilton-Miller, 1995; Yam *et al.*, 1998) or red wine (Cozzolino *et al.*, 2008). Nevertheless, red wine can extract phenolic compounds like tannic acid from aging in oak wine barrels. In fact, tannic acid is found in the bark and wood of oaks and other plants such as Tara pods (*Caesalpinia spinosa*), gallnuts from *Rhus semialata* or *Quercus infectoria* or Sicilian Sumac leaves (*Rhus coriaria*). It is light brown to yellow in colour powder and is highly soluble in water. It is used also as a process aid in beer and wine clarification and as a colour stabilizer.

According to Cornell University Department of Animal Science, the primary function of tannic acid is to bind and precipitate proteins. It has been shown that tannic acid also acts as a potent inhibitor of phorbol ester-induced nitric oxide generation in rat hepatocytes (Srivastava *et al.*, 2000) and stimulates glucose transport (Liu *et al.*, 2005). In 2009 Bian reported that tannic acid increases the levels of polypyrimidine tract binding (PTB) protein by activating its promoter region (Bian *et al.*, 2009).

1.16 RNAi pathway

Expression of a specific gene can be reduced and silenced by the expression of short-hairpin RNA (shRNA), which is a short sequence of RNA that makes a tight hairpin turn. This process is called RNA interference (RNAi). First evidence of RNAi system was reported by Napoli *et al.* in 1990. The original purpose was to introduce a chimeric petunia CHS gene to overexpress chalcone synthase (CHS) in pigmented petunia petals but instead they recorded the loss of expression of a pigmentation gene (Napoli *et al.*, 1990). Only few years later RNAi was first used in the nematode *Caenorhabditis elegans* proving that, with the introduction of RNA into cells, was possible to regulate gene expression (Fire *et al.*, 1998). Nowadays this innovative technology of introduction of viral vectors into cells is widely applied *in vitro* and through the years is developing and advancing safety and specificity in order to be designated as a safe and plausible therapeutic tool (Blow, 2007; Giering *et al.*, 2008; Castanotto & Rossi, 2009; Zhanga *et al.*, 2008).

RNAi occurs naturally in plants, fungi and animals and can be induced by the introduction of "naked" shRNA or via a viral vector into the cell (Fig. 1.18 and 1.19) and transcription of shRNA after vector integration is performed by either RNA pol II or pol III promoter.

RNA silencing is initiated with the cleavage of double-stranded RNA (dsRNA) into ~20 short interfering RNAs (siRNAs) by an enzyme called Dicer in an ATP-dependent manner (Bernstein *et al.*, 2001; Hamilton & Baulcombe, 1999). The hairpin is cleaved from the stem following nuclear export in a Ran-GTPase dependent manner.

5' phosphorylated siRNA incorporates into the RNA-induced Silencing Complex (RISC) which unwind the siRNA allowing the binding to target mRNAs. siRNAs that lack a 5' phosphate are promptly phosphorylated by an endogenous kinase. The activated RISC complex terminates the

reaction with mRNA cleavage and degradation (Hammond *et al.*, 2000; Zamore *et al.*, 2000; Pham *et al.*, 2004) (Fig. 1.18).

Lentiviruses are commonly used to deliver siRNAs to cells in order to block the expression of a specific gene through the RNA interference (RNAi) pathway. They are used to introduce a gene product into *in vitro* systems or animal models to replace or knock out mutated genes that causes disease or are functioning improperly. This novel technology is useful for a better understanding of the effects of a given gene in a model system and represents one of the most promising discoveries in molecular biology (Fig. 1.19).

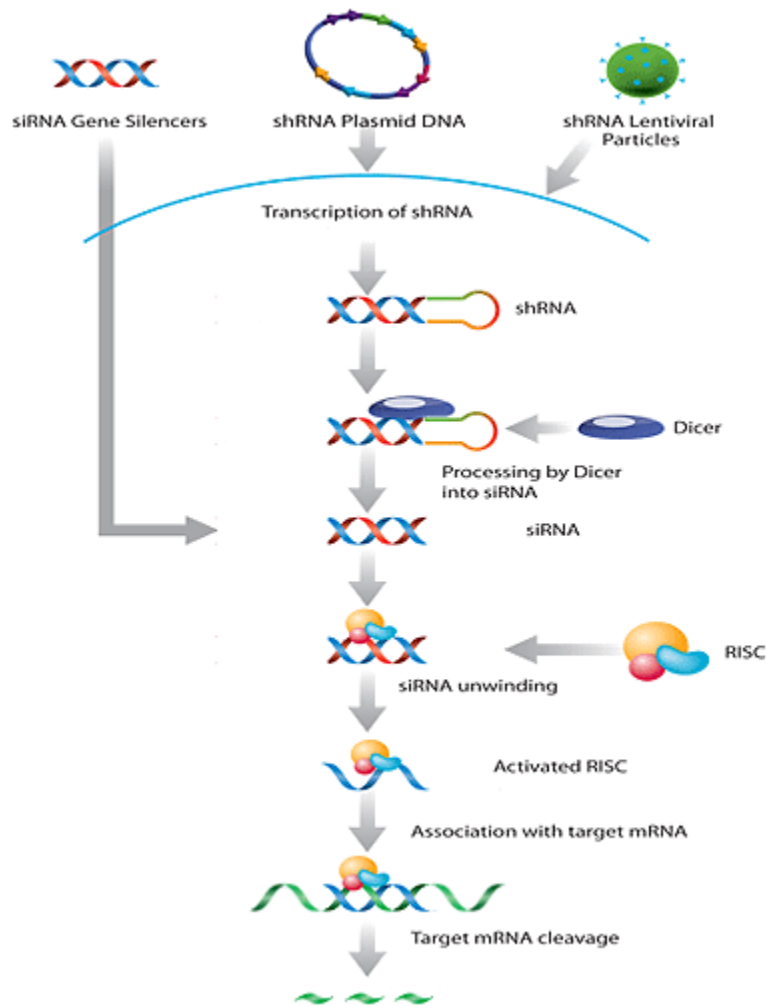


Figure 1.18 RNAi mechanisms

Three individual ways to knock-down target gene expression are shown: siRNA, shRNA Plasmid, shRNA Lentiviral Particle. siRNA refers to small interfering or short interfering RNA. shRNA Plasmid is a plasmid encoding short-hairpin RNA (shRNA). Both require transfection with a lipid-based transfection reagent for delivery into a cell and are useful for a transient knock-down. shRNA Lentiviral Particle are viruses carrying a shRNA that are delivered into the cells through transduction and stably integrate in the cellular genome. Useful for either transient or stable knock-down of a target gene. siRNA Dicer and complex RISC are shown in the figure.

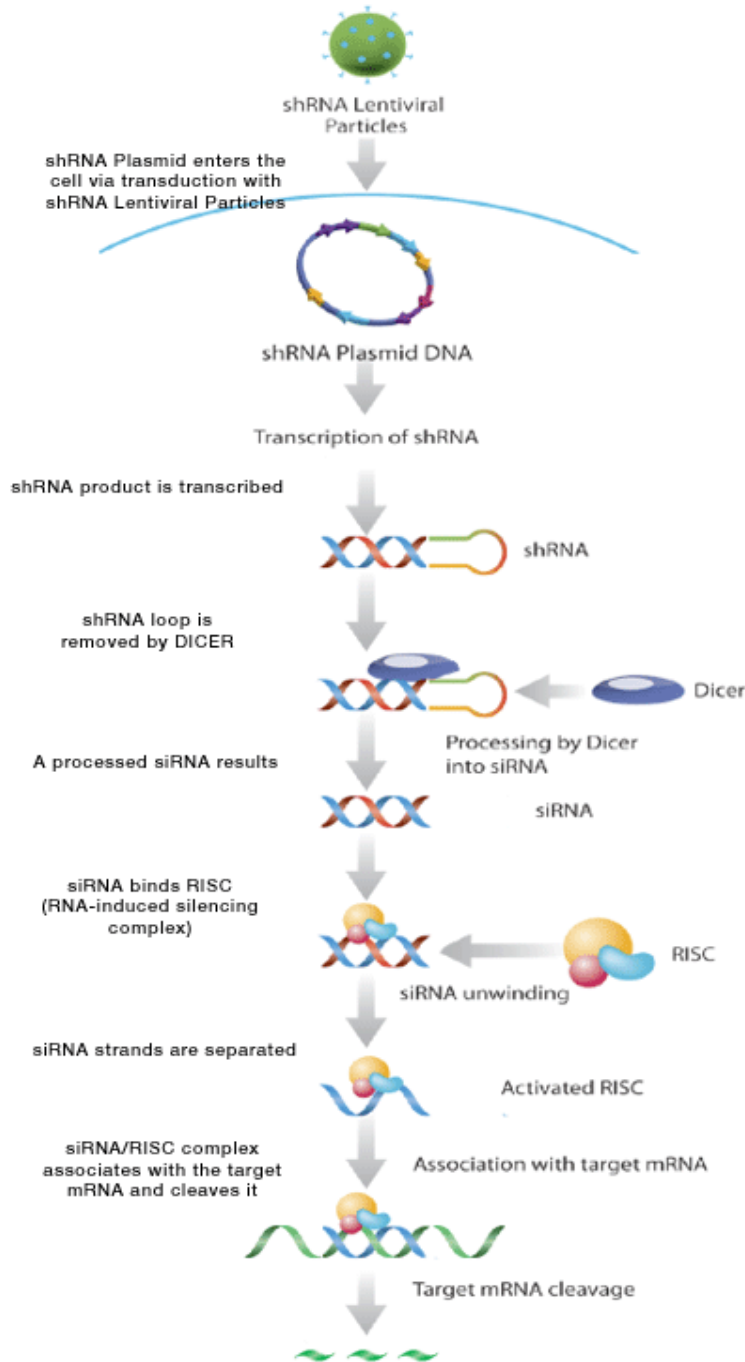


Figure 1.19 shRNA Lentiviral particle delivery

shRNA delivered into the cells through transduction of Lentiviruses and following integration in the cellular genome. Useful for either transient or stable knock-down of a target gene. Transcription of shRNA by the cellular machine. siRNA Dicer and complex RISC are shown in the figure. The association of the RISC complex with target mRNA and following cleavage is shown.

1.17 Immunohistochemistry

Immunohistochemistry is the process in which antigens or proteins are localized in tissue sections exploiting the usage of labeled antibodies that bind specifically to antigens in biological tissues. The interaction of the antibody-antigen is visualized by an enzyme such as peroxidase, that can catalyse a colour-producing reaction or alternatively different fluorescent dyes are available, as for example FITC, rhodamine or Texas Red. In 1941 Coons was the first to introduce a direct fluorescent method in order to localise substances in tissues (Taylor *et al.*, 2006). To improve detection sensitivity different methods were then developed such as the horseradish peroxidase, the peroxidase anti-peroxidase technique (Lee, 2010) and the use of the Avidin-Biotin complex (Hsu *et al.*, 1981) but the basic procedure has not changed since its invention. In immunohistochemistry preservation of antigen and lack of background staining are standard conditions required (Polack & Van Noorden, 1997).

Formalin or other aldehyde fixation can alter protein biochemistry and form protein cross-links which mask the antigenic sites in that tissue, causing faint or false negative staining. Two methods are used to unmask the antigens and epitopes in formalin-fixed and paraffin embedded tissue sections: enzymatic and heat-mediated (also known as heat-induced epitope retrieval, or HIER) methods: the first consists of the application of an enzyme solution, including trypsin, chymotrypsin, pronase and protease, directly to the tissue and the second of boiling the sections in solutions of heavy metal salts in microwave ovens or pressure cookers. Currently citrate buffer, EDTA buffer or Tris-EDTA Buffer are widely used for antigen retrieval enhancing staining intensity of antibodies.

1.17.1 Avidin-Biotin Conjugate (ABC) Method

Avidin-Biotin Conjugate (ABC) method is a standard IHC technique extensively used for immunohistochemical staining. Biotin, a low molecular weight vitamin, can be conjugated to the primary or the secondary antibody so that immunohistochemistry can be performed either as direct or an indirect method. In the indirect method the primary antibody binds first to the antigen and then the biotinylated secondary antibody reacts with the first antibody introducing a large mass of biotin into the section at the location of the primary antibody. The avidin-biotinylated enzyme complex (ABC) is then added and binds to the biotinylated secondary antibody. Avidin is a glycoprotein that can be labeled with peroxidase or fluorescein and has a great affinity for biotin. The peroxidase is then developed by 3,3-Diaminobenzidine (DAB) or other substrate to produce different colorimetric end products (Fig. 1.20).

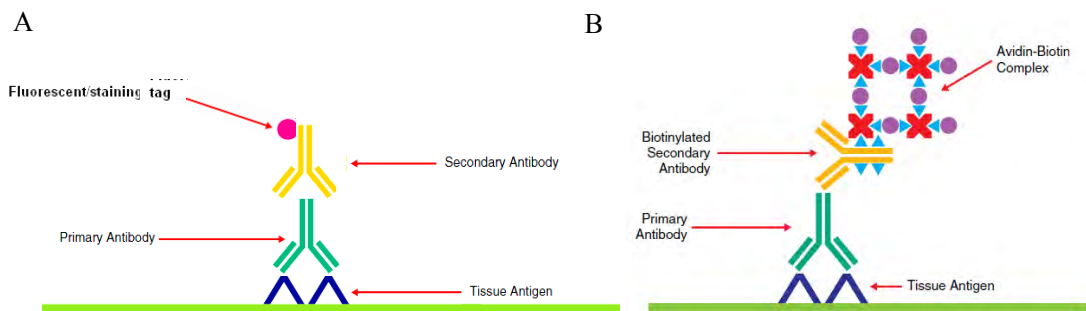


Figure 1.20 IHC Methods

(A) Indirect method of immunohistochemical staining; the first antibody against the antigen being probed for, and a second labelled antibody against the first. (B) Avidin-Biotin Conjugate (ABC) Method secondary antibodies are conjugated to biotin and function as links between tissue-bound primary antibodies and an avidin-biotin-peroxidase complex.

1.17.2 Tissue Microarray

Tissue microarrays allow the high throughput analysis of distinct tissue samples by having multiple tissue samples on a single slide. They consist of different tissue samples collected from several tissue blocks and rearranged into a singular paraffin block in an array pattern (Fig. 1.21). The use of tissue microarray slides has many advantages compared to individual tissue section slides: since the analysis happens on a single slide containing different samples the conditions of the treatment, such as incubation time, reagent concentration, washing conditions and other factors remain the same therefore yielding uniform analysis. Many samples from different patients can be analyzed simultaneously and using less reagents resulting in a wider and cost effective analysis.

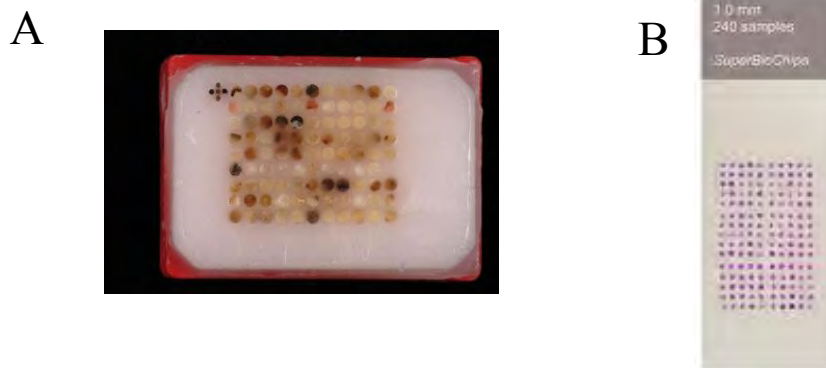


Figure 1.21 Representation of tissue microarrays

(A) Tissue microarray paraffin block (B) Tissue microarray slide.

1.17.3 p16INK4a

p16INK4a, a prototypic INK4 protein, has been identified as a tumor suppressor in many human cancers. It has been shown that the p16INK4a gene is altered by various mechanisms in a vast range of human cancers (Ruas & Peters, 1998) such as melanoma, gastric adenocarcinoma, Hodgkin and non-Hodgkin lymphomas (O'Neill & McCluggage, 2006). It is a member of the CDK inhibitor family of proteins and it has been shown to arrest the progression of the cell cycle and to be dramatically increased in cells undergoing senescence (Ohtani *et al.*, 2004). Under normal condition cell cycle initiation and progression is controlled through a complex mechanism involving many regulatory pathways within the cell, including several classes of cyclin-dependent kinases (CDKs) and the Retinoblastoma protein (pRB). To initiate cell cycle progression pRB is phosphorylated by a holoenzyme complex containing cyclin D and a cyclin-dependent kinase (CDK4 or CDK6). p16INK4a binds and induces an allosteric conformational change in CDK4/CDK6 inhibiting the binding of ATP to the CDK complex. This antagonizes activation of CDK and cyclin binding, therefore preserving pRB in its hypophosphorylated state leading to G1 cell cycle arrest (Fig. 1.22) (Sherr & Roberts, 1999). p16INK4a is a useful biomarker for the identification of cervical intraepithelial lesions because it is a surrogate measure of active HPV gene expression. Its over-expression is associated with the oncogenic transformation caused by persistent high-risk HPV infection (Benevolo *et al.*, 2008). In fact, under normal conditions pRB inhibits the transcription of p16INK4a but inactivation of pRB by HPV E7 results in the accumulation of p16INK4a.

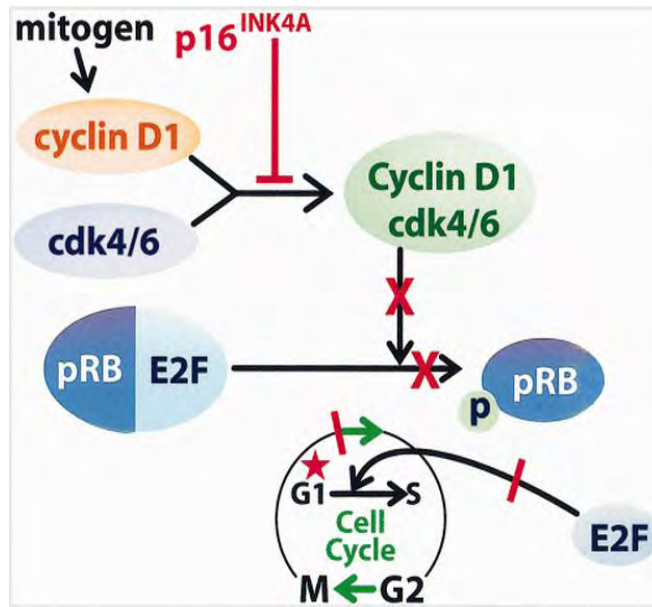


Figure 1.22 p16INK4a pathway

Mechanism of action of p16INK4a as a cell cycle inhibitor causing G1 arrest (from Natarajan *et al.*, 2005).

1.18 Aims

Overarching Aim

Elucidate the mechanism of HPV-16 gene regulation, primarily regulation of late gene expression

Individual Aims

- Generate reporter plasmids based on HPV-16 genome with the L1 gene replaced by an easily measurable reporter gene
- Generate stable cell lines with the reporter plasmids integrated in the cellular genome to ensure long term gene expression
- Identify small molecules that can induce late gene expression and that could potentially be used to treat persistent HPV infections
- Identify novel cellular factors involved in regulation of HPV-16 late gene expression.

2. Chapter 2 - Material and Methods

2.1 Generation of subgenomic HPV-16 reporter plasmids

2.1.1 Plasmid constructions

pBEL and pBELM, have been described previously (Fig. 2.1) (Zhao *et al.*, 2004). In order to construct the reporter plasmids PCR was first performed using the plasmids, annealing temperatures, primers, components and general PCR cycling conditions as outlined in Table 2.1, 2.2, 2.3 and 2.4. For example the poliovirus IRES sequence was PCR amplified from pKSPOLIO:S at 59°C annealing with oligonucleotides IRESs and IRESa. The IRES set primers have at each 5' end restriction sites useful for cloning into pBEL and pBELM and subsequently insertion of a reporter gene. The IRES sense has a restriction site for BamHI and the IRES antisense primer has a short polylinker with restriction sites for XhoI, MluI and HpaI restriction enzymes (Table 2.1). The PCR fragment was subcloned into pCR2.1-TOPO according to the manufacturers protocol (Invitrogen) and transformation into *E. coli* and analysis of recombinant was carried out. The IRES sequence was then transferred as a BamHI-XhoI fragment into pBEL and pBELM generating pBEL-IRES and pBELM-IRES (Fig. 2.2 A), respectively. The chloramphenicol acetyltransferase (CAT) sequence was first PCR amplified at 65°C annealing with primers CATs and CATa (Table 2.1, 2.2, 2.3 and 2.4). CATs and CATa primers have at each 5' end MluI and XhoI restriction sites, respectively (Table 2.1). The CAT PCR product was first cloned into pCR2.1-TOPO, the cloning reaction transformed and recombinant plasmids analyzed. The CAT sequence was excised from TOPO vector with MluI and XhoI and inserted downstream of the IRES sequence in pBEL-IRES and pBELM-IRES to generate pBEL-IRES-CAT and pBELM-IRES-CAT, renamed pBELCAT and pBELMCAT for simplification (Fig. 2.2

B). pBspliceCAT and pBspliceMCAT (Fig. 2.3 A) were generated by digestion of pBELCAT and pBELMCAT, respectively with BssHII and Sall followed by filling in of overhangs with T4 DNA polymerase and religation. To construct pMt1sdCAT (Fig. 2.3 A), a HindIII-ApaI fragment from plasmid pMt1sd (Somberg *et al.*, 2008) was inserted into pBELMCAT.

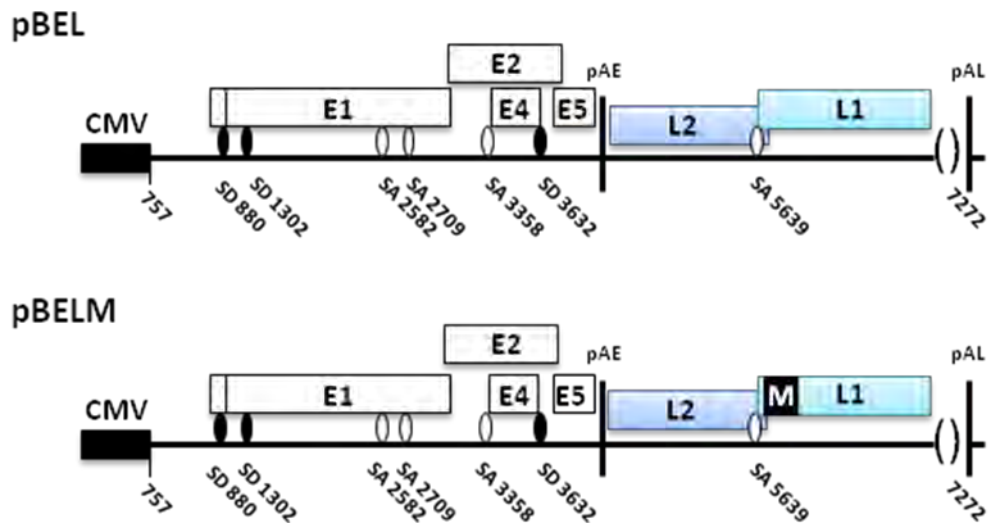


Figure 2.1 pBEL and pBELM plasmids

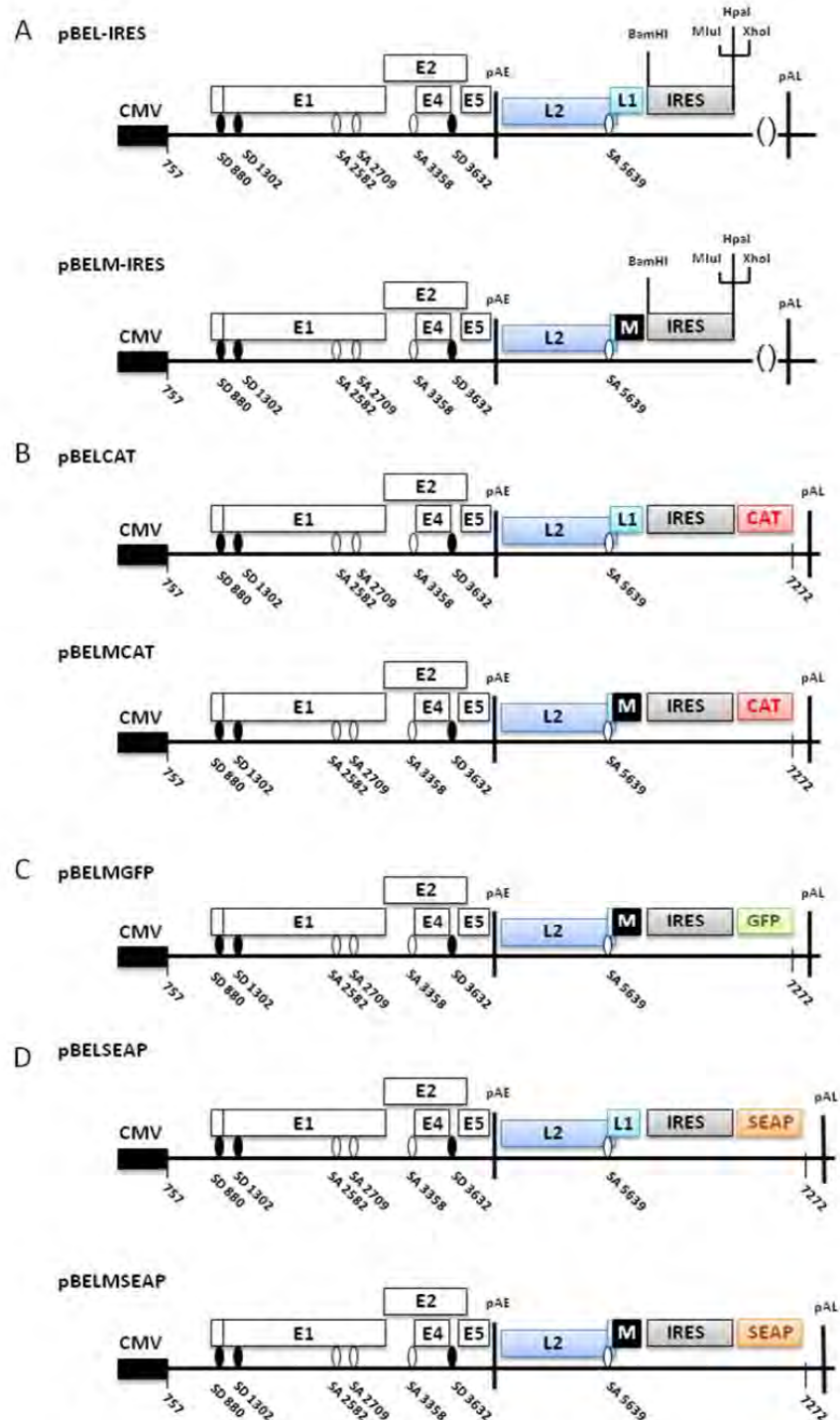


Figure 2.2 Representation of the reporter plasmids generated in this study

(A) pBEL-IRES, pBELM-IRES (B) pBELCAT, pBELMCAT, (C) pBELMGFP (D) pBELSEAP, pBELMSEAP.

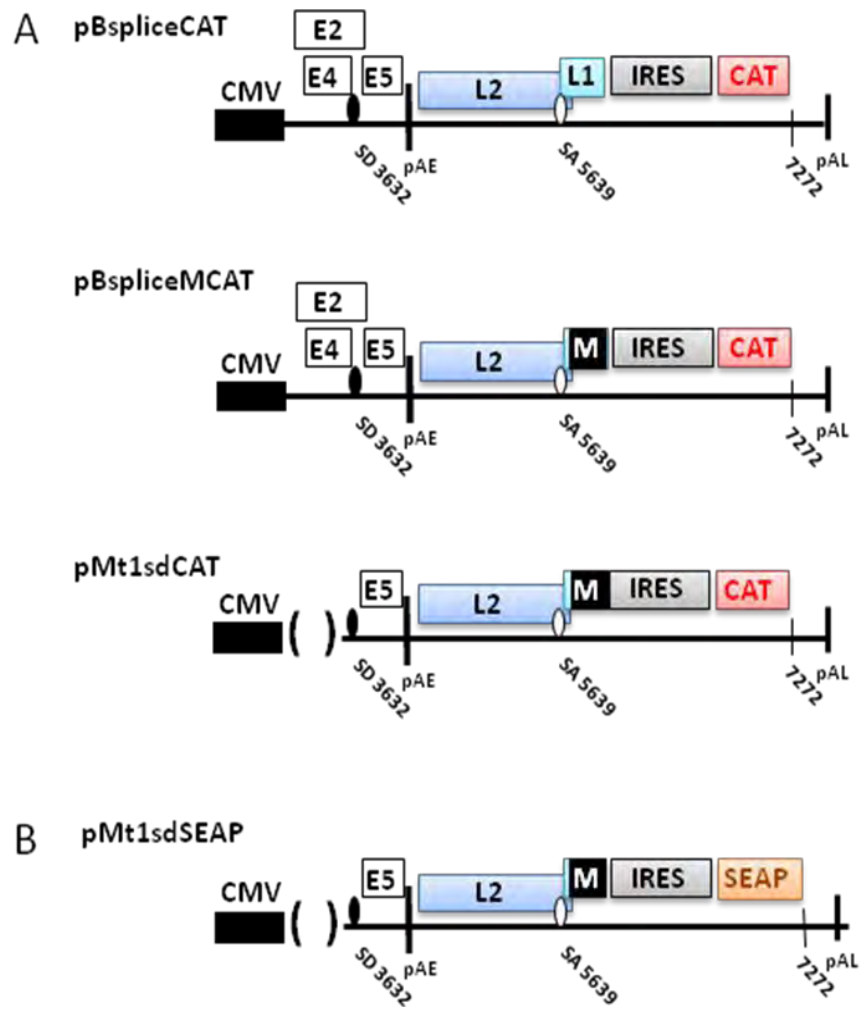


Figure 2.3 Representation of the reporter plasmids carrying only 2 splice sites SD3632 and SA5639

(A) pBspliceCAT, pBspliceMCAT, pMt1sdCAT (B) pMt1sdSEAP.

In order to construct pBELM-IRES-GFP (Fig. 2.2 C), pUC57/tat/LTR/GFP and pBELM-IRES-CAT plasmids were digested with MluI and XhoI restriction enzymes and religated. As a positive control for GFP the pdE1-sp1AhCMV EGFP plasmid was used (7923bp).

To generate pBEL-IRES-SEAP and pBELM-IRES-SEAP the SEAP reporter gene was first PCR amplified at 60.2°C using SEAPs and SEAPa primers. SEAPs and SEAPa primers have at each 5' end MluI and XhoI restriction sites, respectively (Table 2.1), useful for insertion into pBEL-IRES and pBELM-IRES. The PCR product was inserted into pCR2.1-TOPO (Invitrogen) and then transformation and analysis of recombinant plasmids was performed. The SEAP sequence from TOPO vector was digested with MluI and XhoI and inserted downstream the IRES sequence in pBEL-IRES and pBELM-IRES resulting in pBEL-IRES-SEAP and pBELM-IRES-SEAP, which hereafter will be called pBELSEAP and pBELMSEAP, respectively (Fig. 2.2 D). To generate pMt1sdSEAP (Fig. 2.3 B), pMt1sdCAT and pBELMSEAP were digested with HindIII and ApaI restriction enzymes and religated.

The LacZ gene was PCR amplified at 59°C using LacZ primers with the Expand long template PCR system kit according to the protocol of the manufacturer (Roche). The LacZ primers both have a restriction site for XhoI enzyme at the 5' end. The PCR product was cloned into pCR2.1-TOPO vector and then transformation and analysis of recombinant plasmids was performed.

Plasmids pCMVE4orf4 (Somberg *et al.*, 2009), pCMVPTB (Somberg *et al.*, 2008), pCMVSRp30c (Somberg *et al.*, 2011) and pASF/SF2 (Somberg & Schwartz, 2010) express E4orf4, PTB, SRp30c and ASF/SF2 respectively and have been described previously.

Plasmid	Primers	Annealing temp.	PCR product	Cloning enzymes
1/10000 pKSPOLIO:S	IRESs/IRESa	59°C	752 bp	BamHI, XhoI, MluI, HpaI
1/50000 pcSEAP	SEAPs/SEAPa	60.2°C	1530 bp	MluI, XhoI
1/2000 pCMVsport β-gal	LacZs/LacZa	59°C	3141 bp	XhoI, XhoI
1/2000 pCCKH1DXK	CATs/CATa	65°C	660 bp	MluI, XhoI

Table 2.1 Plasmids dilution, primers pairs, annealing temperature, PCR product and restriction enzymes used for the cloning steps.

Oligonucleotide	Sequence (5'-3')
IREs	GGGATCCTTAAAACAGCTCTGGGGTTG
IREs _a	CCTCGAGTTAACACGCGTAGGTAATTCCAATAGGTGTGAGTG
SEAPs	AACGCGTGAAATGCTGCTGCTGCTGCTGCTGCTGGGCC
SEAP _a	CCTCGAGTTAACCCGGGTGCGCGGCGTCGGTGGTGCCGG
LacZs	CCTCGAGATGTCGTTTACTTTGACCAAC
LacZ _a	CCTCGAGTTTTTTGACACCAGACCAACTG
CATs	GACGCGTACCATGAGTAAAGGAGAAGAAGACTTTTCACTGGA
CAT _a	CCTCGAGCTATTTGTATAGTTCATCCATGCC
GFPs	ATTCTCGTGGAAGTGGATGG
GFP _a	CACTGCACGCCATAAGAGAA

Table 2.2 Sequences of oligonucleotides used in this study for PCR

Solution	Concentration	1x
DNA	-	1µl
Primer mix	100ng/µl	1µl
dNTPs	10mM	0.5µl
PCR buffer (10mM Mg ⁺⁺)	5x	5µl
Taq polymerase	5u/µl	0.2µl
H ₂ O	-	17.3µl
Total		25µl

Table 2.3 PCR reaction mix

Temperature	Time	Cycles
95°C	5 min	1
95°C	30 sec	35
Annealing temperature required (Table 2.1)	30 sec	35
72°C	1 min	35
72°C	5 min	1

Table 2.4 PCR cycling condition

2.1.2 TOPO TA Cloning

2.5µl of fresh PCR product was mixed with 1µl of salt solution. 1µl of TOPO vector and 1.5µl of H₂O. The reaction was mixed gently and incubated for 5 min at room temperature.

2.1.3 Transformation into TOP10 Chemically competent *E. coli*

2.5µl of the TOPO[®]Cloning reaction was added into a vial of One Shot[®] Chemically Competent *E. coli* and mixed gently. After incubation on ice for 25 min the cells were heated-shocked for 30 sec at 42°C and then placed for 2 additional minutes on ice. 250µl of S.O.C. medium was added and the tube was let shake horizontally (225 rpm) at 37°C for 1 hour. 200µl from transformation was spread on a pre-warmed selective agar plate (Appendix I) containing 40µl of 40mg/ml X-gal in dimethylformamide (DMF) and then the plates were incubated at 37°C overnight.

2.1.4 Purification of PCR product and vector reactions

10-15µl of pBEL and the PCR product of interest inserted into TOPO vector were digested in the same reaction tube with 1µl of each restriction enzyme required, 2.5µl of Buffer and H₂O up to 25µl. After incubation for more than 3 hours at the required temperature a phenol/chloroform purification was performed. 25µl of H₂O was added into the tube and mixed. 50µl of phenol/chloroform/isoamyl alcohol was added and mixed for 5 sec. After centrifugation at 14000 rpm for 30 sec the aqueous phase was removed to a fresh tube and 50µl of chloroform was added and the mixture was vortexed. After an additional centrifugation at 14000 rpm for 30 sec the aqueous phase was removed to a fresh tube and 5µl of 5M NaCl and 120µl of 100% ethanol were added. The mixture was vortexed and incubated at -70°C for 10 min. After centrifugation at

14000 rpm for 15 min the supernatant was aspirated and 200µl of ice cold 70% ethanol was added. An additional centrifugation at 14000 rpm for 5 min was performed and the supernatant was aspirated and the pellet was air-dried completely. The pellet was then resuspended in 8µl of H₂O and ligation was performed.

2.1.5 Ligation Reaction

The ligation was performed using the Rapid DNA Dephos & Ligation Kit according to the protocol of the manufacturer (Roche). 2µl of DNA dilution Buffer was added into the tube containing 8µl of purified reaction and mixed thoroughly. Then 10µl of T4 DNA Ligation buffer and 1µl of T4 DNA ligase were added and mixed thoroughly. The reaction was incubated for 5 min at room temperature and 2.5µl was then transformed into One Shot[®] Chemically Competent *E. coli* as described above.

2.1.6 Analysis of recombinant plasmids

Single colonies were picked from the agar plate and inoculated into 3ml LB medium (Appendix I) containing 3 μ l of ampicillin and incubated overnight at 37°C in an orbital incubator. 1.5ml of the overnight culture was poured into tubes and centrifuged at 14000 rpm for 3 min. The supernatant was aspirated and the pellet resuspended in 100 μ l of solution I (Appendix I), then 200 μ l of solution II (Appendix I) was added and the mixture was incubated at room temperature for 5 min. 150 μ l of solution III (Appendix I) was added, vortexed and incubated on ice for 5 min, then centrifuged at maximum speed for 5 min. 400 μ l of the supernatant was transferred into a fresh tube containing 400 μ l of phenol/chloroform/isoamyl alcohol (Appendix I), and the mixture was vortexed and centrifuged for 30 sec. Then the aqueous (top) phase was transferred into a fresh tube and 400 μ l of chloroform was added. After vortexing and centrifugation for a further 30 sec the aqueous phase was transferred into a fresh tube containing 800 μ l of 100% ethanol (Merck). The mixture was then vortexed and incubated on ice for 5 min. The tubes were centrifuged at maximum speed for 5 min to pellet the DNA and the supernatant was aspirated. 1ml of cold 70% ethanol (Appendix I) was then added and centrifuged at maximum speed for 5 min. The supernatant was carefully aspirated and the pellet let to air-dry. 25 μ l of TE buffer (Appendix I) containing RNase A (22 μ g/ml) (Appendix I) were added.

2.1.7 Restriction of mini preparation plasmid DNA

In order to identify the recombinant plasmid of interest each mini preparation plasmid was subjected at restriction. 10 μ l of plasmid was mixed with 1.5 μ l Buffer solution, 0.25 μ l of each restriction enzyme (Table 2.1) and milliQ-water up to 15 μ l according to the protocol of the manufacturer (Roche). The mixture was incubated at the required temperature of each restriction

enzyme for at least 2 hours and then analyzed with electrophoresis on a 1% agarose gel with ethidium bromide staining (0.5µg/ml ethidium bromide dissolved in gel).

2.1.8 Isolation of high quality plasmid

Isolation of purified plasmid DNA from bacterial cultures was carried out using the High Pure Plasmid Isolation kit (Roche). The bacterial culture with the plasmid of interest was grown overnight at 37°C in 8ml LB medium containing 8µl of 100mg/ml ampicillin. 3ml of overnight culture were poured into a tube. After centrifugation at 6000 × g for 30 sec the supernatant was discarded and the pellet was resuspended with 250µl suspension buffer containing RNase and mixed well. 250µl of Lysis buffer was added and mixed gently, the mixture was then incubated for 5 min at room temperature. 350µl of chilled Binding buffer was added and mixed gently and incubated on ice for 5 min. The tube was then centrifuged for 10 min at 13000 x g. The supernatant was then transferred into a high pure filter tube, previously inserted into a collection tube, and it was then centrifuged for 1 min at 13000 x g. The flowthrough liquid was discarded and 500µl of Wash buffer I was added to the filter tube. After centrifugation at maximum speed for 1 min the flowthrough was discarded and 700µl of Wash buffer II were added to the filter tube followed by centrifugation for an additional minute at maximum speed. The filter tube was then inserted into an eppendorf tube. 100µl of Elution buffer was added and the tube was centrifuged 1 min at maximum speed. 10µl of purified plasmid DNA was first subjected at restriction (Table 2.1) and bands analyzed with electrophoresis on a 1% agarose gel with ethidium bromide staining and then sent for sequencing.

In order to prepare purified plasmid DNA in large quantities all plasmids were purified with the Genopure Plasmid Midi kit (Roche). The bacterial culture with the plasmid of interest was grown

overnight at 37°C in 100ml LB medium containing 100µl ampicillin, with vigorous shaking (approx. 200 rpm). To pellet the bacteria the culture medium was centrifuged for 10 min at 3000 × g at 4°C and then the supernatant was discarded. The pellet was let air-dry for few minutes and thereafter carefully resuspended in 8ml Suspension buffer containing RNase. 8ml Lysis buffer was added and mixed gently by inverting the tube and the mixture was then incubated 2-3 min at room temperature. 8ml chilled Neutralization buffer were added and mixed immediately by inverting gently until a homogenous mix and then incubated 5 min on ice. The bacteria lysate was first cleared by filtration and then loaded into a column and let empty by gravity flow. The flowthrough was loaded into the column once more and then it was discarded. The column was thereafter washed with 4ml of Wash buffer and let to flow through. This step was repeated 2 times more. The column was then inserted into a collection tube capable of withstanding high speed centrifugation and 2,5ml of pre-warmed Elution buffer (50°C) were added and let to empty by gravity flow. This step was repeated one more time and then 3,6ml of isopropanol was added into the tube to precipitate the eluted plasmid. The plasmid DNA was centrifuged for 30 min at 15000 × g at 4°C. The supernatant was then carefully discarded and the plasmid DNA was washed with 3ml of chilled 70% ethanol and centrifuged for an additional 10 min at 15000 × g at 4°C. The supernatant was then carefully removed and the tube was then let to air-dry for about 10 min. The DNA plasmid pellet was then re-dissolved in 50µl TE buffer and stored at -20°C.

2.2 Analysis of the reporter genes

2.2.1 Cell culture and transfections

HeLa cells were cultured in RPMI-1640 medium containing 10% fetal bovine calf serum, glutamine and penicillin-streptomycin. Transfections were carried out by using GeneJuice[®] Transfection reagent according to the protocol of the manufacturer (Novagen). In short, 2×10^6 HeLa cells were plated per well into 35mm plates the day before transfection and incubated at 37°C. 3µl of GeneJuice[®] Transfection reagent and 100µl of serum-free RPMI-1640 medium were mixed and incubated for 5 min. Then the required amount of plasmid DNA was added, and the mixture was incubated at room temperature for 10 min. Thereafter, the mixture was added drop-wise into each well at approximately 70% confluence, and incubated at 37°C for 24 hours.

2.2.2 CAT ELISA assay

Cells were harvested 24 hours post-transfection, lysed, and the CAT levels were determined using a CAT ELISA assay based on the sandwich ELISA principle according to the protocol of the manufacturer (Roche). Briefly, cells were lysed in 1ml lysis buffer and the cell extract containing the CAT enzyme was centrifuged at 13000 x g for 10 min at 4°C. 200µl of cell extract was added per well into a coated microplate and incubated for 1h at 37°C. The wells were rinsed 5 times with a washing buffer and an Anti-CAT-DIG working dilution was added and incubated for 1h at 37°C. After the washing steps an Anti-DIG-POD working dilution was added in each well. After 1 hour at 37°C the wells were washed 5 times and a POD substrate was added into

each well and incubated at room temperature. Absorption was measured at 405nm using a 96-well microplate ELISA reader (Labsystems Multiskan[®]Plus).

2.2.3 SEAP assay

The medium from the transfected cells was analyzed based on the SEAP Assay Protocol (Roden/Bossis). The medium was heat inactivated at 65°C for about 10-15 min to reduce background and 200µl was then poured into a 96-well microplate. 20µl of 0.05% CHAPS (Appendix I) in PBS was added to each well and then 40µl of substrate (Appendix I) was added and incubated at room temperature in darkness and measured at 405nm in a 96-well microplate ELISA reader (Labsystems Multiskan[®]Plus).

2.2.4 GFP assay

For the analysis of CMVGFP and pBELMGFP expression plasmids, a cover slip was placed into each well of a 35mm plate and then 2×10^6 cells were added. Transfection was carried out as previously described. After 24 hours the cells were washed twice with PBS and the cover slip was removed from the well and mounted on a slide. GFP fluorescence intensities were detected and quantitated with a fluorescence microscope from the living cells (CELL[^]F software, Olympus Soft Imaging solutions GmbH).

2.2.5 β -galactosidase assay

A second reporter vector, pCH110, was included in all transfection experiments and assayed so that the expression of the experimental reporter gene may be normalized. The normalizing cells

were transfected with each plasmid in a minimum of three independent experiments with similar results. β -Galactosidase was assayed by adding 150 μ l of 2x assay buffer (200mM sodium phosphate buffer, pH7.3, 2mM MgCl, 100mM β -ME, 1.33mg/ml ONPG) to 150 μ l of cell extract. The mixture was incubated at 37°C for 3 hours and the reaction was stopped adding 500 μ l of 1M sodium carbonate. The absorbance was monitored at 420nm in a spectrophotometer using plastic cuvettes.

2.3 Establishment of stable cell lines

2.3.1 Stable transfection and clones selection

2x10⁶ HeLa cells were plated per well into 35mm plates and incubated overnight at 37°C. 1.5 μ g of DNA reporter plasmid was co-transfected respectively with 0.25 μ g of pSVNEO into cells and 24 hours post-transfection cell were transferred into 100mm dishes. Stable clones were selected with 200 μ g/ml of G418 and after approximately 3 weeks individual clones were picked and plated in new culture dishes. For DNA extraction the cells were trypsinised, pelleted and resuspended in 200 μ l of Phosphate-buffered saline (PBS). The cells were lysed with 300 μ l of cell lysis solution (Appendix I). 100 μ l of a 10M ammonium acetate solution was then added to precipitate the proteins followed by vortexing for 15 sec. The mixture was centrifuged for 5 min at 13000 x g and 300 μ l of isopropanol was added to the supernatants followed by centrifugation for 5 min to pellet the DNA. The DNA pellets were washed with 300 μ l of 70% ethanol and resuspended in 25 μ l TE buffer (10mM Tris.Cl pH8.0, 1mM EDTA). The DNA was subjected to PCR amplification with CATs and CATa primers (Table 2.2) for CAT stable cell lines or GFPs and GFPa (Table 2.2) for pBELMGFP stable cell line. A further PCR amplification was

performed with IRESs and IRESa primers (Table 2.2) in both CAT and GFP positive clones to confirm the integration of the entire sequence. 15µl of each PCR product were analysed on a 1% agarose gel.

2.3.2 RNA extraction and reverse transcription (RT)-PCR

(i) RNA extraction

Total RNA was isolated using the High pure RNA isolation kit according to the protocol of the manufacturer (Roche). The cells were trypsinised and centrifuged at 1500 x rpm for 3 min. The pellet was washed with PBS and centrifuged for 5 additional minutes. The cells were resuspended with 200µl of PBS and 400µl of Lysis/-Binding Buffer was added. After vortexing for 15 sec the entire sample was transferred into the upper reservoir of the filter tube inserted in one collection tube. The tube was centrifuged for 15 sec at 8000 × g and the flowthrough liquid discarded. 500µl of Wash Buffer I was added to the upper reservoir of the filter tube and centrifuged for 15 sec at 8000 × g.

After discarding the flowthrough 500µl of Wash Buffer II was added. After a centrifuge for 15 sec at 8000 × g 200µl of Wash Buffer II was added and the tube centrifuged at maximum speed for 2 additional minutes. The RNA was then eluted with 80µl of Elution Buffer and the tube centrifuge for 1 min at 8000 × g.

(ii) RT-PCR

Total RNA was then reverse transcribed at 42°C by using Superscript II and gene specific primers (Table 2.5) according to the manufacturer's instructions (Invitrogen). In short, 2µg of total RNA was added into a RNase free tube. Then 1µl (100ng) of gene-specific primer (GSP), 1µl of dNTP Mix (10mM each) and RNase free water up to 10µl were added. The mixture was heated at 65°C for 5 min and quick chilled on ice. The content of the tube was collected by brief centrifugation and a reaction mix of 4µl of 25mM MgCl₂, 2µl of 5X First-Strand Buffer, 2µl of 0.1M DTT and 1µl of RNaseOUT™ (40 units/µl) was added into the tube and mixed gently. The tube was incubated at 42°C for 2 min and then 1µl of SuperScript™ II RT was added and mixed by pipetting gently up and down. The tube was then incubated at 42°C for 50 min and the reaction was inactivated by heating at 70°C for 15 min. From the synthesised cDNA 2µl was taken and PCR-amplified in a 25µl reaction volume using oligonucleotides listed in Table 2.1. An annealing temperature of 55°C was used to PCR amplify cDNAs with 757s and E4a primers, whereas an annealing temperature of 57°C was used with 757s and L1a primers or 757s and L1Ma primers. cDNA synthesized with oligonucleotide CATa was PCR amplified at 68°C with primers 757s and CATa (Table 2.5).

