UPPSALA
UNIVERSITET

Overexpression of E2 and ADAR Individually Induce Human Papillomavirus Type 16 Late Gene Expression

Tove Samuelsson

Degree project in biology, Master of science (1 year), 2008
Examensarbete i biologi 30 hp till magisterexamen, 2008
Biology Education Centre and the Department of Medical Biochemistry and Microbiology, BMC, Uppsala University
Supervisors: Stefan Schwartz and Monika Somberg

TITEL PAGE

1 Summary 2

2 Introduction 3

1.1 The HPV-16 genome....3
1.2 Splicing and polyadenylation....4
1.3 The E2 protein....5
1.4 ADAR....6
1.5 How to turn on the late genes?....6

2. Aim 7

3. Results 8

3.1 Determining RNA quality....8
3.2 Northern Blot analysis....8

4 Discussion 16

4.1 E2 and ADAR regulate late gene expression in HPV-16....16
4.2 The role of E2....16
4.3 Future perspective....16

5. Materials and methods 18

5.1 Cell line, bacterial strain and plasmids....18
5.2 Plasmids used or constructed in this project....18
5.3 Polymerase chain reaction....20
5.4 Gel electrophoresis....21
5.5 Transfection....21
5.6 Extraction of cytoplasmic ribonucleic acid....22
5.7 Northern Blot analysis....23

6 Acknowledgements 25

7 References 26

Summary

Cervical cancer is the second most common cancer in women today and is in 99.7 % of all cases caused by human papillomaviruses (HPVs). More than 100 types of this virus have been identified, but it is primarily HPV type 16 that is seen in cervical cancer. The HPV genome can be divided into early genes and late genes. The early genes express regulatory proteins, whereas the late genes express capsid proteins. These capsid proteins would be an easy target for recognition by the immune system. However, the virus suppresses the late genes in differentiating cells and does not express these genes until very late in the infection, when the cells burst and capsid proteins are necessary for the formation of new virus particles. It would be advantageous to be able to turn on the late gene expression earlier in the infection, so that the infection could be cleared by the immune system before it had a chance to persist and cancer could develop.

It has been shown previously that the polypyrimidine tract binding protein (PTB) induces late gene expression in human papillomavirus type 16 (HPV-16). It does so by interfering with splice site inhibitors surrounding the splice donor at position 3632, a splice site used only by late mRNAs. In this study two other proteins that could be involved in the regulation of late gene expression of HPV-16, namely the regulatory protein E2 and the RNA editing protein ADAR, were investigated. E2 is particularly interesting since it is encoded by one of the early genes in HPV-16. ADAR was investigated because of the results from another study that showed that ADAR plays a role in polyoma infection. In this project I also investigated with what part of the genome E2 could be interfering.

When E2 or ADAR were overexpressed and co-transfected with various plasmids, an increase in L1 late gene mRNA was seen. This study showed that E2 and ADAR individually induce late gene expression of the human papillomavirus type 16.

1 Introduction

Human papillomaviruses (HPVs) are small, double-stranded, circular DNA viruses that infect epithelial cells and mucous membranes (Murray *et al.* 2005). They are pathogens of humans, other mammals and birds (Zheng & Baker 2006). More than 100 different types of HPV have been identified and sequenced (Murray *et al.* 2005, Ault 2006, Zheng & Baker 2006) and the virus is spread by skin-to-skin contact (Ault 2006). It gives rise to warts, such as foot warts or genital condyloma, but it has also been seen in more than 99.7 % of all cases of cervical cancer (Doorbar 2006). Cervical cancer is the second most common cancer among women (Schiffman *et al.* 2007) and is caused by so called oncogenic high-risk HPV-types, such as 16, 18, 31 or 45 (Murray *et al.* 2005). In fact, HPV-16 has been found in approximately 50 % of all cervical cancers and is, therefore, of special interest (Stone *et al.* 2002).

In developed countries women can regularly, preferably every third year, do a Pap smear screening. This test was developed 80 years ago by Georgios Papanikolaou and is done to screen for the early development of cervical cancer. It is a screening for abnormal cells in the cervix and it is performed without previous indication of HPV infection (Blumenthal & Gaffikin 2005). It is important to take these tests every third year, since it takes 5-10 years before a pre-stage to cancer can be seen after an HPV-infection. The establishment of cancer then can take up to ten years after a pre-stage has been seen (Schiffman *et al.* 2007).

A real study of how the current vaccines will affect the incidents of cervical cancer will not be obtainable until 30 years from now (Adams *et al.* 2007) and even though the vaccines might be good, they are still too expensive for usage in developing countries. This is a problem, since cases of cervical cancer are highest in countries that cannot afford a welfare system that gives the women access to regular Pap smear screening. Thus, it is important to continue the research of this virus and find new ways to treat the diseases it is causing.

1.1 The HPV-16 genome

The genome of HPV-16 is divided into an early region, a late region and a long control region (LCR), figure 1. These three regions make up the basis of the 8 kb large genome and there are two polyadenylation sites that separate them, the early polyadenylation site (pAE) and the late polyadenylation site (pAL) (Zheng & Baker 2006). The early region codes for six regulatory proteins, E1, E2, E4, E5, E6 and E7, with a promoter called p97. The late region takes up almost 40 % of the genome and codes for the capsid proteins L1 and L2, with the p670 promoter located in the early region. The control region is non-coding but has important sites, such as an origin of replication and transcription factor binding sites. (Zheng & Baker 2006, Murray *et al.* 2005)

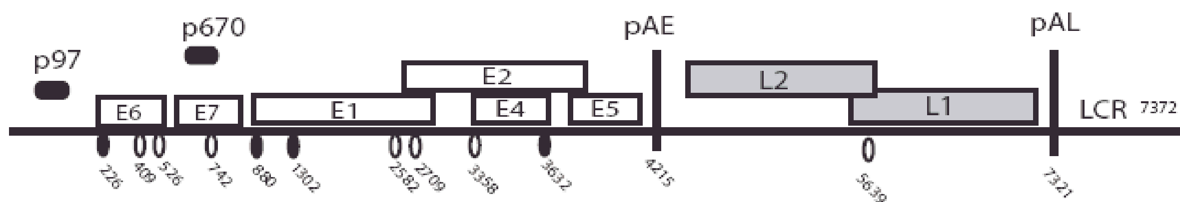


Figure 1. The HPV-16 genome. Boxes represent the eight genes of the HPV-16 genome. The bright boxes code for early regulatory proteins, while the dark boxes code for the late capsid proteins. Underneath the boxes there are splicing donors (black ovals) and splicing acceptors (white ovals). Splice sites and polyadenylation sites are indicated with numbers at their nucleotide position.

The regulatory genes are expressed early in the infection and these proteins are isolated to the nucleus. It is, therefore, hard for the host's immune system to recognize an infection at first. Since L1 and L2 are capsid proteins one would assume that L1 and L2 proteins would be an easy target for the immune defence. However, the fact is that the late genes are downregulated, mainly by alternative splicing and polyadenylation, during the initiation of infection, figure 2. They are not upregulated until very late in the infection, just prior to the disruption of epithelial cells. Thus, the regulation of the HPV-16 genes depends on epithelial differentiation. (Tindle 2002, Milligan *et al.* 2007).

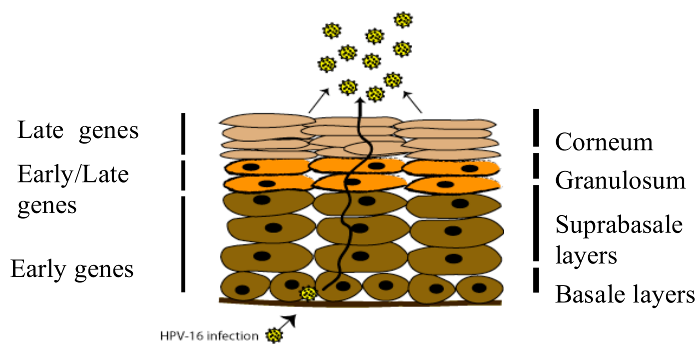


Figure 2. Lifecycle. The virus expresses its genes dependent on the state of differentiation of epithelial cells.

HPV is the most commonly spread sexually transmitted disease and most infections are cleared within two years (Shiffman *et al.* 2007). Why however some infections persist and cause cervical cancer in some individuals is still unclear. What is known is that no late gene expression occurs in cervical cancer cells that contain HPV-16 DNA (Doorbar 2005). Due to the fact that the immune defence has such trouble recognizing an infection when only the early genes are expressed, it would be of great advantage if one could turn on the expression of the late genes earlier in the infection. This would lead to recognition and clearance of the infection by the immune system before cancer had the chance to develop.

1.2 Splicing and polyadenylation

HPV-16 infected cancer cells do not express the late genes, L1 and L2, but show constant expression of the oncogenic early genes, E6 and E7 (Schwartz 2008). This allows the virus to persist, without being recognized by the immune system of the host. Thus, the inhibition of late gene expression is thought to be a prerequisite for cancer progression (Rush *et al.* 2005). The expression of viral transcripts in HPV-16 is highly regulated by means of posttranscriptional regulations, such as splice site selection and polyadenylation (Schwartz 2008).

The eight genes of HPV-16 are generally referred to as open reading frames (ORFs). This means that its proteins are encoded, not from single genes, but from sequences in the genome containing start codons followed by at least 100 bases without interruption of a stop codon (Lodish *et al.* 2004). Thus, the genome is transcribed as bicistronic or polycistronic templates, i.e. with two or several ORFs in each transcription. To become mature mRNAs, these sequences have to undergo alternative splicing (Milligan *et al.* 2007).

Alternative splicing takes place in the spliceosome and can allow a single gene to produce an array of different gene products. What this means is that pre-mRNA that is made up by exons and introns can be alternatively spliced, where different parts will be removed and therefore give rise to different gene products, figure 4. The spliceosome is made up by splicing factors

and also four small nuclear ribonucleoproteins (snRNPs U1, U2, U4/U6 and U5). (Cáceres & Kornblihtt 2002).