Oligonucleotide	Sequence (5'-3')
CATa	CCTCGAGCTATTTGTATAGTTCATCCATGCC
757s	GTCGACGGTATCGATCGGTTGTGCGTACAAAGCACACAG
E4a	CCGCGCGCTGCCTAATAATTCAGGAGAGG
L1a	CCGTGCTTACAACCTTAGATACTGGGACAG
L1Ma	CGCTGGGCAGCCACAGGC

Table 2.5 Sequences of oligonucleotides used in this study for RT-PCR

2.4 Treatment with small molecules drug

2.4.1 Stable transfection with small molecules drugs

2×10^6 cells from derived stables cell lines were plated per well into 35mm plates containing 2ml RPMI-1640 medium. Next day medium was replaced with 2ml complete medium containing different concentrations of small molecules: TPA (phorbol 12-myristate 13-acetate, Sigma), emetine (emetine dihydrochloride hydrate, Sigma), cycloheximide (Sigma), ceramide (N-Hexanoyl-D-sphingosine, Sigma), valproic acid (Sigma), scriptaid (Sigma), amiloride (Sigma) or tannic acid (Sigma). Cells were then incubated at 37°C and harvested at different time points and analysed.

2.4.2 Transient transfection with small molecules drugs

2×10^6 HeLa cells were plated per well into 35mm plates containing 2ml RPMI-1640 medium. Transfections were carried out by using GeneJuice[®] Transfection reagent according to the protocol of the manufacturer (Novagen). 0.5µg of plasmid DNA was added, and the mixture was incubated at room temperature for 10 min. Thereafter, the mixture was added drop-wise into each well and incubated at 37°C. 24 hours posttransfection medium was replaced with 2ml RPMI-1640 medium containing 0.5µl H₂O, 800ng TPA, 0.5mM valproic acid or 10µM tannic acid and cells were incubated at 37°C for further 24 hours.

2.5 Protein analysis

2.5.1 Protein extraction

Cells were lysed from the 6 well-plate directly with the lysis buffer from the CAT ELISA KIT or by trypsination as follows: cells were harvested with 200µl trypsin and then 800µl of complete

medium was added in each well. After centrifuging the cells in media for 3 min at 1500 rpm at room temperature, the supernatant was aspirated off and the pellet resuspended in 1ml of cold PBS. After a further centrifugation at 1500 rpm for 5 min, the PBS was aspirated off and cells were resuspended in 40-80 μ l RIPA buffer and vortexed vigorously. The mixer was incubated on ice for 20 min with occasional mixing followed by centrifugation at 14000 rpm for 30 min at 4°C. The supernatant was transferred into a fresh tube and protein concentration was measured in duplicate with Bradford assay using a BSA standard curve. Samples were stored at -80°C.

2.5.2 Western blot

15 μ l of each sample was heated at 95°C for 5 min and then was loaded with 5 μ l of 4x Laemmli sample buffer onto on an SDS-PAGE minigel (Appendix I). Gels were let run at 100V for 3 hours (dye about the bottom of the gel).

Proteins were then transferred onto a nitrocellulose membrane. All components of the blot were pre-wet in cold transfer buffer pH 8.3 (Appendix I) and the entire stack was placed in the transfer apparatus with cold transfer buffer using an ice-pack to cool down the apparatus. Transfer was carried out at 100V for 60-70 minutes.

After successful transfer, the membrane was blocked in 5% non-fat dry milk for 1 hour at room temperature on a shaker. The membrane was then washed with TBS + 1% Tween20, 5 times for 5 minutes.

A dilute solution of primary antibody (for specific dilutions see Table 2.6) was incubated with the membrane at 4°C for 16 hours on a nutator shaker under gentle agitation. After 5 washes in TBS + 1% Tween20 for 5 min the membrane was then incubated with secondary antibody for 1 hour at room temperature in 10ml 1x TBS + 5% milk + 0.1% Tween20 on a shaker under gentle

agitation. The membrane was then washed 5 times for 5 min each in ~20ml 1xTBS + 0.1% Tween20.

Protein detection was carried out with the ECL (enhanced chemiluminescence) kit. ECL solutions were mixed in a 1:1 ratio and poured onto membranes for 5 min. The ECL solutions were drained and membrane was exposed to film.

Antibody	Dilution	Low fat Milk solution	Molecular weight
PTB	1:1000	5% milk + 0.1% Tween 20	61-64 kDa
ASF/SF2	1:500	5% milk + 0.1% Tween 20	34 kDa
hnRNP A2/B1	1:1000	5% milk	37 kDa
β -actin	1:5000	5% milk + 0.1% Tween 20	42 kDa

Table 2.6 Antibodies used in this study for Western blot

2.6 Apoptosis-program

2.6.1 Serum deprivation

2×10^6 cells were seeded per well into 35mm plates and incubated overnight at 37°C with complete medium. Next day medium was removed and cells washed once with serum free medium and then incubated in medium with or without serum for 24, 48 and 72 hours. Cells were harvested for CAT at all time points.

2.6.2 Heat shock

2×10^6 cells were plated per well into 35mm plates and incubated overnight at 37°C with complete medium. Medium was then discarded and replaced with medium preheated to 45°C. Cells incubated at 37°C were also included in the experiment as a control for all time points. Cells were analysed after 30min, 1, 2 and 3 hours.

2.7 Infection of cells with Lentiviruses expressing SiRNA-GFP constructs

2.7.1 Transduction

5×10^5 cells were plated in the 35mm culture dishes and incubated overnight. After 24 hours the medium was replaced with 0.5ml serum free medium and cells were infected at different multiplicity of infection (MOI) (Table 2.7) with each lentivirus carrying short hairpin RNA (shRNA) ASF/SF2, PTB or scramble. 4 hours post-transduction 1.5ml of complete medium was added to the well and then cells were incubated at 37°C for 48 hours. Selection for infected cells was carried out with 1 µg/ml puromycin (Sigma) for further 5 days.

Lentiviruses	Number of Cell/well	TU/ μ l	MOI	Time (h)	Puromycin μ g/ml
ASF/SF2	5×10^5	3.38×10^5	0.3, 1, 2, 5, 10	48	1
PTB	5×10^5	4.32×10^5	0.3, 1, 2, 5, 10	48	1
scramble	5×10^5	3.75×10^5	0.3, 1, 2, 5, 10	48	1

Table 2.7 Lentiviruses used in this study

2.7.2 Detection of GFP fluorescence

Cells were plated in 35mm Glass Bottom Culture Dishes and cultured in 1.5ml complete RPMI-1640 medium. The next day the complete medium was removed and replaced with medium without serum. GFP fluorescence intensity was detected with a confocal laser scanning microscope, Zeiss LSM 510.

2.8 Tissue microarray preparation and analysis

2.8.1 Samples collection and tissue microarray preparation

Cervical samples were collected from the Rotunda Hospital (samples sent between 2005 and 2007). Ethical approval was obtained from the Rotunda Hospital to collect patient information and formalin fixed paraffin embedded (FFPE) cervical tissue samples, and from Dublin Institute of Technology Ethics Committee. Samples were arbitrarily numbered to maintain confidential all patient information.

Tissue microarrays (TMA) were prepared using a tissue arrayer (Beecher Instruments Manual Tissue Arrayer MTA-1) following the instructions of the manufacturer. Cores of 1.0mm in size were produced using MP10 needles, produced by Beecher Instruments. Each tissue microarray contains 4 cases of non-neoplastic cervical epithelium, 12 cases of low-grade squamous intraepithelial lesion, 12 cases of high-grade squamous intraepithelial lesion and 12 cases of cervical carcinoma. Each case was sampled in duplicate for a total of 80 cores per TMA.

2.8.2 Immunohistochemistry (Avidin-Biotin Complex Immuno-peroxidase method)

Five micrometer thick cervix TMA was cut from formalin-fixed paraffin embedded tissue blocks and melted in the oven at 56°C for 1 hour. The tissue microarray slides were then dewaxed in xylene for 10 min and rehydrated in absolute alcohol for 10 min and further 5 min in spirit. To remove alcohol slides were washed in water for 1 min.

Slides were then submerged in 10mM/l Citrate Buffer (pH 6) (Appendix I) and heated in the microwave for 12 minutes. After antigen retrieval the slides were allowed to cool in the citrate buffer for a further 20 min.

After fixation with 3% methyl hydrogen peroxide for 5 min, the sections were rinsed 3 times in phosphate buffered saline (PBS) pH 7 and then stained using the PK-6200 universal VECTASTAIN ABC kit as follows. Normal horse serum was applied to the sections for 5 min. Sections were washed with PBS three times and approximately 300µl of specific primary antibody was applied and incubated at room temperature for 1 hour. hnRNP A2/B1 and p16INK4a were used as primary antibodies (see Table 2.8 for dilutions). Following three washing with PBS the biotinylated secondary antibody was applied and allowed to conjugate for

15 min. Sections were washed three times with PBS and then covered with ABC reagent for 15 min. Peroxidase labeling was visualized using 0.02% hydrogen peroxide and 0.2% 2,4-diaminobenzidine. Mayer's hematoxylin was applied to the slides for 40 sec and the slides were then blued in warm water for 3 min. The tissue arrays were then dehydrated in spirit and absolute alcohol through xylene. A cover slip was mounted on the slide using distrene plasticiser xylene (DPX).

Antibody	Dilution	Company	Raised in
hnRNP A2/B1	1:1000	Abcam	Mouse
P16INK4a	1:500	Abcam	Mouse

Table 2.8 Antibodies used in this study for IHC

3. Chapter 3 - Results

Generation and validation of HPV-16 late gene reporter constructs

The life cycle of HPV-16 is strictly associated with the differentiation program of the infected cells (Doorbar, 2005). At the molecular level, a persistent HPV infection is characterised by continued expression of the early viral genes, and a complete shut-down of the late viral genes encoding the highly immunogenic viral structural proteins L1 and L2. Similarly high grade, premalignant lesions containing HPV-16 primarily express early viral genes, and not L1 or L2. Inhibition of HPV-16 late gene expression in the early stage of life cycle is probably a prerequisite for persistence of infection, therefore activation of L1 and L2 late gene expression in the persistently infected cells would alert the immune system of the host of the presence of the virus and it might clear the infection. As the L1 RNA is difficult to detect it would be useful to generate a novel and easy system to facilitate the investigation of HPV-16 late gene regulation.

3.1 Generation of reporter plasmids

In order to generate a simple and reliable bioassay to study HPV-16 late gene expression two previously described plasmids, named pBEL and pBELM, were used (Fig. 3.1) (Zhao *et al.*, 2004). Both plasmids carry viral early and late genes, except E6 and E7 and the weak viral promoters are replaced by the strong human cytomegalovirus (CMV) immediate early promoter. Early and late splice sites are present in the plasmid allowing the examination of the splicing events that can occur in HPV-16. Similar to the HPV-16 genome during an infection, pBEL transfected into proliferating cells expresses high level of the early genes, primarily E4, whereas expression of late genes is undetectable. In pBELM, the splicing silencer elements adjacent to

the SA5639 have been mutated to reduce the silencing on this splice acceptor therefore activating late gene expression (Fig. 3.1) (Zhao *et al.*, 2004).

This study aimed to examine the regulation of HPV-16 L1 gene by replacing it with an easily detectable reporter such as CAT, GFP, SEAP, Luciferase and β -galactosidase. In order to achieve this, the plasmids pBEL and pBELM were modified to incorporate the poliovirus IRES and reporter gene as described in figure 3.2 and 3.3.

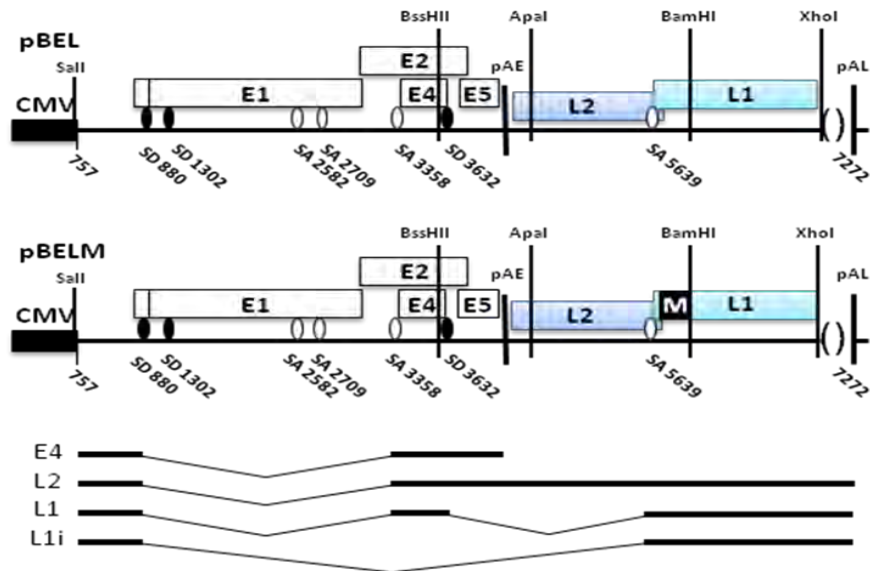


Figure 3.1 Schematic representation of the subgenomic HPV-16 expression plasmids, pBEL and pBELM, and the major potential mRNAs that can be produced

The position of the human cytomegalovirus immediate early (CMV) promoter is shown. Numbers refer to nucleotide positions of 5' splice sites (black circles), 3' splice sites (white circles). The mutation site in pBELM (M) is indicated. The early and late polyadenylation signals pAE and pAL are indicated. Major potential mRNAs produced by pBEL and pBELM are indicated below plasmids.

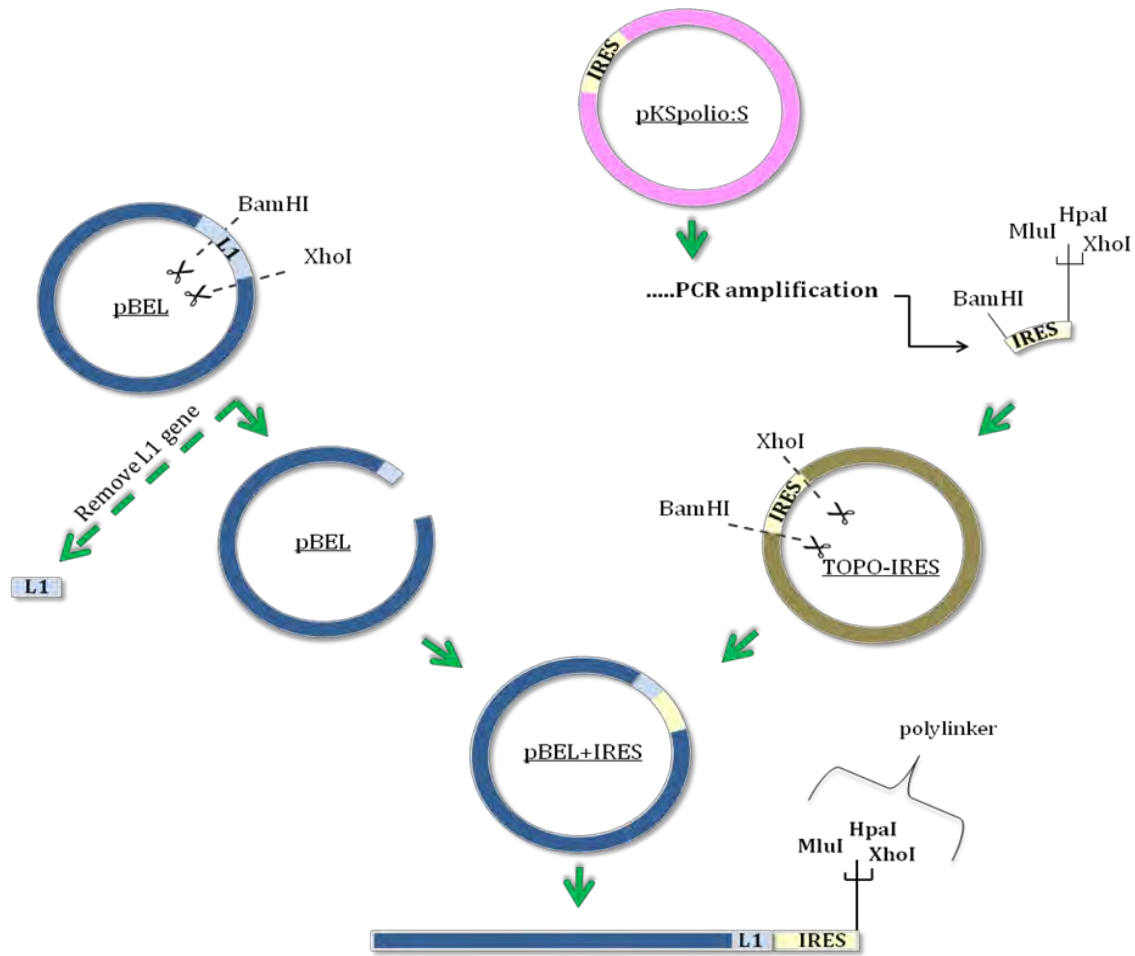


Figure 3.2 Schematic representation of cloning steps for insertion of IRES sequence and reporter gene into pBEL plasmid

To release the first part of L1 gene, pBEL was digested with BamHI and XhoI restriction enzymes at position 6150nt in HPV-16 L1 and position 7272nt in the HPV-16 late UTR. The IRES sequence from the poliovirus plasmid was PCR amplified using IRESs and IRESa primers (Table 2.2) and cloned into TOPO followed by insertion as a BamHI/XhoI fragment into pBEL using T4 DNA ligase. The polylinker present in the IRES sequence was used for insertion of reporter genes into the plasmid. Each step was reproduced for the cloning of IRES into pBELM.

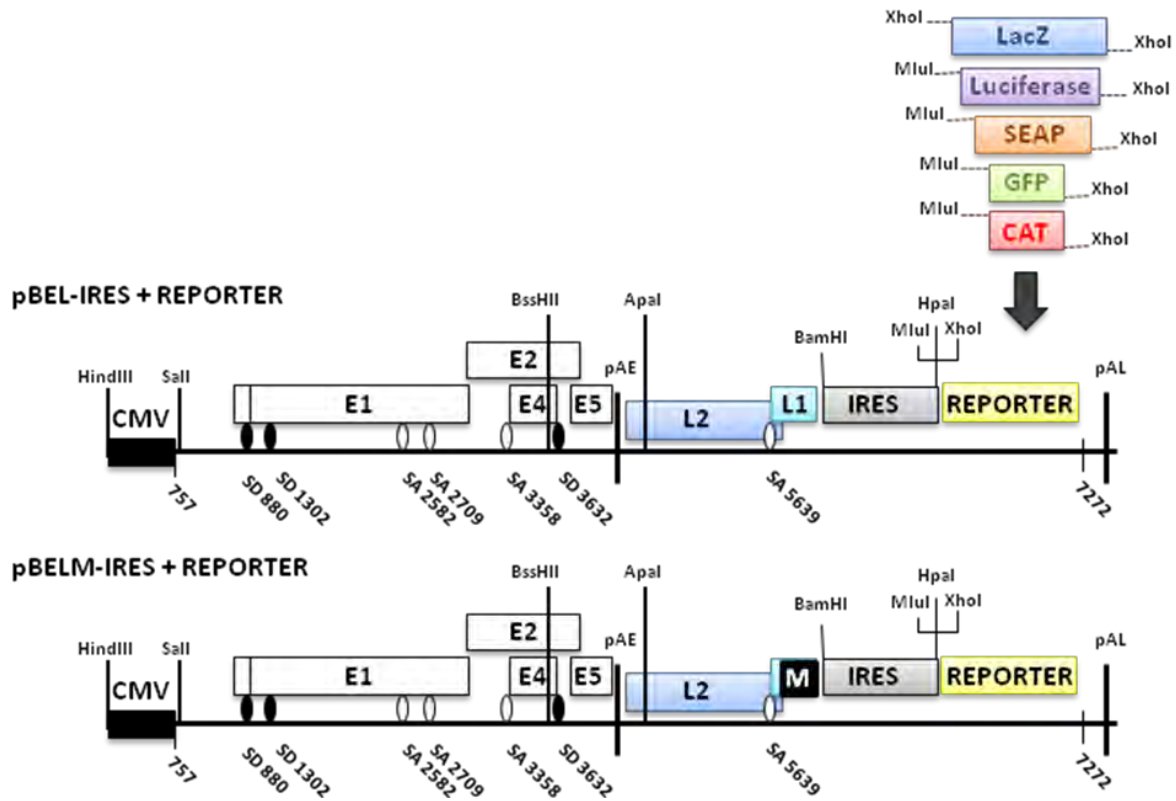


Figure 3.3 Generation of reporter plasmids

Chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), secreted embryonic alkaline phosphatase (SEAP), Luciferase and β -galactosidase (LacZ) were selected as reporter genes to be inserted into pBEL and pBELM plasmids downstream the IRES sequence. All restriction sites used for the cloning steps are shown in the figure. Numbers refer to nucleotide positions of 5' splice sites (black circles), 3' splice sites (white circles).

The L1 region in pBEL and pBELM has unique restriction sites for BamHI and XhoI enzymes necessary for the cloning. When electrophoresed on an agarose gel the uncut plasmid showed a

supercoiled band. The plasmid digested with BamHI or XhoI resulted on a linearised band of 10004 bp and cut with BamHI and XhoI together it released a 1 Kb fragment of L1 (Fig. 3.4). Cutting with BamHI or XhoI at position 6150 and 7272, respectively, will release L1 from nucleotide 514 and allow insertion of the IRES and reporter sequences.

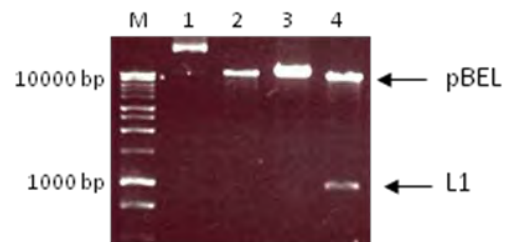


Figure 3.4 Restriction enzymes in pBEL plasmid

10 Kb DNA ladder (M), lane 1 undigested pBEL plasmid, lane 2 pBEL plasmid digested with BamHI restriction enzyme (10004 bp), lane 3 pBEL plasmid digested with XhoI restriction enzyme (10004 bp), lane 4 pBEL plasmid digested with BamHI and XhoI restriction enzymes.

The first 514-nucleotides of the L1 coding sequence contain RNA elements that regulate splicing to the 3' -splice site named SA5639. These sequences therefore must be present on the reporter plasmids and since this L1 sequence contains the L1 ATG and multiple internal methionines, the poliovirus 2A internal ribosome entry site (IRES) is required upstream of the reporter gene to ensure that the reporter gene expression is a true measure of late mRNA levels. A serial dilution of the pKSPOLIO:S plasmid (initially approximately 1 μ g/ μ l) containing the IRES sequence was carried out and PCR amplification was performed at 59°C annealing as described in Material and Methods, with IRESs and IRESa primers (Table 2.2, 2.3 and 2.4). The PCR product resulted in a 752 bp band (Fig. 3.5).

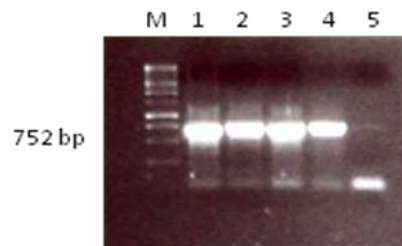


Figure 3.5 PCR amplification of IRES sequence with IRES set primers

10Kb DNA ladder (M), lane 1 PCR amplification of 1/10 pKSPOLIO:S plasmid, lane 2 1/100 pKSPOLIO:S plasmid, lane 3 1/1000 pKSPOLIO:S plasmid, lane 4 1/10000 pKSPOLIO:S plasmid, lane 5 blank.

The oligonucleotide primers used for IRES PCR amplification have at each 5' end restriction sites useful for cloning into pBEL and pBELM and subsequently insertion of a reporter gene. The 5' end of the IRES sense has a restriction site for BamHI and a short polylinker (XhoI, MluI and HpaI) is present on the IRES antisense primer (Table 2.1 and 2.2) (Fig. 3.2 and 3.3). The PCR product was inserted into pCR2.1-TOPO vector (3931 bp) and the recombinant vector was transformed into competent *E. coli*. Several colonies were picked and plasmids were isolated and purified from bacterial cultures. In order to identify the recombinant plasmid of interest each purified plasmid DNA was subjected at restriction digest with BamHI and XhoI and then analyzed on a 1% agarose gel. Only two recombinant plasmids, lanes 2 and 4, have the correct IRES sequence inserted (Fig. 3.6).

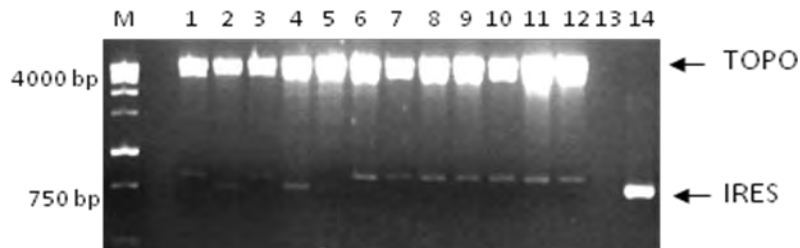


Figure 3.6 Analysis of recombinant plasmids TOPO + IRES

10Kb DNA ladder (M), lanes 1-13 different recombinant plasmids, lane 2 and lane 4 TOPO + IRES (4000 bp + 752 bp), lane 14 IRES PCR product band 752 bp. 1% agarose gel ethidium bromide-stained.

The TOPO vector carrying the IRES gene and pBEL were digested together with BamHI and XhoI restriction enzymes. By cutting the two plasmids with the same restriction enzymes, the resulting molecular fragments terminate with complementary sticky ends and in order to recombine the fragments into recombinant DNA molecules T4 DNA ligase was used. The ligation was then transformed into competent *E. coli* and plasmid DNA was isolated and purified from bacterial cultures and recombinant plasmids analyzed as mentioned above. The same was repeated for pBELM. The result of this ligation can be the original TOPO-IRES, the original pBEL or pBELM, TOPO-L1 and pBEL-IRES or pBELM-IRES. Only two colonies seemed to have the right recombinant pBEL + IRES (Fig. 3.7) and pBELM + IRES (Fig. 3.8). These 2 new plasmids have the original pBEL and pBELM sequence with part of the L1 gene (approx. 1000 bp) replaced by the IRES sequence (752 bp). Lane 4 (Fig. 3.7) below appears to have the correct pBEL backbone and insert and lane 6 (Fig. 3.8) the correct pBELM backbone and insert. High quality plasmid was generated for each correct recombinant and sequenced to confirm the clone, (see appendix I for alignment with poliovirus sequence).

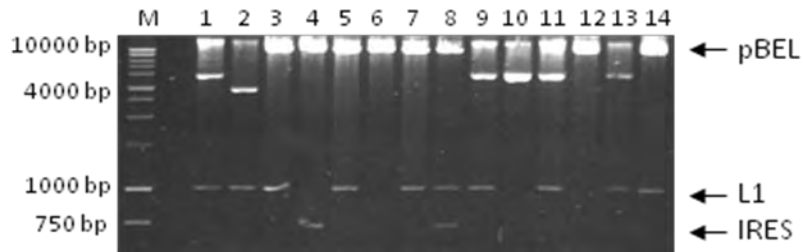


Figure 3.7 Analysis of recombinant plasmids pBEL + IRES

10Kb DNA ladder (M), lanes 1-14 different recombinant plasmids, lane 4 pBEL + IRES (10000 bp + 752 bp). 1% agarose gel ethidium bromide-stained.

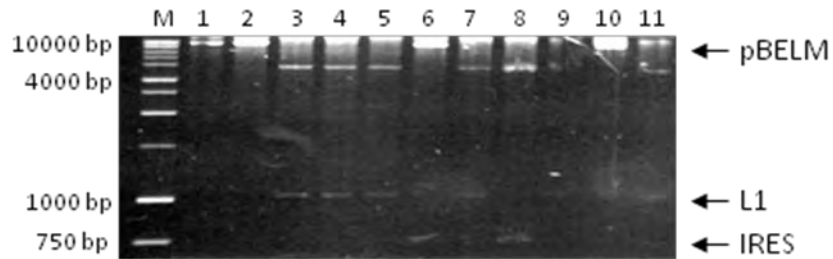


Figure 3.8 Analysis of recombinant plasmids pBELM + IRES

10Kb DNA ladder (M), lanes 1-11 different recombinant plasmids, lane 6 pBELM + IRES (10000 bp + 752 bp). 1% agarose gel ethidium bromide-stained.

In order to generate an easily detected reporter plasmid, 5 different reporter genes were selected to replace the L1 gene: chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), secreted alkaline phosphatase (SEAP), Luciferase and β -galactosidase (LacZ) (Fig. 3.9).

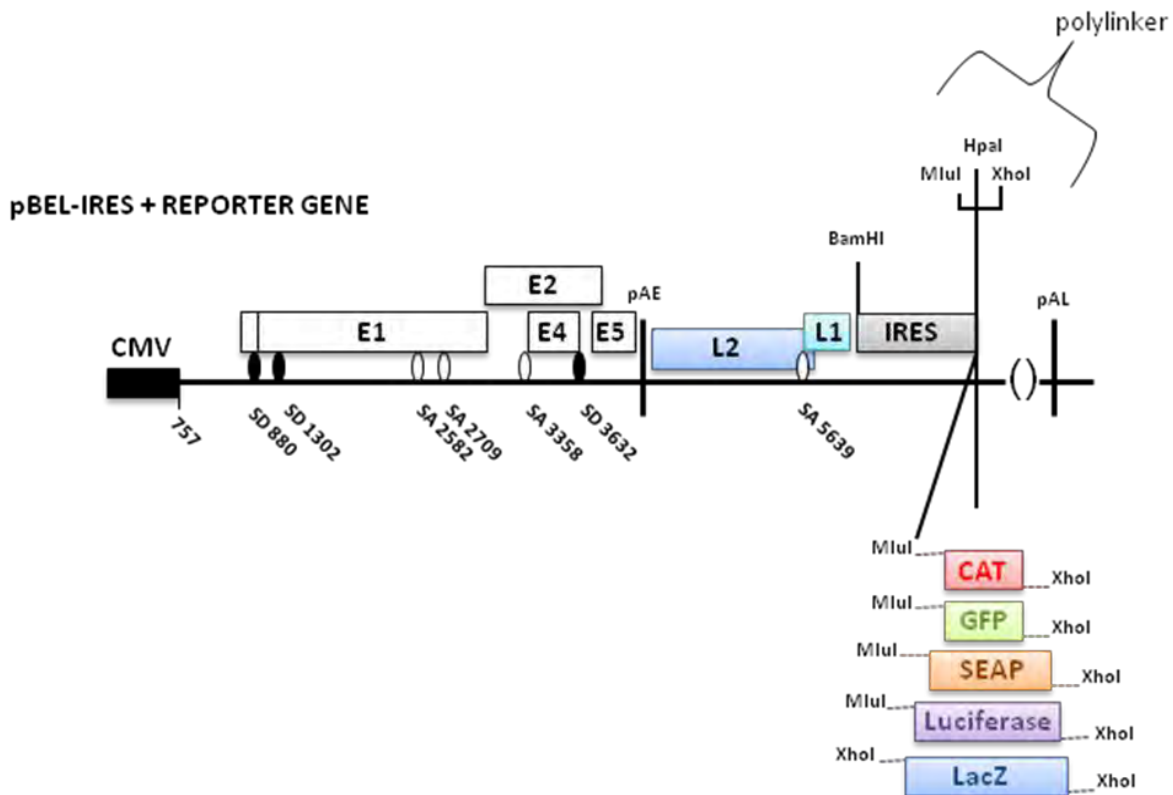


Figure 3.9 Schematic representation of the reporter plasmids

Chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), secreted alkaline phosphatase (SEAP), Luciferase and β -galactosidase (LacZ) were chosen as reporter gene to clone pBEL-IRES and pBELM-IRES. The polylinker is shown in the diagram.

CAT was PCR amplified at 65°C with CATs and CATa primers (Table 2.1 and 2.2) resulting in a PCR product of 660 bp (Fig. 3.10). SEAP was PCR amplified at 60.2°C with SEAPs and SEAPa primers (Table 2.1 and 2.2) resulting in a PCR product of 1530 bp (Fig. 3.10). LacZ was PCR amplified at 59°C with LacZs and LacZa primers (Table 2.1 and 2.2) resulting in a PCR product of 3141 bp (Fig. 3.10). Each oligonucleotide used for these PCR amplifications has a 5' restriction site required for the cloning into pBEL-IRES and pBELM-IRES. For example as shown in Table 2.1 CATs and CATa primers have MluI and XhoI restriction site, respectively. SEAPs primer has at the 5' end restriction site for MluI restriction enzyme and SEAPa for XhoI. LacZ primers both have at each 5' end restriction site for XhoI.

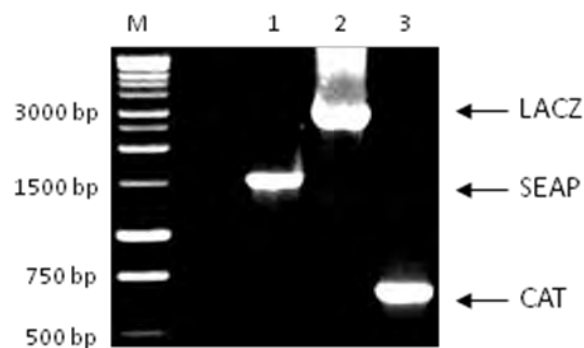


Figure 3.10 PCR amplification of secreted alkaline phosphatase (SEAP), β -galactosidase (LacZ) and chloramphenicol acetyltransferase (CAT) reporter genes

10Kb DNA ladder (M), lane 1 PCR amplification of SEAP gene (1530 bp) using the SEAPs/SEAPa primers. Lane 2 PCR amplification of LacZ gene (3141 bp) using the LacZs/LacZa primers. Lane 3 PCR amplification of CAT gene (660 bp) using the CATs/CATa primers.

The CAT PCR product was first ligated with pCR2.1-TOPO vector (3931 bp) and the recombinant vector was transformed into competent *E. coli*. Several colonies were picked and plasmids were isolated and purified from bacterial cultures. In order to identify the recombinant plasmid of interest each purified plasmid DNA was digested with MluI and XhoI restriction enzyme and then analyzed on a 1% agarose gel stained with ethidium bromide-stained. As shown in figure 3.11 all recombinant plasmids, except one (lane 2), resulted in TOPO + CAT sequence.

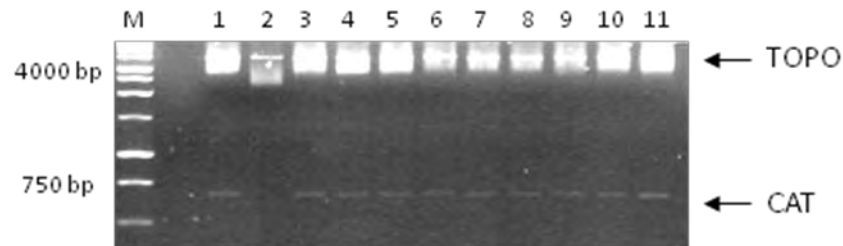


Figure 3.11 Analysis of recombinant plasmids TOPO + CAT

10Kb DNA ladder (M), lanes 1-11 different recombinant plasmids TOPO + CAT (4000 bp + 660 bp). 1% agarose gel ethidium bromide-stained.

In order to clone the CAT sequence into pBEL-IRES and pBELM-IRES, the TOPO vector carrying the CAT reporter gene and pBEL-IRES and pBELM-IRES plasmids were digested separately with MluI and XhoI restriction enzymes and then ligated in 2 different reactions: pBEL-IRES and TOPO-CAT one, and pBELM-IRES and TOPO-CAT the other. Ligations were then transformed into competent *E. coli* and plasmid DNA was isolated and purified from bacterial cultures and recombinant plasmids analyzed as mentioned above. The result of this ligation can be the original TOPO-CAT, the original pBEL-IRES or pBELM-IRES and pBEL-IRES-CAT or pBELM-IRES-CAT. Only two recombinant plasmids showed to have the CAT reporter gene inserted into pBEL-IRES (Fig. 3.12, lane 6) and pBELM-IRES (Fig. 3.13, lane 8). The CAT reporter gene is inserted downstream of the IRES sequence resulting in pBEL-IRES-CAT and pBELM-IRES-CAT that for simplification we will call pBELCAT and pBELMCAT, respectively.

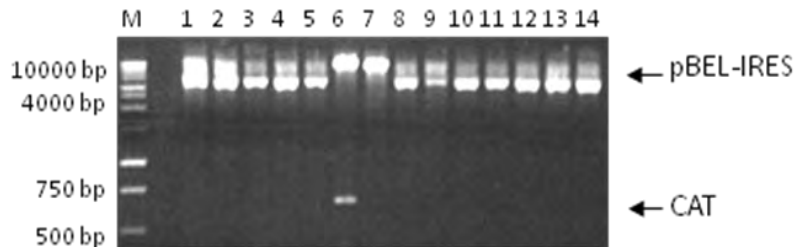


Figure 3.12 Analysis of recombinant plasmids pBEL-IRES + CAT

10Kb DNA ladder (M), lanes 1-14 different recombinant plasmids, lane 6 pBEL-IRES-CAT (bands 10000 bp + 660 bp). 1% agarose gel ethidium bromide-stained.

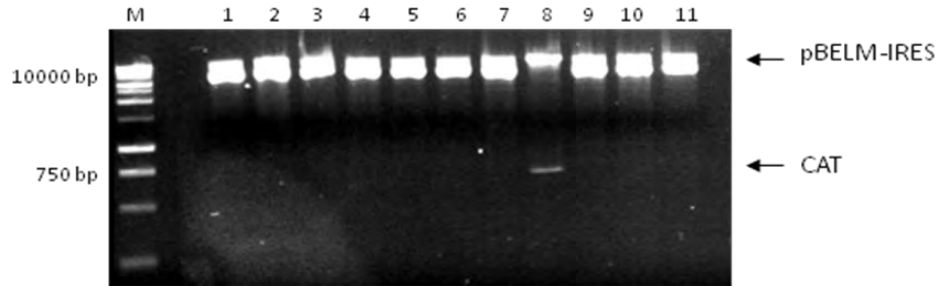


Figure 3.13 Analysis of recombinant plasmids pBELM-IRES + CAT

10Kb DNA ladder (M), lanes 1-11 different recombinant plasmids, lane 8 pBELM-IRES-CAT (bands 10000 bp + 660 bp). 1% agarose gel ethidium bromide-stained.

The SEAP PCR product was first inserted into pCR2.1-TOPO vector (3931 bp) as mentioned above. In order to identify the recombinant plasmid of interest each purified plasmid DNA was digested with MluI and XhoI restriction enzymes and then analyzed on a 1% agarose gel. Five recombinant plasmids resulted in TOPO + SEAP sequence (Fig. 3.14, lanes 10, 12, 14, 15 and 17).

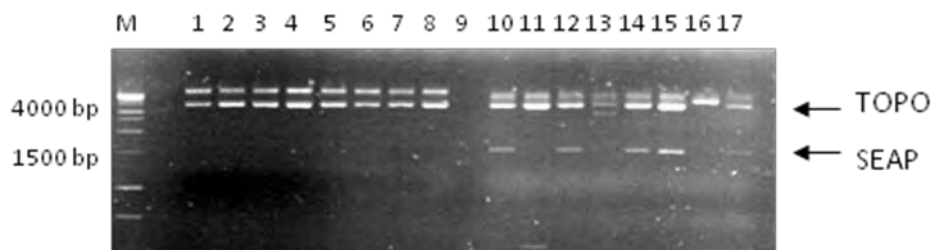


Figure 3.14 Analysis of recombinant plasmids TOPO + SEAP

1 10Kb DNA ladder (M), lanes 1-17 different recombinant plasmids for TOPO-SEAP, lanes 10, 12, 14, 15 and 17 positive TOPO + SEAP (4000 bp + 1530 bp). 1% agarose gel ethidium bromide-stained.

The SEAP sequence was then inserted into pBEL-IRES and pBELM-IRES, by digestion with MluI and XhoI restriction enzymes of the TOPO vector carrying SEAP reporter gene and pBEL-IRES and pBELM-IRES plasmids and ligation. Ligations were then transformed into competent *E. coli* and plasmid DNA was isolated and purified from bacterial cultures and recombinant plasmids subjected at restriction. As shown in figure 3.15 many recombinant plasmids showed to have the SEAP reporter gene inserted into pBEL-IRES (lanes 1, 2) and pBELM-IRES (lanes 10, 14, 15, 16, 17 and 18), respectively, resulting in pBELSEAP and pBELMSEAP reporter plasmids.



Figure 3.15 Analysis of recombinant plasmids pBEL-IRES + SEAP and pBELM-IRES + SEAP

10Kb DNA ladder (M), lanes 1-9 different recombinant plasmids for pBEL-IRES-SEAP, lanes 10-18 different recombinant plasmids for pBELM-IRES-SEAP (10000 bp + 1530 bp). 1% agarose gel ethidium bromide-stained.

In order to generate the GFP reporter plasmid a pUC57-tat-LTR-GFP plasmid has been generated with restriction sites MluI and XhoI in the GFP gene. The CAT reporter gene from pBELMCAT was replaced with the GFP reporter gene by digestion with MluI and XhoI of pBELMCAT and pUC57-tat-LTR-GFP and religation. The result of this ligation can be the original pUC57-tat-LTR-GFP, pBELM-IRES, pBELMCAT or pBELM-IRES-GFP. As shown in figure 3.16 two recombinant plasmids resulted to be pBELM-IRES-GFP (lane 12 and 13) that from now on will be called pBELMGFP. We reasoned that GFP levels produced from the pBEL-backbone would be too low to detect, and therefore constructed only pBELMGFP.

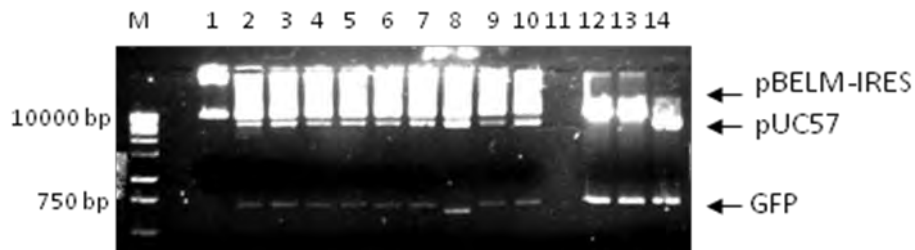


Figure 3.16 Analysis of recombinant plasmids pBELM-IRES + GFP

10Kb DNA ladder (M), lanes 1-14 different recombinant plasmids, lane 12 and lane 13 pBELM-IRES + GFP (10000 bp + 650 bp). 1% agarose gel ethidium bromide-stained.

The LacZ reporter gene was inserted into TOPO vector as shown in figure 3.17 (lanes 1, 4, 5 and 10). Cloning into pBEL and pBELM was unsuccessful, probably due to the considerable size of the reporter gene itself (3141bp) and the plasmid which is approximately 10Kb.

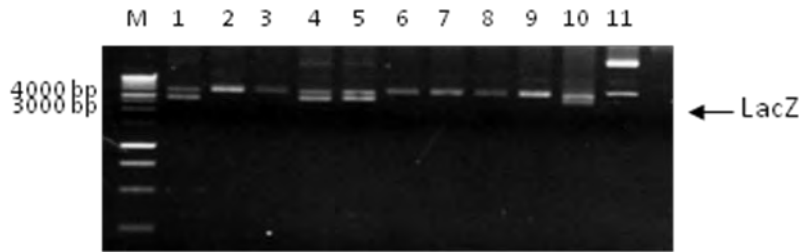


Figure 3.17 Analysis of recombinant plasmids TOPO + LacZ

10Kb DNA ladder (M), lanes 1-11 different recombinant plasmids for TOPO-LacZ, lanes 1, 4, 5 and 10 positive TOPO + LacZ (4000 bp + 3111 bp). 1% agarose gel ethidium bromide-stained.

3.2 Determination of CAT expression in pBELCAT and pBELMCAT

The pBELCAT and pBELMCAT reporters were chosen as the first constructs to analyse. These were transfected into HeLa cells in culture to examine if they behave in a similar manner to the L1 gene in previous transfections with pBEL and pBELM (Fig. 3.18).

pBELCAT and pBELMCAT (Fig. 3.19) were transfected into HeLa cells and CAT expression was monitored 24 hours posttransfection. Similar to previously Northern blot analysis, pBELCAT produced barely detectable levels of CAT (0.13 CAT units), whereas pBELMCAT showed higher CAT expression (29 CAT units) as expected, and produced over 200-fold more CAT than pBELCAT (Fig. 3.20). However, these levels were relatively low compared to CAT expression from the positive control plasmid CMVCAT, which produced 10,600 CAT units (Fig. 3.20). These results demonstrated that only a minority of the mRNAs produced by pBELCAT and pBELMCAT were late mRNAs, suggesting that late gene expression was relatively inefficient from both plasmids. These results demonstrate that the reporter plasmids are functional and

express CAT protein. Multiple nucleotide substitutions in the first 514 nucleotides of L1 enhanced CAT expression as a result of the inactivation of splicing silencers as previously described (Zhao *et al.*, 2004). These results also indicate that the pBELCAT and pBELMCAT reporters can be used in future studies to examine proteins that regulate L1 gene expression.

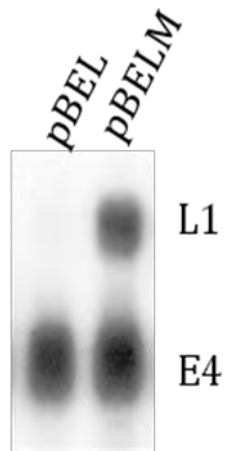


Figure 3.18 Northern blot analysis on pBEL and pBELM

pBEL and pBELM in previous transfections produced high levels of the early mRNAs, primarily E4 mRNA, whereas late mRNAs are undetectable. pBELM showed to produce late mRNAs without interfering on the production of the E4 mRNAs. (Zhao *et al.*, 2004).

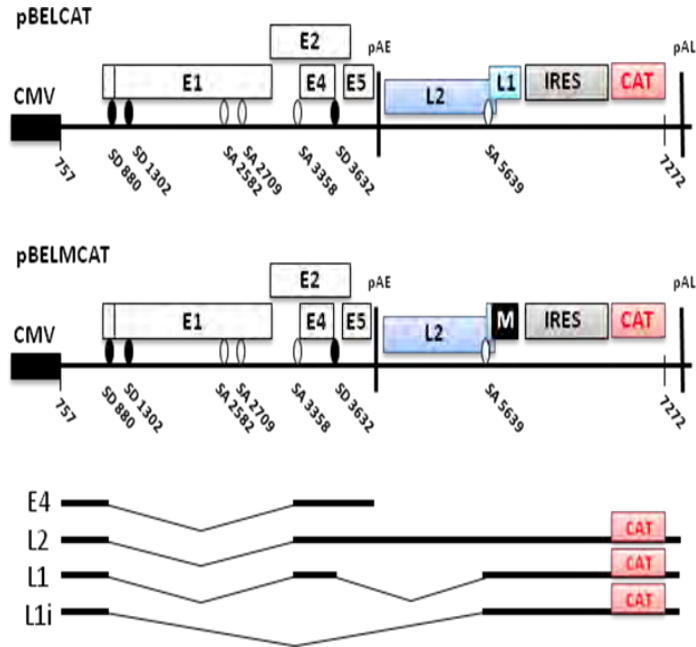


Figure 3.19 Schematic representation of the subgenomic HPV-16 expression plasmids, pBELCAT and pBELMCAT, and the major potential mRNAs that can be produced

The position of the human cytomegalovirus immediate early (CMV) promoter is shown. Numbers refer to nucleotide positions of 5' splice sites (black circles), 3' splice sites (white circles). The early and late polyadenylation signals pAE and pAL are indicated. Major potential mRNAs produced by pBELCAT and pBELMCAT are indicated below plasmids.

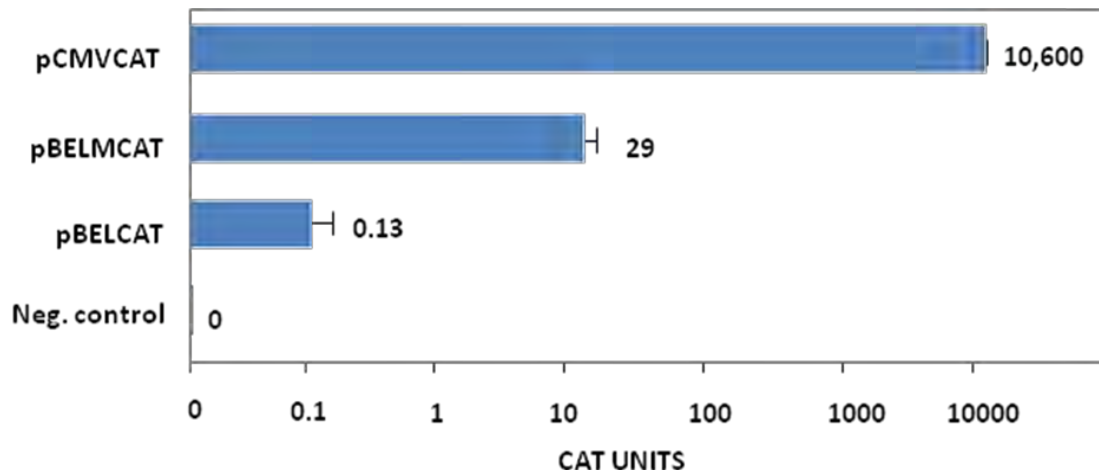


Figure 3.20 Determination of CAT levels in transient transfection

Transfection of pBELCAT, pBELMCAT and CMVCAT expression plasmids into HeLa cells. Logarithmic scale. Cells were transfected with each plasmid in a minimum of three independent experiments. Error bars indicate standard error for each sample.

3.3 Adenovirus E4orf4, polypyrimidine tract binding protein, ASF/SF2 and SRp30c induce CAT expression from pBELCAT and pBELMCAT reporter plasmids

HPV-16 late gene expression is regulated at the level of RNA processing (Graham, 2008; Schwartz, 2008; Zheng & Baker, 2006). Overexpression of some viral and cellular proteins, e. g. adenovirus E4orf4 protein (E4orf4), polypyrimidine tract binding protein (PTB), alternative splicing factor/splicing factor 2 (ASF/SF2) and arginine/serine-rich SRp30c protein (SRp30c), have been shown previously to induce HPV-16 late gene expression from pBEL and pBELM (Somberg *et al.*, 2011; Somberg & Schwartz, 2010; Somberg *et al.*, 2009; Somberg *et al.*, 2008). In order to validate the functionality of the novel reporter plasmids pBELCAT and pBELMCAT, each reporter plasmid was cotransfected with E4orf4, PTB, SRp30c and ASF/SF2 into HeLa cells. Overexpression of E4orf4, PTB, ASF/SF2 or SRp30c caused an increase in CAT expression from both pBELCAT and pBELMCAT (Fig. 3.21). E4orf4, which regulates the switch from early to late gene expression in the adenovirus life cycle (Akusjärvi & Stevenin, 2003), efficiently induced CAT production from both plasmids but in particular from pBELCAT, resulting in a 300 fold induction (Fig. 3.21). A titration of the E4orf4-, PTB-, SRp30c or ASF/SF2-plasmids sub-genomic HPV-16 expression plasmid in pBELMCAT, revealed that induction of CAT was dose-dependent (Fig. 3.22). While CAT induction caused by PTB and SRp30c was increased with higher levels of PTB or SRp30c plasmid, the highest concentrations of E4orf4 and ASF/SF2 plasmids inhibited CAT production, suggesting other effects on the cells (Fig. 3.22). We concluded that E4orf4, PTB SRp30c and ASF/SF2 induced CAT expression from pBELCAT and pBELMCAT in a dose dependent manner and that the reporter plasmids respond to known inducers of late gene expression in a manner similar to the native L1 gene.

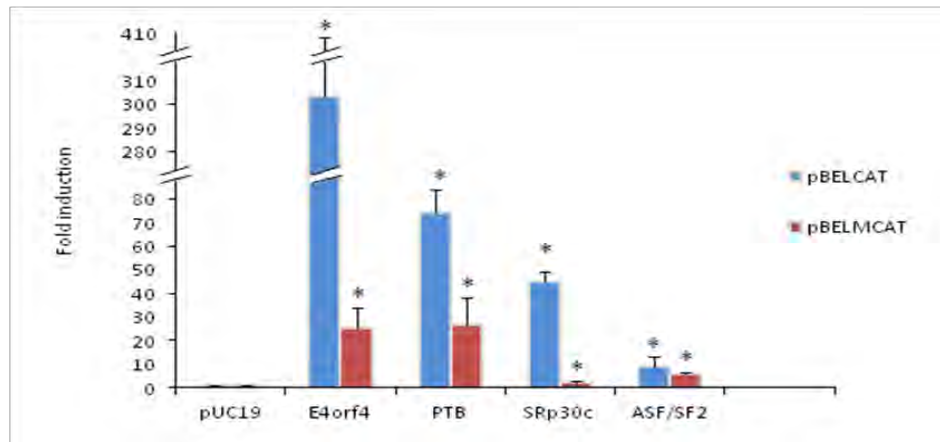


Figure 3.21 Cotransfection of CAT reporter plasmids with E4orf4, PTB, SRp30c or ASF/SF2 expression vector

pBELCAT or pBELMCAT were transfected into HeLa cells in the presence and absence of cotransfected expression vectors. Cells were transfected with each plasmid in a minimum of three independent experiments. Error bars indicate standard error for each sample. * indicates significance in relation to pBELCAT or pBELMCAT cotransfected with the empty vector (pUC19) ($p < 0.05$).

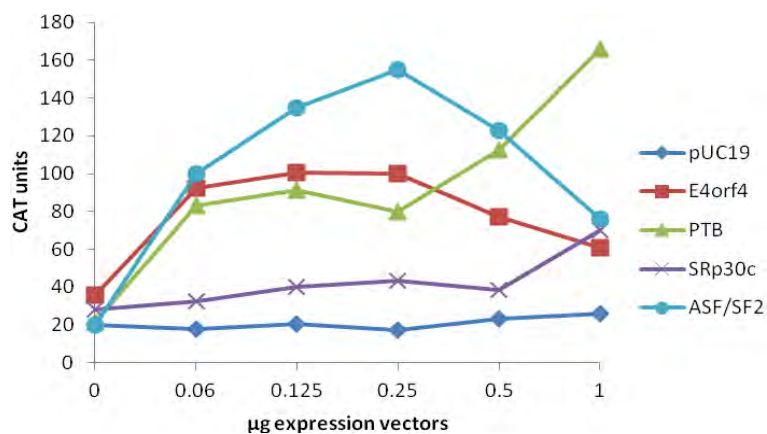


Figure 3.22 Titration of E4orf4, PTB, SRp30c or ASF/SF2 expression vector in pBELMCAT reporter plasmid

A titration up to 1µg of E4orf4, PTB, SRp30c or ASF/SF2 was carried out in pBELMCAT. pUC19 plasmid was used as an empty vector to equalize the concentration in each single transfection.

3.4 Determination of GFP levels in pBELMGFP

Green fluorescent protein (GFP) is a reporter gene widely used as a tool to study gene expression. It is a reliable reporter of gene expression in individual eukaryotic cells and its fluorescence can be measured and quantified from living cells.

pBELMGFP (Fig. 3.23) and a CMV driven GFP expression plasmid (CMVGFP) were transfected into HeLa cells. Analysis of the cells post-transfection revealed that pBELMGFP produced low detectable levels of GFP (Fig. 3.24). Quantification of GFP expression in a number of different GFP positive cells revealed that GFP levels produced from pBELMGFP were approximately 200-fold lower than those produced by CMVGFP (Fig. 3.25).

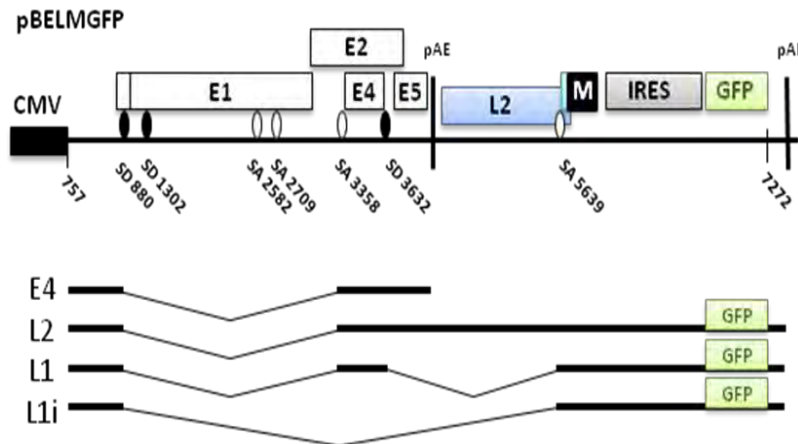


Figure 3.23 Schematic representation of subgenomic expression plasmid pBELMGFP

The position of the human cytomegalovirus immediate early (CMV) promoter is shown. Numbers refer to nucleotide positions of 5' splice sites (black circles), 3' splice sites (white circles). The early and late polyadenylation signals pAE and pAL are indicated.

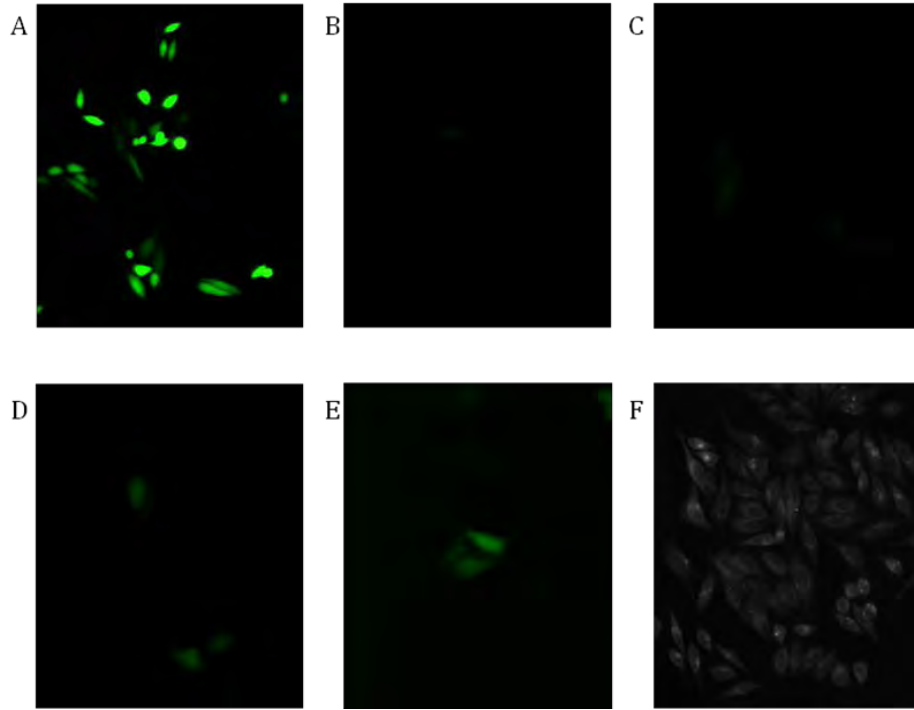


Figure 3.24 Examination of GFP expression after 24 hours posttransfection

HeLa cells were transfected with (A) CMVGFP plasmid as a control, (B) pBELMGFP reporter plasmid and (C) pBELMGFP + E4orf4 (D) pBELMGFP + PTB (E) pBELMGFP + SRp30c (F) control HeLa cells. Untransfected cells are shown as a comparison.

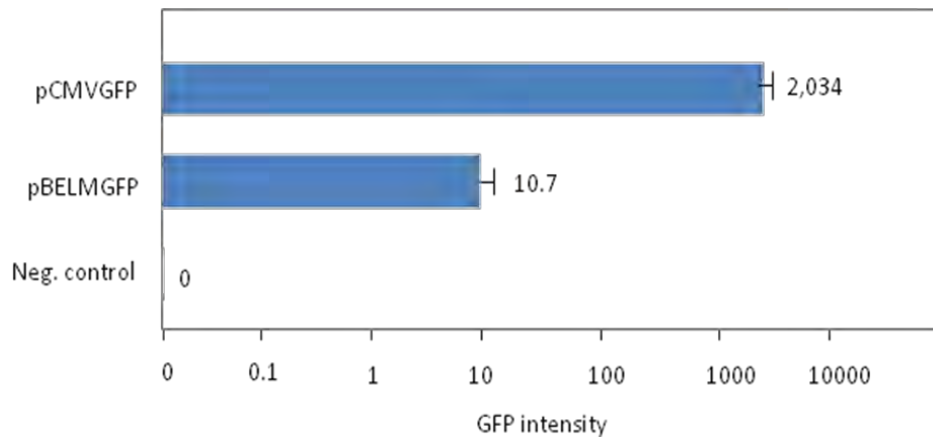


Figure 3.25 Determination of GFP levels in transient transfection

Transfection of CMVGFP and pBELMGFP expression plasmids into HeLa cells. GFP fluorescence intensities were detected with a fluorescence microscope from the living cells. GFP fluorescence intensities were detected and quantitated in 10 randomly chosen cells as described in material and methods. Cells were transfected with each plasmid in a minimum of three independent experiments. Logarithmic scale. Error bars indicate standard error for each sample.

Cotransfection with E4orf4 showed a small but significant induction compared to the other cotransfected expression vectors and to the experiments with pBELCAT and pBELMCAT (Fig. 3.26). However, cotransfection of pBELMGFP with a PTB and SRp30c expression plasmids resulted in induction of GFP expression (Fig. 3.24 and 3.26). PTB and SRp30c increased the levels of GFP of almost 6- and 30- fold, respectively (Fig. 3.26 B).

Taken together these results showed that pBELMGFP reporter plasmid is functional and its expression is fast and easily detectable when transfected into proliferating cells.

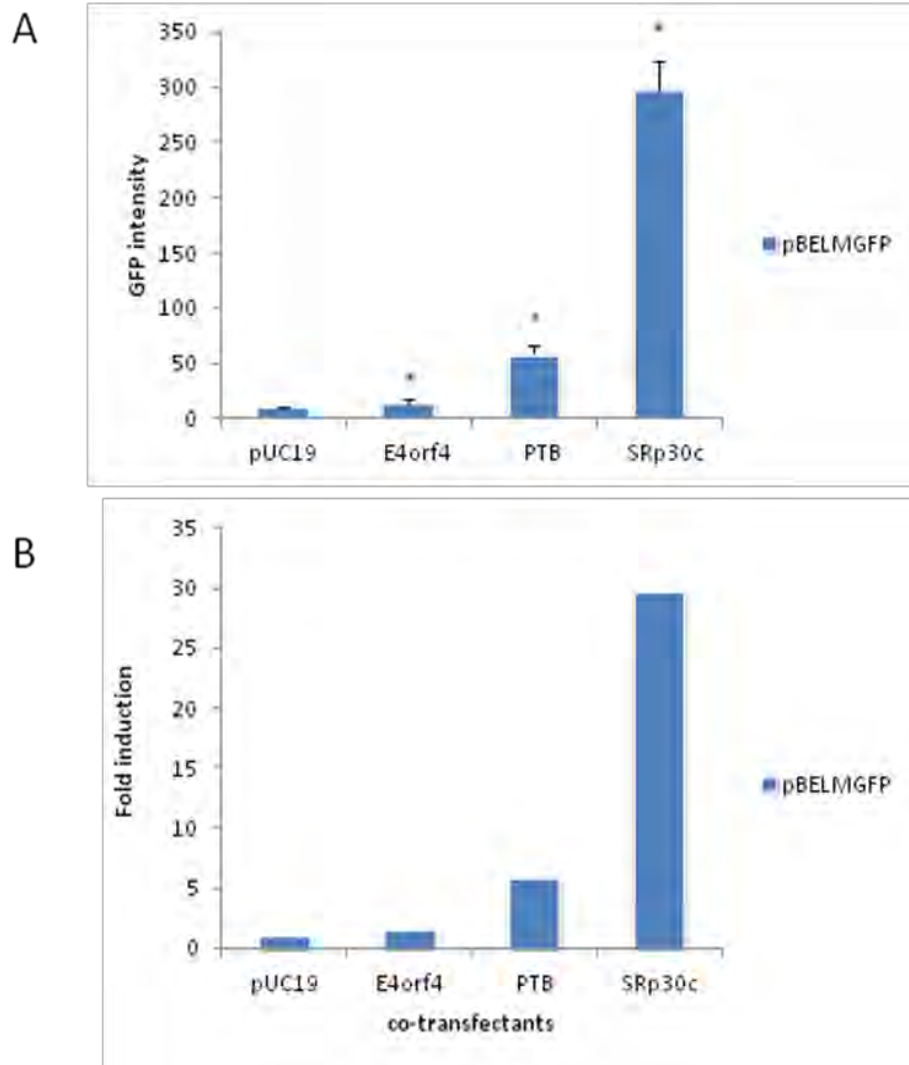


Figure 3.26 Determination of GFP levels in pBELMGFP transient transfection

(A) GFP intensities in transfection of pBELMGFP reporter plasmid into HeLa cells in the absence or presence of cotransfected expression vectors. (B) Induction of GFP expression by E4orf4, PTB and SRp30c in pBELMGFP. GFP fluorescence intensities were detected and quantitated in 10 randomly chosen cells as described in material and methods. Cells were transfected with each plasmid in a minimum of three independent experiments. Error bars indicate standard error for each sample. * indicates significance in relation to pBELMGFP alone ($p < 0.05$).

In order to detect higher levels of GFP expression HeLa cells were transfected with plasmids pBELMGFP and E4orf4, PTB or SRp30c and cells were analyzed 48 hours posttransfection. Still low GFP expression was detected from cells transfected with pBELMGFP. A moderate increase of GFP levels was observed in cells transfected with pBELMGFP and PTB or SRp30c as seen in previous experiment (Fig. 3.27).

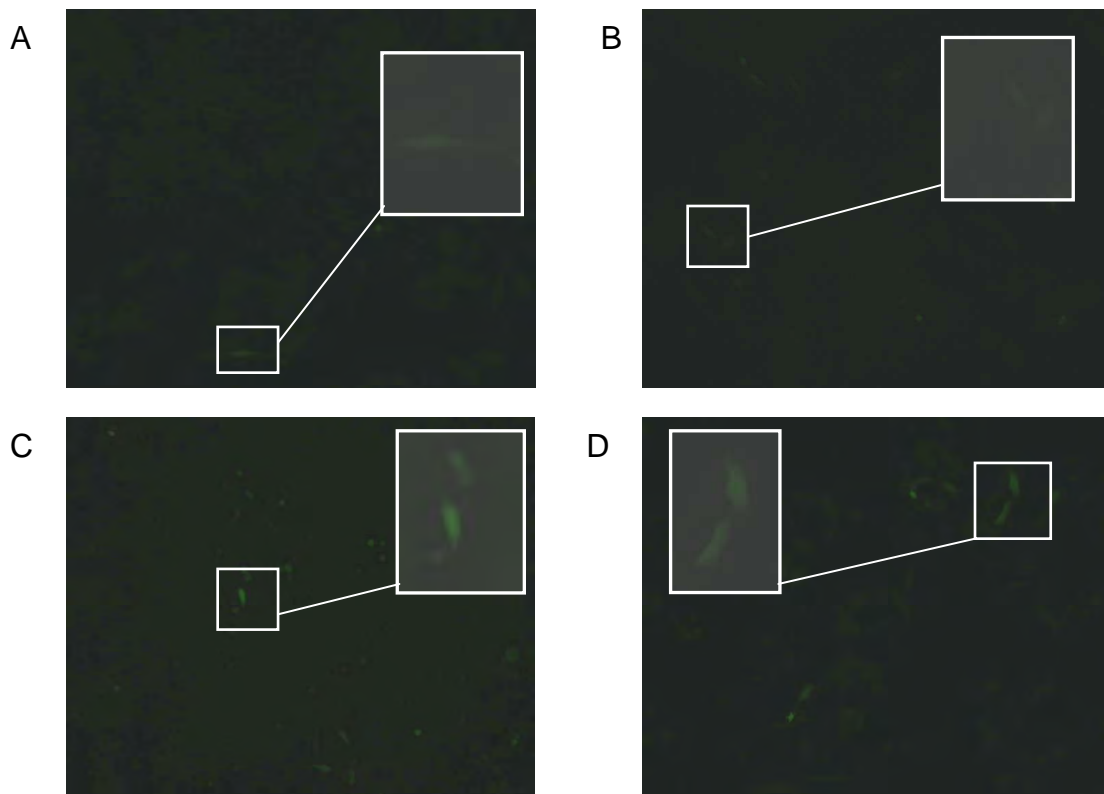


Figure 3.27 Examination of GFP expression after 48 hours posttransfection

HeLa cells were transfected with (A) pBELMGFP reporter plasmid and (B) pBELMGFP + E4orf4 (C) pBELMGFP + PTB (D) pBELMGFP + SRp30c.

3.5 Determination of SEAP levels in pBELSEAP and pBELMSEAP

The secreted embryonic alkaline phosphatase (SEAP) assay is a very simple method to analyze gene expression since the SEAP gene product is secreted from mammalian cells and therefore it is easily detected in a sample of culture medium. pBELSEAP, pBELMSEAP (Fig. 3.28) and a CMV driven SEAP expression plasmids were transfected into HeLa cells. Analysis of SEAP expression was performed after 24 and 48 hours post-transfection. Very high background levels were observed and only SEAP expression from the CMVSEAP was detected (Fig. 3.29).

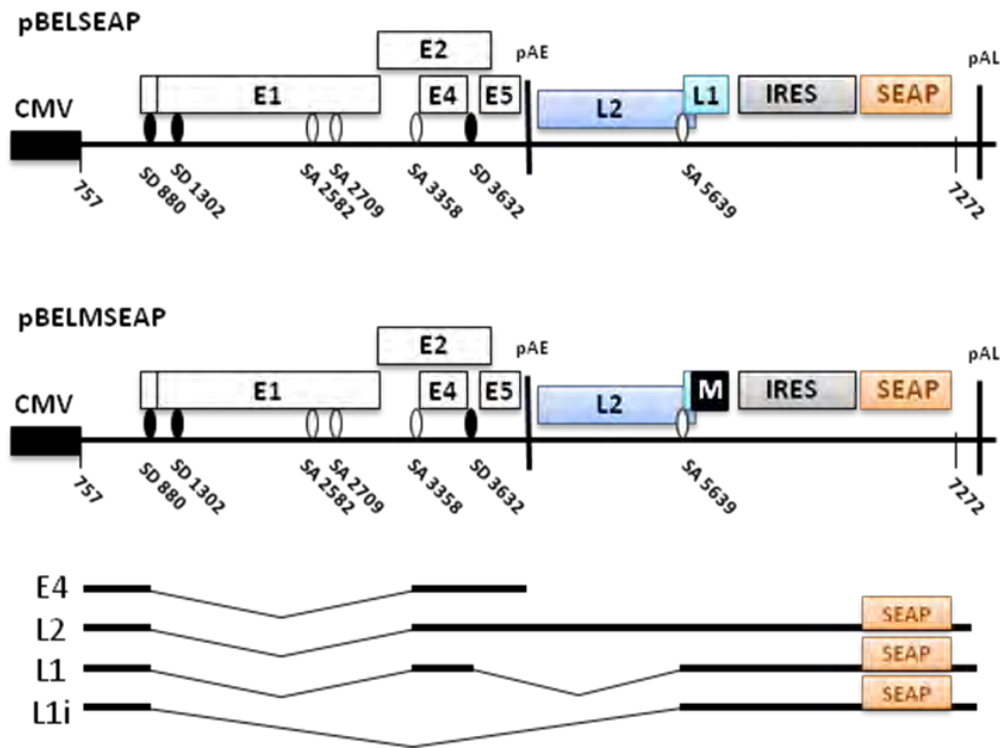


Figure 3.28 Schematic representation of subgenomic expression plasmids pBELSEAP and pBELMSEAP

Numbers refer to nucleotide positions of 5' splice sites (black circles), 3' splice sites (white circles). The early and late polyadenylation signals pAE and pAL are indicated.

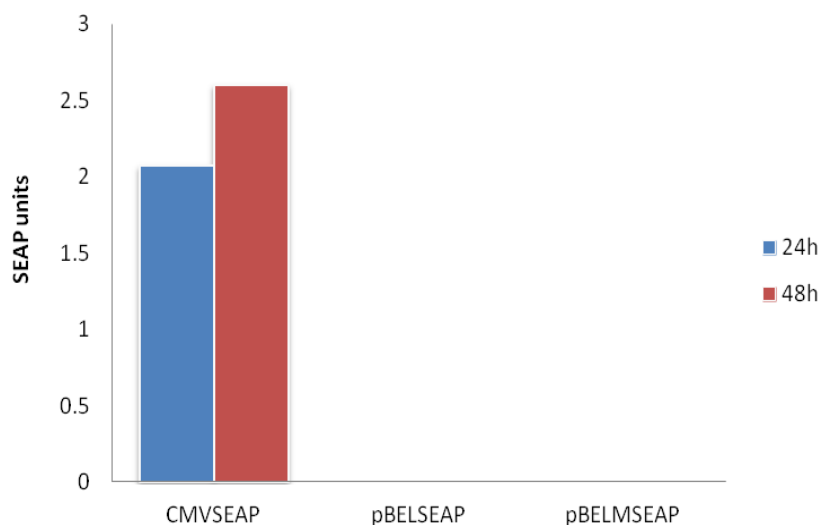


Figure 3.29 Determination of SEAP levels in SEAP reporter plasmids transient transfection
 Analysis of SEAP expression in CMVSEAP, pBELSEAP and pBELMSEAP 24 and 48 hours post-transfection.

The fetal calf serum (FCS) present in the medium may contain alkaline phosphatase (AP) that can interfere with AP quantification, therefore, in order to optimize the assay, samples from cultured HeLa cells were analyzed after 24 and 48 hours at room temperature or after a heat-treatment at 65°C for 30 min to eliminate non-SEAP phosphatase activity (from the FCS used in the culture medium) (Fig. 3.30 A). The secreted embryonic alkaline phosphatase (SEAP) is insensitive to heat, in contrast to the endogenous alkaline phosphatase (AP). To evaluate the presence of alkaline phosphatase in the fetal calf serum, samples of only medium containing serum and serum-free medium were also tested at room temperature or after heat-treatment at 65°C for 30 min (Fig. 3.30 B). The heat-treatment at 65°C reduced the background level, but after 48 hours of incubation it resulted still very high. The second experiment demonstrated that the medium without serum reduces the background levels when compared with the medium containing serum, but heat-treatment had any effect on it. However the serum in the medium is

required for the growth of the cell thus the heat-treatment proved to be the best way to reduce the background level.

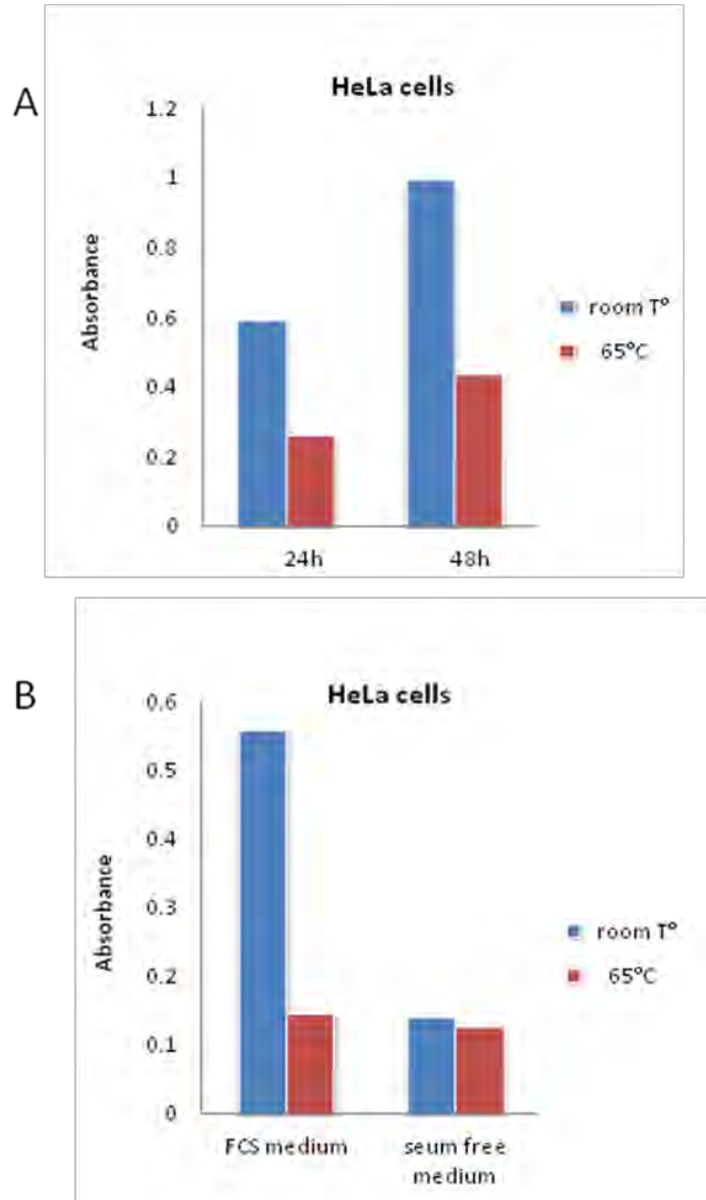


Figure 3.30 Optimization of SEAP assay

(A) HeLa cells were incubated at 37°C for 24 and 48 hours and the medium samples were analyzed at room temperature and after a heat-treatment at 65°C for 30 min. (B) Analysis of the medium in the presence or absence of serum at room temperature or after heat-treatment at 65°C for 30 min.

Next, a transfection of the SEAP reporter plasmids in the presence and in the absence of E4orf4 expression plasmid was carried out and the samples were pre-heated at 65°C but no detection of SEAP expression was observed (Fig. 3.31). In this figure it is possible to see the equivalent production of SEAP from HeLa cells and the SEAP reporter plasmids. Only the CMVSEAP plasmid showed to produce SEAP however, SEAP levels revealed to be much lower when compared with the levels of CAT produced by the CMVCAT. In conclusion we were not able to detect any SEAP expression from SEAP reporter plasmids due probably to the high background levels.

Since the CAT assay proved to be the most sensitive, reliable and suitable for large scale screening, this reporter was chosen for the generation of the stable cell lines and continued study.

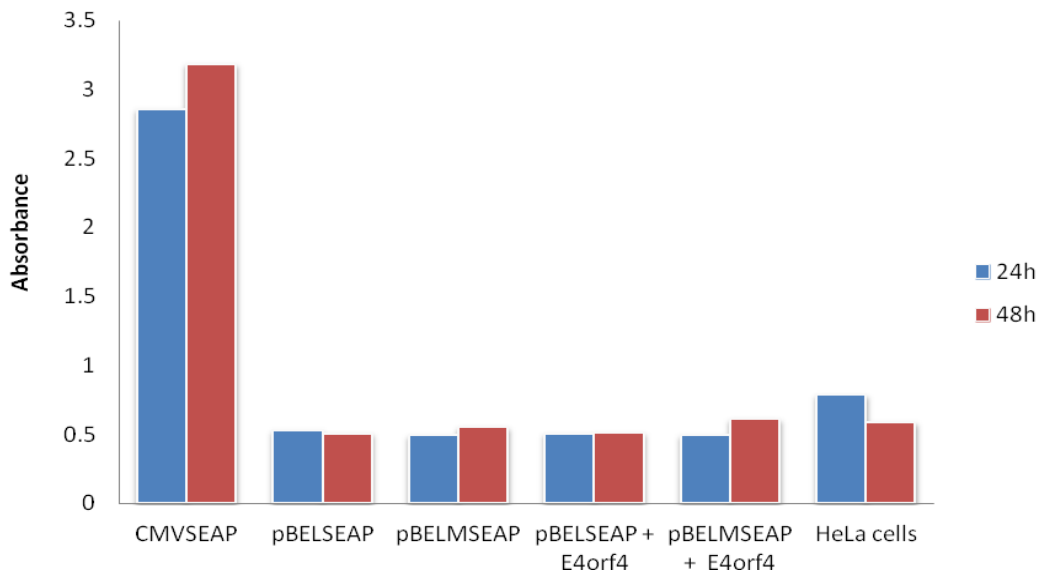


Figure 3.31 SEAP levels of SEAP reporter plasmids transiently transfected in HeLa cells

Analysis of SEAP levels in CMVSEAP, pBELSEAP, pBELMSEAP, pBELSEAP + E4orf4, pBELMSEAP + E4orf4 and HeLa cells pre-heated at 65°C for 30 min, 24 and 48 hours post-transfection. Cells were transfected with each plasmid in three independent experiments.

3.6 Identification and characterization of cis-acting regulatory RNA elements in HPV-16 CAT and SEAP reporter plasmids

Two viral splice sites are used exclusively by HPV-16 late mRNAs: late 5'-splice site SD3632 (located between the E4 and E5 coding regions) and late 3'-splice site SA5639 (located in the L1 coding region). Both splice sites are suppressed by multiple splicing silencer elements (Schwartz *et al.*, 2007). To further investigate how the reporter plasmids presented here could be used to study HPV-16 late gene expression, a set of simpler plasmids were generated containing the late region with the CAT reporter gene but only the two late splice sites SD3632 and SA5639, named pBspliceCAT and pBspliceMCAT (Fig. 3.32). Transfection of these plasmids into HeLa cells revealed that pBspliceCAT failed to produce detectable levels of CAT whereas pBspliceMCAT expressed barely detectable levels of CAT (1.3 CAT units) (Fig. 3.33 A). These results confirmed that SD3632 and SA5639 are efficiently suppressed in proliferating cancer cells. Transfection of pBspliceMCAT in the presence of plasmids expressing E4orf4, PTB or SRp30c resulted in an induction of CAT production, but this induction was lower than the induction of CAT from pBELMCAT by the same proteins (Fig. 3.33 B). In contrast, deletion of a previously described sequence (Rush *et al.*, 2005; Somberg *et al.*, 2008) that suppressed SD3632 as in plasmid pMt1sdCAT (Fig. 3.32), resulted in efficient production of high levels of CAT (155 CAT units) (Fig. 3.33 A), demonstrating that this set of plasmids could be used to study cis-acting, splicing regulatory elements at the HPV-16 late splice sites.

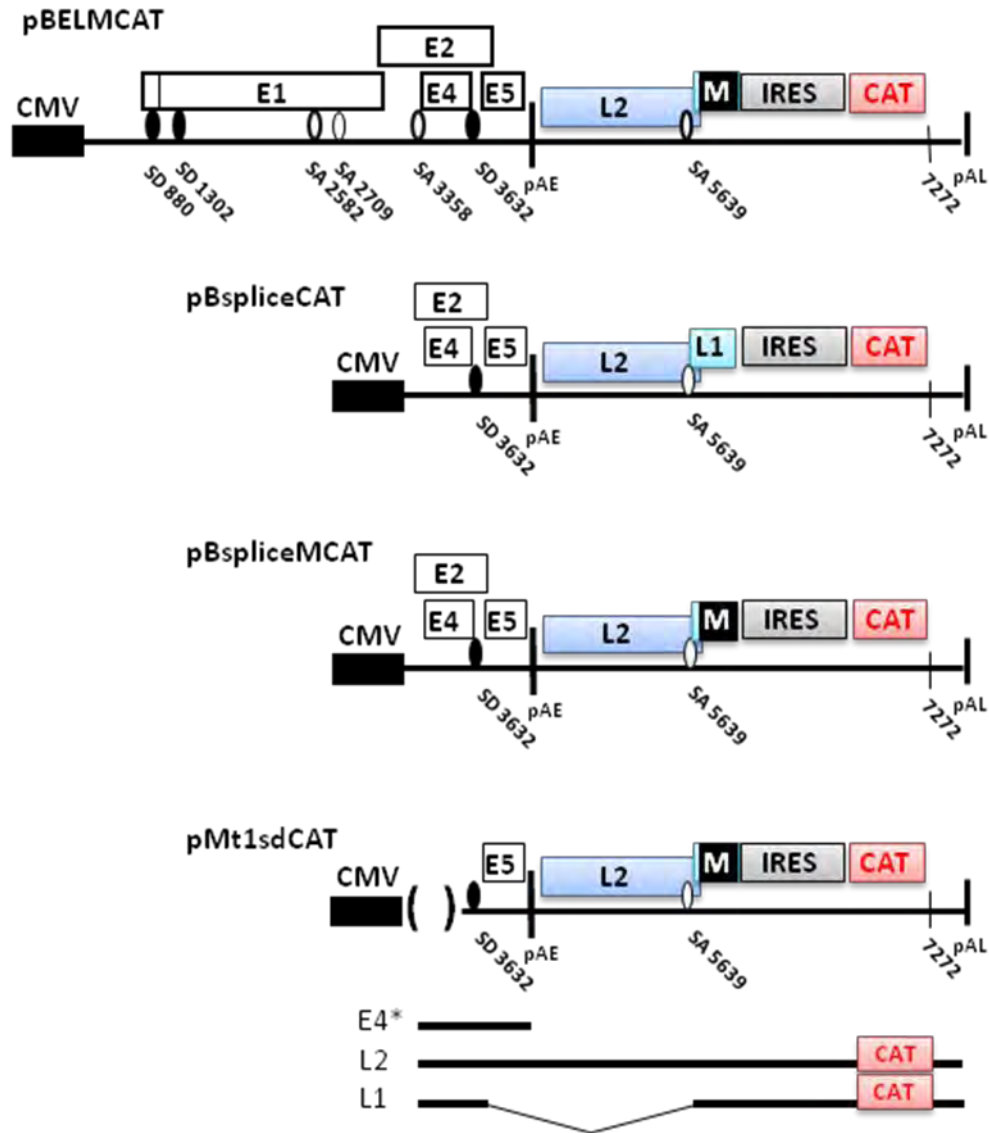


Figure 3.32 Schematic representation of subgenomic expression plasmids pBspliceCAT, pBspliceMCAT and pMt1sdCAT

Numbers refer to nucleotide positions of 5' splice sites (black circles), 3' splice sites (white circles). The early and late polyadenylation signals pAE and pAL are indicated. Major potential mRNAs produced by pBspliceCAT, pBspliceMCAT and pMt1sdCAT are indicated below. E4* refers to a short mRNA encoding E4 and E5 that is unspliced due to the absence of SD880 and SA3358.

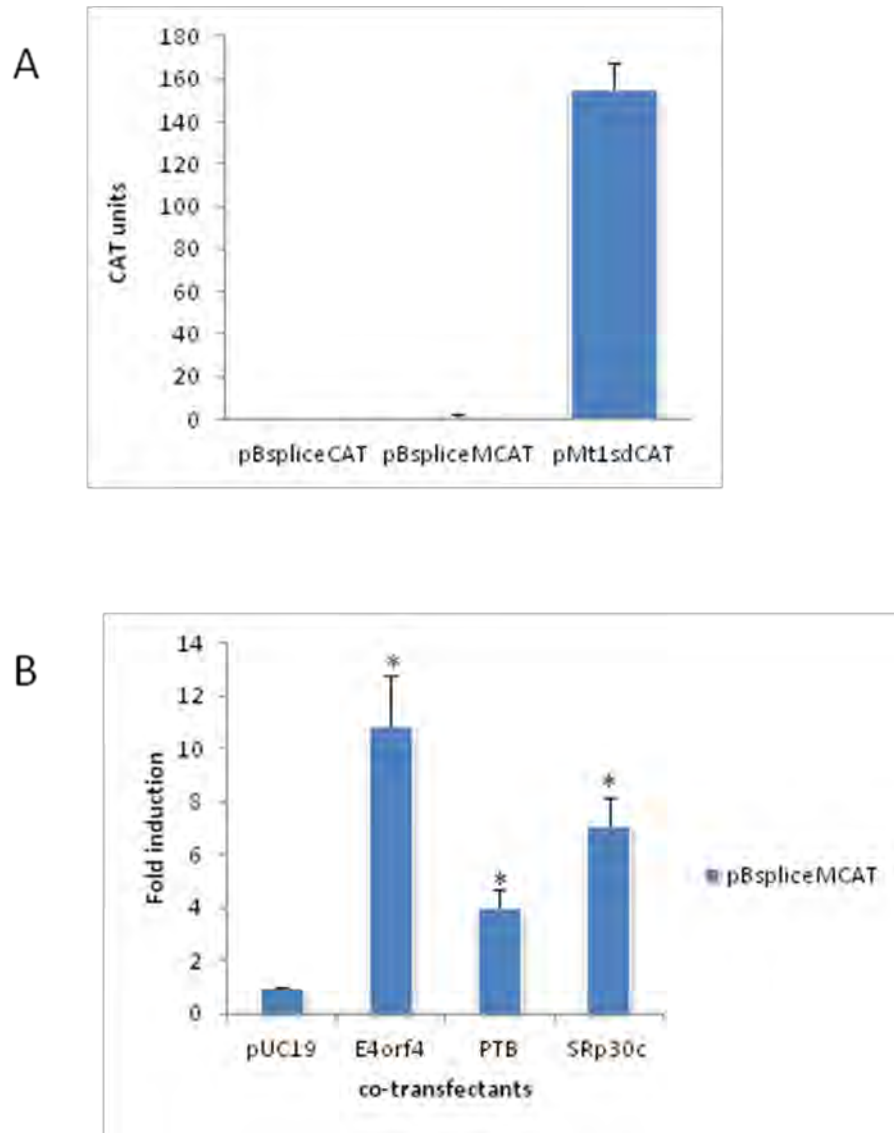


Figure 3.33 Determination of CAT levels in CAT reporter plasmids

(A) pBspliceCAT, pBspliceMCAT and pMt1sdCAT were transfected into HeLa cells and levels of CAT expression were determined 24h posttransfection. (B) pBspliceMCAT was transfected into HeLa cells in the presence and absence of plasmid expressing E4orf4, PTB and SRp30c. Transfections were carried out in triplicate. Error bars indicate standard error for each sample. * indicates significance in relation to pBspliceMCAT cotransfected with the empty vector (pUC19) ($p < 0.05$).

Since pMt1sdCAT showed to efficiently induce CAT expression, this experiment was reproduced with the SEAP reporter plasmid pMt1sdSEAP (Fig. 3.34). In previous experiments the SEAP expression was undetectable from pBELSEAP and pBELMSEAP due to high background levels but as shown in figure 3.35 deletion of sequences upstream SD3632 (pMt1sdSEAP), resulted in a modest production of SEAP (0.5 SEAP units). This result demonstrated that these sequences upstream SD3632 strongly suppress this splice site therefore inhibiting late gene expression. In conclusion pMt1sdSEAP plasmid could be used to study cis-acting, splicing regulatory elements at the HPV-16 late splice sites.

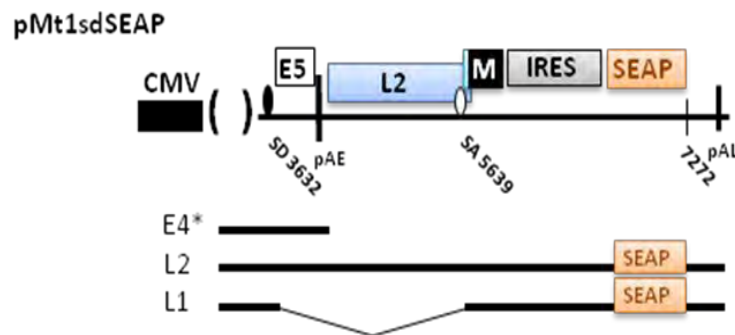


Figure 3.34 Schematic representation of subgenomic expression plasmid pMt1sdSEAP

Numbers refer to nucleotide positions of 5' splice sites (black circles), 3' splice sites (white circles). The early and late polyadenylation signals pAE and pAL are indicated. Major potential mRNAs produced by pMt1sdSEAP are indicated below. E4* refers to a short mRNA encoding E4 and E5 that is unspliced due to the absence of SD880 and SA3358.

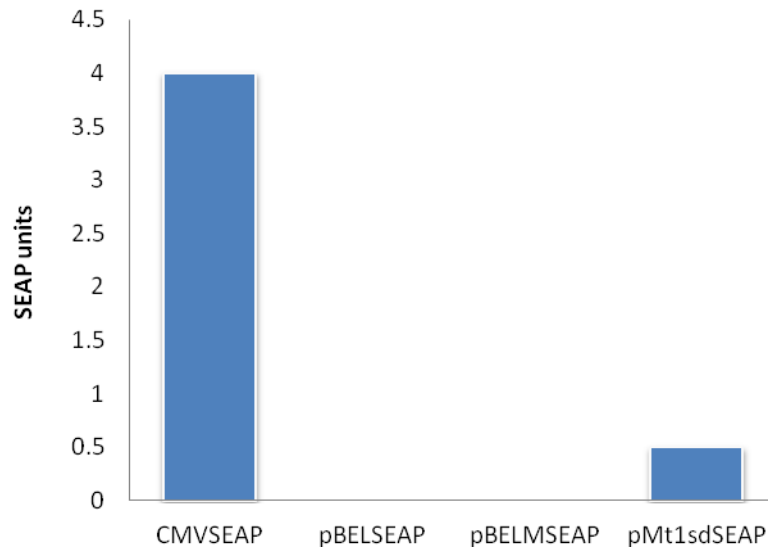


Figure 3.35 Determination of SEAP levels in SEAP reporter plasmid

CMVSEAP, pBELSEAP, pBELMSEAP and pMt1sdSEAP were transfected into HeLa cells and levels of SEAP expression were determined 24h posttransfection. Transfections were carried out in triplicate.

4. Chapter 4 – Results

Generation and characterization of stable cell lines

The capability to integrate genes into the full DNA sequence of a mammalian cell has a significant impact on biomedical research. Stable transfection of mammalian cells is a widely used technique to study gene expression and it ensures long term gene expression.

4.1 Establishment of two novel HPV-16 reporter cell lines that can serve as tools to investigate regulation of HPV-16 gene expression

Having demonstrated the use of transient transfections of the CAT reporter constructs, it was decided to generate stable cell lines of the pBELCAT and pBELMCAT constructs to aid in large scale screening. To establish stable cell lines containing the reporter plasmids pBELCAT or pBELMCAT, these plasmids were separately introduced into HeLa cells in the presence of plasmid pSVneo, which encodes the neomycin-resistant gene under control of the SV40 promoter. The transfected cells were propagated in G418 to select for cells with integrated plasmid DNA. More than 100 neomycin-resistant colonies were picked and DNA was first extracted and PCR amplified with CAT and IRES set primers (Table 2.2) (Fig. 4.1). The level of CAT expression of each positive clone was then monitored with a CAT ELISA assay. Five positive clones for pBELCAT (pBELCAT-24, 51, 102, 47 and 67) and two for pBELMCAT (pBELMCAT-31 and 25) stable cell lines expressed CAT. As shown in figure 4.2, CAT expression was barely detectable in pBELCAT cell lines (between 0.002 and 0.3 CAT units). In contrast, significant levels of CAT were detected in both pBELMCAT-derived cell lines (Fig. 4.2). In particular one, pBELMCAT-31, produced more than 500 CAT units. These results

demonstrated that pBELCAT- and pBELMCAT-derived stable cell lines were functional and expressed detectable levels of CAT. The expression of CAT from the stable cell lines mimicked that seen in transient transfections using pBELCAT and pBELMCAT and also of HPV-16 late genes confirm the suitability of the cell lines for further study.