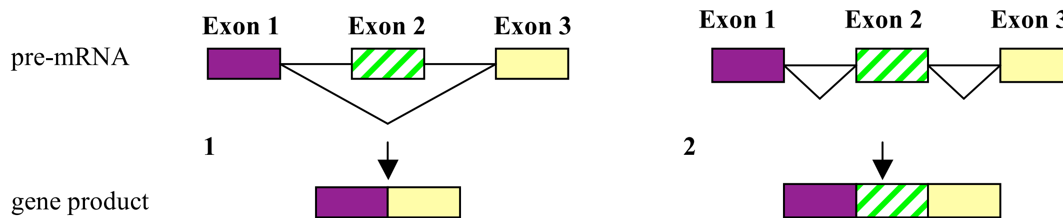


Figure 4. Two examples of alternative splicing.

An early and a late polyadenylation site, with the sequence AAUAAA, are found in HPV-16 downstream of the ORFs E5 and L1 respectively (Terhune *et al.* 1999). To get expression of the early genes in the viral life cycle the early polyadenylation site has to be polyadenylated efficiently and the switch from early to late genes is very much dependent on the early polyadenylation site (Oberg *et al.* 2005). To allow transcription of the late genes, the early polyadenylation site has to be ignored, so that the transcript can read through it (Terhune *et al.* 1999). The regulation of this early polyadenylation site is therefore very interesting.

1.3 The E2 protein

There are several factors that regulate transcription and replication in HPV-16. One protein that seems to be involved in these events is the E2 protein, which is encoded by one of the early genes in HPV-16. It is a trans-activating protein and increasing amounts appear to down-regulate the early promoter, p97 (Hernandez-Ramon *et al.* 2008). The integration of viral DNA into the host genome often leads to disruption of E2 in oncogenic types of HPV. This allows transcription from the oncogenes E6 and E7, leading to down-regulation of host tumour suppressor genes p53 and pRb (Hernandez-Ramon *et al.* 2008, Motoyama *et al* 2004).

E2 binds to a palindromic sequence, ACCN$_6$GGT, and has several binding sites in the long control region (Gammoh *et al.* 2006). It is about 45 kD and is divided into three regions, the transactivating domain (TAD), a hinge region and a DNA binding domain (DBD) (Hernandez-Ramon *et al.* 2008), figure 3. Both the TAD and the DBD are constant and well conserved among different papilloma viruses. The hinge region, however, varies in length and the function of this part remains unclear.

Figure 3. The three domains of the E2 protein.

The binding of E2 to DNA does not by itself allow replication. Another protein, E1, which is a helicase is also encoded in the early region and needed to initiate this process (White *et al.* 2003). Therefore, E2 attracts E1, which in turn binds to cellular DNA polymerase. It is this complex that can bind DNA and initiate replication (Wilson *et al.* 2002).

1.4 ADAR

Another protein of interest regarding the expression of late genes in HPV-16 is ADAR (adenosine deaminase acting on RNA), an RNA editing protein. ADAR acts posttranscriptionally on doublestranded RNA such as secondary structures like stem loops or bulges (Valente & Nishikura 2005). The target sequence of ADAR is an AAUAAA repeat, and poly (A) signals are very efficiently edited (Gu *et al.* 2007). The editing by ADAR occurs when it deaminates adenosine to inosine, which allows for the production of a variety of final protein sequences due to inosine's less stringent base pairing to tRNA. This is a site-specific change, whereas regulations such as polyadenylation or splicing influence greater parts of the transcript (Valente & Nishikura 2005).

Gu *et al.* 2007 has shown that there is more than 50 % conversion of adenosines to inosines in persistent polyoma-infection, and it is, therefore, interesting to see if this has an affect on late gene expression in HPV-16.

1.5 How to turn on the late genes?

It has been shown previously that the polypyrimidine tract binding protein (PTB) promotes late gene expression by the activation of a 5′-splice site referred to as SD3632. Splicing inhibitors that normally down-regulate this particular late gene splice site are in some way interfered with in the presence of PTB (Somberg *et al.* 2008). They investigated what parts of the HPV-16 genome that could be deleted without changing the effect of PTB and found a fascinating fact. If a 198 nt long sequence downstream of the SD3632 site was deleted, the late genes were expressed not only in the presence, but also in the absence of PTB. The conclusion was that overexpression of PTB may interfere with this sequence through cellular factors that interact with the same sequence. (Somberg *et al.* 2008).

2 Aims

Based on the results with PTB (Somberg *et al.* 2008), it was interesting to investigate if E2, the regulatory protein, works in the same way as PTB. This would mean that one of HPV´s own early genes, if overexpressed, might upregulate the expression of the late genes. It would also be of interest to make new plasmids to examine this inhibitory sequence.

Thus, the main questions to be answered in this study were whether E2 and ADAR induced late gene expression in HPV-16 and also with what part of the HPV-16 genome E2 could interfere.

The effect of E2 on different parts of the HPV-genome was investigated by plasmid transfections, with following RNA extraction and Northern blots with a probe for the HPV-16 late mRNA.