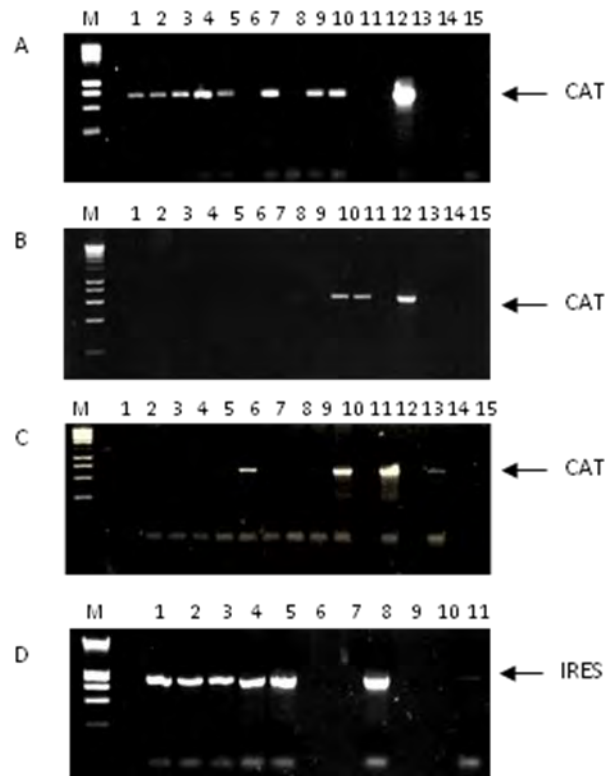


Figure 4.1 PCR amplification of DNA extracted from pBELCAT and pBELMCAT stable cell lines using CAT or IRES set primers

(A) 1KB DNA ladder (M), lanes 1-11 PCR amplification of DNA extracted from pBELCAT stable cell lines using CAT primers, lane 13 positive control, lane 15 blank. (B) 1KB DNA ladder (M), lanes 1-11 PCR amplification of DNA extracted from pBELCAT stable cell lines using CAT primers, lane 13 positive control, lane 15 blank. (C) 1KB DNA ladder (M), lanes 1-11 PCR amplification of DNA extracted from pBELMCAT stable cell lines using CAT primers, lane 14 positive control, lane 15 blank. (D) 1KB DNA ladder (M), lanes 1-5 PCR amplification of DNA extracted from pBELCAT and pBELMCAT stable cell lines using IRES primers, lane 8 positive control, lane 11 blank.

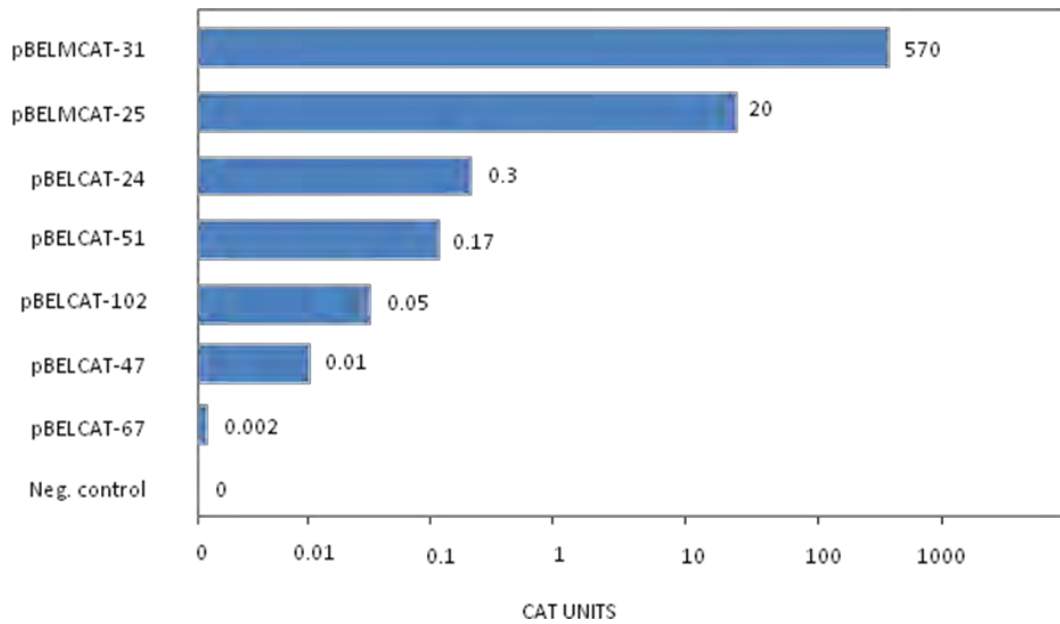


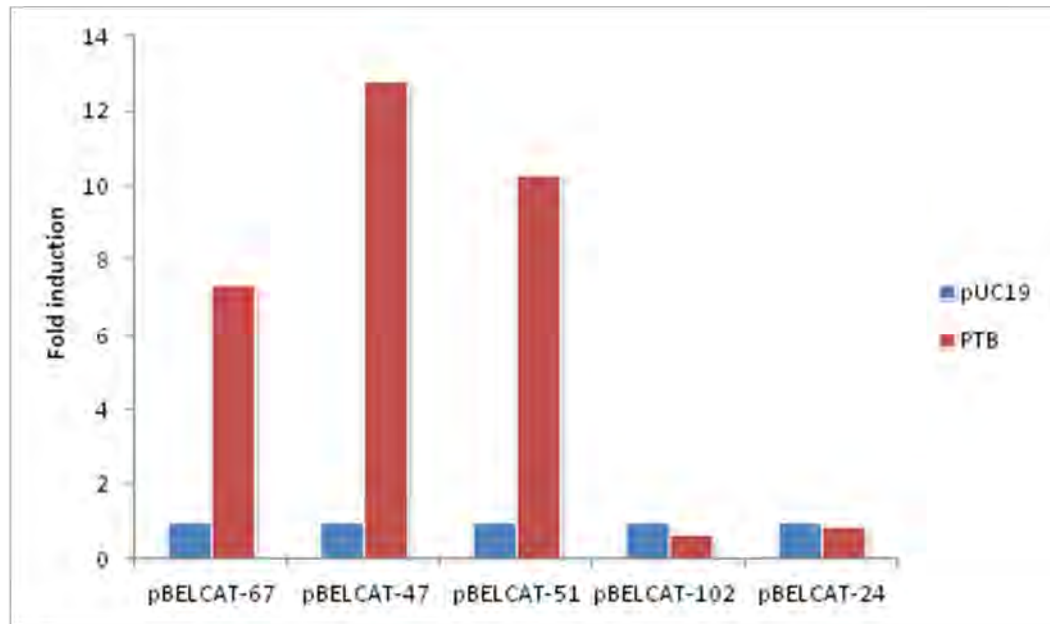
Figure 4.2 Determination of CAT levels in stable transfection

CAT levels were determined from pBELCAT- and pBELMCAT-derived cell lines. 5 pBELCAT stable cell lines produced little CAT, and 2 pBELMCAT produced high levels of CAT. Logarithmic scale.

4.2 Analysis of overexpression of E4orf4, PTB and SRp30c in pBELCAT and pBELMCAT stable cell lines

To evaluate the functionality of these novel stable cell lines we wished to examine if they responded to inducers of late gene expression in a similar manner to HPV-16 L1 and to the transient transfections of the reporters pBELCAT and pBELMCAT. pBELCAT- and pBELMCAT-derived cell lines were screened for functionality by transfecting them with adenovirus E4orf4, PTB or SRp30c expression plasmids (Fig. 4.3). CAT expression was observed in three pBELCAT-derived cell lines (pBELCAT-51, -47 and -67) whereas the other two (pBELCAT-24 and -102) showed no induction of CAT expression (Fig. 4.3 A). Overexpression of E4orf4, PTB and SRp30c resulted in an increase of CAT expression in both pBELMCAT-derived cell lines (Fig. 4.3 B). In particular pBELMCAT-31, which showed really high level of CAT when untransfected, in the presence of the viral or cellular proteins showed an induction of 2.6-, 1.7- and 2.2-fold respectively. These results showed a lower induction of CAT compared to the transient transfections, which could be explained in two ways: either the levels of CAT produced by the pBELMCAT-derived cell line were too high and expression could only be modestly enhanced, or low transfection efficiency resulted in relatively fewer CAT producing cells transfected with adenovirus E4orf4, PTB or SRp30c, resulting in a relatively low increase in CAT. In order to demonstrate that CAT expression is dependent on the level of these proteins a titration up to 4 μ g of E4orf4 and/or PTB was carried out in each stable cell line. Three pBELCAT cell lines (pBELCAT-51, -47 and -67) showed an induction of CAT expression in a PTB dose-dependent manner, whereas CAT induction failed in the other two stable cell lines (pBELCAT-102 and -24) (Fig. 4.4). Induction of CAT in an E4orf4 dose-dependent manner was observed only in pBELCAT-47 and -67 but not in pBELCAT-51 (Fig. 4.5).

A



B

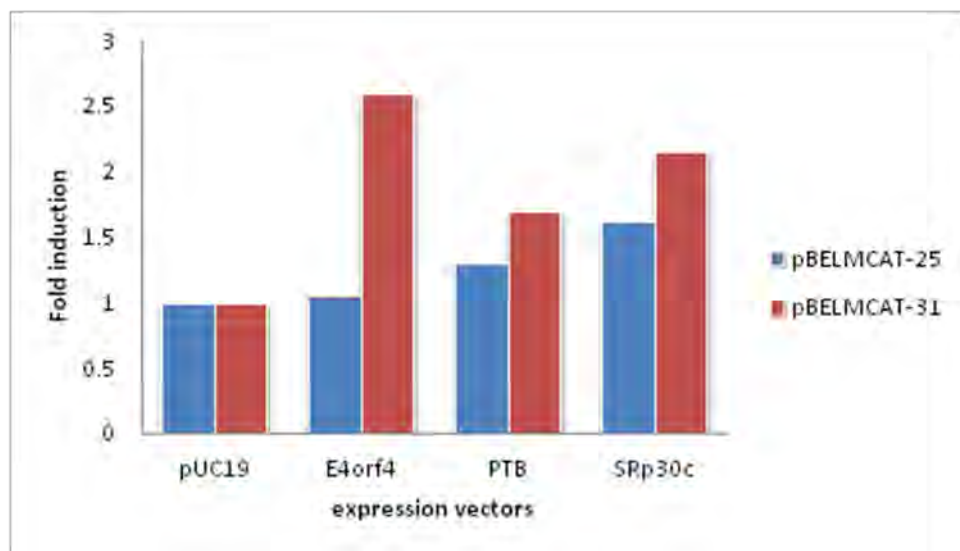


Figure 4.3 Induction of CAT expression in pBELCAT and pBELMCAT derived-cell lines

(A) Induction of CAT expression by overexpression of PTB expression plasmid in pBELCAT-derived cell lines. (B) Induction of CAT expression by overexpression of E4orf4, PTB and SRp30c expression plasmid in pBELMCAT-derived cell lines.

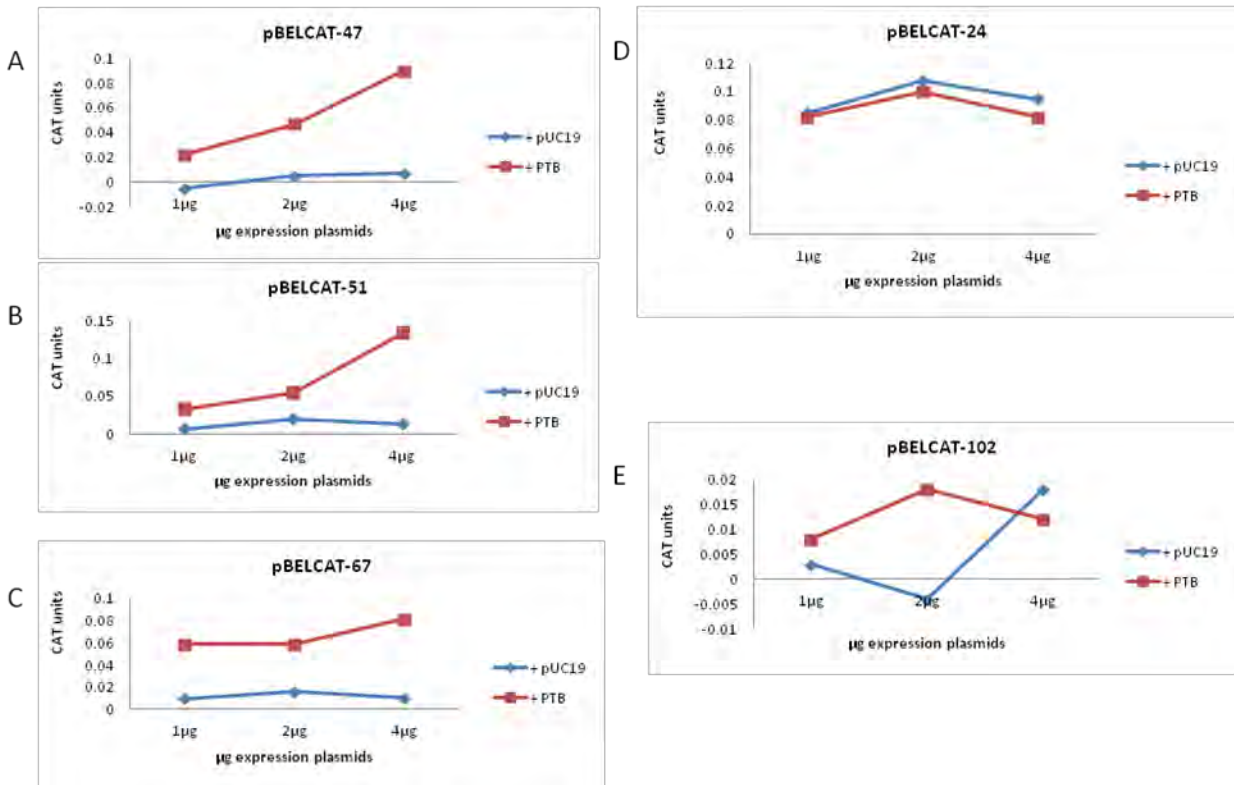


Figure 4.4 Titration of PTB expression plasmid in pBELCAT stable cell lines

Titration up to 4µg of PTB expression plasmid into (A) pBELCAT-47, (B) pBELCAT-51, (C) pBELCAT-67, (D) pBELCAT-24 and (E) pBELCAT-102 stable cell lines. pUC19 plasmid was used as an empty vector to equalize the concentration in each single transfection.

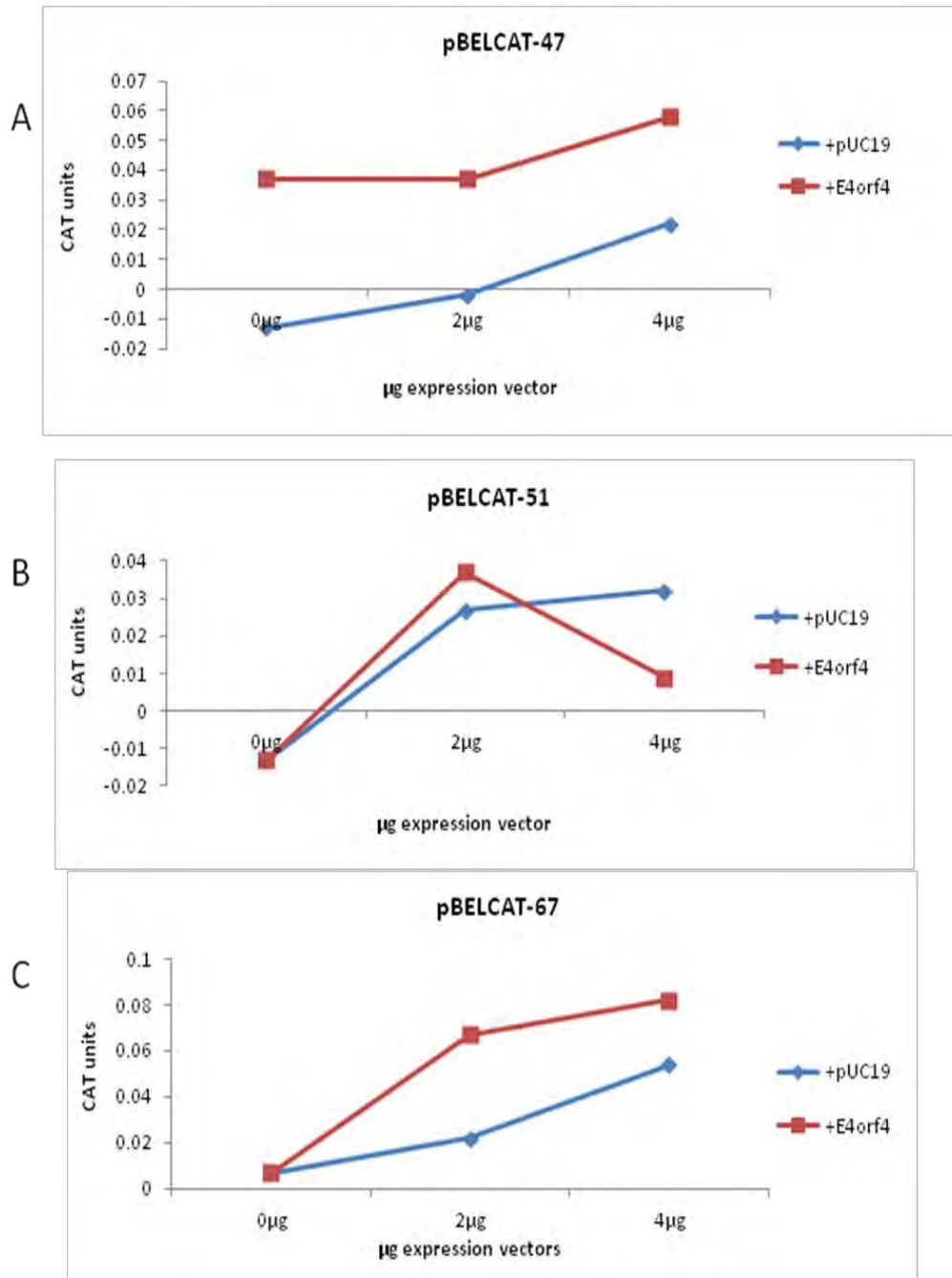
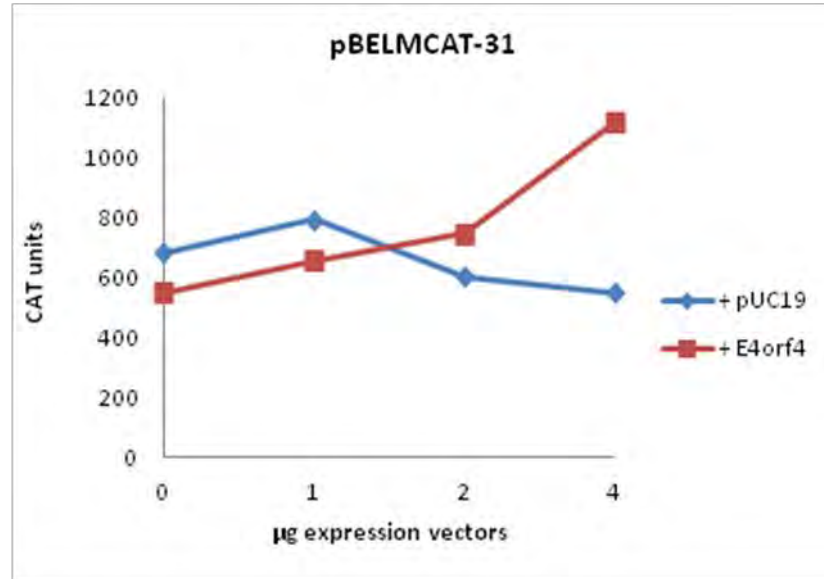


Figure 4.5 Titration of E4orf4 expression plasmid in pBELCAT stable cell lines

Titration up to 4μg of E4orf4 expression plasmid into (A) pBELCAT-47, (B) pBELCAT-51 and (C) pBELCAT-67 stable cell lines. pUC19 plasmid was used as an empty vector to equalize the concentration in each single transfection.

As pBELMCAT-31 produced higher levels of CAT when compared with pBELMCAT-25 and showed to respond well to the effect of the viral and cellular proteins, it was used for continued study. With a titration up to 4 μ g of E4orf4 or PTB plasmid into pBELMCAT-31 higher levels of CAT were observed at higher concentrations of transfected E4orf4 or PTB plasmid, demonstrating that CAT induction in pBELMCAT cell line was dependent on the amount of transfected plasmid (Fig. 4.6). These results also indicated that the relatively low induction of CAT in the stable cell lines compared to transient co-transfections with CAT reporter plasmids and E4orf4, PTB and SRp30c, was most likely the result of low transfection efficiency.

A



B

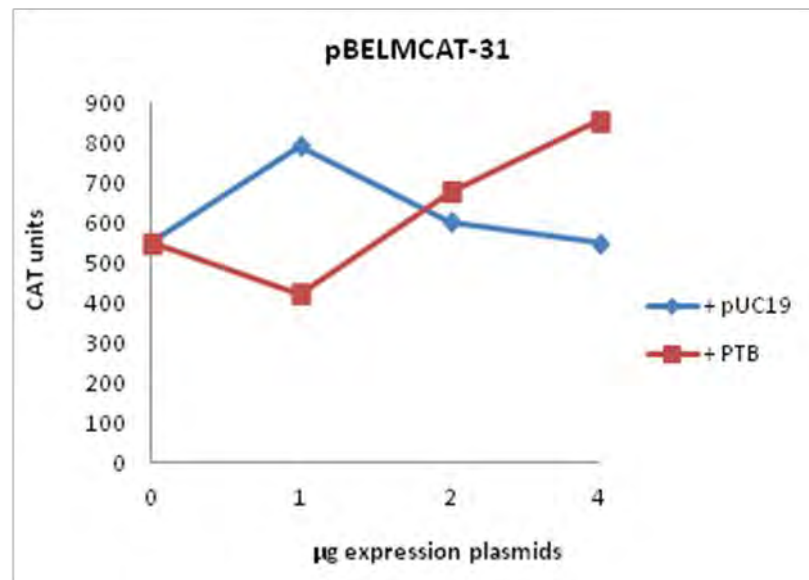


Figure 4.6 Titration of inducers of late gene expression in pBELMCAT stable cell lines

A titration up to 4µg of (A) E4orf4 or (B) PTB was carried out in pBELMCAT-31 derived-stable cell line. pUC19 plasmid was used as an empty vector to equalize the concentration in each single transfection.

Transfections with 4µg of E4orf4, PTB or SRp30c were then reproduced in pBELCAT-67 and pBELMCAT-31 stable cell lines in a minimum of three independent experiments and fold induction was calculated. As previously seen in transient transfection, higher levels of CAT were observed in the wild type-derived cell line (pBELCAT-67) in the presence of E4orf4 expression plasmid resulting in 8 fold induction but also PTB or SRp30c efficiently modulated CAT expression inducing 6 and 4 fold, respectively (Fig. 4.7). A lower induction was observed in pBELMCAT-31 in the presence of E4orf4, PTB or SRp30c, resulting in 1.7-, 1.6- and 1.5-fold, respectively (Fig. 4.7). The results indicate that pBELCAT and pBELMCAT stable cell lines respond to inducers of later gene expression in a manner similar to HPV-16 L1 and to that seen in transient transfections.

In conclusion, we have generated functional stable reporter cell lines that will be useful for analyzing HPV-16 late gene expression and for the identification of cellular factors that regulate HPV-16 gene expression. Much of the further work on the cell lines focussed on pBELCAT-67 and pBELMCAT-31.

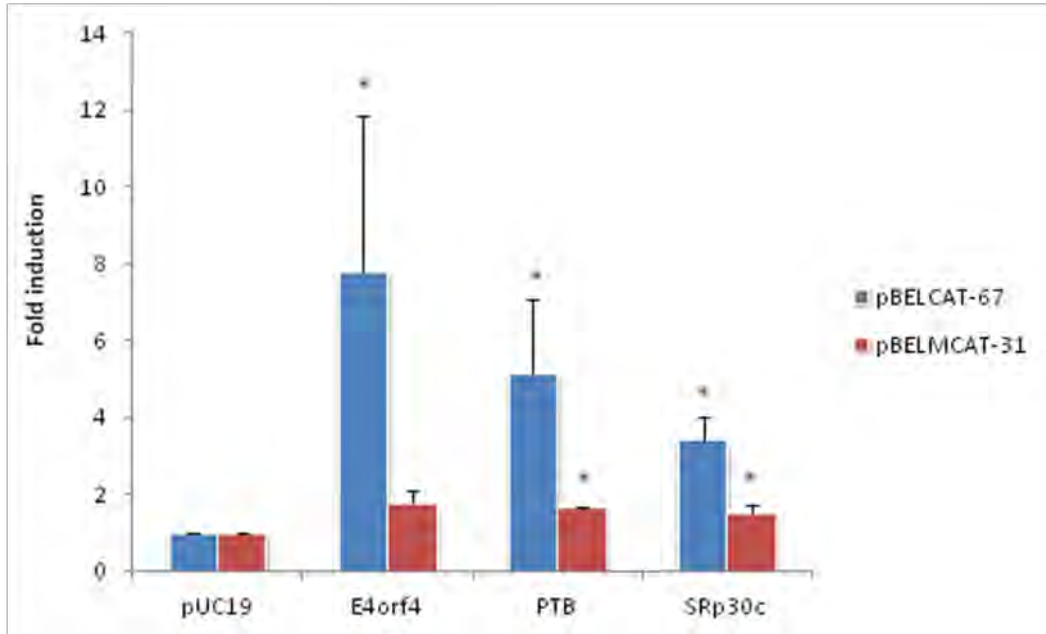


Figure 4.7 Induction of CAT expression in pBELCAT-67 and pBELMCAT-31 stable cell lines

Induction of CAT expression by overexpression of E4orf4, PTB and SRp30c expression plasmid in pBELCAT and pBELMCAT-derived cell lines. Cells were transfected with each plasmid in a minimum of three independent experiments. Error bars indicate standard error for each sample. * indicates significance in relation to pBELCAT-67 or pBELMCAT-31 transfected with the empty vector (pUC19) ($p < 0.05$).

4.3 RT-PCR analysis

To confirm that the integrated pBELCAT and pBELMCAT plasmids produced the expected early and late mRNAs, RNA was extracted and subjected to RT-PCR. cDNA was synthesized from RNA extracted from pBELCAT or pBELMCAT-derived cell lines using gene specific primers: E4a, L1a or L1Ma and CATa primers (Table 2.5). RT-PCR on cDNA from selected pBELCAT-67 cell line with primers 757s and E4a (Table 2.5) yielded a single 190bp-band, representing E4 mRNAs produced by splicing from SD880 to SA3358 (Fig. 4.8 A), whereas RT-PCR performed using 757s and L1a or 757s and CATa primers (Table 2.5), revealed that none of the late mRNAs could be detected, as expected (Fig. 4.8 A). cDNA synthesized with E4a primer from pBELMCAT-31 was then PCR amplified with primers 757s and E4a (Table 2.5) yielding a single 190bp-band, representing splicing from SD880 to SA3358 (Fig. 4.8 B), as expected. PCR on the cDNA synthesized with L1Ma using primers 757s and L1Ma (Table 2.5) resulted in two splice products; one 430bp cDNA representing the L1 mRNA that contains the sequence between SA3358 and SD3632 in the E4 coding region, and one 160bp cDNA representing the L1i mRNA that is directly spliced from SD880 to SA5639 (Fig. 4.8 C). cDNA synthesized with CATa primer was then PCR amplified with primers 757s and CATa primers (Table 2.5) yielding a major product at 2000bp. The 2000bp band resulted in a direct splice from SD880 to SA5639, representing the PCR product of roughly 500bp (L1) + 752bp (IRES) + 660 (CAT) (Fig. 4.8 D). These results demonstrated that correctly spliced HPV-16 early and late mRNAs were produced, corroborating the integrity of the reporter plasmids in the stable reporter cell lines.

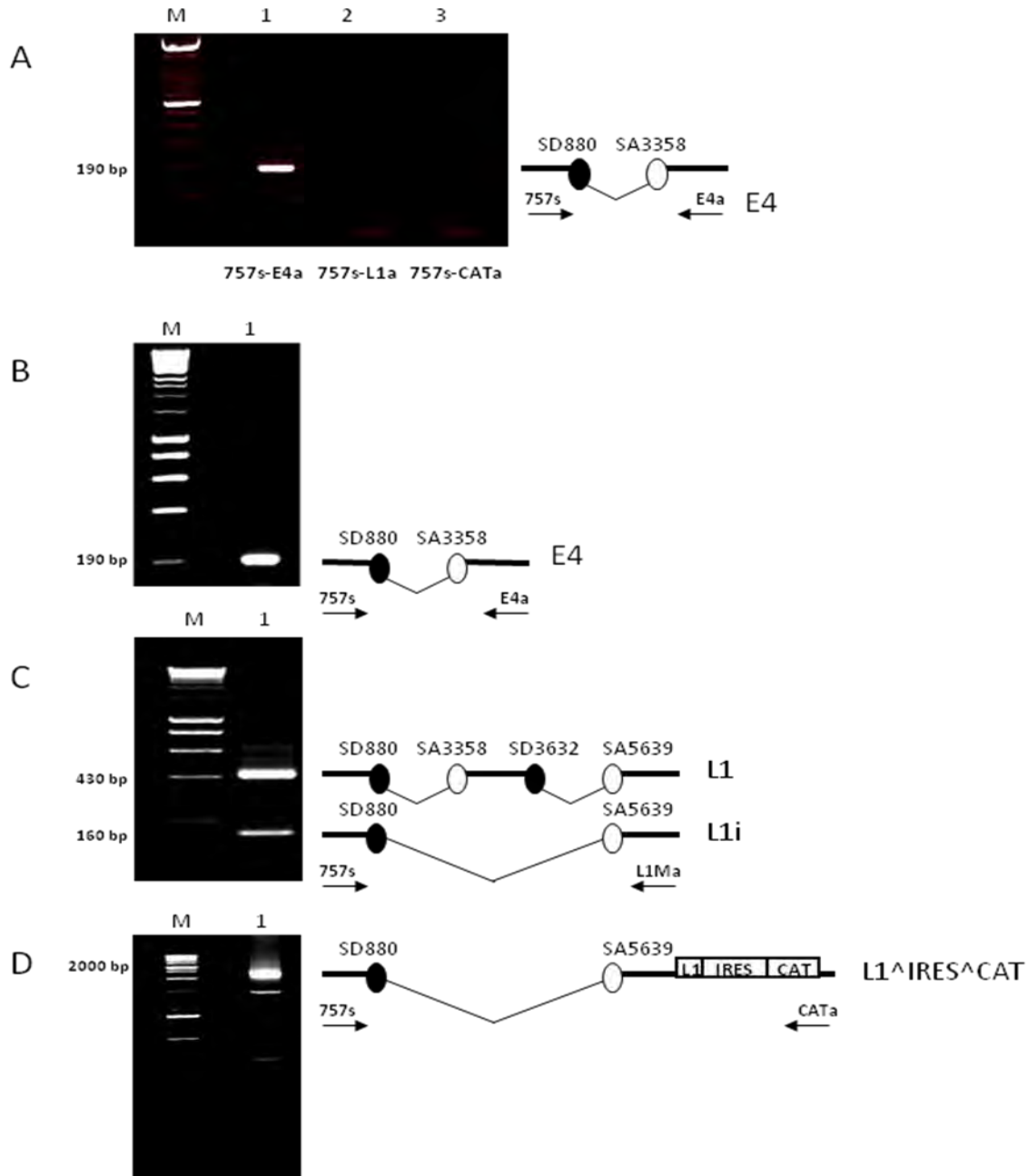


Figure 4.8 Characterization of pBELCAT-67 and pBELMCAT-31 stable cell lines

Analysis of mRNAs produced by stable cell lines pBELCAT-67 and pBELMCAT-31. (A) RT-PCR amplification of the early CMV→ E4 region with 757s-E4a primers and the late region CMV→L1 with 757s-L1a primers and CMV→CAT with 757s-CATa primers in pBELCAT-67. (B) RT-PCR amplification of the early CMV→ E4 region with 757s-E4a primers, (C) the late region CMV→L1 with 757s-L1Ma primers and (D) CMV→CAT with 757s-CATa primers in pBELMCAT-31.

4.4 Treatment of CAT reporter cell lines with small molecules

(i) TPA

In order to demonstrate that these cell lines could be used to identify small molecules that induce HPV-16 late gene expression the cells were treated with TPA (phorbol 12-myristate 13-acetate), a small molecule that has been shown to induce HPV-31 late gene expression (Meyers *et al.*, 1992). A number of pBELCAT or pBELMCAT cell lines were treated with different concentrations of TPA and an increase of CAT expression in a dose-dependent manner was observed in pBELCAT-67, pBELMCAT-31 and -25 stable cell lines, whereas the others cell lines, as an example pBELCAT-47 in figure 4.9, did not show any increase in CAT levels (Fig. 4.9), further validating the functionality of these stable cell lines.

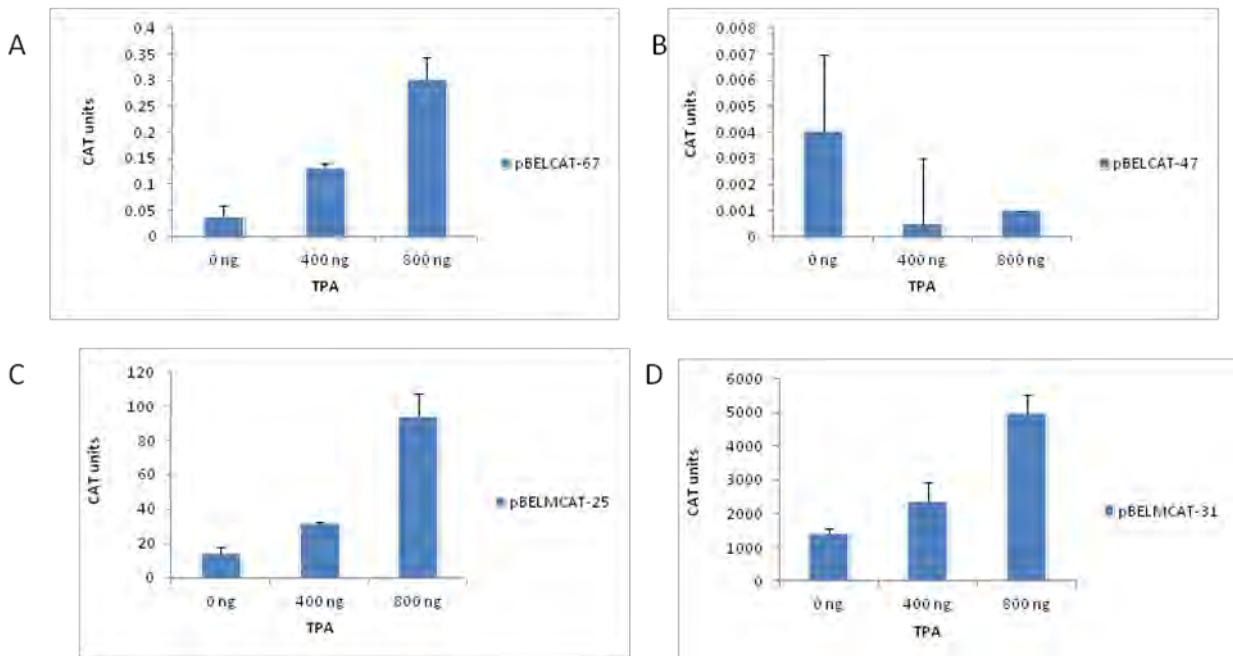


Figure 4.9 Treatment of pBELCAT- and pBELMCAT-derived cell lines with TPA

Induction of CAT by TPA in (A) pBELCAT-67 (B) pBELCAT-47, (C) pBELMCAT-25, (D) pBELMCAT-31 stable cell lines. Error bars indicate standard error for each sample.

(ii) Emetine, ceramide and cycloheximide

In order to identify other small molecules that could be used to treat HPV-16 infections we tested a number of substances which have been shown to induce apoptosis by interfering with splicing such as emetine, ceramide and cycloheximide (Boon-Unge *et al.*, 2007). First a time course experiment up to 6 hours was performed in pBELCAT-67 and pBELMCAT-31 stable cell lines but CAT induction was not observed at any time point in the presence of emetine, ceramide or cycloheximide (Fig. 4.10). A serial dilution of emetine and ceramide revealed that CAT expression is not affected by the presence of these substances at any concentration (Fig. 4.11).

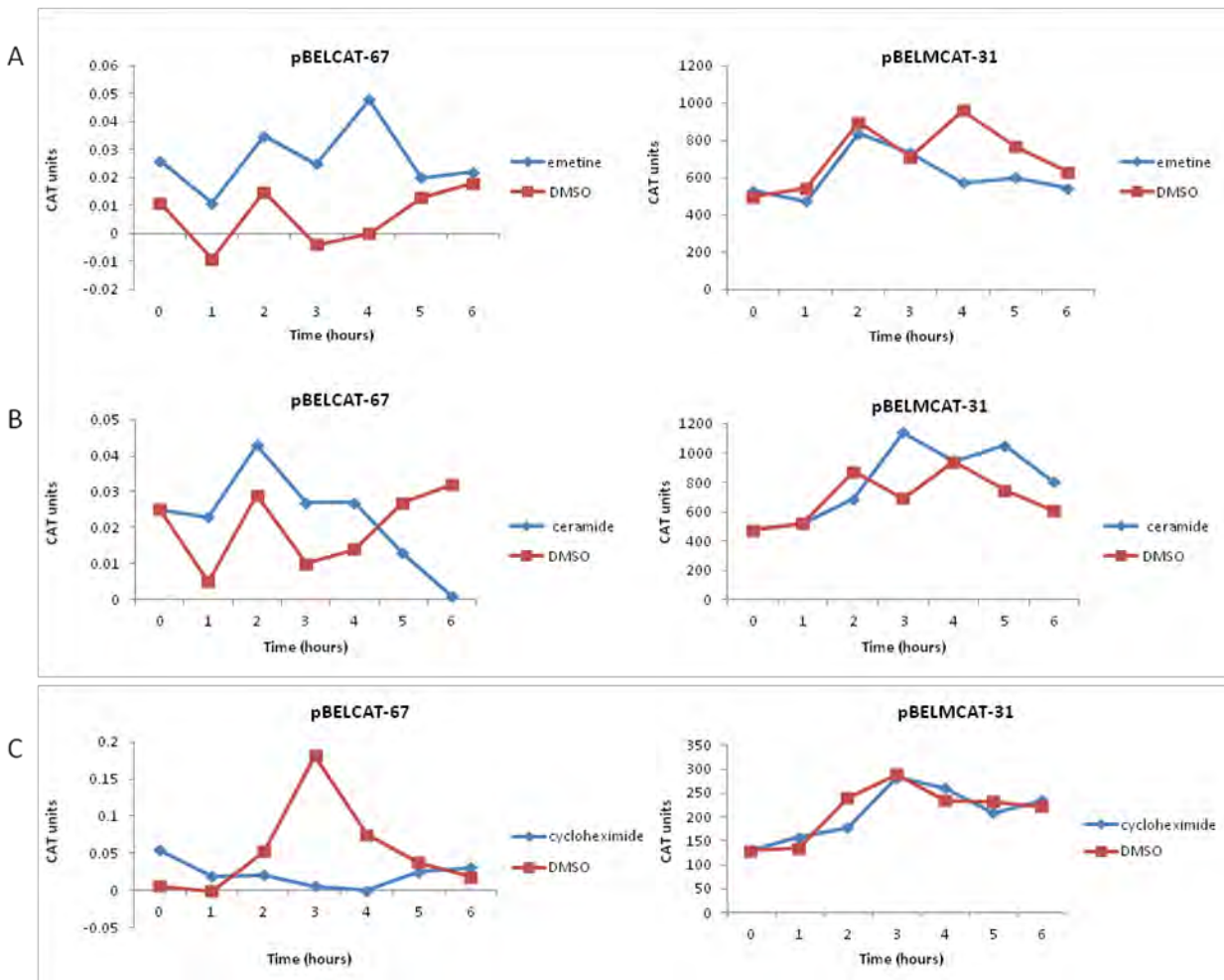


Figure 4.10 Treatment of CAT reporter cell lines with emetine, ceramide or cycloheximide
 Time course experiment of pBELCAT-67 and pBELMCA31 stable cell lines treated with (A) 10 μ M emetine, (B) 20 μ M ceramide or (C) 400 ng cycloheximide.

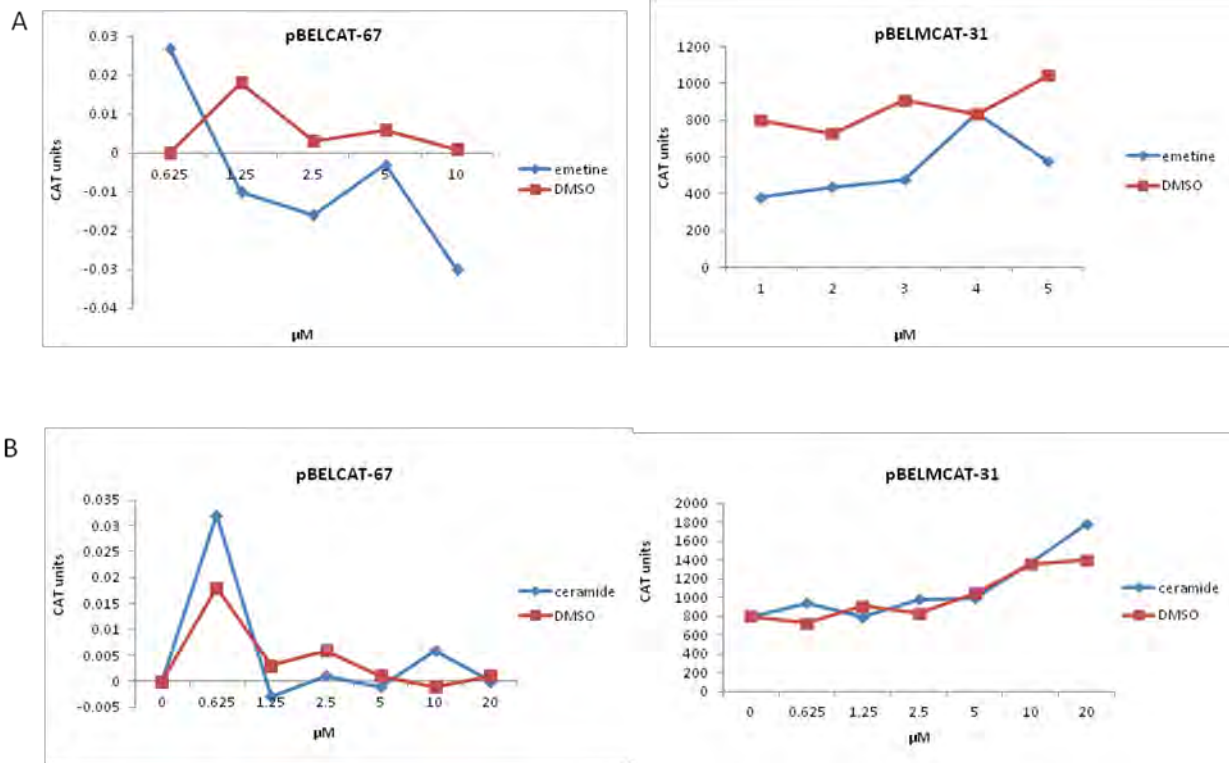


Figure 4.11 Serial dilution of emetine and ceramide

Serial dilution of (A) emetine or (B) ceramide in pBELCAT-67 and pBELMCAT-31 stable cell lines.

Treatment of pBELCAT-67 and pBELMCAT-31 with TPA, emetine, cycloheximide or ceramide at optimal concentrations was then reproduced and fold induction was calculated. Emetine and cycloheximide decreased CAT levels in both stable cell lines possibly due to their inhibition of protein synthesis. Any effect on CAT expression was observed in the presence of ceramide in both cell lines. Induction of CAT in a TPA-dose-dependent manner was consistent, resulting in an increase of CAT expression of 10 and 5 fold in pBELCAT-67 and pBELMCAT-31, respectively (Fig. 4.12 A). To control for possible effects of these substances on cellular protein synthesis, a western blot for actin was performed and demonstrated that they had no effect on actin levels (Fig. 4.12 B). It was concluded that induction of CAT from pBELCAT-67 and pBELMCAT-31 was specific for the small molecule TPA and that the cell lines can be used to study the effect of small molecules and drugs on late gene expression.

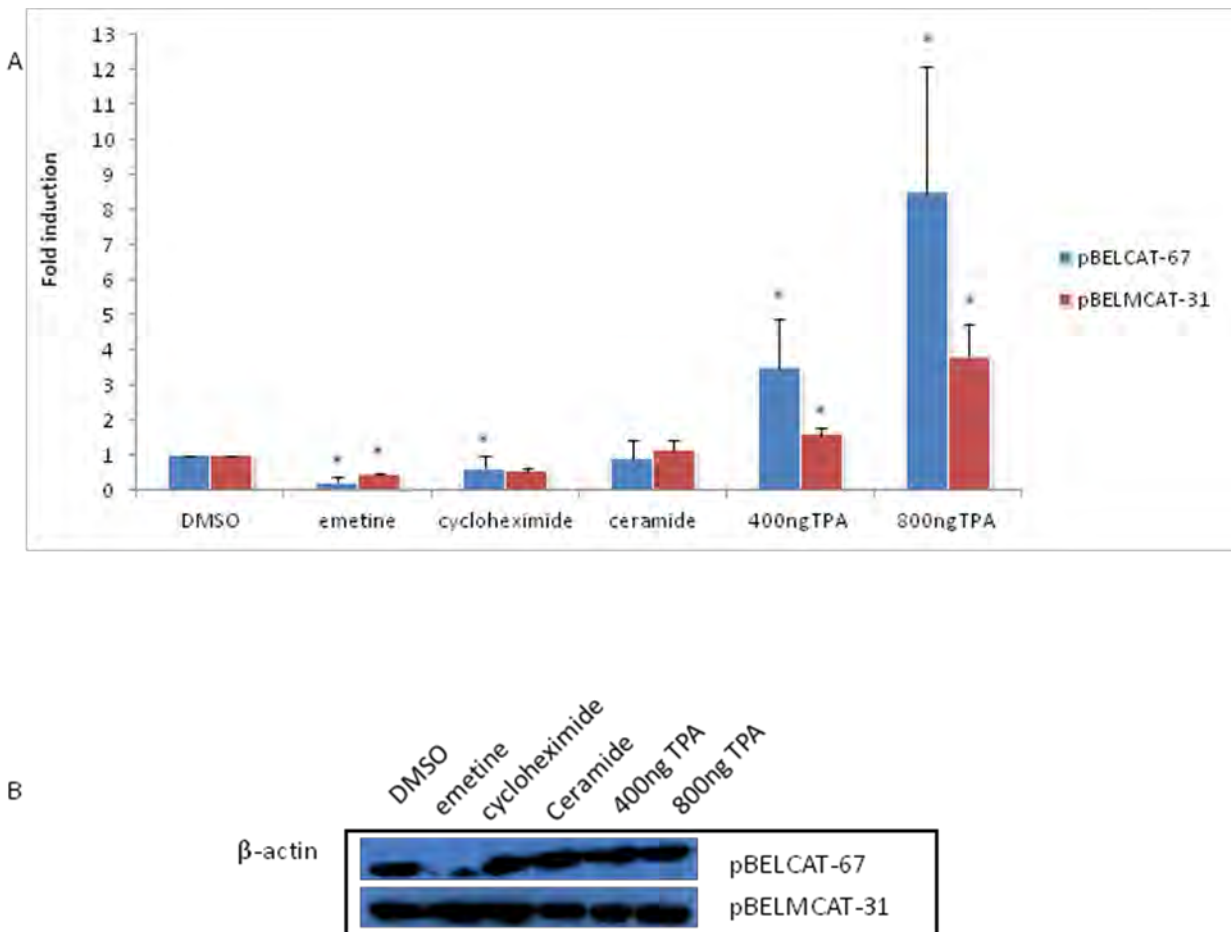


Figure 4.12 CAT induction in stable cell lines treated with TPA, emetine, cycloheximide or ceramide and Western blot analysis

(A) CAT induction in pBELCAT-67 and pBELMCAT-31 by TPA, emetine, cycloheximide or ceramide. Cells were treated with each small molecule in a minimum of three independent experiments. A control of 0.5 μ l DMSO was included. Error bars indicate standard error for each sample. * indicates significance in relation to pBELCAT-67 or pBELMCAT-31 untreated ($p < 0.05$). (B) Expression levels of β -actin were determined in pBELCAT-67 and pBELMCAT-31 cell lines in the absence and presence of TPA, emetine, cycloheximide or ceramide by Western Blot.