3 Results

To investigate whether or not the proteins E2 or ADAR induce late gene expression, Northern blots were performed. To further investigate what part of the HPV-16 genome is affected by E2, plasmids with different deletions or replacements were used for transfection into HeLa cells in the absence or presence of E2. Before the Northerns were run, though, the quality of the extracted RNA had to be determined.

3.1 Determining RNA quality

After extracting cytoplasmic RNA, its quality was checked on an agarose gel containing ethidium bromide to make sure the RNA had not been degraded during the process and was good enough for a Northern blot. The 28S ribosomal RNA is present in greater quantities than the 18S and 5S ribosomal RNAs and should therefore appear brighter. When RNA is degraded, the 5S ribosomal RNA will appear more strongly, whereas the 28S and the 18S bands will be very hard to distinguish. With RNA quality as shown in figure 5, the experiment could proceed and Northern blots could be performed.

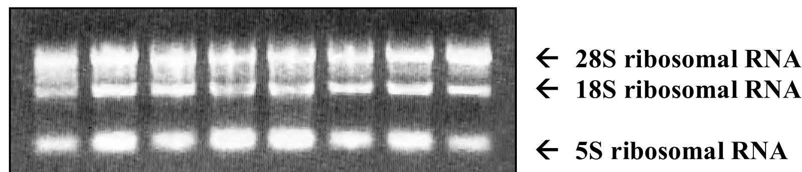


Figure 5. RNA quality. RNA run on a 1 % agarose gel. The three bands represent ribosomal RNA.

3.2 Northern blot analysis

Extracted RNA from HeLA cells transfected with plasmids encoding different parts of HPV DNA in the presence or absence of the plasmids CMV-E2 or ADAR was run for Northern blot analysis. CMV-E2 and ADAR promotes late gene expression in HPV-16.

3.2.1 Plasmids permitting usage of the splicing site within L1

pBEL is a plasmid where the early and late promoters, p97 and p670 respectively, and also the early genes E6 and E7 have been replaced with a cytomegalovirus promoter (CMV). pBELM is constructed in the same way as pBEL with the exception of a mutation in the late gene L1, figure 6A. This mutation allows efficient usage of the splice site at position 5639, which normally is inhibited by splicing regulatory sequences (Zhao & Schwartz 2008).

Northern blots of RNA from cells transfected with pBEL and pBELM in the absence of CMV-E2 did not show any expression of late gene mRNAs. Transfection with pBEL in the presence of CMV-E2 did not show any significant increase in the expression of late gene mRNA. The presence of CMV-E2 had an effect on the expression of late gene mRNA when co-tranfected with pBELM, though, figure 6B. The same result was obtained when pBEL and pBELM were co-transfected with or without ADAR. This could also be seen in the RT-PCR performed on the same RNA as in the Northern blot of cells transfected with pBEL and pBELM in the absence or presence of ADAR, figure 6C. Thus, late gene mRNA increase could only be seen in the Northern blots from cells transfected with pBELM in the presence of CMV-E2 or ADAR.