(iii) Valproic acid

Since it has been shown that valproic acid (VPA) increases the expression of ASF/SF2, which has previously shown to induce CAT levels, we wished to investigate if this small molecule can induce HPV-16 late gene expression. First a serial dilution of VPA was performed in pBELCAT-67 and pBELMCAT-31 cell lines resulting in a moderate increase of CAT at lower concentration of VPA in both cell lines (Fig. 4.13). Treatment of pBELCAT-67 and pBELMCAT-31 cell lines at optimal concentration of VPA resulted in a high induction of CAT in pBELCAT-67 whereas in pBELMCAT-31 CAT induction was only slightly enhanced (Fig. 4.14). These results demonstrated that VPA could be a candidate as a novel drug for treatment of HPV infections. Western blot confirmed an increase of ASF/SF2 protein expression by VPA and actin levels were unaffected in both cell lines (Fig. 4.15). The result showed also a slight increase of PTB levels in cells treated with valproic acid (Fig. 4.15).

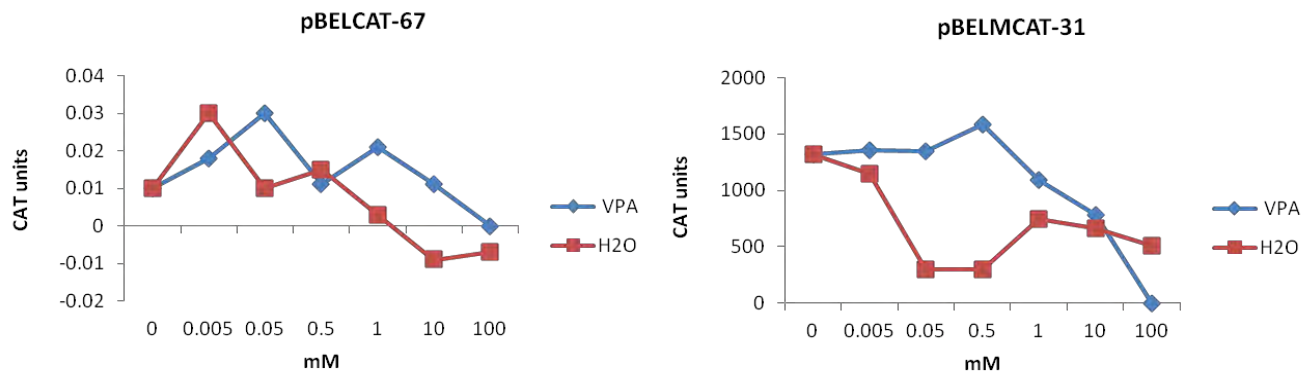


Figure 4.13 Serial dilution of valproic acid

pBELCAT-67 and pBELMCAT-31 stable cell lines were treated with different concentrations of valproic acid.

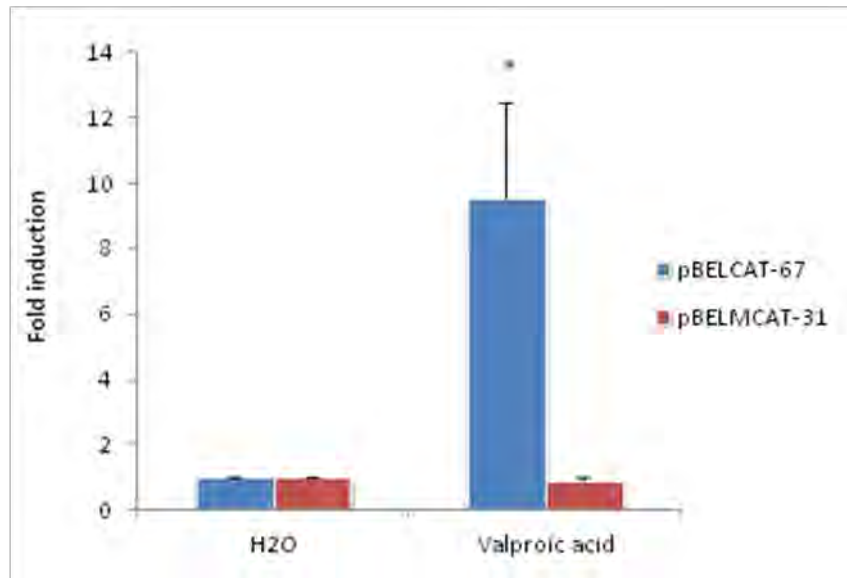


Figure 4.14 Treatment of CAT stable cell lines with valproic acid

Induction of CAT by VPA in pBELCAT-67 and pBELMCAT-31 cell lines. Cells were treated with valproic acid in a minimum of three independent experiments. A control of 0.5µl H₂O was included. Error bars indicate standard error for each sample. * indicates significance in relation to pBELCAT-67 or pBELMCAT-31 untreated ($p < 0.05$).

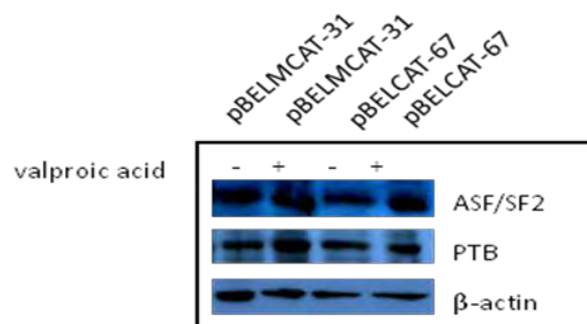


Figure 4.15 Western blot analysis on stable cell lines treated with valproic acid

Expression levels of ASF/SF2, PTB and β-actin proteins were determined in both cell lines in the absence and presence of VPA by Western Blot.

(iv) Scriptaid and amiloride

Next, we wanted to examine the role of two small molecules scriptaid and amiloride that in recent studies have been shown to be useful for treatment of a variety of cancers. pBELCAT-67 and pBELMCAT-31 were treated at different concentrations of scriptaid or amiloride and CAT levels were monitored 24 hours post-treatment. Scriptaid had no effect on late gene expression in both cell lines (Fig. 4.16) whereas a slight increase of CAT expression was observed in pBELCAT-67 but not in pBELMCAT-31 in the presence of amiloride.

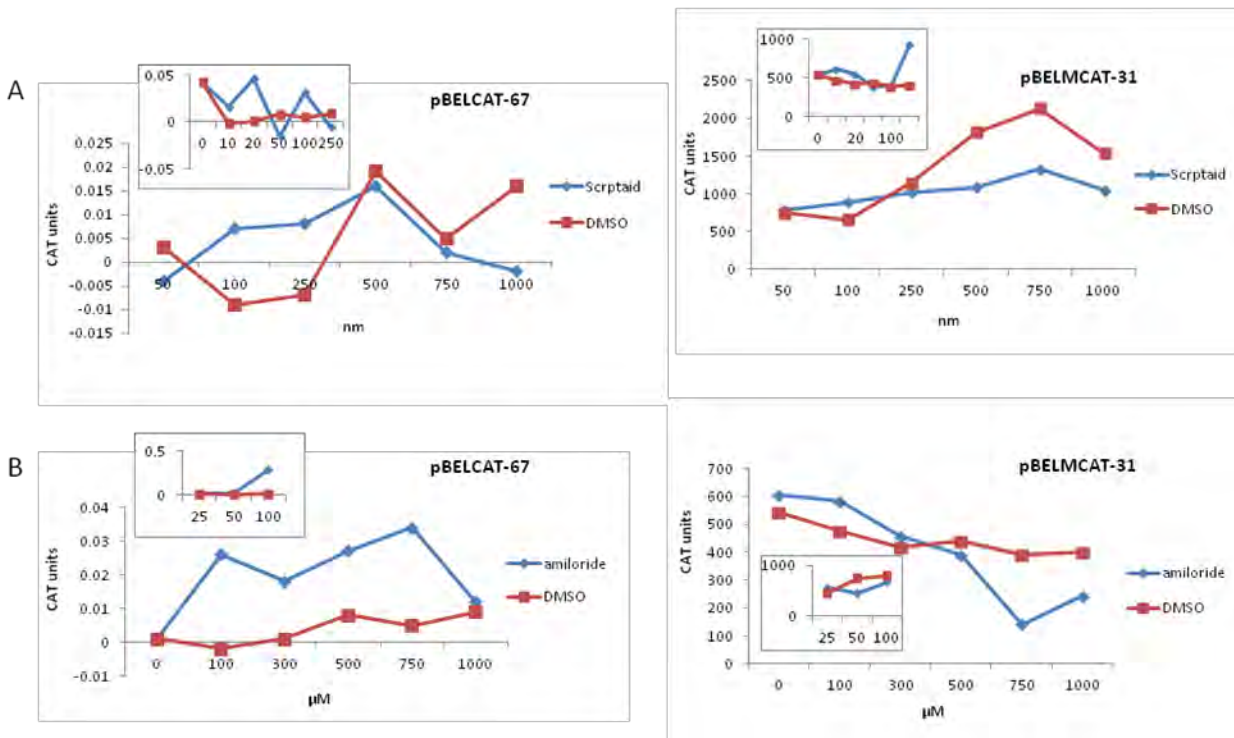


Figure 4.16 Treatment of CAT stable cell lines with scriptaid or amiloride

Treatment of pBELCAT-67 and pBELMCAT-31 stable cell lines with different concentrations of (A) scriptaid or (B) amiloride. Small graphs represent treatment of stables cell lines with lower concentration of amiloride or scriptaid.

(v) Tannic acid

Since in previous experiments was demonstrated that overexpression of PTB efficiently induces CAT expression in transient and stable transfections, it was investigated if the small molecule tannic acid, which it is believed to increase levels of polypyrimidine tract binding protein, has any effect on CAT expression. A serial dilution of tannic acid was performed in both stable cell lines and cells were analysed 24 hours post-treatment. Lower concentrations of tannic acid up to 1 μM did not modulate CAT expression and at high concentrations (100 μM) was toxic to cells (Fig. 4.17). However, 10 μM of tannic acid showed a significant increase of CAT in both cell lines suggesting that this small molecule could be useful to treat persistently HPV infected cells. Western blot analysis confirmed an increase of PTB levels by this substance further proving that this splicing factor plays a critical role in the regulation of late gene expression in HPV-16 and with its overexpression HPV-16 late gene expression can be induced (Fig. 4.18).

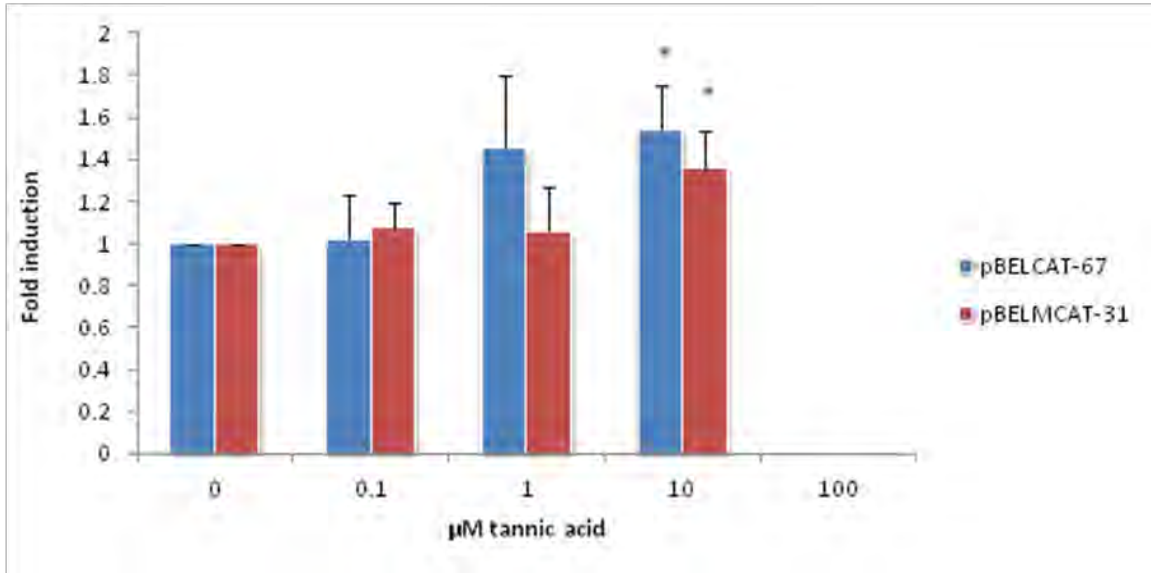


Figure 4.17 Treatment of CAT stable cell lines with tannic acid

Serial dilution of tannic acid in pBELCAT-67 and pBELMCAT-31 stable cell lines. Cells were treated with tannic acid in a minimum of three independent experiments. A control of 0.5μl H₂O was included. Error bars indicate standard error for each sample. * indicates significance in relation to pBELCAT-67 or pBELMCAT-31 untreated (p<0.05).

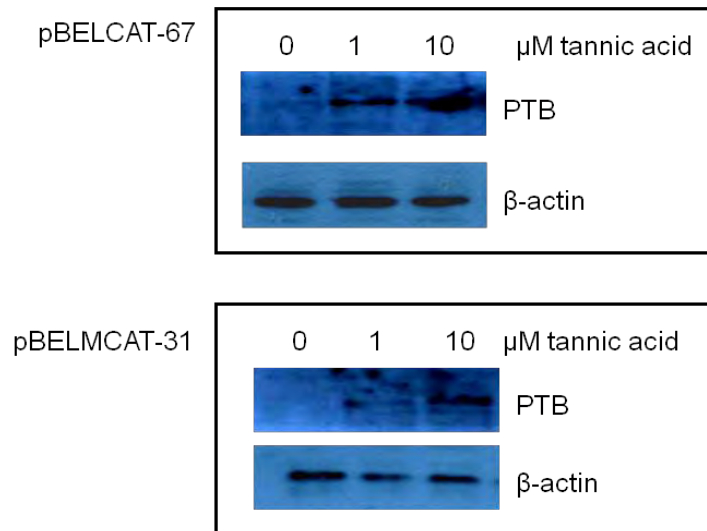


Figure 4.18 Western blot analysis on stable cell lines treated with tannic acid

Expression levels of PTB and β-actin proteins were determined in both cell lines in the absence and presence of tannic acid by Western Blot.

4.5 Treatment of pBELCAT or pBELMCAT plasmids transiently transfected in HeLa cells with TPA, valproic acid or tannic acid

Seeing that treatment of pBELMCAT-31 with VPA had very little effect on CAT induction this substance was tested in cells transiently transfected with pBELCAT or pBELMCAT. Also TPA and tannic acid, which have shown to modulate CAT expression in both cell lines, were analysed in transient transfections. HeLa cells were transfected with pBELCAT or pBELMCAT and 24 hours posttransfection cells were treated with 800ng TPA, 10 μ M tannic acid or 0.05mM valproic acid. An induction of 2 and 3.5 fold was observed in cells transfected with pBELCAT and pBELMCAT in the presence of VPA, respectively (Fig. 4.19). Higher levels of CAT were observed in cells treated with TPA and also tannic acid showed to greatly induce CAT, particularly in pBELCAT (Fig. 4.19). These results suggested that the lower induction observed in pBELMCAT-31 by valproic acid could be due to the high levels of CAT produced by this stable cell line that can be difficult to increase further. Furthermore, all 3 substances proved to induce late gene expression in HPV-16-derived plasmids and stable cell lines suggesting that they could be possibly used as drugs for treatment of HPV-16 infections.

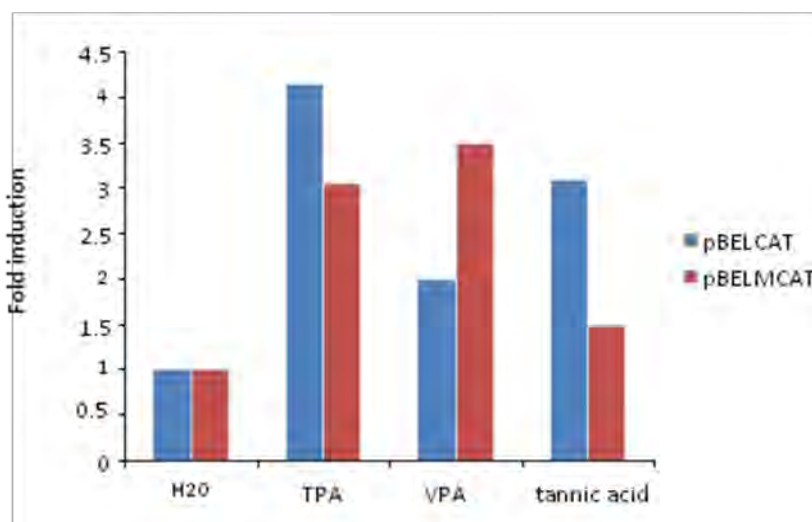


Figure 4.19 Effect of TPA, valproic acid or tannic acid in cells transiently transfected with pBELCAT or pBELMCAT plasmids

Induction of CAT of pBELCAT or pBELMCAT transiently transfected into HeLa cells by TPA, valproic acid or tannic acid. Transfections were carried out in triplicates. A control of 0.5 μ l H₂O was included.

4.6 Treatment of pBELCAT-67 cell line with TPA, valproic acid or tannic acid combined

So far we demonstrated that TPA, valproic acid and tannic acid can induce CAT in transient and stable transfected CAT reporter plasmids, possibly due to their effect on cellular factors involved in splicing. We speculated that using simultaneously one or more substances on CAT stable cell lines the levels of two or more cellular factors would it be induced and, as consequence, higher levels of CAT would it been detected. In order to verify that, pBELCAT-67 cell line was treated with TPA, valproic acid and tannic acid combined. Best result was observed in cells treated with TPA and tannic acid or TPA and valproic acid combined inducing more than 3 and 2 fold, respectively (Fig. 4.20). Lower levels of CAT were observed in the cells treated with a mix of

valproic acid and tannic acid whereas a combination of all 3 substances together did not modulate CAT expression (Fig. 4.20). Taken together these results demonstrated that TPA alone or combined with other substances gives the greater CAT induction, except for the mixture of all three substances that it might be toxic for the cells.

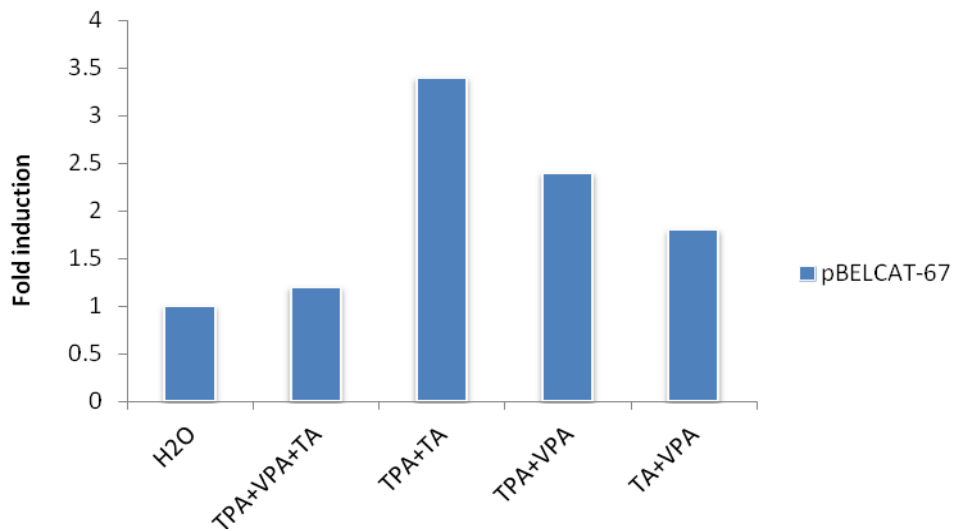


Figure 4.20 Treatment of pBELCAT-67 stable cell line with TPA, valproic acid and tannic acid combined

Induction of CAT in pBELCAT-67 by TPA, valproic acid (VPA) and tannic acid (TA) combined. Treatments of cells were carried out in triplicates. A control of 0.5 μ l H₂O was included.

4.7 APOPTOSIS-program

Many cells initiate an apoptosis program which is similar to what can be seen in epithelial cells that differentiate. This induction also counteracts proliferation and progression to cancer. In order to see if apoptosis could induce CAT production in our cell lines two simple apoptosis-programs were started on them including serum deprivation and heat shock. Both stable cell lines were treated as described in material and methods. None of these treatments seemed to have a positive effect on CAT production and after 2 days of serum deprivation (Fig. 4.21 A) and 2 hours of heat shock (Fig. 4.21 B) the cells looked quite sick and CAT levels start to drop off.

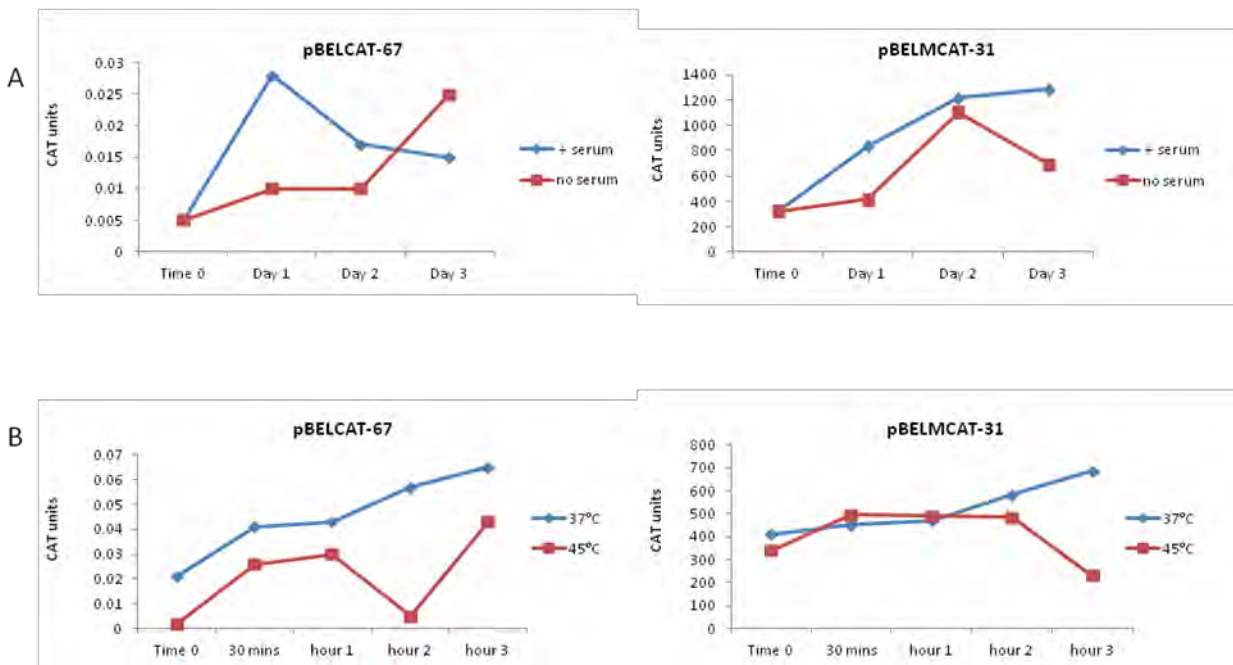


Figure 4.21 Apoptosis program

CAT levels in pBELCAT-67 and pBELMCAT-31 after (A) Serum deprivation or (B) Heat shock treatments.

4.8 Establishment of pBELMGFP reporter cell lines

Since in previous experiments the GFP assay proved to be very easy and inexpensive it was decided next to generate stable cell lines with pBELMGFP reporter plasmid integrated in the HeLa cellular genome. This would allow screening of compounds capable of inducing late gene expression without the need to use expensive reagents. Transfection was carried out as described previously for the establishment of CAT stable cell lines. Over 80 neomycin-resistant colonies were picked and DNA was first extracted and then PCR amplified with GFP or IRES set primers (Fig. 4.22). Three clones (pBELMGFP-31, -61, -69) resulted positive for both GFP and IRES sequences amplifications proving the successful integration of the full reporter plasmid (Fig. 4.22). Cells from each positive clone were analysed with a fluorescence microscope revealing very few glowing cells on each of them. The intensity of these visible positive cells was much lower compared to the cells transfected with pCMVGFP in transient transfections (Fig. 4.23). Even though GFP expression from pBELMGFP reporter plasmid transiently transfected was very low, this result was unexpected considering the high levels of CAT produced by pBELMCAT-31 stable cell line, which produces 20 fold more than pBELMCAT plasmid transiently transfected.

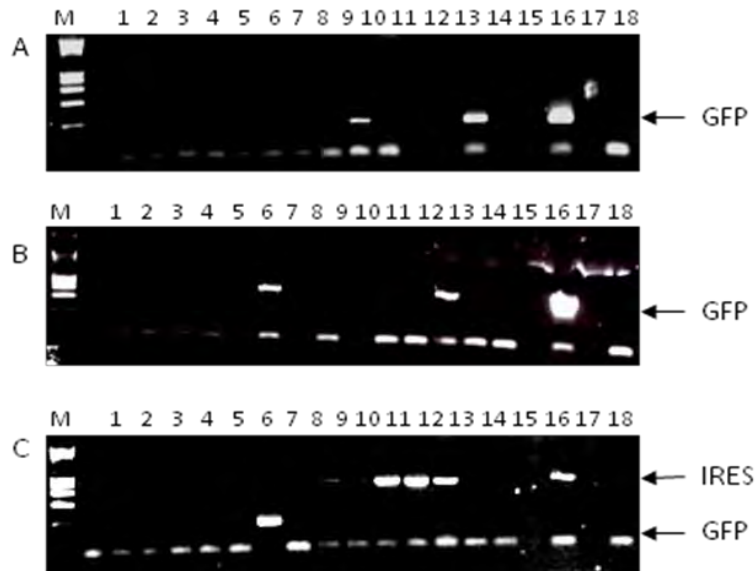


Figure 4.22 PCR amplification of DNA extracted from pBELMGFP-derived cell lines using GFP or IRES set primers

(A) 1KB DNA ladder (M), lanes 1-15 PCR amplification of DNA extracted from pBELMGFP stable cell lines using GFP primers, lane 16 positive control, lane 18 blank. (B) 1KB DNA ladder (M), lanes 1-15 PCR amplification of DNA extracted from pBELMGFP stable cell lines using GFP primers, lane 16 positive control, lane 18 blank. (C) 1KB DNA ladder (M), lanes 1-10 PCR amplification of DNA extracted from stable cell lines pBELMGFP using GFP primers and lane 11-18 using IRES primers.



Figure 4.23 Fluorescence of pBELMGFP stable cell lines

GFP fluorescence intensities were detected with a fluorescence microscope from the living cells in (A) pBELMGFP-31, (B) pBELMGFP-61 and (C) pBELMGFP-69 (CELL[^]F software, Olympus Soft Imaging solutions GmbH).

4.9 Treatment of pBELMGFP with TPA

Since TPA showed to increase levels of CAT in both pBELCAT-67 and pBELMCAT-31 stable cell lines we considered that it would be useful to treat all 3 pBELMGFP-derived cell lines with TPA to see if GFP intensity can be enhanced by this small molecule. In the presence of TPA only the clones pBELMGFP-31 and -61 showed few glowing cells with a relative higher GFP intensity whereas in the clone number 69 fluorescence was not detected.

However, GFP fluorescence intensity from these pBELMGFP-derived cell lines in the absence or in the presence of TPA was very low and only detectable in a few cells (Fig. 4.24). We reasoned that the fluorescence from these pBELMGFP-derived cell lines might be insufficient to be detected under the microscope light. Therefore 2×10^4 pBELMGFP-derived cell lines were plated in a 96 well plate and treated with 0ng, 400ng or 800ng TPA in triplicate. 24 hours post-treatment cells were analysed with a Luminescence Spectrometer (Perkin Elmer LS55B), but GFP fluorescence was not detected. Hence we concluded that these stable cell lines are unusable for analysis of HPV-16 late gene expression.

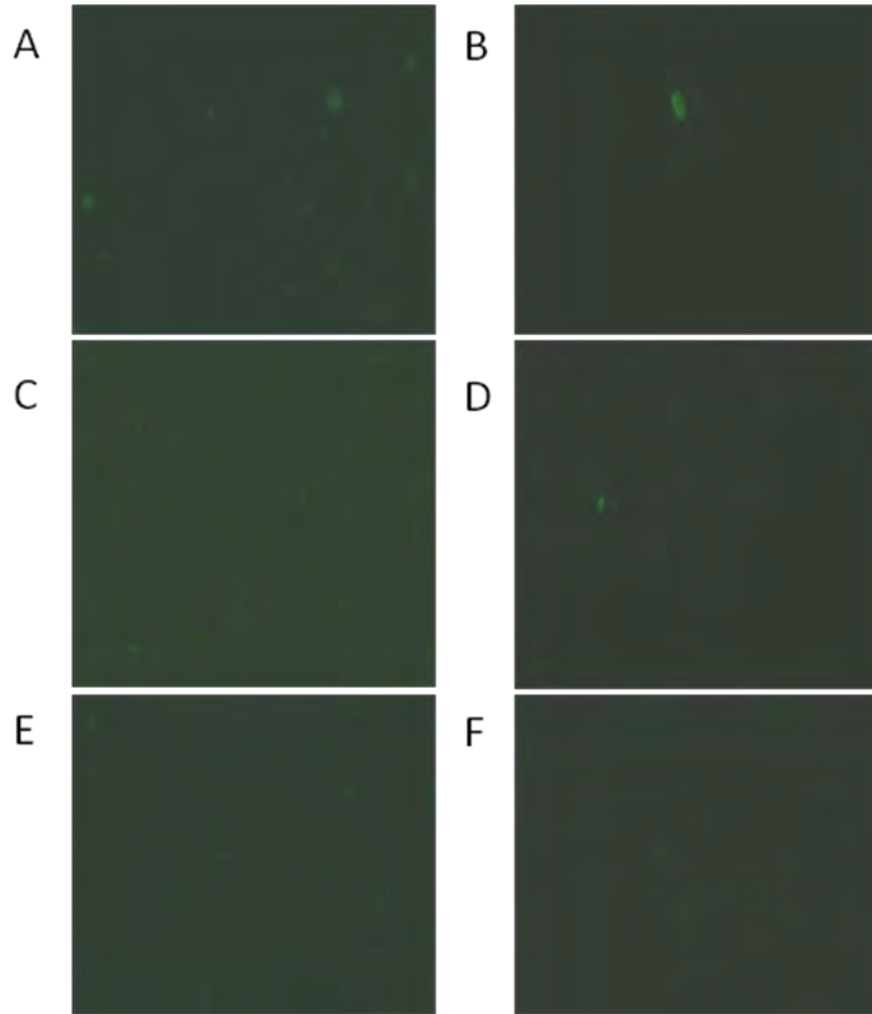


Figure 4.24 Treatment of pBELMGFP with TPA

(A) pBELMGFP-31, (C) pBELMGFP-61 and (E) pBELMGFP-69 stable cell lines untreated and (B) pBELMGFP-31, (D) pBELMGFP-61 and (F) pBELMGFP-69 stable cell lines treated with 800ngTPA. Fluorescence was detected 24 hours posttreatment with a fluorescence microscope (CELL[^]F software, Olympus Soft Imaging solutions GmbH).

4.10 Gene Knockdown

So far we have shown that with overexpression of a number of cellular proteins late gene expression is induced in transient and stable transfected subgenomic HPV-16 plasmids. To further examine the role of these cellular proteins on the regulation of HPV-16 late gene expression, next goal was to see if, decreasing the levels of these cellular proteins in the cells, late gene expression would be affected. In order to do that Lentiviruses carrying a shRNA to knockdown specific genes were used and the pBELMCAT-31 was selected as a cell line, which itself produces high levels of CAT. As genes to be knockdown we opted for ASF/SF2 and PTB, which have shown to efficiently induce late gene expression in HPV-16-derived plasmids. The lentivirus expressed RNAs were targeted against ASF/SF2, PTB or a scrambled non targeting sequence. In these lentiviruses GFP is used as a marker for shRNA expression and thus the use of a fluorescence or confocal microscope can be used to follow shRNA expression in the cells.

Lentiviruses carrying ASF/SF2, PTB or scramble shRNA were transduced at different multiplicity of infectivity (MOI) into pBELMCAT-31 stable cell lines and then cells were analysed under a confocal microscope. At MOI 0.3 and 1 transduction failed, but at MOI 2, 5 and 10 more than 90% of the cells were infected by the virus confirming the successful transduction (Fig. 4.25, 4.26 and 4.27). Higher intensity of GFP was seen in cells infected with lentiviruses at MOI 10. A CAT ELISA was then performed to verify a change in CAT production by pBELMCAT-31 in the presence of the lentiviruses. The lentiviruses definitely dropped down the levels of CAT compared with the levels of CAT produced by pBELMCAT-31 untransduced and infection at MOI 10 resulted in a lower induction of CAT by all 3 lentiviruses, as expected (Fig. 4.28 A). However, levels of CAT expression in the cells infected with lentiviruses carrying shRNA PTB or ASF/SF2 were not significant ($p\text{-value} > 0.05$) compared to

the levels dropped down by the scramble lentivirus (Fig. 4.28 B). Western Blot analysis showed that ASF/SF2 and PTB expression was not knocked down at any MOI utilized for this experiment and actin levels seemed to be unaffected by the presence of the lentiviruses (Fig. 4.29). To assure the highest possibility of modulating gene expression level, multiple shRNA constructs each covering a unique region of the target gene should be used. Therefore we reasoned that this might be the cause of this unsuccessful knockdown.

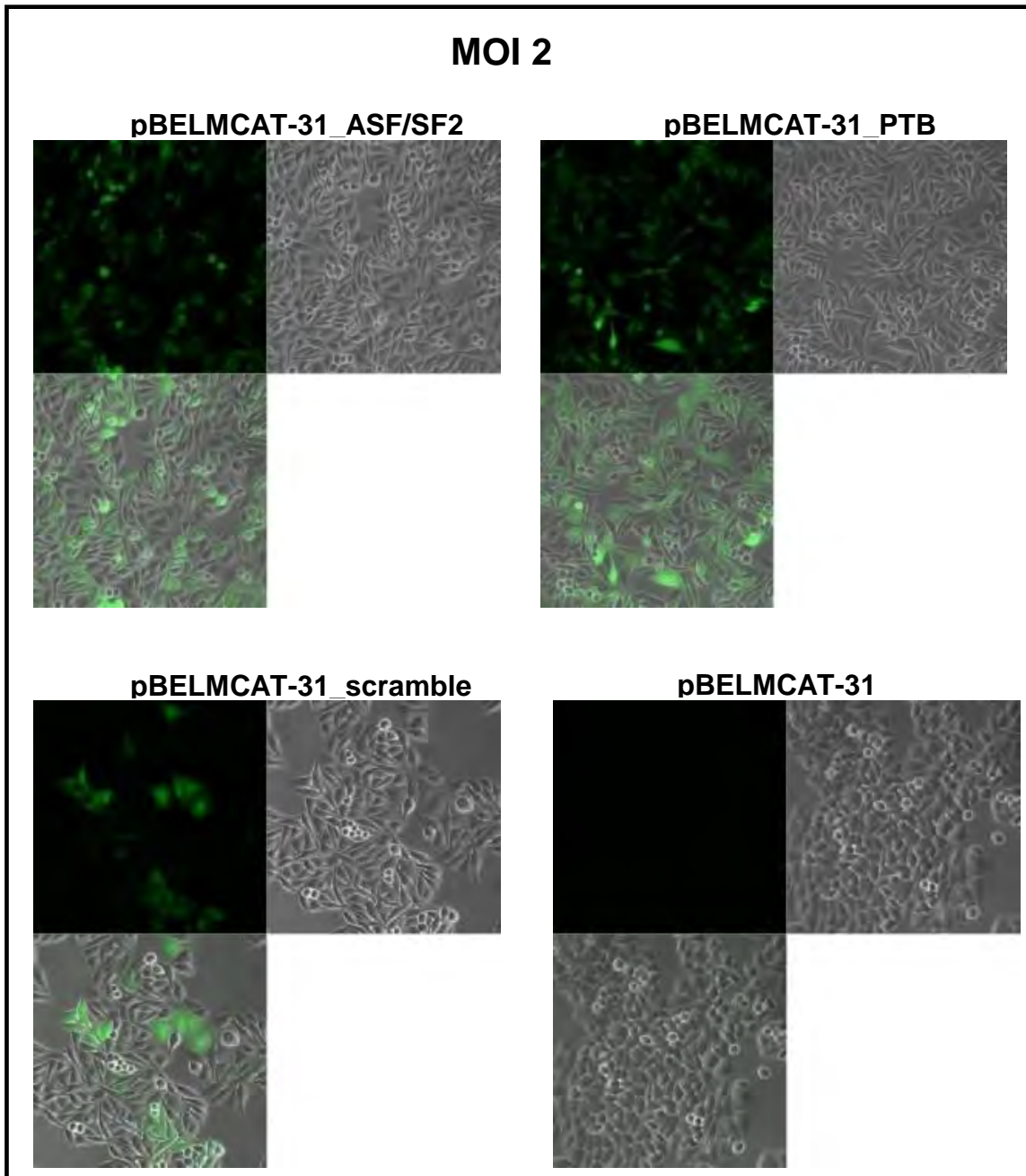


Figure 4.25 Detection of infected cells at MOI 2 with confocal microscope

Detection of GFP expression in cells infected with lentiviruses carrying ASF/SF2, PTB or scramble shRNA at MOI 2. The 3 panels in each image represent fluorescence, light and combined.

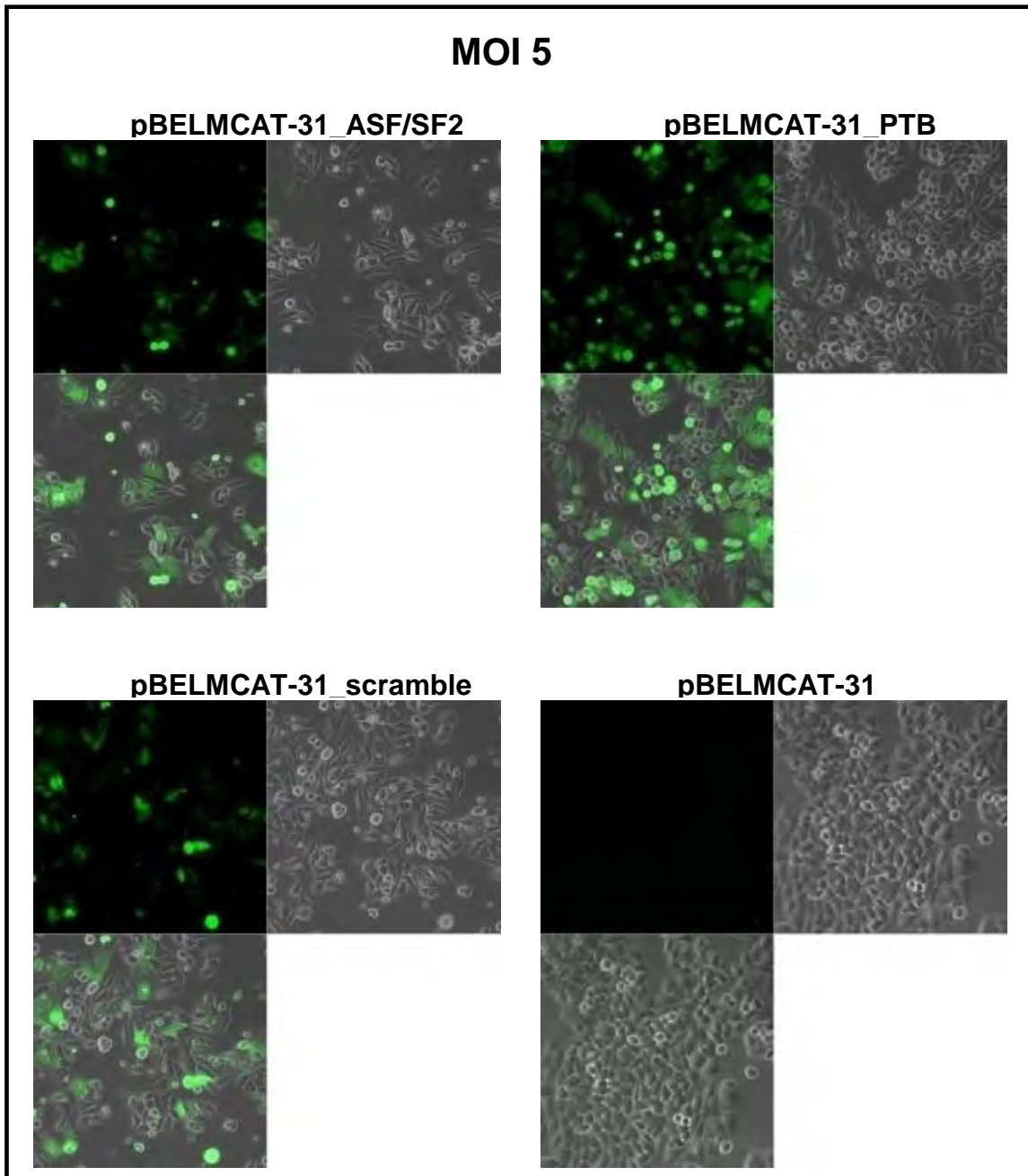


Figure 4.26 Detection of infected cells at MOI 5 with confocal microscope

Detection of GFP expression in cells infected with lentiviruses carrying ASF/SF2, PTB or scramble shRNA at MOI 5. The 3 panels in each section represent fluorescence, light and combined.

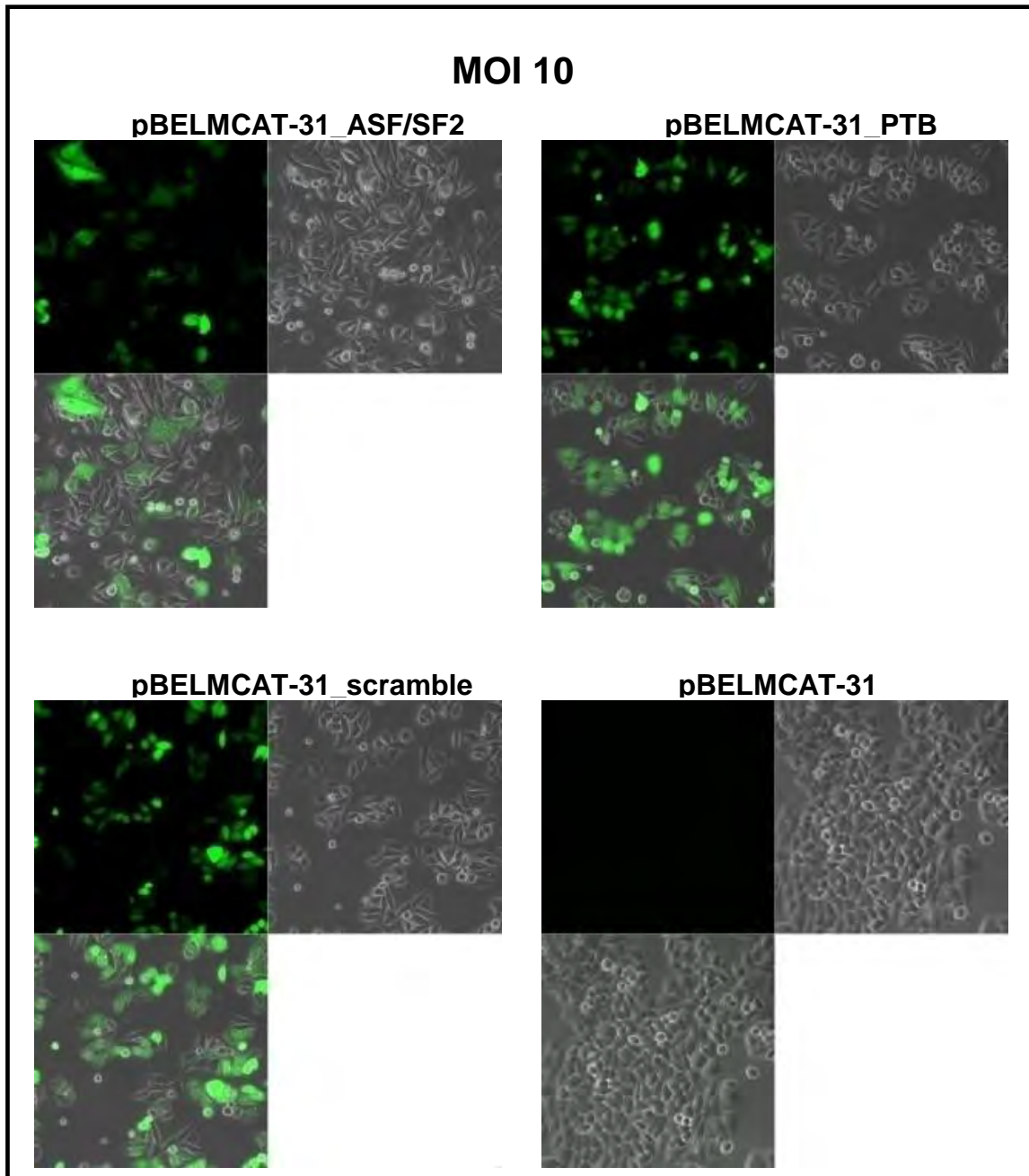


Figure 4.27 Detection of infected cells at MOI 10 with confocal microscope

Detection of GFP expression in cells infected with lentiviruses carrying ASF/SF2, PTB or scramble shRNA at MOI 10. The 3 panels in each image represent fluorescence, light and combined.

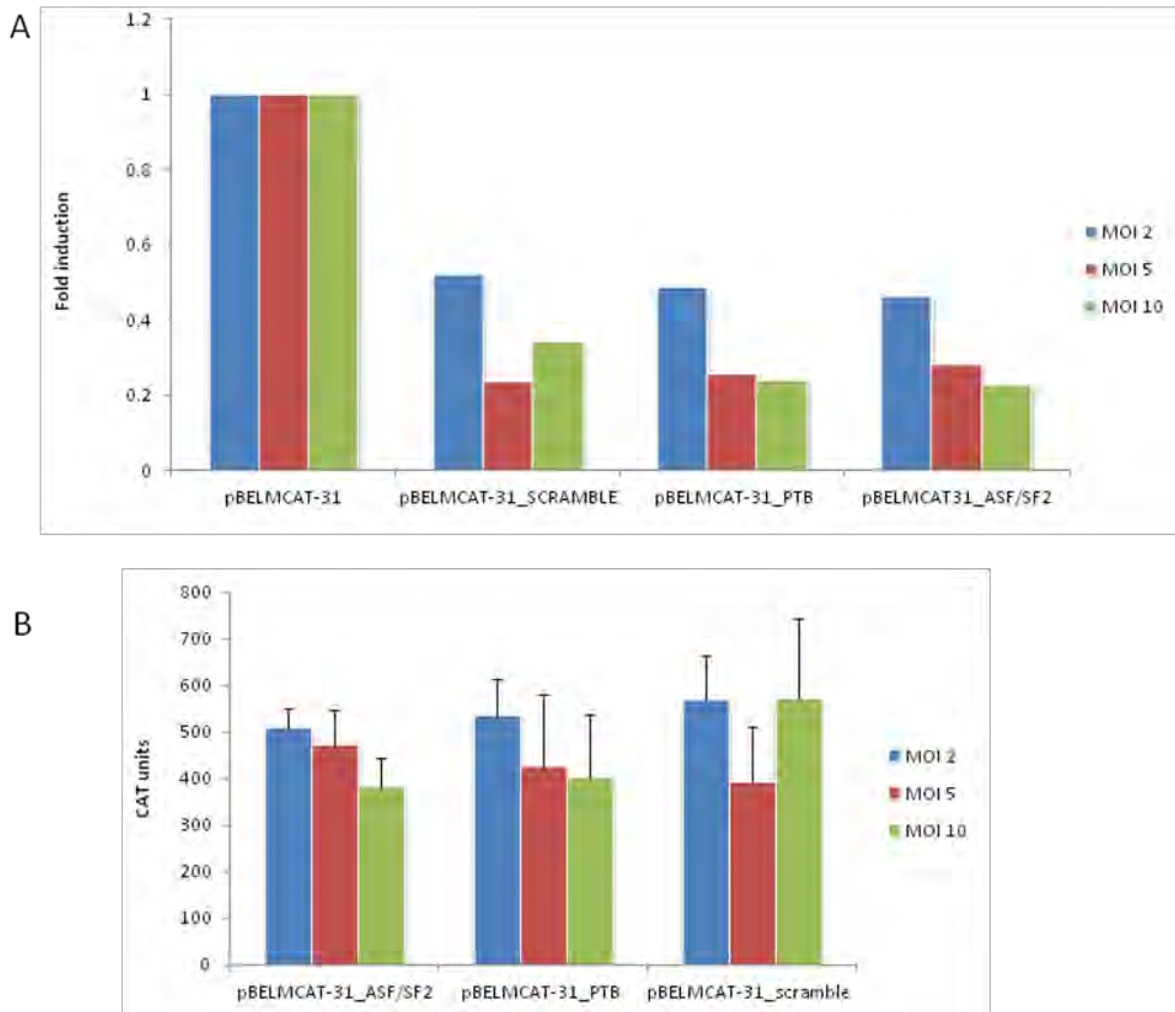


Figure 4.28 CAT expression in pBELMCAT-31 stable cell line transduced with lentiviruses carrying shRNA ASF/SF2, PTB or scramble

(A) Induction of CAT expression in pBELMCAT-31 stable cell lines uninfected and infected by lentiviruses carrying shRNA ASF/SF2, PTB or scramble. (B) CAT units produced by pBELMCAT-31 stable cell lines infected by lentiviruses carrying shRNA ASF/SF2, PTB or scramble.