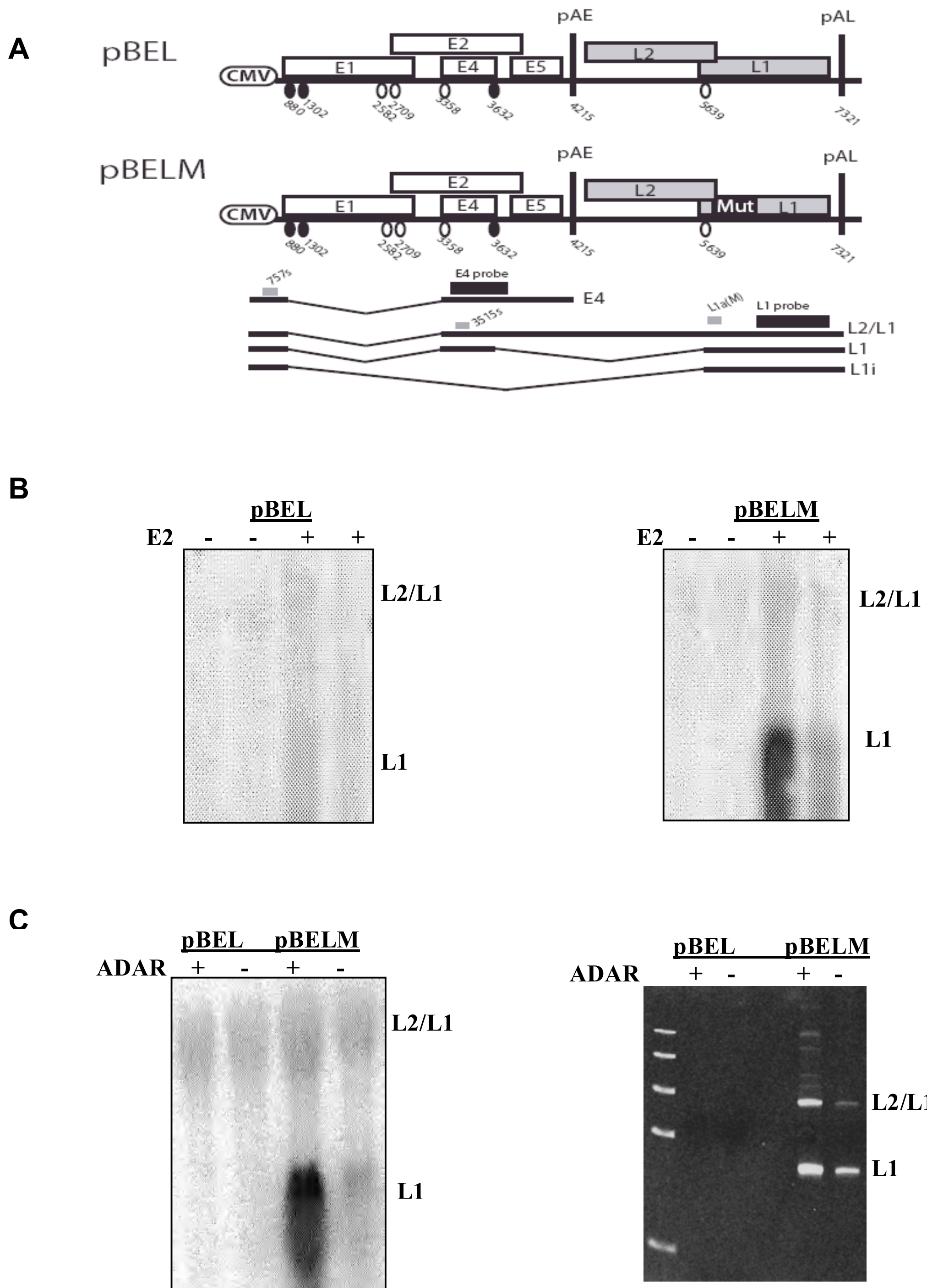


Figure 6. Late gene expression from pBEL and pBELM in HeLa cells. (A) Plasmids pBEL and pBELM. Filled circles represent splice donors and empty circles represent splice acceptors. The early poly (A) site, pAE, and the late poly (A) site, pAL, are indicated. Numbers represent nucleotide positions for both splice sites and polyadenylation sites. The mutation in pBELM enables the use of the splice site at position 5639, which is normally inhibited by splice site regulators. Major mRNAs, such as E4 and L1 are marked underneath. Blots were probed with an L1 probe. 757s and L1a(M) are locations of primers used for RT-PCR. (B) Northern blot representing cytoplasmic RNA extracted from HeLa cells transfected with pBEL and pBELM in the presence or absence of CMV-E2. (C) Northern blot representing cytoplasmic RNA extracted from HeLa cells transfected with pBEL and pBELM in the presence or absence of an ADAR-encoding plasmid. The RNA used in the Northern blot was then used for RT-PCR with primers 757s and L1a(M) (see 6A).

3.2.2 Removal of all splice sites upstream of position 3395
pBSplice is a plasmid where only one splice donor, at position 3632 (SD3632), and one splice acceptor, at position 5639 (SA5639), still remains, figure 7A. This was obtained by deleting the first ~3000nt and thereby remove the other splice sites. SD3632 is exclusively used by late mRNAs and has been investigated previously, where it turned out that it is surrounded by splice site inhibitors. These inhibitors can be downregulated by PTB and it was therefore interesting to see if E2 had a similar effect.

pBSpliceM was constructed as pBSplice except for a mutation in the L1 region, such as in pBELM.

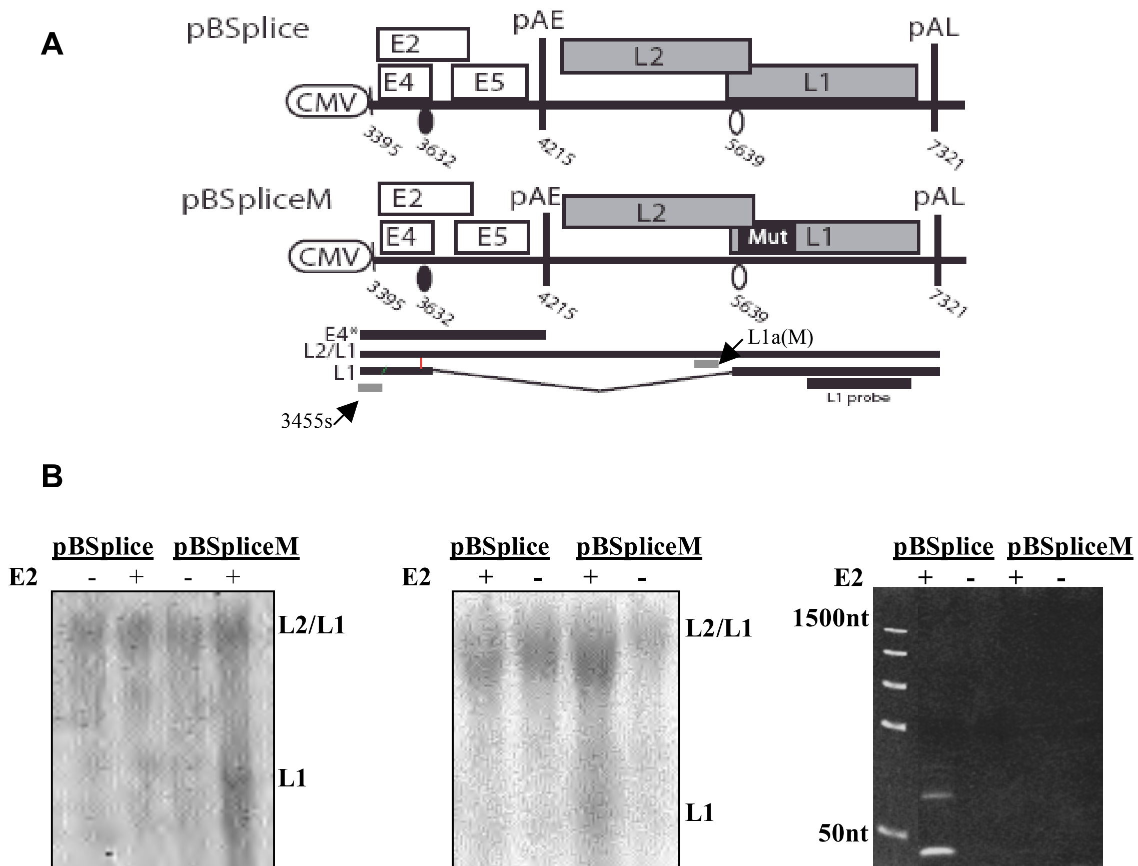


Figure 7. Late gene expression from pBSplice and pBSpliceM. (A) Plasmids pBSplice and pBSpliceM. The filled circle represent splice donor and the empty circle represent splice acceptor. The early poly (A) site, pAE, and the late poly (A) site, pAL, are indicated. Numbers represent nucleotide positions for both splice sites and polyadenylation sites. The mutation in pBSpliceM enables the use of the splice site at position 5639, which is normally inhibited by splice site regulators. Major mRNAs, such as E4 and L1 are marked underneath. Blots were probed with an L1 probe. The grey boxes marked as 3455s or L1a(M) are locations of primers used for RT-PCR. (B) Northern blot representing cytoplasmic RNA extracted from HeLa cells transfected with pBSplice and pBSpliceM in the presence or absence of CMV-E2. The RNA used in the Northern blot was then used for RT-PCR with primers 3455s and L1a(M) (see 6A).

pBSplice did not show any expression of late gene mRNA, neither in the presence nor in the absence of CMV-E2, in the Northern blots. However, in the RT-PCR for pBSplice, bands appeared in the presence of CMV-E2, but the L1 band should have been 372nt, which was not the size of any of the bands in the RT-PCR, figure 7B. pBSpliceM showed some indication of late gene mRNA expression in the presence of CMV-E2, figure 7B.

3.2.3 Plasmids with replacements and deletions of important sites further investigates the role of E2

In p2xpAL the early polyadenylation site has been deleted and replaced with a late polyadenylation site, figure 8A. This plasmid was made from pBSpliceM and it was now interesting to see if p2xpAL has the same properties as pBSpliceM or if the change affected the expression of late gene mRNA.

pMT22 was also based on the pBSpliceM construct, but here the whole early region, upstream of the early polyadenylation site, was deleted. This means that it has lost its splice donor at position 3632, figure 9A.