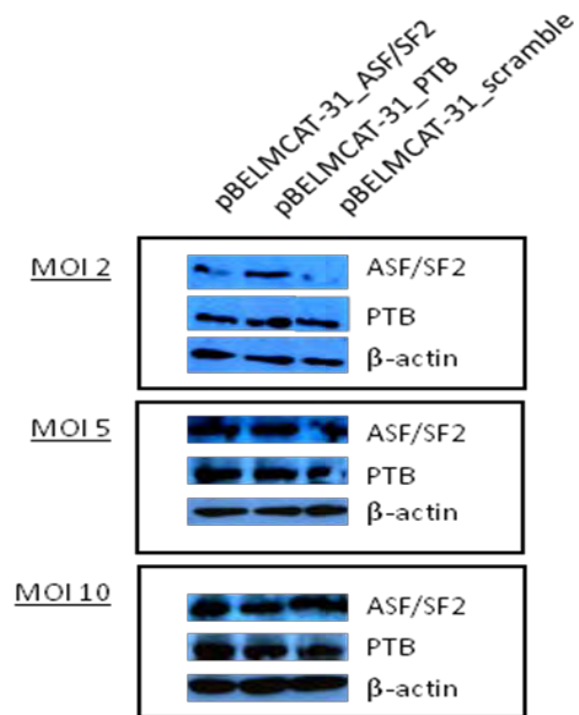


Figure 4.29 Western blot analysis on pBELMCAT-31 stable cell line infected with lentiviruses

Expression levels of ASF/SF2, PTB and β -actin proteins were determined by Western Blot.

5. Chapter 5 – Results

Identification and localization of a RNA-binding protein involved in the regulation of splicing in normal and neoplastic cervical epithelium

RNA processing has a pivotal role in HPV-16 gene regulation (Johansson *et al.*, 2011; Mole, *et al.*, 2006; Schwartz *et al.*, 2007). Therefore the identification of novel cellular proteins or RNA binding factors that might influence RNA processing events, such as alternative splicing and polyadenylation, has a major importance in the study of HPV-16 gene regulation.

5.1 Identification of a novel splicing factor involved in HPV-16 late gene regulation

Previous results demonstrated that the small molecule TPA efficiently modulates CAT expression in transient and stable transfections. Since it has been shown that TPA induces expression of various genes, including the splicing factors arginine/serine 2 (SC35) and hnRNP A2/B1 (Zheng *et al.*, 2002), it was investigated if these two cellular proteins are involved in the splicing events that can occur in HPV-16. Transient transfections of pBELCAT or pBELMCAT plasmids with vectors expressing SC35 or hnRNP A2/B1 were carried out into HeLa cells resulting in a considerable increase of CAT levels in both reporter plasmids in the presence of hnRNP A2/B1. A small CAT induction was observed in pBELCAT by overexpression of SC35

but not in pBELMCAT (Fig. 5.1). These results suggested that the splicing factor hnRNP A2/B1 might have a role in regulation of HPV-16 late gene expression.

In order to prove that hnRNP A2/B1 expression is induced by TPA a western blot on pBELCAT-67 and pBELMCAT-31 stable cell lines treated with TPA, emetine, cycloheximide or ceramide was performed resulting in an increase of hnRNP A2/B1 levels by TPA in both cell lines, whereas hnRNP A2/B1 levels were unaffected by emetine, cycloheximide and ceramide (Fig. 5.2).

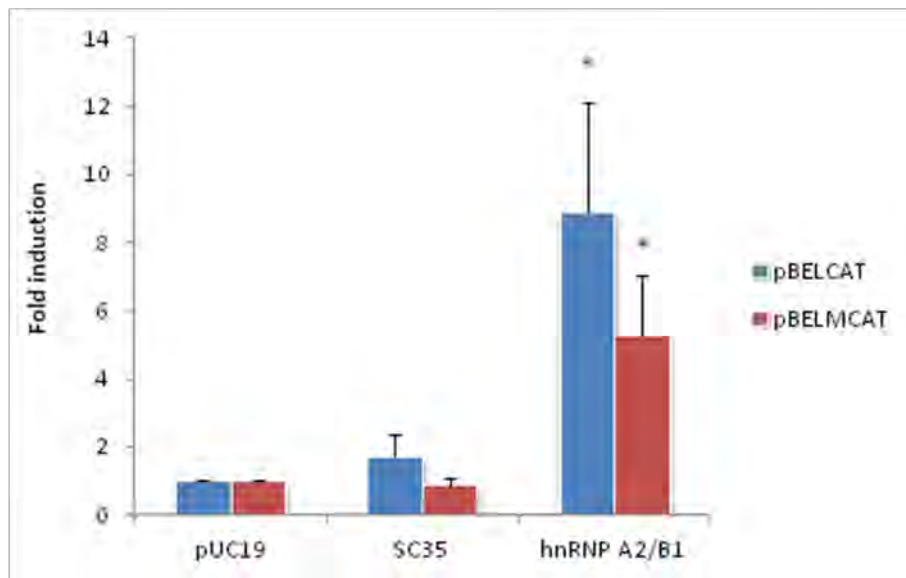


Figure 5.1 Co-transfection of CAT reporter plasmids with SC35 or hnRNP A2/B1 expression vector

pBELCAT pBELMCAT were transfected into HeLa cells in the presence and absence of cotransfected expression vectors. Cells were transfected with each plasmid in a minimum of three independent experiments. Error bars indicate standard error for each sample. * indicates significance in relation to pBELCAT or pBELMCAT cotransfected with the empty vector (pUC19) ($p < 0.05$).

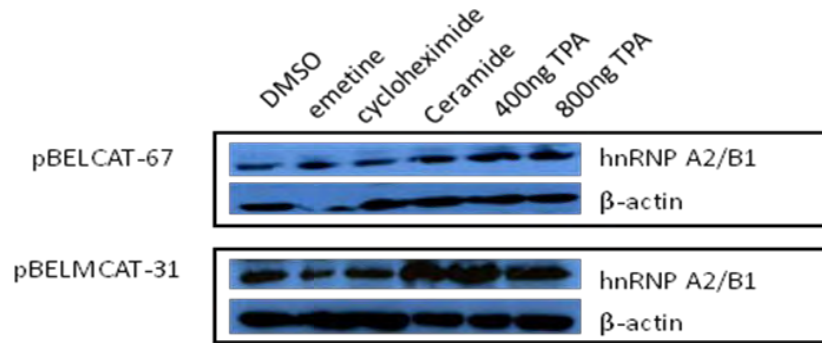


Figure 5.2 Western blot analysis on pBELCAT-67 and pBELMCAT-31 stable cell lines

Expression levels of hnRNP A2/B1 and β -actin proteins were determined in both cell lines in the absence and presence of emetine, cycloheximide, ceramide or TPA by Western Blot.

To further evaluate the role of this cellular protein, a two-fold serial dilution of hnRNP A2/B1 was performed resulting in a dose-dependent induction of CAT from both reporter plasmids. A lower induction was observed at higher hnRNP A2/B1 concentrations whereas low levels of hnRNP A2/B1 efficiently induced CAT expression in both pBELCAT and pBELMCAT plasmids (Fig. 5.3).

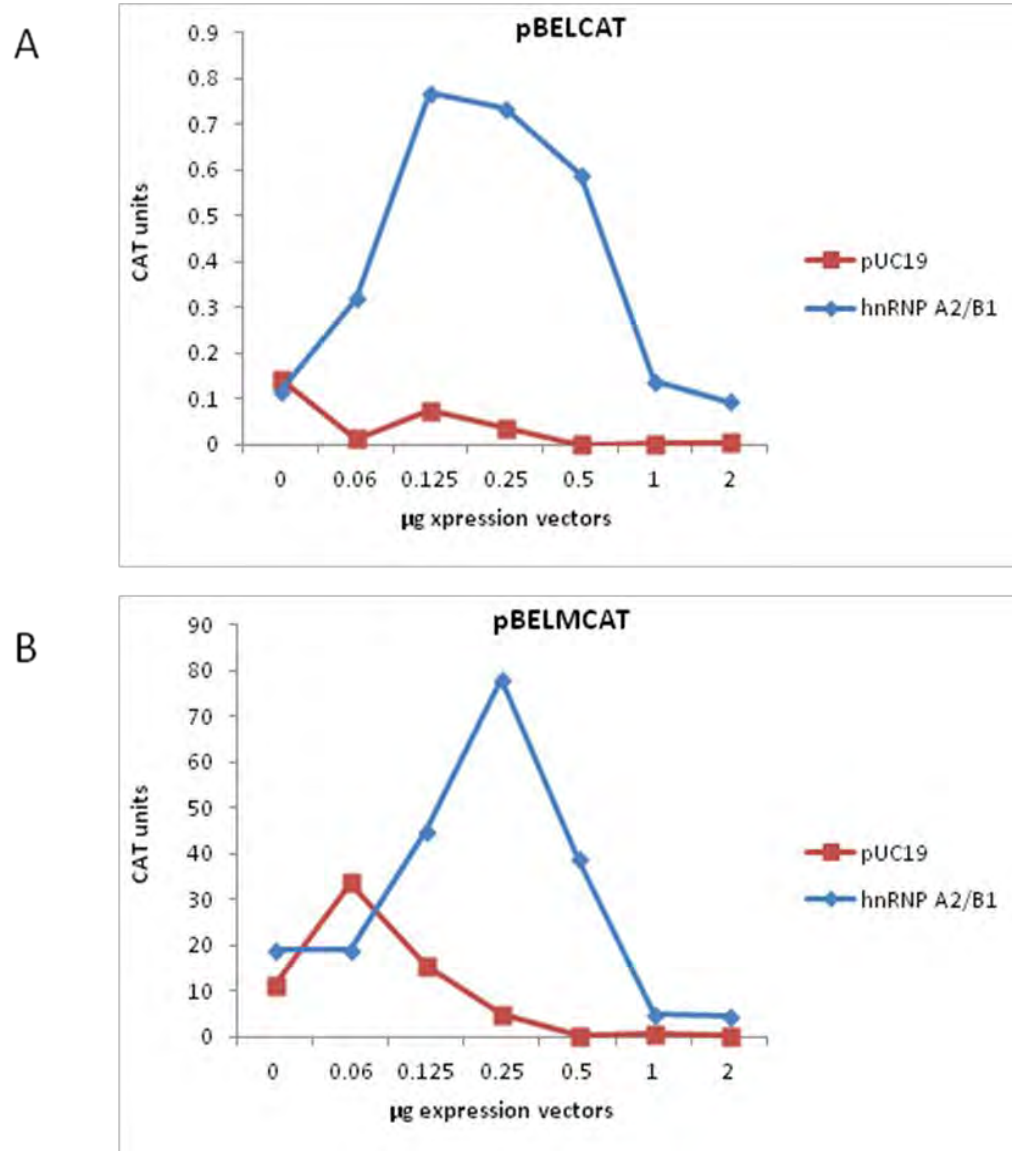


Figure 5.3 Induction of CAT by hnRNP A2/B1

Co-transfection of (A) pBELCAT and (B) pBELMCAT reporter plasmids with a serial dilution of hnRNP A2/B1 expression vector.

To establish if this RNA-binding protein targets sequences located upstream SD3632 and SA5639, reporter plasmids containing only these two splice sites (Fig. 5.4 A) were tested in the presence of hnRNP A2/B1 expressing plasmid. When hnRNP A2/B1 was cotransfected with pBspliceMCAT or pMt1sdCAT (Fig. 5.4 A) an induction of CAT was observed in both reporter plasmids (Fig. 5.4 B) indicating that this RNA-binding factor acts on the splice sites used exclusively by late genes. A lower induction was seen in pMt1sdCAT, in which splicing silencer upstream SD3632 have been deleted, whereas higher induction, up to 7 fold, was seen in pBspliceMCAT as demonstration that hnRNP A2/B1 binds specifically to these sequences that suppress SD3632 (Fig. 5.4 B).

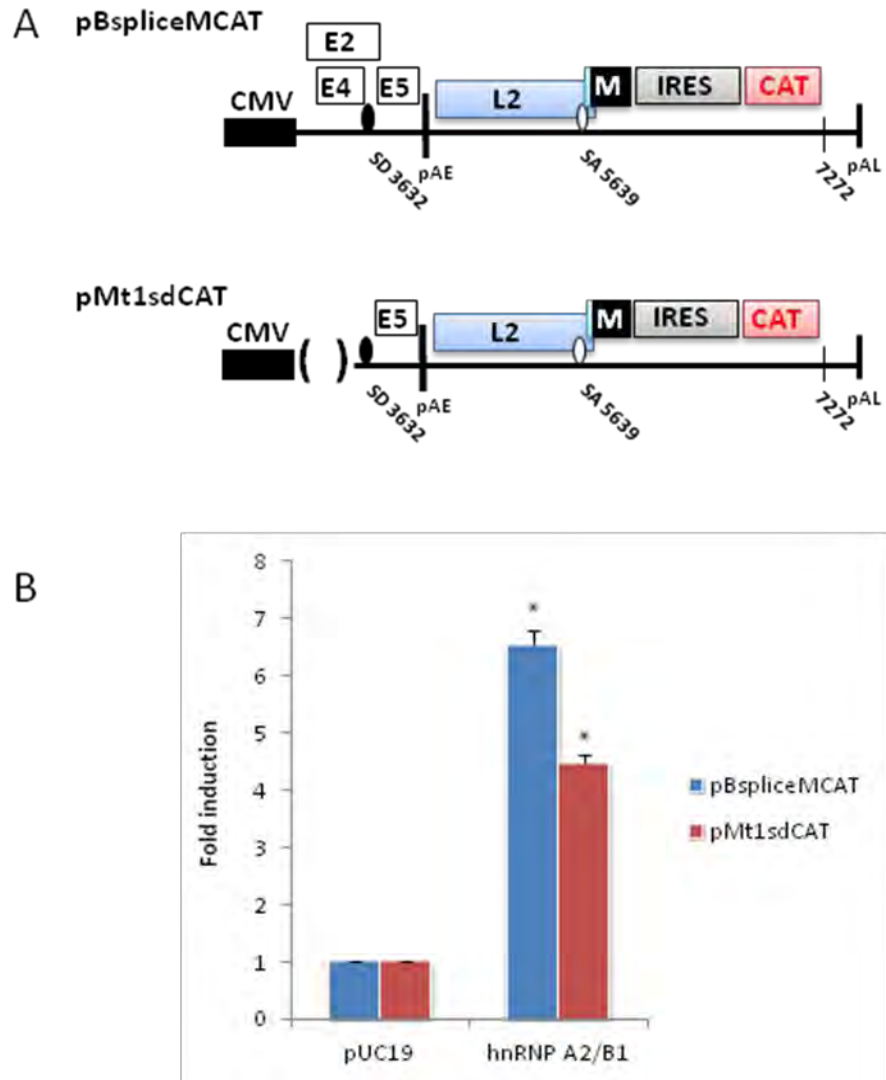


Figure 5.4 Induction of CAT in pBspliceMCAT and pMt1sdCAT by hnRNP A2/B1

(A) pBspliceMCAT and pMt1sdCAT constructs (B) pBspliceMCAT or pMt1sdCAT plasmids were transfected into HeLa cells in the presence and absence of plasmid expressing hnRNP A2/B1. Transfections were carried out in triplicate. Error bars indicate standard error for each sample. * indicates significance in relation to pBspliceMCAT or pMt1sdCAT cotransfected with the empty vector (pUC19) ($p < 0.05$).

Transfection of plasmid expressing hnRNP A2/B1 in pBELCAT-67 and pBELMCAT-31 stable cell lines confirmed that overexpression of hnRNP A2/B1 efficiently modulates CAT expression, inducing 6.9 and 1.7 fold in pBELCAT-67 and pBELMCAT-31 cell lines, respectively (Fig. 5.5).

In conclusion a novel splicing factor involved in the regulation of HPV-16 late gene expression was identified.

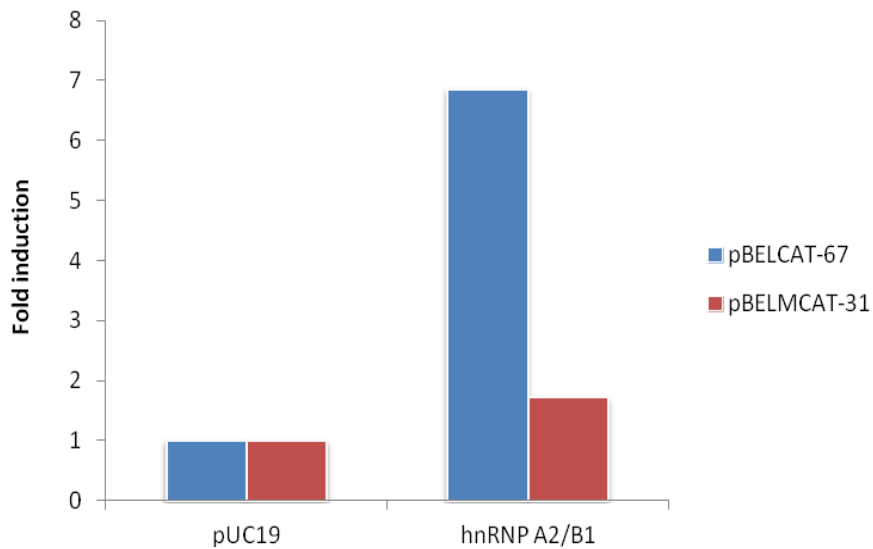


Figure 5.5 Induction of CAT in stable pBELCAT-67 and pBELMCAT-31 cell lines by hnRNP A2/B1

Induction of CAT expression by overexpression of hnRNP A2/B1 expression plasmid in pBELCAT-67 and pBELMCAT-31 stable cell lines. Transfections were carried out in triplicate.

5.2 Detection of hnRNP A2/B1 protein on cervical tissue arrays

Since the hnRNP A2/B1 was identified as a novel cellular factor involved in regulation of HPV-16 late gene expression next it was investigated if the expression level of this RNA-binding protein is altered in cervical premalignant and malignant lesions. Thus an immunohistochemistry assay was performed on cervical tissue arrays in order to detect and localize this protein. The advantage of using tissue microarray slides is that it allows the analysis of multiple histological samples in a single slide. The TMA was designed to have 80 cores in total and out of 80 cores it was possible to read 54 cores. hnRNP A2/B1 immunostaining in normal cervical (NC) tissue cores resulted in high expression of hnRNP A2/B1 protein in the lower layers, especially in the basal layer with a predominant nucleoplasmic staining. In the superficial layers no staining was observed (Fig. 5.6). The strong staining was detected in eight cores out of eight normal cervical tissue cores. In low-grade squamous intraepithelial lesion (LSIL) the staining was also very strong in the basal and intermediate layers (Fig. 5.7). Only two cores out of 15 readable LSIL cores showed a weaker staining. In high-grade squamous intraepithelial lesion (HSIL) and squamous cell carcinoma (SCC) low expression of hnRNP A2/B1 was detected in few scattered nuclei of 15 and 16 cores, respectively (Fig. 5.8 and 5.9).

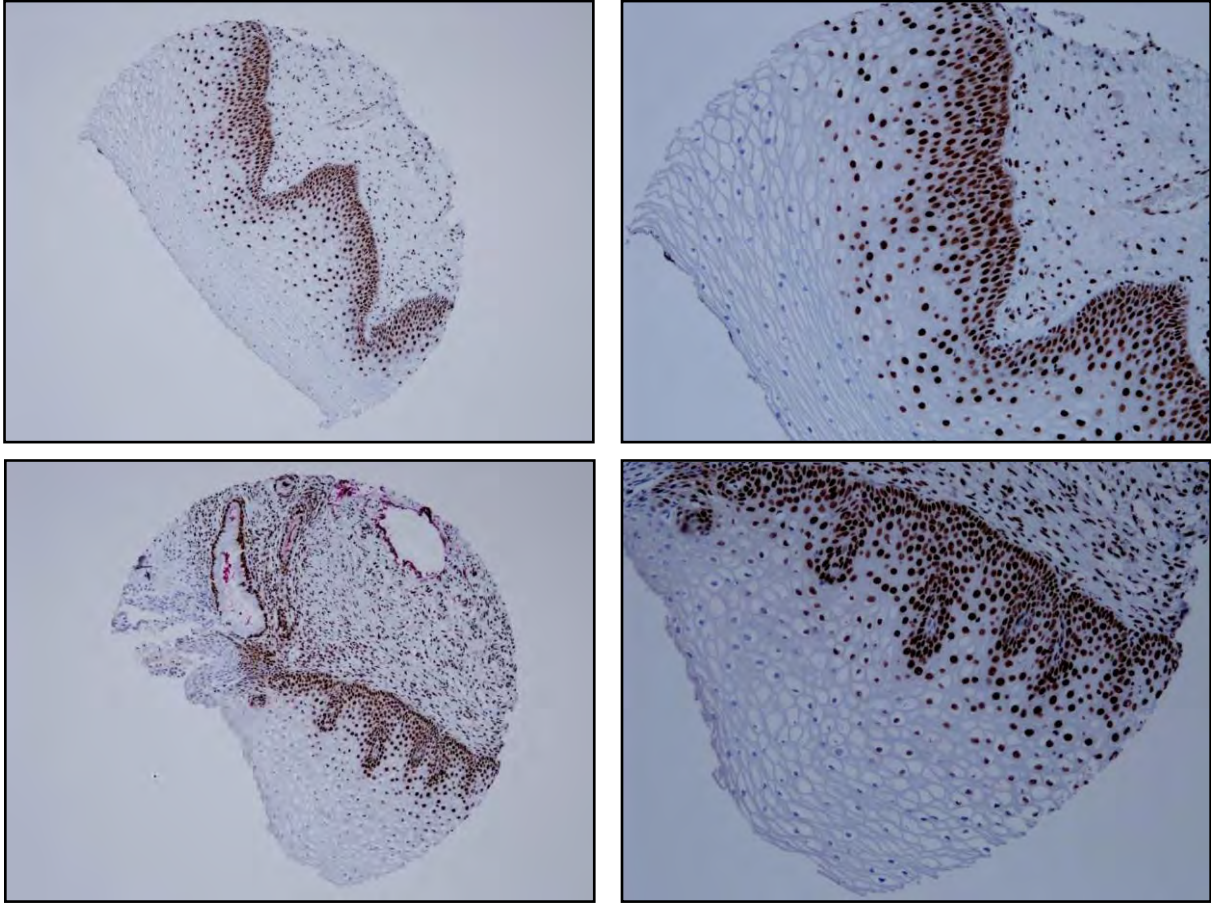


Figure 5.6 hnRNP A2/B1 protein expression in NC

hnRNP A2/B1 protein expression detected by immunohistochemical staining in normal cervical epithelium (NC). Images on the left, magnification 20X; images on the right, magnification 40X.

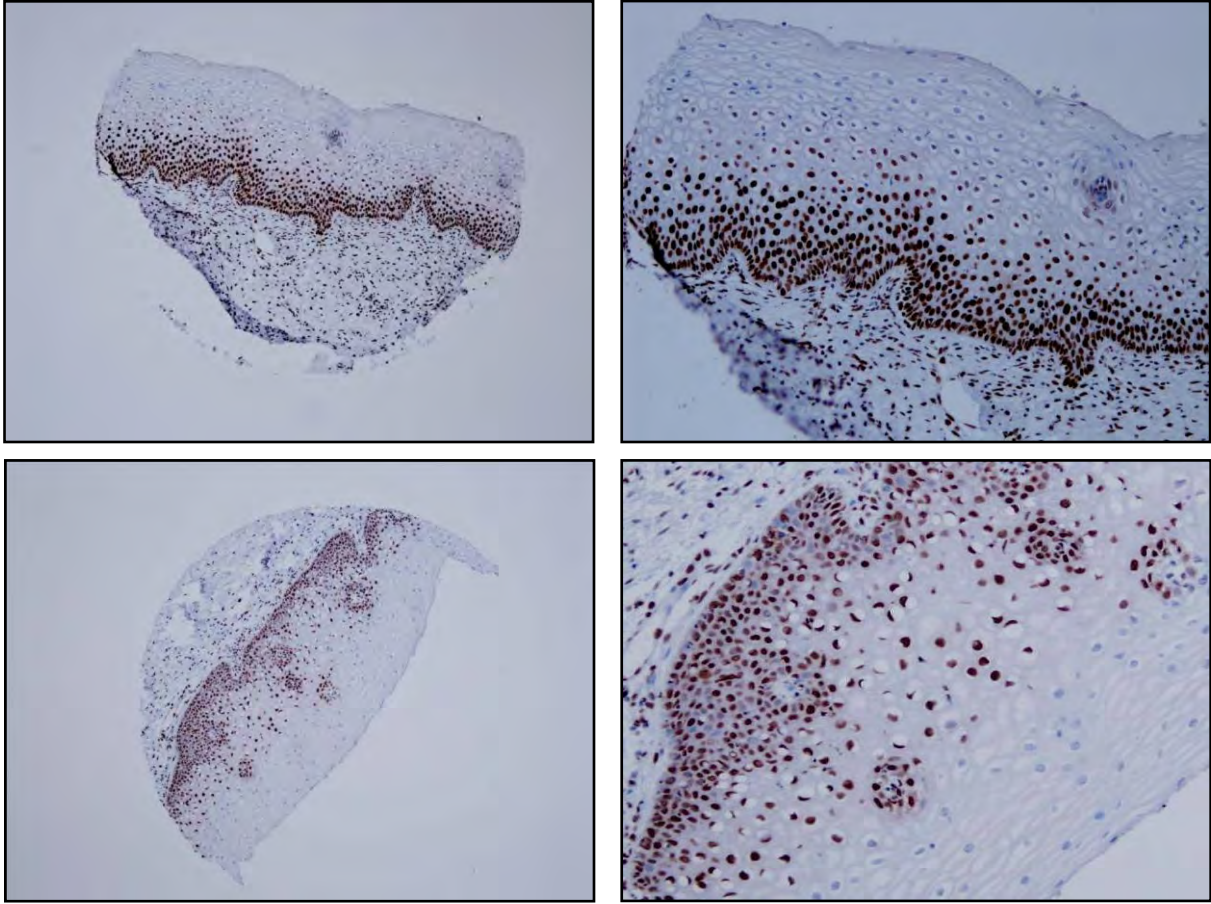


Figure 5.7 hnRNP A2/B1 protein expression in LSIL

hnRNP A2/B1 protein expression detected by immunohistochemical staining in low-grade squamous intraepithelial lesions (LSIL). Images on the left, magnification 20X; images on the right, magnification 40X.

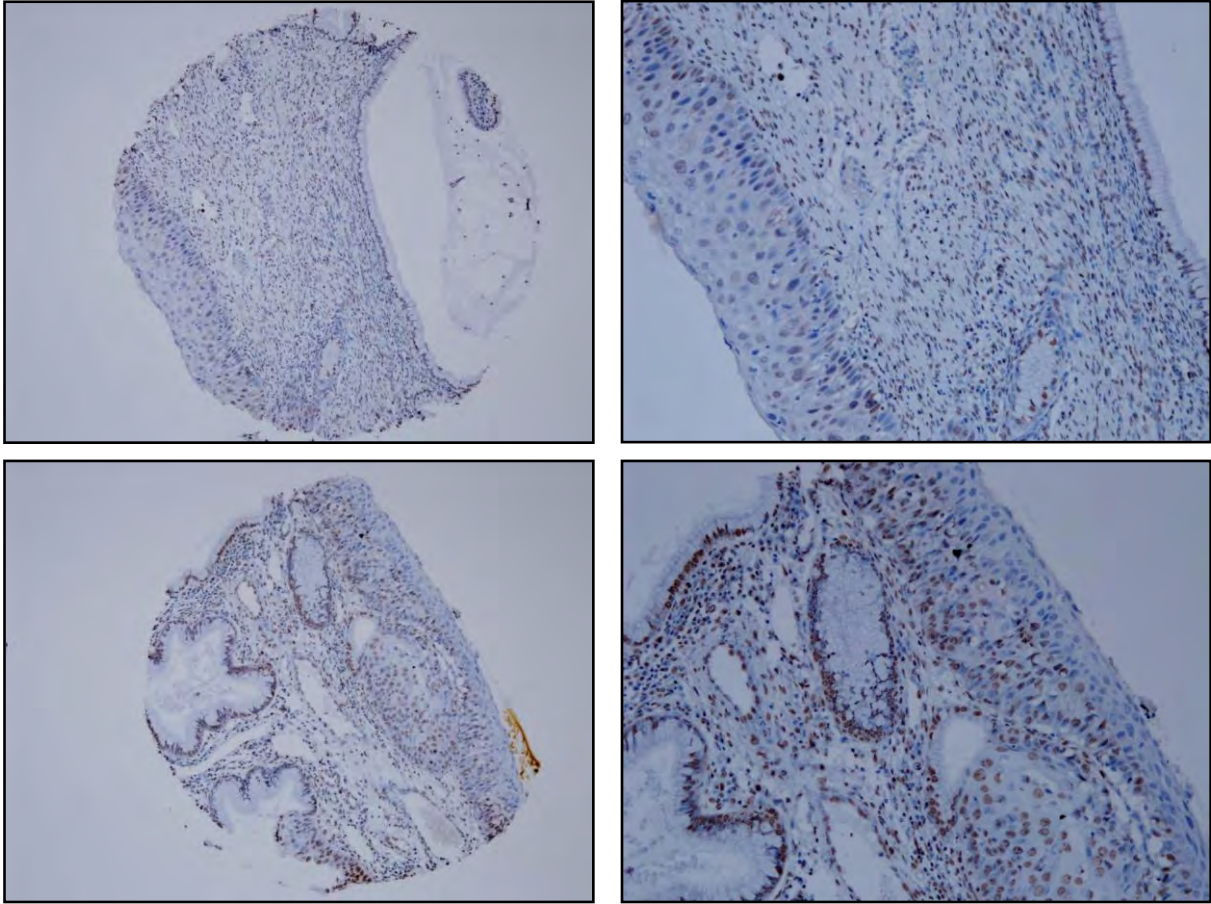


Figure 5.8 hnRNP A2/B1 protein expression in HSIL

hnRNP A2/B1 protein expression detected by immunohistochemical staining in high-grade squamous intraepithelial lesions (HSIL). Images on the left, magnification 20X; images on the right, magnification 40X.

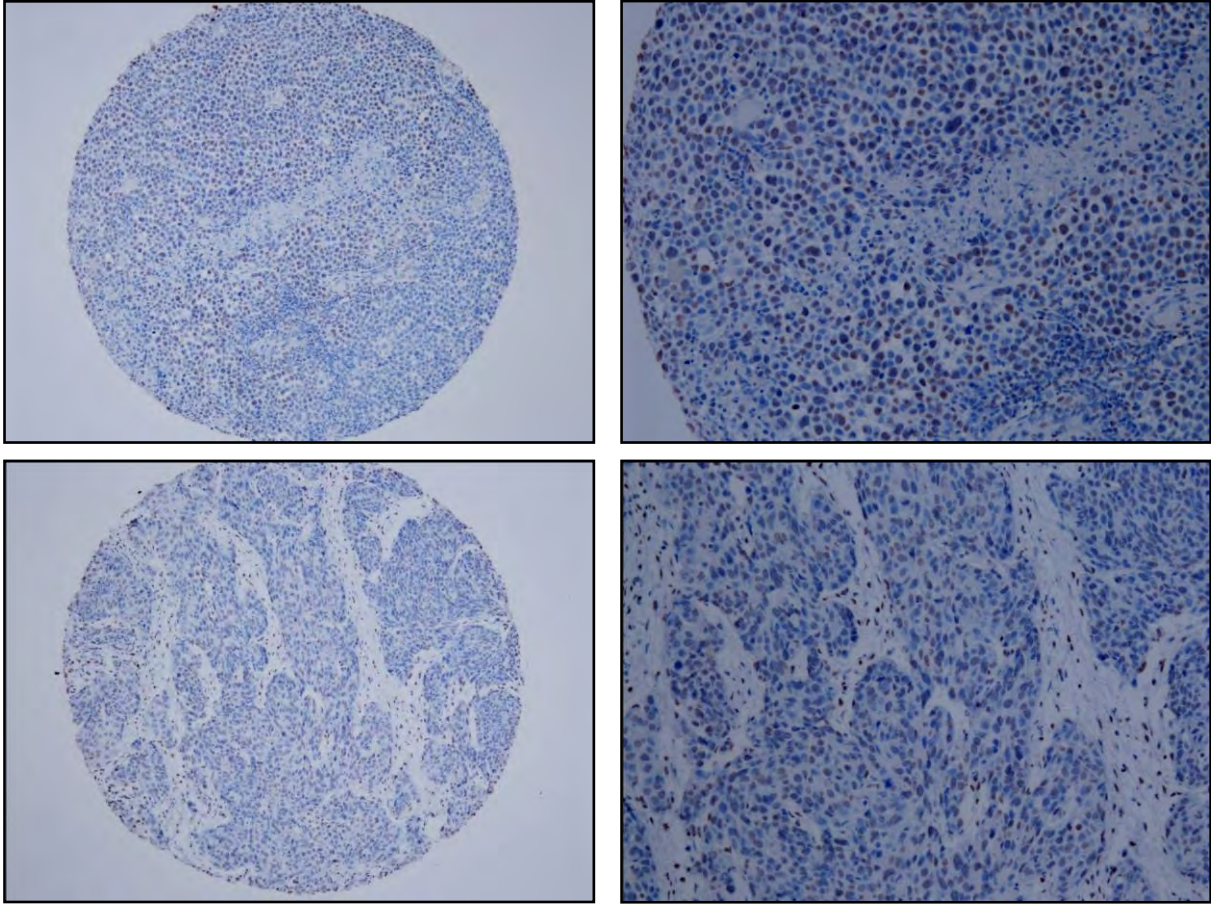


Figure 5.9 hnRNP A2/B1 protein expression in SCC

hnRNP A2/B1 protein expression detected by immunohistochemical staining in squamous cell carcinoma (SCC). Images on the left, magnification 20X; images on the right, magnification 40X.

5.3 Correlation of RNA-binding protein expression in cervical epithelium to p16INK4a, a known marker of differentiation

An immunohistochemistry assay was performed on the same tissue microarrays using p16INK4a, which is a known marker of cervical high-grade lesion and squamous cell carcinoma. In normal cervical (NC) core tissue and low grade lesion p16INK4a was not detected (Fig. 5.10 A and B), whereas in HSIL and SCC it was highly expressed with nuclear and mostly cytoplasmic staining, as expected (Fig. 5.10 C and D).

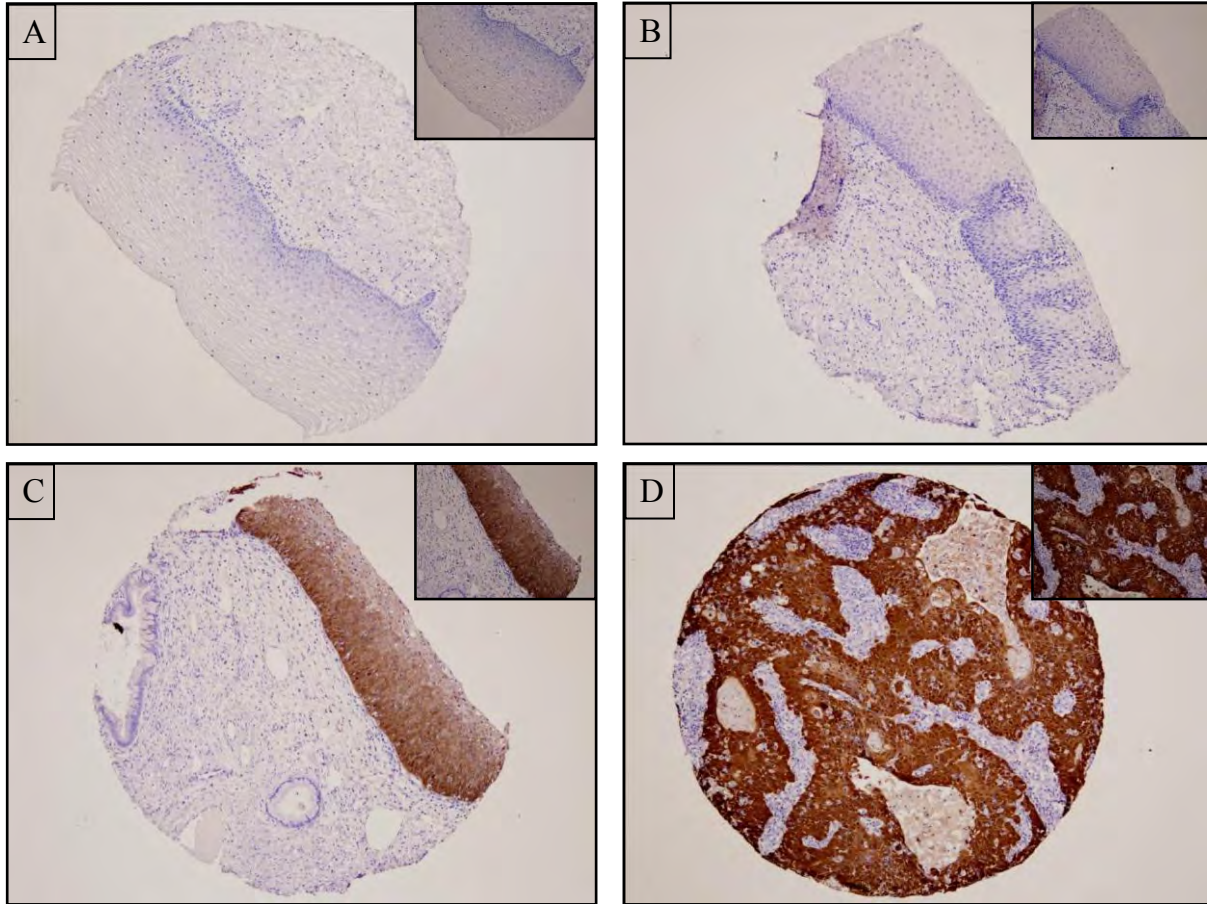


Figure 5.10 p16INK4a protein expression

p16INK4a protein expression detected by immunohistochemical staining in (A) normal cervical epithelium (NC), (B) low-grade squamous intraepithelial lesions (LSIL), (C) high-grade squamous intraepithelial lesions (HSIL), (D) squamous cell carcinoma (SCC). Magnification 20X; small image, Magnification 40X.

In conclusion these results showed that hnRNP A2/B1 is highly expressed in normal cervical tissue (NC) and low-grade squamous intraepithelial lesion (LSIL) and decreases from high-grade squamous intraepithelial lesion (HSIL) to squamous cell carcinoma (SCC) (Fig. 5.11) (Table 5.1). Since we have shown that hnRNP A2/B1 is involved in the regulation of late gene expression in HPV-16 and late genes are not expressed in cervical cancer these results suggested that it might be a correlation with the suppression of late genes and a lower expression of hnRNP A2/B1 in cervical cancer.

	NC	LSIL	HSIL	SCC
hnRNP A2/B1	+++	+++	+	+
p16INK4a	-	-	+++	+++

Table 5.1 Overall expression of proteins in core sections of cervical epithelium

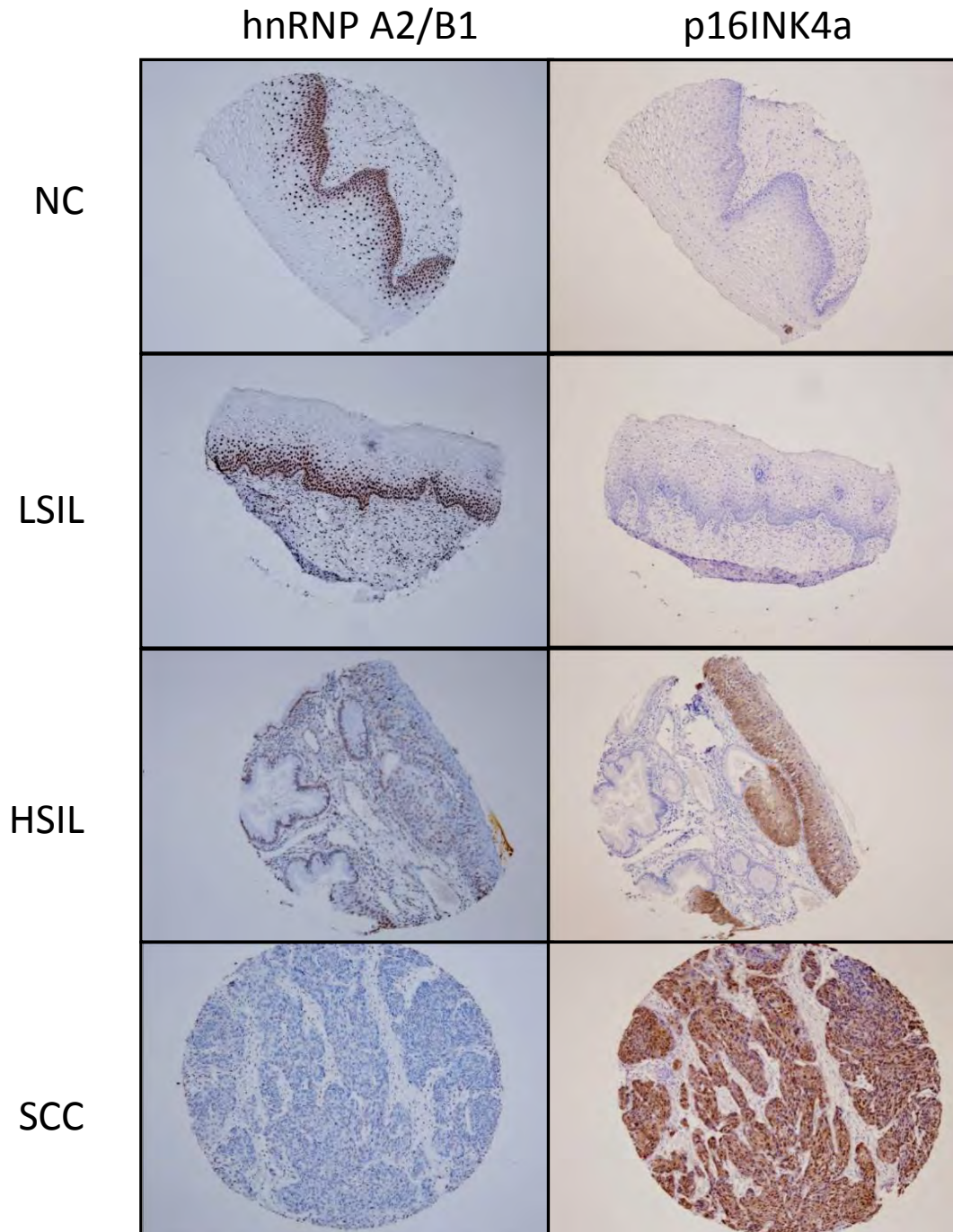


Figure 5.11 Comparison of immunostaining with hnRNP A2/B1 or p16INK4a antibodies on the same tissue microarray slide

Expression levels of hnRNP A2/B1 and p16INK4a in normal cervical tissue (NC), low-grade intraepithelial lesion (LSIL), high-grade intraepithelial lesion (HSIL) and squamous cell carcinoma (SCC).

6. Chapter 6 - Discussion

Cervical cancer is the second leading cause of death by cancer in women worldwide after breast cancer with approximately 250,000 deaths annually. Persistence of HPV and in particular high risk HPV subtype infection is a prerequisite for development of cervical cancer or pre-malignant lesions that could progress to cancer (Lowy & Howley, 2001). Although the lifetime risk for HPV infection in sexually active women is approximately 80% (Baseman & Koutsky, 2005), most infections are cleared spontaneously by the host immune system and only few cases develop cervical cancers after a period of persistent infection.

Two prophylactic vaccines utilizing virus-like particles have been developed to prevent HPV infection, Gardasil against HPV-6, -11, -16, -18, and Cervarix against HPV-16, -18. These vaccines have proved to be efficacious for individuals who have not been exposed to HPV prior to immunization and to prevent up to 70% of all cervical cancer (Batson *et al.*, 2006). Nonetheless, vaccination is not efficacious for individuals who are already infected (Ault, 2006), the duration of efficacy still needs to be confirmed and only few serotypes of HPV are targeted. There are other types not comprised in vaccine, such as HPV-31 and HPV-45, which can become dominant within the population and take over as the major cause of cervical cancer. Furthermore the high cost of vaccination won't be afforded by the developing countries where there is the highest percentage of cervical cancer cases. Therefore research into how the viruses alter the infected cells during infection, what cellular proteins are involved in HPV life-cycle and how gene expression is regulated during an infection will possibly help to develop new therapeutic agents and prevention methods that could prevent cervical cancer.

HPV gene expression is strictly dependent on cellular differentiation (Doorbar, 2005). At the molecular level, high-grade pre-malignant lesions and cervical cancer are characterised by

continued expression of the early viral genes, and a complete shut-down of the late viral genes encoding the viral structural proteins L1 and L2. These proteins are highly immunogenic and they can induce an immune response, therefore with inhibition of L1 and L2 gene expression in the lower layers of the epithelium the virus can escape surveillance by the immune system of the host and establish a persistent infection (Scheurer *et al.*, 2005). Thus it was speculated that activation of L1 and L2 late gene expression in the persistently infected cells would uncover the HPV infected cells for the immune system of the host.

The overall aim of this study was to elucidate the mechanism of HPV-16 gene regulation, primarily regulation of late gene expression. The first goal was to generate a simple reporter assay to facilitate the study of HPV-16 late gene expression; second was to establish useful stable cell lines for the identification of small molecules and cellular factors that can induce HPV-16 late gene expression and third to localize and determine expression levels of identified proteins involved in HPV late gene regulation.

Two HPV-16 genomic constructs were generated where the L1 gene is replaced with an easily detectable reporter gene such as chloramphenicol acetyltransferase (CAT), green fluorescent protein (GFP), secreted alkaline phosphatase (SEAP), Luciferase or β -galactosidase (LacZ). These reporters have different properties but all can be easily measured and quantified when transfected into proliferating cells.

The CAT assay resulted to be the most sensitive compared to the others and thus most of the experiments in this study were focused on that. The two reporters used most commonly were pBELCAT and pBELMCAT where the HPV-16 L1 gene was replaced by the CAT reporter. Evaluation of CAT expression levels showed that pBELCAT, as expected, produces very little CAT whereas pBELMCAT, in which a mutation that reduces the negative regulation on the

splice site present in the late region is introduced, efficiently produce CAT as demonstration that the splice site 5639 has a pivotal role in regulation of late gene expression. These results confirm that the CAT reporter mimics the expression patterns of the L1 gene in the intact viral genome.