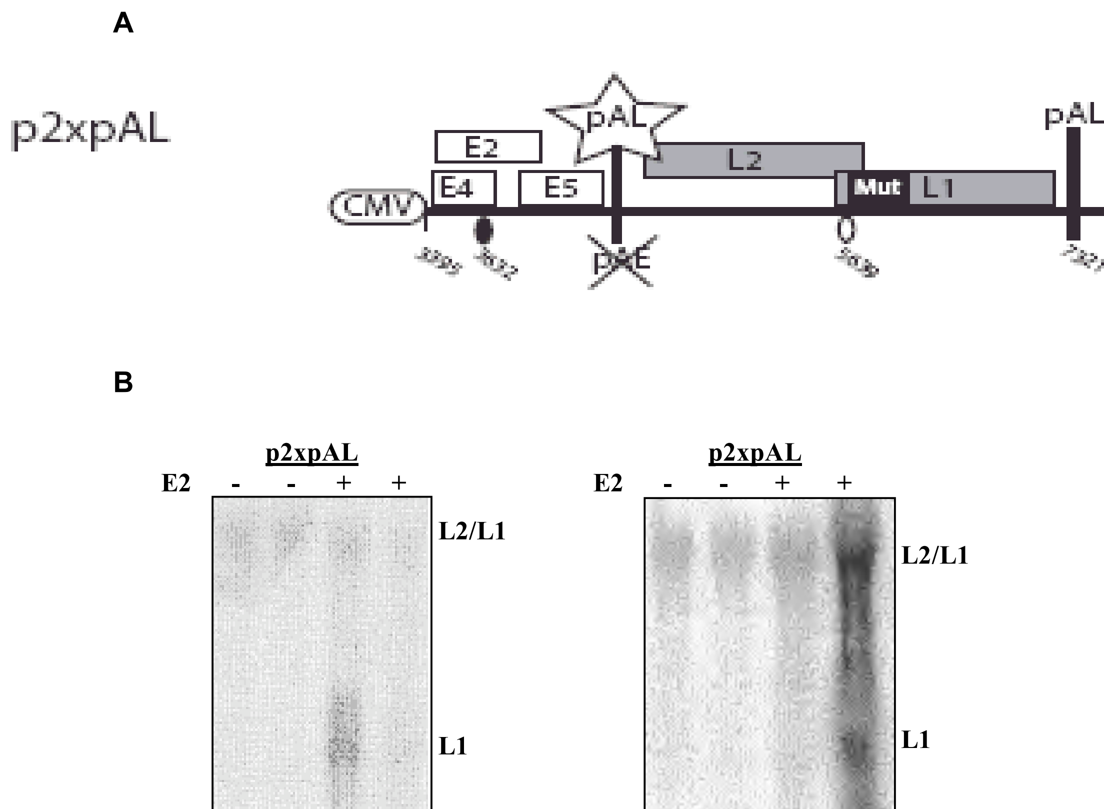


Figure 8. Late gene expression from p2xpAL. (A) Plasmid p2xpAL. The filled circle represent splice donor and the empty circle represent splice acceptor. The early poly (A) site, pAE, gas been deleted and replaced with a late poly (A) site, pAL, indicated by a star. Numbers represent nucleotide positions for both splice sites and polyadenylation sites. The mutation in p2xpAL is the same as in pBSpliceM and enables the use of the splice site at position 5639, which is normally inhibited by splice site regulators. Blots were probed with an L1 probe. (B) Northern blot representing cytoplasmic RNA extracted from HeLa cells transfected with p2xpAL in the presence or absence of CMV-E2.

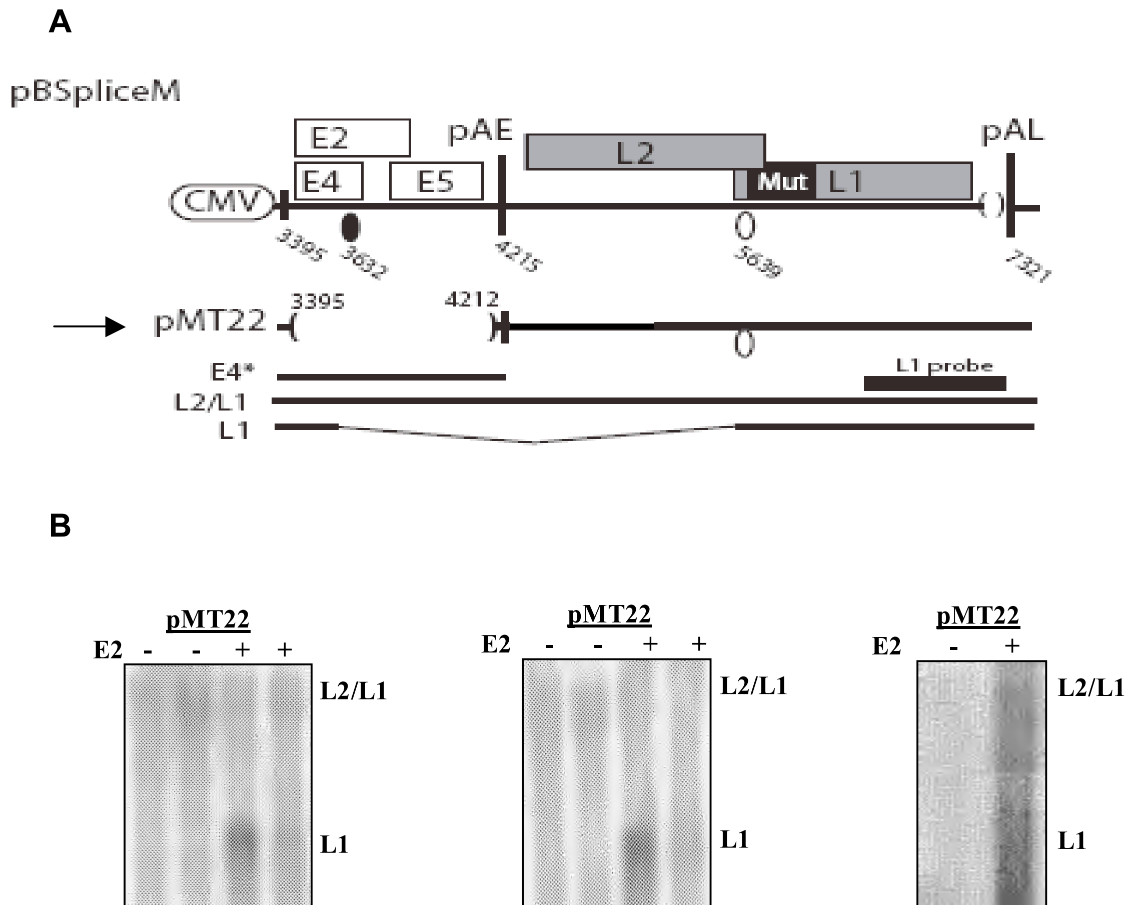


Figure 9. Late gene expression from pMT22. (A) Plasmids pBSpliceM and pMT22. The filled circle represent splice donor and the empty circle represent splice acceptor. In pMT22 the early region seen in pBSpiceM is deleted. Numbers represent nucleotide positions for both splice sites and polyadenylation sites. The mutation in pBSpliceM enables the use of the splice site at position 5639, which is normally inhibited by splice site regulators. Blots were probed with an L1 probe. (B) Northern blot representing cytoplasmic RNA extracted from HeLa cells transfected with pMT22 in the presence or absence of CMV-E2.