In the presence of E4orf4, PTB, SRp30c or ASF/SF2 both reporter plasmids induced an increased expression of CAT. In particular Ad E4orf4, which regulates the switch from early to late gene expression in Adenovirus, showed a higher induction of CAT expression in pBELCAT when compared to the pBELMCAT plasmid, although this difference may be due to the high CAT levels pBELMCAT produces itself when transfected alone. In addition it has been shown that Ad E4orf4 overcomes the negative effect of splicing silencers present in the L1 region (Somberg *et al.*, 2009) and since in pBELMCAT these inhibitory elements had been inactivated by multiple point mutations, the viral protein can no longer target them and positively influence late gene expression as strong as it is observed in pBELCAT. When increasing quantities of the inducers cotransfected with pBELMCAT, higher levels of CAT were observed with the higher concentration of PTB and SRp30c, respectively. E4orf4 decreased CAT expression at higher concentration; this result could be correlated to other studies where adenovirus E4orf4 had shown to induce apoptosis (Roopchand *et al.*, 2001). It was previously shown that low levels of ASF/SF2 induced HPV-16 late gene expression whereas high levels did not (Somberg & Schwartz, 2010). Transfection of pBELMCAT with serially diluted ASF/SF2 plasmid revealed a similar effect, further demonstrating that expression of the CAT reporter gene is an accurate marker for HPV-16 late mRNA production. Unequal rates of ASF/SF2 levels in low-grade lesions and in high-grade lesions have been reported with low and high expression level, respectively (Mole *et al.*, 2009b). This could be due to the many roles that ASF/SF2 undertakes in pre- and post-splicing influencing expression of genes and metabolic activity in the cell. These

data clearly show that induction of CAT expression is dependent on the levels of these proteins, suggesting that it will be possible to identify cellular factors or small molecules that induce late gene expression. Thus functional reporter plasmids were generated that can be used to study HPV-16 late gene expression and that mimic the effect of the viral and cellular proteins on L1 gene expression as demonstrated in previously studies with the original pBEL and pBELM. It was of interest also to investigate the effect of 2 splice sites, SD3632 and SA5639, on these novel reporter plasmids. These splice sites are used exclusively by late mRNA and are suppressed in proliferating cancer cells (Rush *et al.*, 2005; Zhao *et al.*, 2007; Zhao *et al.*, 2004). Deletion of part of the early genes in pBspliceCAT showed no production of CAT whereas pBspliceMCAT, which has a mutation in the late region, modestly induced CAT expression validating the role of the splice site located in the late region for induction of late gene expression. Deletion of sequences located upstream the SD3632, in plasmid pMt1sdCAT resulted in an efficient production of CAT expression supporting that these sequences strongly suppress this splice site, therefore inhibiting late gene expression, as already shown in other studies (Somberg & Schwartz, 2010). Taken together these results demonstrated that these reporter plasmids can be used for identification and characterization of cis-acting regulatory RNA elements in HPV-16.

The second reporter gene was used is the green fluorescent protein (GFP), originated from the jellyfish *Aequorea Victoria*, a protein that auto-fluoresces and its expression is easily detected and quantifiable. At the same time GFP signals can be weak and hard to detect with a low expression. In this study pBELMGFP transfected alone showed low GFP intensities. Cotransfection assays revealed that the viral and cellular proteins Ad E4orf4, PTB and especially

SRp30c induced late gene production from pBELMGFP. Since the GFP intensities are detected from the living cells this assay proved to be very efficient, easy to use and inexpensive.

In order to find the easiest way to analyse HPV-16 late gene expression it was decided to insert the LacZ reporter gene, which encodes for β -galactosidase and seems to be the most versatile reporter gene. Unfortunately cloning of LacZ into pBEL-IRES and pBELM-IRES failed several times possibly it caused by the size of the reporter gene. In fact LacZ gene is 3141 bp in size, a quite big fragment to insert into pBEL-IRES and pBELM-IRES plasmids, which both are already approximately 10000 bp in size. Transformation efficiency decreases as the size of the DNA being transformed goes up therefore it is more difficult to introduce into bacterial cell larger constructs.

Next two reporter plasmids were generated with the SEAP reporter gene which can be simply analyzed in proliferating cells. The SEAP reporter gene product is secreted from mammalian cells and is thus easily detected in a sample of culture medium but unfortunately in this study this assay was not sensitive enough to analyze late gene expression in subgenomic HPV-16 reporter plasmids. Some companies do not recommend HeLa cells as a cell line for experiments with SEAP due to high background. In these experiments in each transfection very high background levels were observed and SEAP expression from the reporter plasmids was undetectable. Only in pMt1sdSEAP plasmid, where sequences that suppress SD3632 are deleted, it is possible to detect SEAP expression confirming once again that this splice site is involved and has an important role in regulation of late gene expression. Therefore it can be reported that this plasmid could be used to study cis-acting, splicing regulatory elements at the HPV-16 late splice sites.

In conclusion subgenomic HPV-16 plasmids with different and versatile reporter genes such as CAT, GFP, Luciferase and SEAP in place of the late L1 gene were generated.

These reporter genes are functional surrogate markers for HPV-16 late gene expression. It was demonstrated that E4orf4, PTB, SRp30c and ASF/SF2 proteins efficiently induce HPV-16 late gene expression and that induction is dependent on the levels of these proteins, confirming the idea that it should be possible to identify cellular factors and small molecules that could induce HPV-16 late gene expression.

The second step was to generate stable cell lines with CAT reporter plasmids integrated into the chromosomal DNA in order to obtain a model for a large scale screening of small molecules and to evaluate levels of CAT expression in these cell lines. Although transient transfections are useful tools for fast analysis of genes, a stable cell line ensures reproducible, long-term and defined gene expression.

The stable cell lines were evaluated by PCR amplification of CAT and IRES sequences from extracted DNA and evaluation of CAT levels. The pBELCAT-derived cell line resulted in 5 positive clones and 2 from pBELMCAT-derived cell lines but they have been demonstrated to behave differently, producing different amount of CAT units. These results revealed that not all the stable cell lines behave at the same way. This may be determined by different integration sites in the genome of the target cell that can affect transcription rate of the gene of interest (Wurm, 2004). A regular expression plasmid is usually integrated into the cellular chromosomal DNA randomly (Murnane *et al.*, 1990). Organization of genomic DNA into euchromatic and heterochromatic regions widely alters the expression of sequences contained within these regions. Integration into active euchromatin allows transgene expression, whereas integration into inactive heterochromatin results in little or no transgene expression and random integration often leads to silencing of the transgene. This might explain why it is possible to see such a difference in CAT expression from the stable cell lines; in particular in pBELMCAT stable cell

lines which, one of them (pBELMCAT-31), produced really high levels of CAT when untransfected, whereas in the other one (pBELMCAT-25) much lower levels of CAT were observed.

In the presence of E4orf4, PTB or SRp30c a slightly increase in CAT expression was observed in pBELMCAT cell lines but these results, when compared to the induction levels seen in transient transfections, showed a lower induction of CAT expression. This could be due to the fact that the levels of CAT produced by the pBELMCAT-derived cell line alone were too high and expression could only be slightly increased, or low transfection efficiency affected CAT production by adenovirus E4orf4, PTB or SRp30c, resulting in a relatively low increase in CAT. If only 5% of the stable cells are transfected with the expression vector then the CAT induction in these cells would be masked by the levels of CAT produced in the 95% of untransfected cells. Two of the five pBELCAT-derived cell lines failed to induce CAT expression in the presence of PTB whereas three showed an increase of CAT levels in a dose dependent manner. Also in this case a lower induction compared to the transient transfections is observed. The transient cotransfection results appear to be more efficient than stable transfection for the demonstration of induction of late gene expression. Transfection efficiency could be increased using different methods of transfection, such as electroporation which have been shown that 80% of the cells can receive the foreign DNA (Nickoloff, 1995). After analysis of CAT expression as described above, two stable cell lines (pBELCAT-67 and pBELMCAT-31) consistently reproduced same effects seen in transient transfection by viral and cellular proteins in CAT reporter plasmids, reliably reproducing L1 gene expression in the original subgenomic HPV-16 plasmids. RNA analysis on pBELCAT and pBELMCAT stable cell lines showed that correctly spliced HPV-16 early and late mRNAs were produced, confirming that we established functional stable cell lines

that could be used for the identification of cellular proteins or microRNAs that may regulate HPV-16 gene expression.

Generation of pBELMGFP-derived cell lines was less successful. Low levels of fluorescence were detected in each clone that had the IRES and GFP inserted in the cellular genome. Also in transient transfection GFP expression from this plasmid was very low and detectable only in few cells. It was expected that with a stable transfection all cells should have the transgene inserted in the cellular genome and, considering the high levels of CAT produced by the pBELMCAT-31 stable cell line higher levels of fluorescence were expected from the GFP stable cell lines. Although the evidence presented here, using GFP instead of CAT appeared as a step forward to update and simplify the system, the signal was too low to be considered a reliable and accurate instrument to study HPV-16 gene regulation. Therefore, CAT stable cell lines, which proved functionality and authenticity, were chosen for further study and screening of small molecules.

The major anticipated use of the stable cell lines was in the screening of small molecule libraries. Libraries of compounds are increasingly becoming commercially available for the use of individual academic laboratories. Small molecules associate with, or bind to, a protein in specific ways and in some cases they modulate the protein's function by inactivating or activating it. Therefore it was of interest to look at a number of small molecules that have been shown to interfere with cellular proteins involved in the regulation of HPV-16 late gene expression. Analysis of compounds on the described stable cell lines could be also useful for the identification of novel regulatory elements of late gene expression in HPV. In addition, many small molecules have being widely used as drugs for the treatment of many types of cancer (Collins & Workman, 2006; Garcia *et al.*, 2008) and for this reason we wished to identify small molecules suitable for treatment of HPV infections. Thus diverse substances were tested on CAT

stable cell lines with an emphasis on molecules that have been demonstrated to have an effect on cellular splicing mechanisms. The system was initially tested with TPA which is a known inducer of late gene expression in HPV-31 (Meyers *et al.*, 1992). A consistent CAT induction was observed after a treatment of these stable cell lines with TPA, proving the utility of pBELCAT-67 and pBELMCAT-31 cell lines as high quality tools to study HPV-16 late gene expression. Treatment of pBELCAT and pBELMCAT-derived cell lines with other substances, such as emetine, ceramide and cycloheximide had no or an inhibitory effect on late gene expression in both cell lines, suggesting that CAT induction was specific for the small molecule TPA. This inhibitory effect could be caused by their role in inhibition in protein synthesis (Grollman, 1968; Grollman, 1966).

Induction of CAT was also observed in both cell lines in the presence of valproic acid and tannic acid, two substances that have been shown to enhance levels of cellular proteins, such as ASF/SF2 and PTB, respectively. Therefore we deduced that this effect is due to the increased levels in the cell of these cellular factors that we have demonstrated to induce expression of late genes in transient and stable transfected subgenomic HPV-16 plasmids.

To date, valproic acid it has been used for 45 years as a chronic therapy for epileptic disorders showing a good tolerability and safety profile (Loscher, 2002). Tannic acid, which for many years has been used to treat burn patients, has been revealed to have hepatotoxic effects (Duffin, 1942; McClure & Lam, 1940). Evidence of an impaired liver function accompanied by central liver necrosis was reported in some cases of patients treated with tannic acid and many other studies in different animals provided further support for its toxicity (Clark & Rossiter, 1943; Erb *et al.*, 1943; Jackson, 1944; Saltonstall *et al.*, 1945; Wells *et al.*, 1942). A relevant toxicity was also seen in our cell lines after treatment with tannic acid at high concentrations. Therefore, we

can assume that valproic acid is the safest small molecule identified in this study that could be easily used for HPV infections treatment, but also tannic acid, that is generally present in red wine, could potentially be a good candidate to cure persistently HPV infected cells by exposing the highly immunogenic L1 protein to the immune system. Indeed, it would be a big breakthrough if a glass of red wine per day would help the fight against cervical cancer.

Two more small molecules, amiloride and scriptaid, were tested in this study. Scriptaid had no effect on late gene expression whereas a slight induction of CAT was observed in both cell lines after treatment with amiloride, but it was not consistent. This might be due to the various effects this substance has on different cellular factors. In fact it has been shown that amiloride treatment of cells results in hypo-phosphorylation of SR proteins, in particular SF2/ASF (Chang *et al.*, 2011), and it has been shown that both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing (Cao *et al.*, 1997). Phosphorylation is essential for SR protein nuclear import and for spliceosome assembly (Roscinno & GarciaBlanco, 1995; Xiao & Manley, 1997; Yeakley *et al.*, 1999; Yun & Fu, 2000; Lai *et al.*, 2001), and partial SR proteins dephosphorylation is also essential for progression of the assembled spliceosome to catalysis (Mermoud *et al.*, 1994; Cao *et al.*, 1997) and for post-splicing events such as interaction with the mRNA transport machinery and SR protein-mediated translational control (Huang *et al.*, 2004; Lai & Tarn, 2004; Lin *et al.*, 2005; Sanford *et al.*, 2005). Therefore, both SR protein hypo- and hyper-phosphorylation inhibits splicing (Prasad *et al.*, 1999). In addition, amiloride upregulates a number of proteins of the hnRNP family, such as hnRNP A1, hnRNP A2/B1 and hnRNP Q1 (Chang *et al.*, 2011). In this study was shown that hnRNP A2/B1 positively modulates late gene expression but in previous studies it has been shown that hnRNP A1 inhibits splicing by interaction with HPV-16 L1 3' splice silencer at 5639 (Zhao *et al.*, 2004). Therefore, even though

amiloride overexpresses hnRNP A2/B1, which should induce late gene expression, other factors are altered in the cell, that might interfere with the positive effect of hnRNP A2/B1 or, in general, with the regulation of late splice sites thereby inhibiting late gene expression.

Scriptaid, like valproic acid, is a histone deacetylase inhibitor but, while valproic acid efficiently induced CAT in both cell lines, scriptaid did not, suggesting that the induction of late gene expression by valproic acid is due to the upregulation of ASF/SF2 in the cell by this substance and not for the action as HDAC inhibitor. Our findings further authenticate the pivotal role of cellular RNA binding factors, like in this particular case ASF/SF2, in HPV-16 gene regulation. In fact, it has been demonstrated that late HPV gene expression is regulated at the transcriptional and post-transcriptional levels (Schwartz, 2008; Zheng & Baker, 2006) and alternative splicing is necessary to produce L1 since the 3' end of L2 and the 5' end of L1 overlap. Inhibitory RNA elements present in the L1 sequence of HPV-16 prevent production of late mRNAs (Collier *et al.*, 2002) and late gene expression is also inhibited in mitotic cells because of a competition between early and late splice sites (Rush *et al.*, 2005). Therefore, since this study and others have demonstrated that cellular RNA binding factors and viral RNA elements are very important for HPV-16 gene regulation (Johansson *et al.*, 2011; Mole *et al.*, 2006; Schwartz *et al.*, 2007) it was investigated further the role of these cellular proteins on the regulation of late gene expression in HPV-16 by knocking down protein expression using lentivirus delivered RNAi. The target genes chosen for knockdown were ASF/SF2 and PTB to better analyse regulation of HPV-16 late mRNAs and to see if the expression of late genes would be effected. So far we have shown that with overexpression in the cell of these two cellular proteins the expression of HPV-16 late genes is increased, therefore with their downregulation it would be expected a decrease of late mRNAs production even though more than 90% of cells were infected with lentiviruses carrying

ASF/SF2, PTB or scramble shRNA, knockdown of these specific genes failed. In this study only one shRNA construct for each gene was used and the possibility of modulating gene expression level rises using multiple shRNA constructs each covering a unique region of the target gene. This might be the reason for an ineffective knockdown. Forasmuch as HPV-16 late gene expression is a complex process that involves many regulatory steps it would be interesting to further investigate or to fully knockdown expression of genes on these stable cell lines to intimately explore their relationship with HPV-16 gene regulation. It would be also of interest to start a screening of antisense oligonucleotides against HPV-16 for ability to induce late gene expression in infected cancer cells.

An interesting finding in this study was the identification of hnRNP A2/B1 as a novel regulator of HPV-16 late gene expression. This RNA-binding factor efficiently induced CAT from transient and stable CAT reporter plasmids, whereas another common splicing factor, such as SC35, had no effect on CAT expression. Sureau *et al.* reported that SC35 accumulation is correlated to changes in the splicing pattern regulating its expression by promoting alternative splicing events changing the stability of its own mRNAs (Sureau *et al.*, 2001). They also demonstrated that endogenous SC35 mRNA levels in HeLa cells are decreased when SC35 is overexpressed and this could be the reason of failure in CAT induction by SC35 in the reporter plasmids. When an hnRNP A2/B1 expressing plasmid was cotransfected with pBspliceMCAT and pMt1sdCAT, plasmids containing the late region with the CAT reporter gene, part of the early region but only the two splice sites SD3632 and SA5639, an induction of CAT was observed in both reporter plasmids, suggesting that this RNA-binding factor acts on sequences adjacent to these two splice sites. The lower induction observed in pMt1sdCAT by hnRNP A2/B1 could either indicate that this splicing factor binds specifically to sequences that suppress

SD3632, that have been deleted in this plasmid, therefore hnRNP A2/B1 can no longer target them, or also can be caused by the higher levels of CAT produced by this reporter plasmids alone, and therefore levels of CAT could be only modestly enhanced. Further studies on hnRNP A2/B1 would be useful to identify specific sequences on the HPV-16 genome that this RNA-binding factor targets. With a serial dilution of hnRNP A2/B1 higher levels of CAT were observed at lower concentrations of hnRNP A2/B1, whereas at higher concentrations CAT induction was inhibited. This could be due to low transfection efficiency, caused by the considerable amount of exogenous DNA and transfection reagent that can be toxic to the cells, or to other metabolic activity this RNA-binding factor can trigger in the cell. Recently, hnRNP A2/B1 has been designated as a putative proto-oncogene due to its overexpression in glioblastoma cells and to its role in the regulation of a number of tumor suppressors and oncogenes, such as c-FLIP, BIN1, Wwox and RON (Golan-Gerstl *et al.*, 2011). Despite the fact that overexpression of hnRNP A2/B1 in different tumors has been reported by other researchers (Tockman *et al.*, 1997; Tockman *et al.*, 1988; Zhou *et al.*, 1996), immunohistochemistry analysis, carried out in this study on cervical tissue microarrays, revealed that hnRNP A2/B1 is highly expressed in normal cervical tissue and low-grade squamous intraepithelial lesion (LSIL), whereas expression levels of this protein decreased from high-grade squamous intraepithelial lesion (HSIL) to squamous cell carcinoma (SCC). Our findings are in line with other studies that proved a down-regulation of hnRNP A2/B1 in tumors, such as the colon, thyroid, small intestine and kidney (Karni *et al.*, 2007).

Since hnRNP A2/B1 regulates and stimulates the production of late mRNAs, it would be expected that in the upper layers of the epithelium, where expression of HPV late genes naturally occurs, the levels of hnRNP A2/B1 would be high, inducing late gene expression. In this study in

epithelium that was infected with HPV but was histologically normal and in low-grade squamous intraepithelial lesion, hnRNP A2/B1 was highly expressed in the basal and intermediate layers, but as the cells differentiate no expression was detected. This may suggest that expression of HPV late genes in the upper layers of the cervical epithelium is dependent on other factors and regulatory steps. In high-grade lesions and tumour samples expression of hnRNP A2/B1 radically decreased, therefore, considering that HPV-16 late genes expression is suppressed in cervical cancer and hnRNP A2/B1 regulates expression of late gene in HPV-16, we suppose that it might be a correlation with the lower expression of hnRNP A2/B1 and prevention of late gene expression in cervical cancer. Other components of the hnRNP family, such as hnRNP H, hnRNP A1 and hnRNP C1/C2, that unlike hnRNP A2/B1 inhibit late gene expression, and hnRNP I have been shown to be highly expressed in HSIL and SCC (Fay *et al.*, 2009).

It was also determined the expression of a known marker of HSIL and SCC (Carozzi, 2007; Hairiri & Oster, 2007; Iaconis *et al.*, 2007), such as p16INK4a, on the same TMA, to validate the integrity of all tissue cores in the TMA slides and to compare their levels of expression in the different cases. As expected p16INK4a expression was not detected in normal cervical tissue cores and LSIL, but a very strong staining was observed in HSIL and SCC proving the integrity of the tissue cores and therefore authenticating the veracity of the immunostainings performed in this study.

In conclusion, CAT and GFP genes in place of the late L1 gene in subgenomic HPV-16 plasmids proved to be functional surrogate markers for HPV-16 late gene expression in HeLa cells. Moreover functional stable cell lines useful for the identification of cellular proteins, small molecules or microRNAs that regulate HPV-16 gene expression were established. Three small molecules such as TPA, valproic acid and tannic acid positively modulated late gene expression

in HPV-16-derived cell lines and these substances, in particular valproic acid, could potentially be used as antiviral drugs to treat persistent HPV infections. Furthermore a RNA-binding factor was identified as a novel regulator of HPV-16 late genes expression and immunohistochemical analysis allowed us to detect and localize this protein in cervical epithelium at the different stages after HPV infection.

Since it has been shown that the prevalence of HPV in the oropharyngeal area in men is increasing significantly and it is suspected that oropharyngeal squamous cell carcinoma will exceed cervical malignancies in women by 2020 (Chatuverdi *et al.*, 2011; Sanders *et al.*, 2012), it would be useful to speculate that the results presented in this thesis might have a major application in this area in the future.

It also would be interesting to extend the experiments to other HPV types, such as HPV-6 or HPV-18, since they have different pathogenic proprieties and although this study focussed exclusively on HPV-16, the most common HPV type found in cervical cancer, the reporters presented herein could be easily modified to examine other HPV types and this would it be useful for a better understanding of the differences in the gene expression mechanism of HPVs.

7. References

Albrecht H. Can Valproic Acid Help Deplete Latent Reservoirs of HIV? *AIDS Clinical Care Massachusetts Medical Society* 2005.

Admirant AM, Hendricks JA, De Witt Hamer PC *et al.* Valproic Acid is toxic to malignant glioma cells and increases sensitivity to irradiation and chemotherapy. Abstracts for the Seventh Congress of the European Association for Neuro-Oncology (EANO). Vienna, Austria,. Neuro-Oncology 2006; Abstract No. P80: 334.

Akusjärvi G. Temporal regulation of adenovirus major late alternative RNA splicing. *Front Biosci* 2008; **13**: 5006-15.

Akusjärvi G, Stevenin J. Remodelling of the host cell RNA splicing machinery during an adenovirus infection. *Curr Top Microbiol Immunol* 2003; **272**: 253–286.

Andreason GL, Evans GA. Introduction and expression of DNA molecules in eukaryotic cells by electroporation. *Biotechniques* 1988; **6**: 650–660.

Arends MJ, Buckley CH, Wells M. Aetiology, pathogenesis, and pathology of cervical neoplasia. *J Clin Pathol* 1998; **51**: 96-103.

Ault KA. Epidemiology and natural history of human papillomavirus infections in the female genital tract. *Infect Dis Obstet Gynecol* 2006; **2006**: 1-5.

Baker C, Calef C. Maps of papillomavirus mRNA transcripts. In SR Billakanti, Ed CE Calef, AD Farmer, AL Halpern, GL Myers. Human papillomaviruses: a compilation and analysis of nucleic acid and amino acid sequences. Theoretical biology and biophysics. Los Alamos National Laboratory, Los Alamos, N. Mex. 1997.

Baseman JG, Koutsky LA. The epidemiology of human papillomavirus infections. *J Clin Virol* 2005; **32**: 16–S24.

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

Bentley DL. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr Opin Cell Biol* 2005; **17**: 251-6.

Benevolo M, Vocaturo A, Mottolese M *et al*. Clinical role of p16INK4a expression in liquid-based cervical cytology: correlation with HPV testing and histologic diagnosis. *Am J Clin Pathol* 2008; **129**: 606–612.

Bernstein E, Caudy AA, Hammond SM, Hannon GJ. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* 2001; **409**: 363-6.

Bernard HU. The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. *J Clin Virol* 2005; **32 Suppl 1**: S1-6.

Bian Y, Masuda A, Matsuura T *et al*. Tannic acid facilitates expression of the polypyrimidine tract binding protein and alleviates deleterious inclusion of CHRNA1 exon P3A due to an hnRNP H-disrupting mutation in congenital myasthenic syndrome. *Human Molecular Genetics* 2009; **18**: 1229–1237

Birney E, Kumar S, Krainer AR. Analysis of the RNA recognition motif and RS and RGG domains: conservation in metazoan premRNA splicing factors. *Nucleic Acids Res* 1993; **21**: 5803-16.

Blow N. Delivering the future. *Nature* 2007; **450**: 1117-20.

Blobe GC, Stribling S, Obeid LM, Hannun YA. Protein kinase C isoenzymes: regulation and function. *Cancer Surv* 1996; **27**: 213-48.

Blumberg PM. Protein Kinase C as the Receptor for the Phorbol Ester Tumor Promoters: Sixth Rhoads Memorial Award Lecture1. *Cancer Res* 1983; **1988**: 48:1-8.

Blumberg PM. In vitro studies on the mode of action of the phorbol esters, potent tumor promoters: part 1. *Crit Rev Toxicol* 1980; **8**: 153-97.

Boise LH, Gonzalez-Garcia M, Postema CE *et al.* Bcl-x, a Bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. *Cell* 1993; **74**: 597–608.

Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV. The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 2002; **55**: 244–265.

Bosch FX, Muñoz N, Shah KV, Meheus A. Int J. Second International Workshop on the epidemiology of cervical cancer and human papillomavirus. *Cancer* 1992; **52**: 171-173.

Boon-Unge K, Yu Q, Zou T, Zhou A, Govitrapong P, Zhou J. Emetine regulates the alternative splicing of *Bcl-x* through a protein phosphatidase 1- dependent mechanism. *Chem Biol* 2007; **14**: 1386-1392.

Brambilla E, Gazdar. A series “lung cancer”. *Eur Respir J* 2009; **33**: 1485–1497.

Brichta L, Hofmann Y, Hahnen E *et al.* Valproic acid increases the SMN2 protein level: a well-known drug as a potential therapy for spinal muscular atrophy. *Human Molecular Genetics* 2003; **12**: 2481–2489.

Brown TA. Chapter 10: Synthesis and processing of RNA, *In* Genomes 2. Bios Scientific Publishers, New York. 2002: 273-312.

Bull MB, Laragh JH. Amiloride. A potassium-sparing natriuretic agent. *Circulation* 1968; **37**: 45–53.

Burghes AH. When is a deletion not a deletion? When it is converted. *Am J Hum Genet* 1997; **61**: 9–15.

Burk RD, Chen Z, Van Doorslaer K. Human Papillomaviruses: Genetic Basis of Carcinogenicity. *Public Health Genomics* 2009; **12**: 281–290.

Burton BS. On the propyl derivatives and decomposition products of ethylacetoacetate. *Am Chem J* 1882; **3**: 385–395.

Cáceres JF, Misteli T, Sreaton GR, Spector DL, Krainer AR. Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J Cell Biol* 1997; **138**: 225-238.

Caceres JF, Sreaton GR, Krainer AR. A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes Dev* 1998; **12**: 55-66.

Castagnag M, Takai Y, Kaibuchi K, Sano K, Kikkawa U, Nishizukag Y. Direct Activation of Calcium-activated, Phospholipid-dependent Protein Kinase by Tumor-promoting Phorbol Esters. *J Biol Chem* 1982; **257**: 7847-7851.

Cao W, Jamison SF, Garcia-Blanco MA. Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. *RNA* 1997; **3**: 1456–1467.

Carozzi FM. Combined analysis of HPV DNA and p16INK4a expression to predict prognosis in ASCUS and LSIL pap smear. *Coll Antropol* 2007; **2**: 103-106.

Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nature Reviews Genetics* 2002; **3**: 285–298.

Castanotto D, Rossi JJ. The promises and pitfalls of RNA-interference-based therapeutics. *Nature* 2009; **457**: 426-33.

Centers for Disease Control and Prevention (CDC). FDA licensure of quadrivalent human papillomavirus vaccine (HPV4, Gardasil) for use in males and guidance from the Advisory Committee on Immunization Practices (ACIP). *Morb Mortal Wkly Rep (MMWR)* 2010; **59**: 630-2.

Chalfant CE, Ogretmen B, Galadari S, Kroesen BJ, Pettus BJ, Hannun YA. FAS Activation Induces Dephosphorylation of SR Proteins. Dependence on the *de novo* generation of ceramide and activation of protein phosphatase 1. *J Biol Chem* 2001; **276**: 44848–44855

Chalfant CE, Rathman K, Pinkerman RL, Wood RE, Obeid LM, Ogretmen B, Hannun YA. De novo ceramide regulates the alternative splicing of caspase 9 and Bcl-x in A549 lung adenocarcinoma cells. Dependence on protein phosphatase-1. *J Biol Chem* 2002; **277**: 12587–12595.

Chandler SD, Mayeda A, Yeakley JM, Krainer AR, Fu XD. RNA splicing specificity determined by the coordinated action of RNA recognition motifs in SR proteins. *Proc Natl Acad Sci* 1997; **94**: 3596–3601.

Chang BS, Kelekar A, Harris MH, Harlan JE, Fesik SW, Thompson CB. The BH3 domain of Bcl-x(S) is required for inhibition of the antiapoptotic function of Bcl-x(L). *Mol Cell Biol* 1999; **19**: 6673–6681.

Chang JG, Yang DM, Chang WH *et al.* Small molecule amiloride modulates oncogenic RNA alternative splicing to devitalize human cancer cells. *PLoS One* 2011; **6**: e18643.

Chang WH, Liu TC, Yang WK *et al.* Amiloride modulates alternative splicing in leukemic cells and resensitizes *Bcr-AbIT315I* mutant cells to imatinib. *Cancer Res* 2011; **71**: 383-392.

Chatuverdi AK, Engels EA, Pfeiffer RM *et al.* Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol* 2011 online early.

Chew SL, Liu HX, Mayeda A, Krainer AR. Evidence for the function of an exonic splicing enhancer after the first catalytic step of pre-mRNA splicing. *Proc Natl Acad Sci USA* 1999; **96**: 10655-10660.

Chen J, Ghazawi FM, Bakkar W, Li Q. Valproic acid and butyrate induce apoptosis in human cancer cells through inhibition of gene expression of Akt/protein kinase B. *Molecular Cancer* 2006; **5**: 71.

Cho YL, Lee KS, Lee SJ *et al.* Amiloride potentiates TRAIL-induced tumor cell apoptosis by intracellular acidification-dependent Akt inactivation. *Biochem Biophys Res Commun* 2005; **326**: 752-8.

Chong T, Apt D, Gloss B, Isa M, Bernard HU. The enhancer of human papillomavirus type 16: binding sites for the ubiquitous transcription factors oct-1, NFA, TEF-2, NF1 and AP-1 participate in epithelial cell-specific transcription. *J Virol* 1991; **65**: 5933-43.

Ciaran B, Woodman J, Stuart I, Collins & Lawrence S Young. The natural history of cervical HPV infection: unresolved issues. *Nature Reviews Cancer* 2007; **7**: 11-22.

Clancy JK, Yan L, Mack KM *et al.* A novel histone deacetylase inhibitor, Scriptaid, enhances expression of functional estrogen receptor α (ER) in ER negative human breast cancer cells in combination with 5-aza 2-deoxycytidine. *Breast Cancer Research and Treatment* 2003; **81**: 177–186.

Clark EJ, Rossiter RJ. Liver function in rabbits after injection of tannic acid. *Lancet* 1943; **21**: 222-3.

Colgan DF, Manley JL. Mechanism and regulation of mRNA polyadenylation. *Genes Dev* 1997; **11**: 2755-66.

Collier B, Oberg D, Zhao X, Schwartz S. Specific inactivation of inhibitory sequences in the 5' end of the human papillomavirus type 16 L1 open reading frame results in production of high levels of L1 protein in human epithelial cells. *J Virol* 2002; **76**: 2739–2752.

Collins I, Workman P. New approaches to molecular cancer therapeutics. *Nature Chemical Biology* 2006; **2**: 689 – 700.

Collins AS, Nakahara T, Do A, Lambert PF. Interactions with pocket proteins contribute to the role of human papillomavirus type 16 E7 in the papillomavirus life cycle. *J Virol* 2005; **79**: 14.769-80.

Colosimo A, Goncz KK, Holmes AR *et al.* Transfer and expression of foreign genes in mammalian cells. *Biotechniques* 2000; **29**: 314–318.

Committee opinion no. 467: “Human papillomavirus vaccination”. *Obstet Gynecol* 2010; **116**: 800-3.

Cooper, G.M. and R.E. Hausman, Chapter 7: RNA synthesis and processing, in the cell: a molecular approach. *ASM Press* 2007: 253-308.

Covarrubias L, Nishida Y, Terao M, D'Eustachio P, Mintz B. Cellular DNA rearrangements and early developmental arrest caused by DNA insertion in transgenic mouse embryos. *Mol Cell Biol* 1987; **7**: 2243–2247.

Cozzolino D, Cynkar WU, Dambergs RG, Mercurio MD, Smith PA. Measurement of condensed tannins and dry matter in red grape homogenates using near infrared spectroscopy and partial least squares. *J Agric Food Chem* 2008; **56**: 7631–7636.

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

Culp TD, Budgeon LR, Christensen N. Human papillomavirus bind a basal extracellular matrix component secreted by keratinocytes which is distinct from a membrane-associated receptor. *Virology* 2006a; **347**:147-159.

Culp TD, Budgeon LR, Marinkovich P, Meneguzzi CN. Keratinocyte-secreted laminin-5 can function as a transient receptor for human papillomaviruses by binding virions and transferring them to adjacent cells. *J Virol* 2006b; **80**: 8940-8950.

Cumming SA, McPhillips MG, Veerapraditsin T, Milligan SG, Graham SV. 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**: 5167-77.

Dahl E, En-Nia A, Wiesmann F *et al.*, Nuclear detection of Y-box protein-1 (YB-1) closely associates with progesterone receptor negativity and is a strong adverse survival factor in human breast cancer. *BMC Cancer* 2009; **9**: 410.

Dell G, Gaston K. Human papillomaviruses and their role in cervical cancer. *Cell Mol Life Sci* 2001; **58**: 1923-42.

Demers GW, Foster SA, Halbert CL, Galloway DA. Growth arrest by induction of p53 in DNA damaged keratinocytes is by-passed by human papillomavirus 16 E7. *Proc Natl Acad Sci USA* 1994; **91**: 4382-6.

de Villiers EM., Fauquet C, Broker TR, Bernard HU, zur Hausen H. Classification of papillomaviruses. *Virology* 2004; **324**: 17-27.

Dixon EP, Pahel GL, Rocque WJ *et al.* The E1 helicase of human papillomavirus type 11 binds to the origin of replication with low sequence specificity. *Virology* 2000; **270**: 345-57.

Doorbar J, Molecular biology of human papillomavirus infection and cervical cancer. *Clinical Sciences* 2006; **110**: 525-541.

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

Doorbar J, Campbell D, Grand RJ, Gallimore PH. Identification of the human papilloma virus-1a E4 gene products. *Embo J* 1986; **5**: 355-62.

Doorbar J, Ely S, Sterling J, McLean C, Crawford L. Specific interaction between HPV 16 E1-E4 and cytokeratins results in collapse of the epithelial cell intermediate filament network. *Nature* 1991; **352**: 824-827.

Doorbar J, Parton A, Hartley K et al. Detection of novel splicing patterns in a HPV16-containing keratinocyte cell line. *Virology* 1990; **178**: 254-262.

Dredge BK, Polydorides AD, Darnell RB. The splice of life: Alternative splicing and neurological disease. *Nature Reviews Neuroscience* 2001; **2**: 43-50.

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

Drobni P, Mistry N, McMillan N, Evander M. Carboxy-fluorescein diacetate, succinimidyl ester labelled papillomavirus virus-like particles fluoresce after internalization and interact with heparan sulphate for binding and entry. *Virology* 2003; **310**: 163-172.

Duffin JD. Liver necrosis following burns. *Can Med Ass J* 1942; **47**: 138-9.

Durst M, Croce CM, Gissmann L, Schwarz E, Huebner K. Papillomavirus sequences integrate near cellular oncogenes in some cervical carcinomas. *Proc Natl Acad Sci USA* 1987; **84**: 1070-1074.

Dyson NJ, Howley PM, Munger K, Harlow E. The human papilloma virus-16 E7 oncoprotein is able to bind to the retinoblastoma gene product. *Science* 1989; **243**: 934-937.

Elgin SC. Heterochromatin and gene regulation in *Drosophila*. *Curr Opin Genet Dev* 1996; **6**: 193-202.

Erb IH, Morgan EM, Farmer AW. The pathology of burns. *Ann Surg* 1943; **117**: 234-55.

Evander M, Frazer IH, Payne E, Qi YM, Hengst K, McMillan NA. Identification of the $\alpha 6$ -integrin as a candidate receptor for papillomaviruses. *J Virol* 1997; **71**: 2449-56.

- Faustino NA, Cooper TA. Pre-mRNA splicing and human disease. *Genes Dev* 2003; **17**: 419-437.
- Favre M, Ramoz N, Orth G. Human papillomaviruses: general features. *Clin Dermatol* 1997; **24 Suppl3**: S219-25.
- Fay J, Kelehan P, Lambkin H, Schwartz S. Increased expression of cellular RNA-binding proteins in HPV-induced neoplasia and cervical cancer. *J Med Virol* 2009; **81**: 897-907.
- Fehrmann F, Laiminis LA. Human papillomaviruses: targeting differentiating epithelial cells for malignant transformation. *Oncogene* 2003; **22**: 5201-7.
- Feldkotter M, Schwarzer V, Wirth R, Wienker TF, Wirth B. Quantitative analyses of SMN1 and SMN2 based on real-time lightcycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am J Hum Genet* 2002; **70**: 358–368.
- Felgner JH *et al.* Cationic lipid-mediated transfection in mammalian cells: “Lipofection”. *J Tiss Cult Meth* 1993; **15**: 63-68.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998; **391**: 806-11.
- Frattoni MG, Laiminis LA. Binding of the human papillomavirus E1 origin-recognition protein is regulated through complex formation with the E2 enhancer-binding protein. *Proc Natl Acad Sci USA* 1994; **91**: 12398-402.
- Fu XD. Specific commitment of different pre-mRNAs to splicing by single SR proteins. *Nature* 1993; **365**: 82-5.
- Fuchs PG, Pifister H. Transcription of papillomavirus genomes. *Intervirology* 1994; **37**: 159-167.
- Fürstenberger G, Berry DL, Sorg B, Marks F. Skin tumor promotion by phorbol esters is a two-stage process. *Proc Natl Acad Sci U S A* 1981; **78**: 7722-6.
- Gall JG. Small granules in the amphibian oocyte nucleus and their relationship to RNA. *J Biophys Biochem Cytol* 1956; **2**: 393–396.

Garcia AA, PharmD, BCOP. Small Molecules: Big Changes in the Cancer Treatment Paradigm. *Journal of Pharmacy Practice* 2008; **21**: 17-35

Garcia-Blanco MA, Jamison SF, Sharp PA. Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns. *Genes Dev* 1989; **3**: 1874-86.

Ge H, Manley JL. A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA *in vitro*. *Cell* 1990; **62**: 25-34.

Ghetti A, S Pinol-Roma, Michael WM, Morandi C, Dreyfuss G. hnRNP I, the polypyrimidine tract-binding protein: distinct nuclear localization and association with hnRNAs. *Nucleic Acids Res* 1992; **20**: 3671-8.

Giering, JC, Grimm D, Storm TA, Kay MA. Expression of shRNA from a tissue-specific pol II promoter is an effective and safe RNAi therapeutic. *Molecular Therapy* 2008; **16**:1630-6.

Giroglou T, Florin L, Schafer F, Streeck RE, Sapp M. Human papillomavirus infection requires cell surface heparan sulfate. *J Virol* 2001; **75**: 1565-70.

Glover DJ, Lipps HJ, Jans DA. Towards safe, non-viral therapeutic gene expression in humans. *Nature Rev Gen* 2005; **6**: 299-310.

Golan-Gerstl R, Shilo A, Cohen M *et al*. Splicing factor hnRNP A2/B1 regulates tumor suppressor gene splicing and is an oncogenic driver in glioblastoma. *Cancer Res* 2011; **71**: 4464-72.

Graham SV. Papillomavirus 3' UTR regulatory elements. *Front Biosci* 2008; **13**: 5646-5663.

Graham FL, van der Eb AJ. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 1973; **52**: 456-67.

Graveley BR, Maniatis T. Arginine/serine-rich domains of SR proteins can function as activators of pre-mRNA splicing. *Mol Cell* 1998; **1**: 765-771.

Grimm S. The art and design of genetic screens: mammalian culture cells. *Nature Rev Gen* 2004; **5**: 179-189.

Grollman AP. Structural basis for inhibition of protein synthesis by emetine and cycloheximide based on an analogy between ipecac alkaloids and glutarimide antibiotics. *Proc Natl Acad Sci USA*, 1966; **56**: 1867-1874.

Grollman AP. Inhibitors of protein biosynthesis V. Effects of emetine on protein and nucleic acid biosynthesis in HeLa cells. *J Biol Chem* 1968; **243**: 4089-4094.

Hagerman AE. Tannin Chemistry. 2002. database:
<http://www.users.muohio.edu/hagermae/>

Hagopian JC, Ma CT, Meade BR *et al.* Adaptable molecular interactions guide phosphorylation of the SR protein ASF/SF2 by SRPK1. *J Mol Biol* 2008; **382**: 894-909.

Hairiri J, Oster A. The negative predictive value of p16INK4a to assess the outcome of cervical intraepithelial neoplasia 1 in the uterine cervix. *Int Gynecol Pathol* 2007; **26**: 223-228.

Hamada T, Sasaki H, Seki R, Sakaki Y. Mechanism of chromosomal integration of transgenes in microinjected mouse eggs: sequence analysis of genome–transgene and transgene–transgene junctions at two loci. *Gene* 1993; **128**: 197–202.

Hamilton AJ, Baulcombe DC. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 1999; **286**: 950-952.

Hamilton-Miller JM. Antimicrobial properties of tea (*Camellia sinensis* L.). *Antimicrob Agents Chemother* 1995; **39**: 2375–2377.

Hammond SM, Bernstein E, Beach D, Hannon GJ. An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature* 2000; **404**: 293-6.

Han SP, Kassahn KS, Skarshewski A, Ragan MA, Rothnagel JA, Smith R. Functional implications of the emergence of alternative splicing in hnRNP A/B transcripts. *RNA* 2010; **16**: 1760–1768.

Han SP, Tang YH, Smith R. Functional diversity of the hnRNPs: past, present and perspectives. *Biochem J* 2010; **430**: 379-92.

Han ZT, Zhu XX, Yang RY *et al.* Effect of intravenous infusions of 12-O-tetradecanoylphorbol-13-acetate (TPA) in patients with myelocytic leukemia: Preliminary studies on therapeutic efficacy and toxicity *Proc Natl Acad Sci USA Pharmacology* 1998; **95**: 5357–5361.

Harper D, Gall S, Naud P *et al.* Sustained immunogenicity and high efficacy against HPV 16/18 related cervical neoplasia: Long-term follow up through 6.4 years in women vaccinated with Cervarix (GSK's HPV-16/18 AS04 candidate vaccine). *Gynecol Oncol* 2008; **109**: 158–159.

Hatakeyama H, Kondo T, Fujii K *et al.* Protein clusters associated with carcinogenesis, histological differentiation and nodal metastasis in esophageal cancer. *Proteomics* 2006; **6**: 6300-6316.

Hatfield JT, Rothnagel JA, Smith R. Characterization of the mouse hnRNP A2/B1/B0 gene and identification of processed pseudogenes. *Gene* 2002; **295**: 33–42.

He YW, Smith R. Nuclear functions of heterogeneous nuclear ribonucleoproteins A/B. *Cell Mol Life Sci* 2009; **66**: 1239–1256.

He YW, Brown MA, Rothnagel JA, Saunders NA, Smith R. Roles of heterogeneous nuclear ribonucleoproteins A and B in cell proliferation. *J Cell Sci* 2005; **118**: 3173–3183.

Hebner CM, Laiminis LA. Human papillomaviruses: basic mechanism of pathogenesis and oncogenicity. *Rev Med Virol* 2006; **16**: 83-97.

Helt AM, Funk JO, Galloway DA. Inactivation of both the retinoblastoma tumor suppressor and p21 by the human papillomavirus type 16 E7 oncoprotein is necessary to inhibit cell cycle arrest in human epithelial cells. *J Virol* 2002; **76**: 10559-68.

Hitti M. "Cervical Cancer Vaccines: More Promise". WebMD Medical News. 2006
<http://www.webmd.com/content/artic,e/121/114328?printing=true>

Hoffman DW, Query CC, Golden BL, White SW, Keene JD. RNA-binding domain of the A protein component of the U1 small nuclear ribonucleoprotein analyzed by NMR spectroscopy is structurally similar to ribosomal proteins. *Proc Natl Acad Sci USA* 1991; **88**: 2495-9.

Hofmann Y, Lorson CL, Stamm S, Androphy EJ, Wirth B. Htra2-beta 1 stimulates an exonic splicing enhancer and can restore fulllength SMN expression to survival motor neuron 2 (SMN2). *Proc Natl Acad Sci USA* 2000; **97**: 9618–9623.

Hofmann Y, Wirth B. hnRNP-G promotes exon 7 inclusion of survival motor neuron (SMN) via direct interaction with Htra2-beta1. *Hum Mol Genet* 2002; **11**: 2037–2049.

Howley PM and DR Lowy, Chapter 65: Papillomaviruses and their replication, *In* Field Virology. DM Knipe and PM Howley, Ed. Lippincott, Williams & Wilkins. 2001: 2197-2230.

Hsu SM, Raine L, Fanger H. Use of avidin-biotin peroxidase complex (aBC) in immunoperoxidase techniques: a comparison between aBC and unlabeled antibody (PaP) procedures. *J Histochem Cytochem* 1981; **29**: 577-80.

Hu Y, Benedict MA, Wu D, Inohara N, Nunez G. Bcl-xL interacts with Apaf-1 and inhibits Apaf-1-dependent caspase-9 activation. *Proc Natl Acad Sci USA* 1998; **95**: 4386–4391.

Huang Y, Yario TA, Steitz JA. A molecular link between SR protein dephosphorylation and mRNA export. *Proc Natl Acad Sci* 2004; **101**: 9666–9670.

Hubert WG, Laimins LA. Human papillomavirus type 31 replication modes during the early phases of the viral life cycle depend on transcriptional and posttranscriptional regulation of E1 and E2 expression. *J Virol* 2002; **76**: 2263-2273.

Human Papillomavirus (HPV) Vaccines, 2011. Retrieved November 11, 2009 from National Cancer Institute (NCI) database:
<http://www.cancer.gov/cancertopics/factsheet/prevention/HPV-vaccine>

Iaconis L, Hyjek E, Ellenson LH, Pirog EC. 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**: 1343-1349.

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

Jackson AV. Liver necrosis in burns treated with tannic acid. *Med J Australia, Sydney* 1944; **2**: 352-4.

Jain A, Wordinger RJ, Yorio T, Clark AF. Spliceosome protein (SRp) regulation of glucocorticoid receptor isoforms and glucocorticoid response in human trabecular meshwork cells. *Arvo Journal* 2011; **53**: 857-866.

Jennings HR, Romanelli F. The use of valproic acid in HIV-positive patients. *Ann Pharmacother* 1999; **33**: 1113-1116.

Jimenez A, Carrasco L, Vazquez D. Enzymic and nonenzymic translocation by yeast polysomes. Site of action of a number of inhibitors. *Biochemistry* 1977; **16**: 4727-4730.

Johansson C, Somberg M, Schwartz S. Proteins involved in HPV-16 pre-mRNA processing. In "Current Topics in Virology", *Research Signpost* 2011; **8**: 17-27.

Jondorf WR, Abbott BJ, Greenberg NA, Mead JAR. Antileukemic effectiveness of (-)-emetine and some of its derivatives. *Pharmacologist* 1970; **12**: 282.

Joyce JG, Tung JS, Przysiecki CT *et al.* The L1 major capsid protein of human papillomavirus type 11 recombinant virus-like particles interacts with heparin and cell surface glycosaminoglycans on human keratinocytes. *J Biol Chem* 1999; **274**: 5810-22.

Junqueira LC, Carneiro J. Chapter 22: The female reproductive system. In *Basic Histology*. The McGraw-Hill Companies. 2005: 435-455.

Karni R, de Stanchina E, Lowe SW, Sinha R, Mu D, Krainer AR. The gene encoding the splicing factor SF2/ASF is a protooncogene. *Nat Struct Mol Biol* 2007; **14**: 185-93.

Kamma H, Fujimoto M, Fujiwara M *et al.* Interaction of hnRNP A2/B1 isoforms with telomeric ssDNA and the in vitro function. *Biochem Biophys Res Commun* 2001; **280**: 625-30.

Kamma H, Horiguchi H, Wan LL *et al.* Molecular characterization of the hnRNP A2/B1 proteins: Tissue-specific expression and novel isoforms. *Exp Cell Res* 1999; **246**: 399-411.

Kanopka A, Muhlemann O, Akusiarvi G. Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature* 1996; **381**: 535-538.

Kanopka A, Mühlemann O, Petersen-Mahrt S, Estmer C, Öhrmalm C, Akusjärvi G. Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins. *Nature* 1998; **393**: 185–187.

Kato S, Anderson RA, Camerini-Otero RD. Foreign DNA introduced by calcium phosphate is integrated into repetitive DNA elements of the mouse L cell genome. *Mol Cell Biol* 1986; **6**: 1787–1795.

Kim YD, Lee JY, Oh KM *et al.* NSrp70 is a novel nuclear speckle-related protein that modulates alternative pre-mRNA splicing in vivo. *Nucleic Acids Res* 2011; **39**: 4300-14.

Kirnbauer R, Taub J, Greenstone H *et al.* Efficient self-assembly of human papillomavirus type 16 L1 and L1-L2 into virus-like particles. *J Virol* 1993; **67**: 6929–6936.

Kitchner HC, Castle PE, Cox JT. Chapter 7: Achievement and limitations of cervical cytology screening. *Vaccine* 2006; **24 Suppl 3**: S63-70.

Krainer AR, Conway GC, Kozak D. Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. *Genes Dev* 1990; **4**: 1158-1171.

Krawczak M, Reiss J, Cooper DN. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human gene: causes and consequences. *Hum Genet* 1992; **90**: 41-45.

Krecic AM, Swanson MS. hnRNP complexes: Composition, structure, and function. *Curr Opin Cell Biol* 1999; **11**: 363–371.

Kresge KJ. "Cervical Cancer Vaccines." IAVI Report. 2006.
<http://www.iavireport.org/Issues/Issues9-5/vaccines>

Kusuma Harahap IS, Saito T, Poh San L, *et al.* Valproic acid increases SMN2 expression and modulates SF2/ASF and hnRNPA1 expression in SMA fibroblast cell lines. *Brain & Development xxx–xxx* 2011.

Lai MC, Lin RI, Tarn WY. Transportin-SR2 mediates nuclear import of phosphorylated SR proteins. *Proc Natl Acad Sci* 2001; **98**: 10154–10159.

Lai MC, Tarn WY. Hypophosphorylated ASF/SF2 binds TAP and is present in messenger ribonucleoproteins. *J Biol Chem* 2004; **279**: 31745–31749.

- Lai MC, Lin RI, Huang SY, Tsai CW, Tarn WY. A human importin-beta family protein, transportin-SR2, interacts with the phosphorylated RS domain of SR proteins. *J Biol Chem* 2000; **275**: 7950-7957.
- Lambert P *et al.* Papillomaviruses – Human (Papovaviridae): Molecular Biology, *In Encyclopedia of Virology*, Ed Granoff A, Webster RG. Academic Press. 1999: 1115-1120.
- Lee WT. Department of Anatomy YUMC nd, Introduction to Immunohistochemistry and Immunocytochemistry, *History of Immunohistochemistry*. 2010.
- Leach BE, Ford JH, Whiffen AJ. Actidione, an antibiotic from *Streptomyces griseus*. *J Amer Chem Soc* 1947; **69**: 474.
- Leechanachai P, Banks L, Moreau F, Matlashewski G. The E5 gene from human papillomavirus type 16 is an oncogene which enhances growth factor-mediated signal transduction to the nucleus. *Oncogene* 1992; **7**: 19-25.
- Lehrman G, Hogue I, Palmer S, *et al.* Depletion of latent HIV-1 infection in vivo: a proof-of-concept study. *Lancet* 2005; **366**: 549–55.
- Levine AJ. The Origins of Virology, *In Fundamental Virology*, Knipe DM, Howley PM, Editors, Lipincott, Williams & Wikins, Philadelphia. 2001; 3-18.
- Li CY, Chu JY, Yu JK, *et al.* Regulation of alternative splicing of Bcl-x by IL-6, GM-CSF and TPA. *Cell Res* 2004; **14**: 473–479.
- Li X, Coffino P. High-risk human papillomavirus E6 protein has two distinct binding sites within p53, of which only one determines degradation. *J Virol* 1996; **70**: 4509-16.
- Liang F, Jasin M. Ku80-deficient cells exhibit excess degradation of extrachromosomal DNA. *J Biol Chem* 1996; **271**: 14405–14411.
- Liang Y, Shi SL, Li QF *et al.* The localization of hnRNP A2/B1 in nuclear matrix and the aberrant expression during the RA-induced differentiation of human neuroblastoma SK-N-SH cells. *J Cell Biochem* 2011; **112**: 1722-9.

Lietman PS. Mitochondrial protein synthesis: Inhibition by emetine hydrochloride. *Mol Pharmacol* 1970; **7**: 122-128.

Lin S, Xiao R, Sun P, Xu X, Fu XD. Dephosphorylation- dependent sorting of SR splicing factors during mRNP maturation. *Mol Cell* 2005; **20**: 413–425.

Lindenboim L, Kringel S, Braun T, Borner A, Stein R. Bak but not Bax is essential for Bcl-xS-induced apoptosis. *Cell Death and Differentiation* 2005; **12**: 713–723.

Lindenboim L, Yuan J, Stein R. Bcl-xS and Bax induce different apoptotic pathways in PC12 cells. *Oncogene* 2000; **19**: 1783–1793.

Liu JP. Protein kinase C and its substrates. *Mol Cell Endocrinol* 1996; **116**: 1-29.

Liu JS, Kuo SR, Broker TR, Chow LT. The functions of human papillomavirus type 11 E1, E2 and E2C proteins in cell-free DNA replication. *J Biol Chem* 1995; **270**: 27283-91.

Liu R, Page C, Beidler DR, Wicha MS, Nunez G. Overexpression of Bcl-x(L) promotes chemotherapy resistance of mammary tumors in a syngeneic mouse model. *Am J Pathol* 1999; **155**: 1861–1867.

Liu X, Kim JK, Li Y, Li J, Liu F, Chen X. Tannic acid stimulates glucose transport and inhibits adipocyte differentiation in 3T3-L1 cells. *J Nutr* 2005; **135**: 165-71.

Long JC, Caceres JF. The SR protein family of splicing factors: master regulators of gene expression. *Biochem J* 2009; **417**: 15-27.

Longworth MS, Laiminis LA. Pathogenesis of human papillomaviruses in differentiating epithelia. *Microbiol Mol Rev* 2004; **68**: 362-72.

Loscher W. Basic pharmacology of valproate: a review after 35 years of clinical use for the treatment of epilepsy. *CNS Drugs* 2002; **16**: 669-669.

Lowy DR, Howley PM. Papillomaviruses: In DM Knipe & PM Howley, Fields Virology, Ed Lippincott, Williams & Wilkins, Philadelphia. 2001: 2231-2264.

- Lowy DR, Schiller JT. Prophylactic human papillomavirus vaccines *J Clin Invest* 2006; **116**: 1167–1173.
- Maggipinto M, Rabiner C, Kidd GJ, Hawkins AJ, Smith R, Barbarese E. Increased expression of the MBP mRNA binding protein HnRNP A2 during oligodendrocyte differentiation. *J Neurosci Res* 2004; **75**: 614-23.
- Mahon KA, Overbeek PA, Westphal H. Prenatal lethality in a transgenic mouse line is the result of a chromosomal translocation. *Proc Natl Acad Sci USA* 1988; **85**: 1165–1168.
- Mark WH, Signorelli K, Blum M, Kwee L, Lacy E. Genomic structure of the locus associated with an insertional mutation in line 4 transgenic mice. *Genomics* 1992; **13**: 159–166.
- Marks PA, Jiang X. Histone deacetylase inhibitors in programmed cell death and cancer therapy. *Cell Cycle* 2005; **4**: 549-551.
- Martin P, Vass WC, Schiller JT, Lowy DR, Velu TJ. The bovine papillomavirus E5 transforming protein can stimulate the transforming activity of EGF and CFS-1 receptors. *Cell* 1989; **59**: 21-32.
- Martini FH, Bartholomew EF. The reproductive system, in essential of anatomy and physiology. *Pearson Education Inc.: San Fransisco* 2007; **19**: 612-643.
- Massiello A, Salas A, Pinkerman RL, Roddy P, Roesser JR, Chalfant CE. Identification of Two RNA cis-Elements That Function to Regulate the 5' Splice Site Selection of Bcl-x Pre-mRNA in Response to Ceramide*. *The Journal of Biological Chemistry* 2004; **279**: 15799-15804.
- Massiello A, Roesser JR, Chalfant CE. SAP155 Binds to ceramide-responsive RNA cis-element 1 and regulates the alternative 5' splice site selection of Bcl-x pre-mRNA. *The FASEB Journal* 2006; **20**: 921-29.
- Matsui M, Horiguchi H, Kamma H, Fujiwara M, Ohtsubo R, Ogata T. Testis- and developmental stage-specific expression of hnRNP A2/B1 splicing isoforms, B0a/b. *Biochim Biophys Acta* 2000; **1493**: 33–40.

- Mayeda A, Sreaton GR, Chandler SD, Fu XD, Krainer AR. Substrate specificities of SR proteins in constitutive splicing are determined by their RNA recognition motifs and composite pre-mRNA exonic elements. *Mol Cell Biol* 1999; **19**: 1853–1863.
- McClure RD, Lam CR. Problems in the treatment of burns; liver necrosis as a lethal factor. *Southern Surg* 1940; **9**: 223-34.
- McFarlane M, Wilson JB. A model for the mechanism of precise integration of a microinjected transgene. *Transgenic Res* 1996; **5**: 171–177.
- McKay SJ, Cooke H. hnRNP A2/B1 binds specifically to single stranded vertebrate telomeric repeat TTAGGGn. *Nucleic Acids Res.* 1992; **20**: 6461–6464.
- McMurray HR, Nguyen D, Westbrook TF, Mcance DJ. Biology of human papillomaviruses. *Int J Exp Pathol* 2001; **82**: 15-33.
- McPhillips MG, Veerapraditsin T, Cumming SA. 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**: 10598-605.
- Mermoud JE, Cohen PT, Lamond AI. Regulation of mammalian spliceosome assembly by a protein phosphorylation mechanism. *EMBO J.* 1994; **13**: 5679–5688.
- Mermoud JE, Cohen P, Lamond AI. Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. *Nucleic Acids Res* 1992; **20**: 5263-5269.
- Meyers C, Frattini MG, Hudson JB, Laimins LA. Biosynthesis of human papillomavirus from a continuous cell line upon epithelial differentiation. *Science* 1992; **257**: 971-973.
- Minn AJ, Kettlun CS, Liang H *et al.* Bcl-xL regulates apoptosis by heterodimerization-dependent and -independent mechanisms. *The EMBO Journal* 1999; **18**: 632–643.
- Mole S, McFarlane M, Chuen-Im T, Milligan SG, Millan D, Graham SV. RNA splicing factors regulated by HPV16 during cervical tumour progression. *J Pathol* 2009a; **219**: 383-391.
- Mole S, Milligan SG, Graham SV. Human papillomavirus type 16 E2 protein transcriptionally activates the promoter of a key cellular splicing factor, SF2/ASF. *J Virol* 2009b; **83**: 357-67.

Mole S, Veerapraditsin T, McPhilips MG, Graham SV. Regulation of splicing-associated SR proteins by HPV-16. *Biochem Soc Trans* 2006; **34**: 1145-1147.

Moller M, Herzer K, Wenger T, Herr I, Wink M. The alkaloid emetine as a promising agent for the induction and enhancement of drug-induced apoptosis in leukemia cells. *Oncol Rep* 2007; **18**: 737-744.

Mondrek B, Lee C. A genomic view of alternative splicing. *Nat Genet* 2002; **30**: 13-19.

Moog C, Kuntz-Simon G, Caussin-Schwemling C, Obert G. Sodium Valproate, an Anticonvulsant Drug, Stimulates Human Immunodeficiency Virus Type 1 Replication Independently of Glutathione Levels. *J Gen Virol* 1996; **77**: 1993-1999.

Moscufo N, Sverdrup F, Breiding DE, Androphy EJ. Two distinct regions of the BPV1 E1 replication protein interact with the activation domain of E2. *Virus Res* 1999; **65**: 141–154.

Munger K, Phelps WC. The human papillomavirus E7 protein as a transforming and transactivating factor. *Biochim Biophys Acta* 1993; **1155**: 111-23.

Munoz N, Bosch FX, de Sanjose S *et al.* Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 2003; **348**: 518-527.

Munster PN, Marchion DC, Schmitt M, *et al.* Phase I/II trial combining the HDAC inhibitor, valproic acid (VPA) and FEC100 (5-fluorouracil, epirubicin and cyclophosphamide) in locally advanced/metastatic breast cancer. *Journal of Clinical Oncology, ASCO Annual Meeting Proceedings (Post-Meeting Edition)* 2007; **25**: 1065.

Murnane JP, Yezzi MJ, Young BR. Recombination events during integration of transfected DNA into normal human cells. *Nucleic Acids Res* 1990; **18**: 2733-2738.

Napoli C, Lemieux C, Jorgensen R. Introduction of a Chimeric Chalcone Synthase Gene into *Petunia* Results in Reversible Co-Suppression of Homologous Genes in trans. *The Plant Cell* 1990; **2**: 279-289.

Natarajan E, Omobono JD, Jones JC, Rheinwald JG. Co-expression of p16^{INK4A} and laminin 5 by keratinocytes: a wound-healing response coupling hypermotility with growth arrest that goes

awry during epithelial neoplastic progression. *J Invest Dermatol Symp Proceed* 2005; **10**: 72–85.

National Cancer Institute (NCI). Human Papillomavirus (HPV) Vaccines, 2011.
<http://www.cancer.gov/cancertopics/factsheet/prevention/HPV-vaccine>

Ngo JC, Giang K, Chakrabarti S *et al.* A sliding docking interaction is essential for sequential and processive phosphorylation of an SR protein by SRPK1. *Mol Cell* 2008; **29**: 563-76.

Nickoloff JA. Humana Press: Preface. *In* Electroporation Protocols for Microorganisms. Ed JA Nickoloff, Totowa, New Jersey: 1995: v-vi.

Niedel JE, Kuhn LJ, Vanderbark GR. Phorbol diester receptor copurifies with protein kinase C. *Proc Natl Acad Sci USA* 1983; **80**: 36-40.

Oberg D, Collier B, Zhao X, Schwartz S. A downstream polyadenylation element in human papillomavirus type 16 encodes multiple GGG-motifs and interacts with hnRNP H. *J Virol* 2005; **79**: 9254-9269.

Oberg D, Collier B, Zhao X, Schwartz S. 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**: 11674-84.

Oh YL, Hahm B, Kim YK *et al.* Determination of functional domains in polypyrimidine-tract-binding protein. *Biochem J* 1998; **331**: 169-75.

Ohtani N, Yamakoshi K, Takahashi A, Har E. The p16INK4a-RB pathway: molecular link between cellular senescence and tumor suppression. *The Journal of Medical Investigation* 2004; **51**: 146-53.

O'Neill CJ, McCluggage WG. p16 expression in the female genital tract and its value in diagnosis. *Adv Anat Pathol* 2006; **13**: 8-15.

Paavonen J, Naud P, Salmeron J, Wheeler CM, Chow SN, Apter D, for the HPV PATRICIA Study Group. Efficacy of human papillomavirus (HPV)-16/18 AS04adjuvanted vaccine against cervical infection and precancer caused by oncogenic HPV types: Final analysis of a double-blind, randomized study in young women. *Lancet* 2009; doi:10.1016/S0140-6736(09)61248-4.

Pan G, O'Rourke K, Dixit VM) Caspase-9, Bcl-xL and Apaf-1 form a ternary complex. *J Biol Chem* 1998; **273**: 5841.

Park TW, Fujiwara H, Wright TC. Molecular biology of cervical cancer and its precursors. *Cancer* 1995; **76**: 1902-1913.

Patton JG, Mayer SA, Tempst P, Nadal-Ginard B. Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing. *Genes Dev* 1991; **5**: 1237-51.

Pham JW, Pellino JL, Lee YS, Carthew RW, Sontheimer EJ. A Dicer-2-dependent 80s complex cleaves targeted mRNAs during RNAi in *Drosophila*. *Cell* 2004; **117**: 83-94.

Phelps WC, Yee CL, Munger K, Howley PM. The human papillomavirus type 16 E7 gene encodes transactivation and transforming functions similar to adenovirus E1A. *Cell* 1988; **58**: 539-547.

Polack JM, Van Noorden S. Requirements for immunocytochemistry. *In* Introduction to Immunocytochemistry Chapter 3: (RMS: Microscopy Handbook, Series No. 37), BIOS Scientific Publishers, New York. 1997: 11-35.

Popescu NU, Amsbaugh SC, di Paolo JA. Human papillomavirus type 18 DNA is integrated at a single chromosome site in cervical carcinoma cell line SW756. *J Virol* 1987; **51**: 1682-1685.

Prasad J, Colwill K, Pawson T, Manley JL. The protein kinase Clk/Sty directly modulates SR protein activity: Both hyper- and hypophosphorylation inhibit splicing. *Mol Cell Biol* 1999; **19**: 6991-7000.

Proudfoot NJ, Furger A, Dye MJ, Integrating mRNA processing with transcription. *Cell* 2002; **108**: 501-12.

Raffetseder U, Frye B, Rauen T *et al.* Splicing Factor SRp30c Interaction with Y-box Protein-1 confers Nuclear YB-1 Shuttling and Alternative Splice Site Selection. *J Biological Chemistry* 2003; **278**: 18241-18248

Rebois RV, Patel J. Phorbol Ester Causes Desensitization of Gonadotropin-responsive Adenylate Cyclase in a Murine Leydig Tumor Cell Line. *J Biol Chem* 1985; **260**: 8026-8031.

Rijkers T, Peetz A, Ruther U. Insertional mutagenesis in transgenic mice. *Transgenic Res* 1994; **3**: 203–215.

Robins DM, Ripley S, Henderson AS, Axel R. Transforming DNA integrates into the host chromosome. *Cell* 1981; **23**: 29–39.

Roden RB, Greenstone HL, Kirnbauer R *et al.* *In vitro* generation and type-specific neutralization of a human papillomavirus type 16 virion pseudotype. *J Virol* 1996; **70**: 5875–5883.

Roopchand DE, Lee JM, Shahinian S, Paquette D, Bussey H, Branton PE. Toxicity of human adenovirus E4orf4 protein in *Saccharomyces cerevisiae* results from interactions with the Cdc55 regulatory B subunit of PP2A. *Oncogene* 2001 **20**: 5279-90.

Roscigno RF, Garcia-Blanco MA. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA* 1995; **1**: 692–706.

Rose RC, Reichman RC, Bonnez W. Human papillomavirus (HPV) type 11 recombinant virus-like particles induce the formation of neutralizing antibodies and detect HPV-specific antibodies in human sera. *J Gen Virol* 1994; **75**: 2075–2079.

Ross MH, Pawlina W. Female reproductive system: Chapter 23. *In histology: a text and Atlas*. Lipincott, Williams & Wilkins, Philadelphia. 2006: 773-833.

Roychoudhury P, Chaudhuri K. Evidence for heterogeneous nuclear ribonucleoprotein K overexpression in oral squamous cell carcinoma. *British Journal of Cancer* 2007; **97**: 574–575.

Ruas M, Peters G. The p16INK4a/CDKN2A tumor suppressor and its relatives. *Biochim Biophys Acta* 1998; **1378**: 115-F177.

Rush M, Zhao X, Schwartz S. 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**: 12002-12015.

Salman H, Zbaida D, Rabin Y, Chatenay D, Elbaum M. Kinetics and mechanism of DNA uptake into the cell nucleus. *Proc Natl Acad Sci USA* 2001; **98**: 7247–7252.

Saltonstall H, Walker J, Rhoads JE, Lee WE. The influence of local treatment of burns on liver function. *Ann Surg* 1945; **121**: 291-300.

Sanders AE, Slade GD, Patton LL. National prevalence of oral HPV infection and related risk factors in the U.S. adult population. *Oral Diseases* 2012; **18**: 430-441.

Sanford JR, Ellis J, Caceres JF. Multiple roles of arginine/ serine-rich splicing factors in RNA processing. *Biochem Soc Trans* 2005; **33**: 443-446.

Sanford JR, Ellis JD, Cazalla D, Caceres JF. Reversible phosphorylation differentially affects nuclear and cytoplasmic functions of splicing factor 2/alternative splicing factor. *Proc Natl Acad Sci* 2005; **102**: 15042–15047.

Sapp M, Day PM. Structure, attachment and entry of polyoma- and papillomaviruses. *Virology* 2009; **384**: 400-409.

Sarafi TR, McBride AA. Domains of theBPV-1 E1 replication protein required for origin-specificDNA binding and interaction with the E2 transactivator. *Virology* 1995; **211**: 385–396.

Scheffner M, Huibregtse JM, Vierstra RD, Howley PM. The HPV-16 E6 and E6-AP complex functions as ubiquitin-protein ligase in the ubiquitination of p53. *Cell* 1993; **75**: 495-505.

Scheurer ME, Tortolero-Luna G, Adler-Storthz K. Human papillomavirus infection: biology, epidemiology, and prevention. *International Journal of Gynecological Cancer* 2005; **15**: 727–746.

Schiffman M, Kjaer SK. Natural History of Anogenital Human Papillomavirus Infection and Neoplasia. *J Natl Cancer Inst Monogr* 2003; **31**: 14-19.

Schneider-Maunoury S, Croissant O, Orth G. Integration of human papillomavirus type 16 DNA sequences: a possible early event in the progression of genital tumors. *J Virol* 1987; **61**: 3295-8.

Schwartz C, Palissot V, Aouali N *et al.* Valproic acid induces non-apoptotic cell death mechanisms in multiple myeloma cell lines. *International Journal of Oncology* 2007; **30**: 573–82.

Schwartz S. HPV-16 RNA processing. *Front Biosci* 2008; **13**: 5880–5891.

Schwartz S, Somberg M, Zhao X *et al.* Regulation of HPV gene splicing and expression. In Norrild, B. (Ed.), Human Papillomavirus Gene Regulation and Transformation. *Transworld Research Network*, India. 2007: 47–67.

Scribner JD, Boutwell RK. Inflammation and tumor promotion: selective protein induction in mouse skin by tumor promoters. *European J Cancer* 1972; **8**: 617-621.

Selinka HC, Florin L, Patel HD *et al.* Inhibition of transfer to secondary receptors by heparan sulfate-binding drug or antibody induces noninfectious uptake of human papillomavirus. *J Virol* 2007; **81**: 10970-10980.

Shatkin AJ. Capping of eukaryotic mRNAs. *Cell* 1976; **9**: 645-53.

Shen H, Green MR. A pathway of sequential arginine-serine rich domain-splicing signal interactions during mammalian spliceosome assembly. *Mol Cell* 2004; **16**: 363-73.

Shen H, Kan JL, Green MR. Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly. *Mol Cell* 2004; **13**: 367-376.

Shepard PJ, Hertel KJ. The SR protein family. *Genome Biol* 2009; **10**: 242.

Sherr CJ, Roberts JM. CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev* 1999; **13**: 1501-12.

Simard MJ, Chabot B. SRp30c is a repressor of 3' splice site utilization. *Mol Cell Biol* 2002; **22**: 4001-10.

Smith R. Moving molecules: mRNA trafficking in mammalian oligodendrocytes and neurons. *Neuroscientist* 2004; **10**: 495–500.

Snijders PJ, Steenbergen RD, Heideman DA, Meijer CJ. HPV-mediated cervical carcinogenesis: concepts and clinical implications. *J Pathology* 2006; **208**: 152–164.

Somberg M, Li X, Johansson C, Orru B, Chang R, Rush M, Fay J, Ryan F, Schwartz S. Serine/arginine-rich protein 30c activates human papillomavirus type 16 L1 mRNA expression via a bimodal mechanism. *J Gen Virol* 2011; **92**: 2411-21.

Somberg M, Rush M, Fay J, Ryan F, Lambkin H, Akusjärvi G, Schwartz S. Adenovirus E4orf4 induces HPV-16 late L1 mRNA production. *Virology* 2009; **383**: 279–290.

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

Somberg M, Zhao X, Fröhlich M, Evander M, Schwartz S. PTB induces HPV-16 late gene expression by interfering with splicing inhibitory elements at the major late 5'-splice site SD3632. *J Virol* 2008; **82**: 3665-3678.

Spellman R, Smith CW. Novel modes of splicing repression by PTB. *Trends Biochem Sci* 2006; **31**: 73-6.

Srebrow A, Kornbliht AR. The connection between splicing and cancer. *Journal of Cell Science* 2006; **119**: 2635-2641.

Srivastava RC, Husain MM, Hasan SK, Athar M. Green tea polyphenols and tannic acid act as potent inhibitors of phorbol ester-induced nitric oxide generation in rat hepatocytes independent of their antioxidant properties. *Cancer Lett* 2000; **153**: 1-5.

Stevens A, Lowe J. Chapter 17: Female reproductive system, *In* human histology. Elsevier Mosby: New York. 2005: 345-372.

Strair RK, Schaar D, Goodell L, *et al.* Administration of a phorbol ester to patients with hematological malignancies: preliminary results from a phase I clinical trial of 12-*O*-Tetradecanoylphorbol-13-acetate1. *Clin Cancer Res* 2002; **8**: 2512-2518.

Stunberauch F, Laimins LA. Human papillomavirus life cycle: active and latent phases. *Semin Cancer Biol* 1999; **9**: 379-86.

Stubenrauch F, Hummel M, Iftner T *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**: 1178-86.

Sureau A, Gattoni R, Dooghe Y, Stévenin J, Soret J. SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. *The EMBO Journal* 2001; **20**: 1785–1796.

Suttle CA. Marine viruses -- major players in the global ecosystem. *Nat Rev Microbiol* 2007; **5**: 801–812.

Tate G, Suzuki T, Kishimoto K, Mitsuya T. Novel mutations of *EVER1/TMC6* gene in a Japanese patient with epidermodysplasia verruciformis. *J Hum Genet* 2004; **49**: 223–225.

Tatsuta M, Iishi H, Baba M, Uehara H, Nakaizumi A, Taniguchi H. Inhibition by amiloride of gastric carcinogenesis induced by N-methyl-N'-nitro-N-nitrosoguanidine in Wistar rats *British Journal of Cancer* 1993; **67**: 1011–1014.

Taylor CR *et al.* Chapter 1: Techniques of Immunohistochemistry: Principles, Pitfalls and Standardization. *In* Diagnostic Immunohistochemistry. Ed D Dabbs, Churchill Livingstone Elsevier. 2006: 1-42.

Terhune SS, Milcarek C, Laimins LA. Regulation of human papillomavirus 31 polyadenylation during the differentiation-dependent life cycle. *J Virol* 1999; **73**: 7185-7192.

Thudichum, JLW. Reports of the Medical Officer of Council and Local Government Board, 1876; New Series No. III, 113. 1874; New Series No. VIII, 117.

Tockman MS, Gupta PK, Myers JD, *et al.* Sensitive and specific monoclonal antibody recognition of human lung cancer antigen on preserved sputum cells: a new approach to early lung cancer detection. *J Clin Oncol* 1988; **6**: 1685–1693.

Tockman MS, Mulshine JL, Piantadosi S, *et al.* Prospective detection of preclinical lung cancer: results from two studies of heterogeneous nuclear ribonucleoprotein A2/B1 overexpression. *Clin Cancer Res* 1997; **3**: 2237–2246.

Tronchère H, Wang J, Fu XD. A protein related to splicing factor U2AF35 that interacts with U2AF65 and SR proteins in splicing of pre-mRNA. *Nature* 1997; **388**: 397-400.

Tsutsumi K *et al.* Expression of human papillomavirus (HPV) gene in HPV-positive laryngeal tumors and activity of the HPV long control region in cultured normal laryngeal epithelial cells, *Nippon Jibiinkoka Gakkai Kaiho* 1993; **96**: 767-73.

Turek LP. The structure, function and regulation of papillomavirus gene in infection and cervical cancer. *Adv Viral Res* 1994; **44**: 305-356.

Valentini A, Gravina P, Federici G, Bernardini S. Valproic Acid Induces Apoptosis, p(16INK4A) Upregulation and Sensitization to Chemotherapy in Human Melanoma Cells. *Cancer Biology & Therapy* 2007; **2**: 185–91.

Veldman T, Horikawa I, Barrett JC, Schelegel R. Transcriptional activation of the telomerase hTERT gene by human papillomavirus type 16 E6 oncoprotein. *J Virol* 2001; **75**: 4467-72.

Venables JP. Aberrant and alternative splicing in cancer. *Cancer Res* 2004; **64**: 7647-7654.

Venter JC, Adams MD, Myers EW *et al.* The sequence of the human genome. *Science* 2001; **291**: 1204-51.

Walboomers JM, Jacobs MV, Manos MM *et al.* Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathology* 1999; **189**: 12-9.

Weinberg JB. Tumor cell killing by phorbol ester--differentiated human leukemia cells. *Science* 1981; **213**: 655-657.

Wells DB, Humphrey HD, Coll JJ. The relation of tannic acid to the liver necrosis occurring in burns. *New Engl J Med* 1942; **226**: 629-36.

Wiegrebbe W, Kramer WJ, Shamma M. The emetine alkaloids. *J Nat Prod* 1984; **47**: 397-408.

Wilczynski SP, Pearlman L, Walker J. Identification of HPV 16 early genes retained in cervical carcinomas. *Virology* 1988; **166**: 624-627.

Wilkie TM, Palmiter RD. Analysis of the integrant in MyK-103 transgenic mice in which males fail to transmit the integrant. *Mol Cell Biol* 1987; **7**: 1646–1655.

Wheeler CM. Advances in primary and secondary interventions for cervical cancer: human papillomavirus prophylactic vaccines and testing. *Nature Clinical practice oncology* 2007; **4**: 224 -235.

WHO, Comprehensive cervical cancer control: a guide to essential practice. 2006.

Wright TC Jr, Cox JT, Massad LS, Twiggs LB, Wilkinson EJ. Consensus guidelines for the management of women with cervical cytological abnormalities. *Jama* 2002; **287**: 2120–2129.

Woodman CBJ, Collins SI, Young LS. The natural history of cervical HPV infection: unresolved issues, *Nature Reviews. Cancer* 2007; **7**: 11-22.

Wu JY, Maniatis T. Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* 1993; **75**: 1061-70.

Wurm FM. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnol* 2004; **22**: 1393-1398.

Xiao R, Sun Y, Ding JH *et al.* Splicing Regulator SC35 Is Essential for Genomic Stability and Cell Proliferation during Mammalian Organogenesis. *Mol Cell Biol* 2007; **27**: 5393–5402.

Xiao SH, Manley JL. Phosphorylation of the ASF/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes & Dev* 1997; **11**: 334–344.

Xu Q, Leung DY, Kisich KO. Serine-arginine-rich protein p30 directs alternative splicing of glucocorticoid receptor pre-mRNA to glucocorticoid receptor beta in neutrophils. *J Biol Chem* 2003; **278**: 27112-8.

Yam TS, Hamilton-Miller JM, Shah S. The effect of a component of tea (*Camellia sinensis*) on methicillin resistance, PBP2' synthesis, and beta-lactamase production in *Staphylococcus aureus*. *J Antimicrob Chemother* 1998; **42**: 211–216.

Yeakley JM, Tronchere H, Olesen J, Dyck JA, Wang HY, Fu XD. Phosphorylation regulates in vivo interaction and molecular targeting of serine/arginine-rich premRNA splicing factors. *J Cell Biol* 1999; **145**: 447–455.

Yun CY, Fu XD. Conserved SR protein kinase functions in nuclear import and its action is counteracted by arginine methylation in *Saccharomyces cerevisiae*. *J Cell Biol* 2000; **150**: 707–718.

Zahler AM, Lane WS, Stolk JA, Roth MB. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes Dev* 1992; **6**: 837-847.

Zamore PD, Tuschl T, Sharp PA, Bartel DP. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 2000; **101**: 25-33.

Zerbe LK, Pino I, Pio R *et al.* Relative Amounts of Antagonistic Splicing Factors, hnRNP A1 and ASF/SF2, Change During Neoplastic Lung Growth: Implications for Pre-mRNA Processing. *Molecular carcinogenesis* 2004; **41**: 187–196.

Zhanga Y, Yanga H, Xiaoa B *et al.* Dendritic cells transduced with lentiviral-mediated RelB-specific ShRNAs inhibit the development of experimental autoimmune myasthenia gravis. *Mol Immunol* 2008; **46**: 657-667.

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; **369**: 351-363.

Zhao X, Öberg D, Rush M, Fay J, Lambkin H, Schwartz S. A 57 nucleotide upstream early polyadenylation element in human papillomavirus type 16 interacts with hFIP1, CstF-64, hnRNP C1/C2 and PTB. *J Virol* 2005; **79**: 4270-4288.

Zhao X, Rush M, Schwartz S. 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**: 10888–10905.

Zheng X, Ravatn R, Lin Y *et al.* Gene expression of TPA induced differentiation in HL-60 cells by DNA microarray analysis. *Nucleic Acids Research* 2002; **30**: 4489-99.

Zheng ZM, Baker CC. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front Biosci* 2006; **11**: 2286-302.

Zhou J, Doorbar J, Sun XY, Crawford LV, McLean CS, Frazer IH. Identification of the nuclear localization signal of human papillomavirus type 16 L1 protein. *Virology* 1991; **185**: 625-32.

Zhou J, Mulshine JL, Unsworth EJ *et al.* Purification and characterization of a protein that permits early detection of lung cancer. Identification of heterogeneous nuclear ribonucleoprotein-A2/B1 as the antigen for monoclonal antibody 703D4. *J Biol Chem* 1996; **271**: 10760–10766.

Zhou J, Sun XY, Louis K, Frazer IH. Interaction of human papillomavirus (HPV) type 16 capsid proteins with HPV DNA requires an intact L2 N-terminal sequence. *J Virol* 1994; **68**: 619-25.

Zur Hausen H. Papillomavirus infections—a major cause of human cancers. *Biochim Biophys Acta* 1996; **1288**: F55-78.

A. Appendix I - Solution

Agarose gel

1g agarose (Invitrogen)

100ml of TBE (Invitrogen) (appendix I)

boil until dissolved

let cool to 60 °C

add 5µl of GelRed® (Biotium) or 0.5µg/ml ethidium bromide

mix and pour into gel template

allow to set

TBE buffer

1000ml of 10×TBE (Invitrogen)

mix with 9000ml of distilled water

6×Loading buffer

0.25% bromophenol blue (BDH)

0.25% xylene cyanol (BDH)

30% glycerol (BDH)

5M NaCl

73g Sodium Chloride (Merck)

Diluted in 250ml of distilled water and autoclaved

Phenol/chloroform/isoamyl alcohol

49 ml of phenol (Sigma)

49ml Chloroform (Sigma-Aldich)

2ml isoamyl alcohol (Sigma)

Mix and freeze at -20°C

70% Ethanol

70ml 100% ethanol (Merck)

Made up to 100ml with distilled water

1M TRIS

121.1g Tris (BDH)

Mix with 1000ml of distilled water and adjust pH to 7 using HCl and autoclave

0.5M EDTA

181.1g EDTA (BDH)

Mix with 1000ml of distilled water and adjust pH to 7 using NaOH and autoclave

TE buffer

10mM Tris (appendix I)

1mM EDTA (appendix I)

LB broth

5g Tryptone (Lab M Limited)

5g NaCl (Merck)

2.5g Yeast extract (Lab M Limited)

Mix with 500ml distilled water and autoclave

0.1M CaCl₂

5.5g CaCl₂ (Riedel-de Haën)

Mix with 250ml distilled water and autoclave

SOC medium

4g Bacto-tryptone (Lab M Limited)

1g Yeast extracts (Lab M Limited)

0,1g NaCl (Merck)

Mix with 200ml of distilled water and shake it until dissolved

5ml 1M KCl

Bring the volume up to 200ml adjust pH to 7 using 10M NaOH (appendix I)

4ml 1M glucose solution

100mg/ml Ampicillin

100mg ampicillin (Sigma)

Mix with 1ml sterile water, filter the solution and freeze at -20°C

LB Agar

3g Tryptone (Lab M Limited)

3g NaCl (Merck)

1.5g yeast extracts (Lab M Limited)

4.5g Agar (Lab M Limited)

Mix with 300ml of distilled water and autoclave

Let cool to 55°C and add 150µl ampicillin(100mg/ml)

Pour into Petri dishes and let set

Alkaline Lysis Solution I

25mM Tris pH8

10mM EDTA

50mM Glucose

SDS 10%

10g SDS (BHD)

Make up to 100ml with distilled water

10M NaOH

40g NaOH

Make up to 100ml in a volumetric flask

Solution II

1ml 10% SDS (appendix I)

0.2ml 10M NaOH (appendix I)

Mix with 8.8ml distilled water (make fresh)

Alkaline Lysis solution III

60ml 5M potassium acetate

28.5ml distilled water

Mix and autoclave then add

11.5ml glacial acetic acid

TE buffer containing RNase A

1ml TE buffer

1 μ l RNase(22 μ g/ μ l) (Sigma)

Vortex and centrifuge for 3sec at maximum speed three times

Phosphate Buffered Saline (PBS)

1 PBS tablet (Oxoid) for every 100ml distilled water and autoclave

2.5% Trypsin

50ml trypsin (invitrogen)

1 bottle (500ml) Eare's balance salt (sigma)

0.02% EDTA in PBS

7,4g EDTA (BDH)

Make up to 1 litre with PBS in a volumetric flask

Trypsin with EDTA

10ml 2.5% Trypsin (appendix I)

10ml 0.02% EDTA (appendix I)

Mix by inverting

Sodium phosphate buffer

68,4ml disodium hydrogen phosphate (Na_2HPO_4) (141.96g Na_2HPO_4 (BDH) in 1 litre distilled water)

2× assay β -galactoside buffer

1ml 1M MgCl_2 (appendix I)

10ml 1M sodium phosphate buffer pH 7.2 (appendix I)

350 μ l β -mercaptoethanol (Fluka)

66.5mg *ortho*-Nitrophenyl- β -galactoside (ONPG) (sigma)

Make up to 50ml with distilled water in a volumetric flask

Cell lysis solution

25mM EDTA

2% SDS

Protein precipitation solution

10M ammonium acetate

1M Sodium carbonate

10.5g sodium carbonate (BDH)

Make up to 100ml distilled water in a volumetric flask

2M Magnesium chloride

190,4g magnesium (BDH)

Make up to 1 litre with distilled water in a volumetric flask

0,5M ZnCl₂

68,14g ZnCl₂ (Sigma-Aldrich)

Make up to 1 litre with distilled water in a volumetric flask

2M Diethanolamine

20ml 9,5M Diethanolamine

Make up to 100ml with distilled water in a volumetric flask

15% CHAPS

0.15g CHAPS (Sigma)

1ml of PBS (appendix I)

0,05% CHAPS

33µl 15% CHAPS (appendix I)

9.67ml PBS (appendix I)

Coloring substrate

20ml 2M diethanolamine

20µl 0,5M ZnCl₂

10µl 2M MgCl₂

Mix and add 1 tablet of 4-nitrophenyl phosphate disodium salt hexahydrate (Sigma-Aldrich)

0.01mol/l Citrate Buffer (pH 6)

2.1g citric acid

add approximately 800ml distilled water

adjust pH to pH 6 using 2mol/l NaOH (79.98g NaOH in 1 litre)

Make solution up to 1 litre in a volumetric flask

12% polyacrylamide gel

6.6 ml H₂O

5.2 ml polyacrylamide buffer (appendix I)

8 ml 30% Acrylamide (29:1)

180 µl 10% APS (appendix I)

40 µl Temed

5% polyacrylamide gel

2.1 ml H₂O

0.9 ml Upper buffer (appendix I)

0.55 ml 30% Acrylamide (29:1)

40 µl 10% APS

4 µl Temed

10% APS

0.04g APS

400µl H₂O

Lower buffer

9.08g Tris

0.4% SDS

Bring the volume up to 50ml H₂O and adjust pH to pH 8.8

Upper buffer

3.03g Tris

0.4% SDS

Bring the volume up to 50ml H₂O adjust pH to pH 6.8

Running buffer

15g Tris

72g glycine

5g SDS

Bring the volume up to 500ml H₂O

Transfer buffer

6g Tris

28.82g glycine

Bring the volume up to 200ml H₂O and adjust pH to pH 8.3

10XTBS

12.1g Tris

43.8g NaCl

Bring the volume up to 500ml H₂O

5% milk

5g milk

100ml H₂O (store 4°C)

TBS TWEEN

1/10 TBS

0.1% Tween

Make 1L H₂O

Laemmli 2X buffer

4% SDS

10% 2-mercaptoethanol

20% glycerol

0.004% bromophenol blue

0.125 M Tris HCl

Human poliovirus 2 strain MEF-1 polyprotein gene, complete cds
Length=7440

Score = 1308 bits (708), Expect = 0.0
Identities = 718/723 (99%), Gaps = 0/723 (0%)
Strand=Plus/Plus

```
Query 237 TTAAACAGCTCTGGGGTTGTACCCACCCCAGAGGCCACGTGGCGGCTAGTACTCCGGT 296
      |||
Sbjct 1 TTAAACAGCTCTGGGGTTGTTCACCCCAGAGGCCACGTGGCGGCCAGTACACTGGT 60

Query 297 ATTGCGGTACCTTTGTACGCCTGTTTTATACTCCCTTCCCCCGTAACTTAGAAGCACAAT 356
      |||
Sbjct 61 ATTGCGGTACCTTTGTACGCCTGTTTTATACTCCCTTCCCCCGTAACTTAGAAGCACAAT 120

Query 357 GTCCAAGTTCAATAGGAGGGGGTACAAACCAGTACCACCACGAACAAGCACTTCTGTTCC 416
      |||
Sbjct 121 GTCCAAGTTCAATAGGAGGGGGTACAAACCAGTACCACCACGAACAAGCACTTCTGTTCC 180

Query 417 CCCGGTGAGGCTGTATAGGCTGTTTCACGGCTAAAAGCGGCTGATCCGTTATCCGCTCA 476
      |||
Sbjct 181 CCCGGTGAGGCTGTATAGGCTGTTTCACGGCTAAAAGCGGCTGATCCGTTATCCGCTCA 240

Query 477 TGTACTTCGAGAAGCCTAGTATCACCTTGAATCTTCGATGCGTTGCGCTCAACTCAA 536
      |||
Sbjct 241 TGTACTTCGAGAAGCCTAGTATCACCTTGAATCTTCGATGCGTTGCGCTCAACTCAA 300

Query 537 CCCAGAGTGTAGCTTAGGTCGATGAGTCTGGACGTTCCCTCACCGGCGACGGTGGTCCAG 596
      |||
Sbjct 301 CCCAGAGTGTAGCTTAGGTCGATGAGTCTGGACGTTCCCTCACCGGCGACGGTGGTCCAG 360

Query 597 GCTGCGTTGGCGGCTACCTGTGGCCCAAAGCCACAGGACGCTAGTTGTGAACAAGGTGT 656
      |||
Sbjct 361 GCTGCGTTGGCGGCTACCTGTGGCCCAAAGCCACAGGACGCTAGTTGTGAACAAGGTGT 420

Query 657 GAAGAGCCTATTGAGCTACCTGAGAGTCCTCCGGCCCCTGAATGCGGCTAATCCTAACCA 716
      |||
Sbjct 421 GAAGAGCCTATTGAGCTACCTGAGAGTCCTCCGGCCCCTGAATGCGGCTAATCCTAACCA 480

Query 717 CGGAGCAGGCAGTGGCAATCCAGCGACCAGCCTGTCGTAACGCGCAAGTTCGTGGCGGAA 776
      |||
Sbjct 481 CGGAGCAGGCAGTGGCAATCCAGCGACCAGCCTGTCGTAACGCGCAAGTTCGTGGCGGAA 540

Query 777 CCGACTACTTTGGGTGTCCGTGTTTCCTTTTATTTTACAATGGCTGCTTATGGTGACAA 836
      |||
Sbjct 541 CCGACTACTTTGGGTGTCCGTGTTTCCTTTTATTTTACAATGGCTGCTTATGGTGACAA 600

Query 837 TCATTGATTGTTATCATAAAGCAAATTGGATTGGCCATCCGGTGAGAATTTGATTATTAA 896
      |||
Sbjct 601 TCATTGATTGTTATCATAAAGCAAATTGGATTGGCCATCCGGTGAGAATTTGATTATTAA 660

Query 897 ATTACTCTCTTGTGGGATTGCTCCTTTGAAATCTGTGCACTCACACCTATTGGAATTA 956
      |||
Sbjct 661 ATTACTCTCTTGTGGGATTGCTCCTTTGAAATCTGTGCACTCACACCTATTGGAATTA 720

Query 957 CCT 959
      |||
Sbjct 721 CCT 723
```

Publications

Orrù B, Cunniffe C, Ryan F, Schwartz S.

Development and validation of a novel reporter assay for human papillomavirus type 16 late gene expression.

In Press *Journal of Virological Methods* 2012.

Somberg M, Li X, Johansson C, Orrù B, Chang R, Rush M, Fay J, Ryan F, Schwartz S.
Serine/arginine-rich protein 30c activates human papillomavirus type 16 L1 mRNA expression via a bimodal mechanism.

J Gen Virol 2011 October; 92: 2411-21.

Orrù B, Schwartz S, Ryan F.

Novel factors involved in regulation of HPV-16 late gene expression.

27th International Papillomavirus Conference, Berlin, Germany, September 17-22, 2011.

Li X, Orrù B, Somberg M, Ryan F, Schwartz S.

RNA element supresses HPV-16 splice site SD3632.

27th International Papillomavirus Conference, Berlin, Germany, September 17-22, 2011.

Orrù B, Ryan F, Schwartz S.

Factors influencing HPV-16 late gene expression.

26th International Papillomavirus Conference, Montreal, Canada, July 3-8, 2010.

Li X, Orrù B, Somberg M, Ryan F, Schwartz S.

Identification of inhibitory sequences upstream of SD3632 in the HPV-16 genome.

26th International Papillomavirus Conference, Montreal, Canada, July 3-8, 2010.

Somberg M, Li X, Orrù B, Ryan F, Schwartz S.

Enhancement of HPV-16 early mRNA splicing by ASF/SF2 is antagonised by SRp30c, a specific inducer of L1 mRNA production.

26th International Papillomavirus Conference, Montreal, Canada, July 3-8, 2010.

Li X, Orrù B, Somberg M, Ryan F, Schwartz S.

Upstream sequences suppress HPV-16 late 5'-splice site SD3632.

Molecular Biology of DNA tumor Viruses Conference, Madison, Wisconsin, USA, July 13-18, 2010.

Orrù B, Ryan F, Schwartz S.

A reporter assay for analysis of HPV-16 late gene expression.

4th European Congress of Virology, Lake Como, Italy, April 7-11, 2010.

Schwartz S, Somberg M, Johansson C, Orrù B, Ryan F, Li X.

ASF/SF2 regulates HPV-16 early and late mRNA splicing.

5th Smögen Summer Symposium on Virology, Sweden, August 20-22, 2009.