p2xpAL and pMT22 did not show any expression of the late genes in the absence of CMV-E2. In the presence of CMV-E2 there were indications of expression of L1 mRNA from both plasmids, although the results were not significant, figure 8B and 9B.

3.2.4 To decipher the sequence of interest

Seven plasmids were constructed from the pBSpliceM template. They all have deletion between nucleotide position 3395 and the early polyadenylation site, ending at position 4212, figure 10A. The deletions were made to determine with what part of the genome PTB interfered with, and are now used for the same purpose with E2.

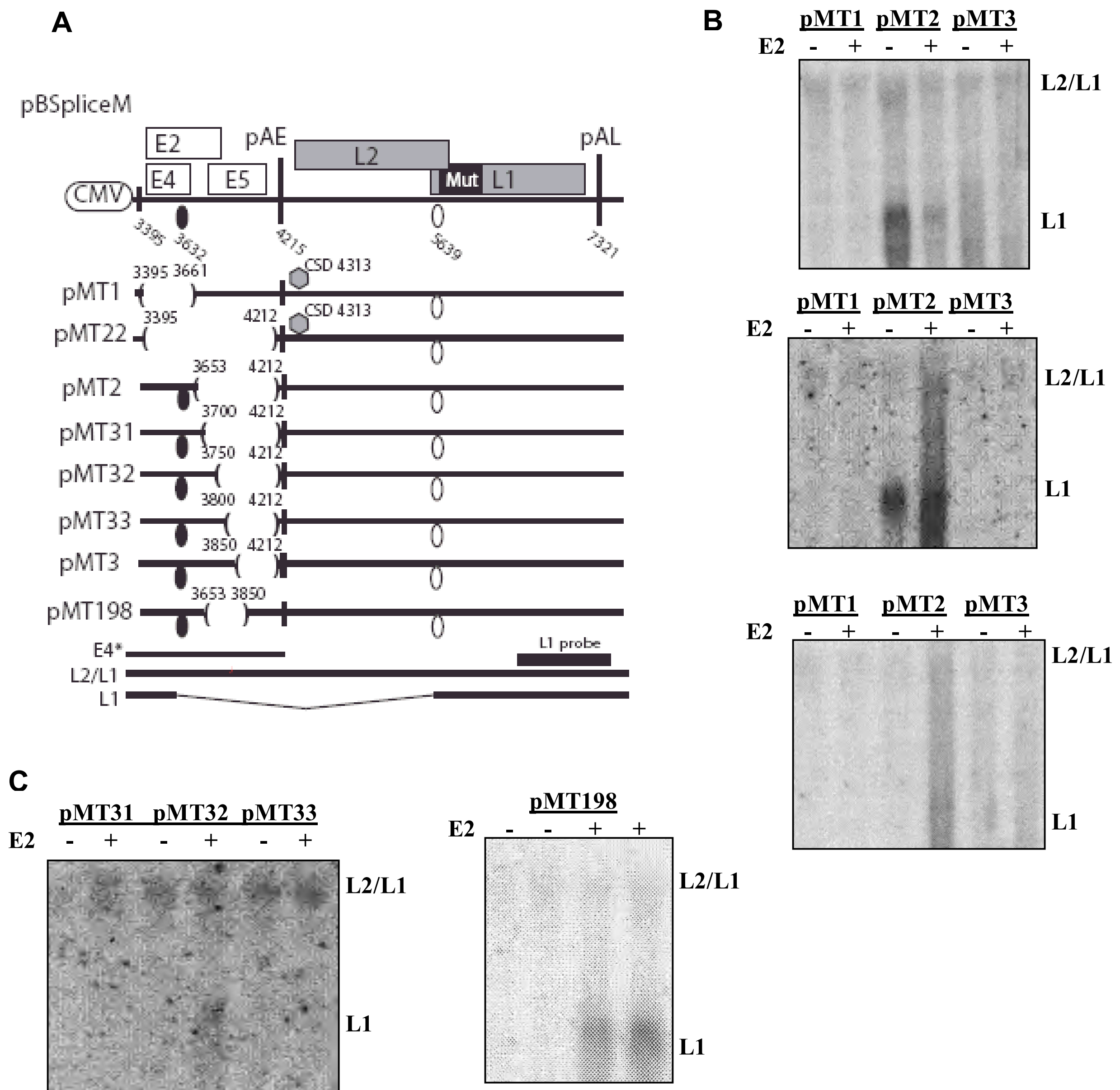


Figure 10. Late gene expression from pMT1, -2, -3, -31, -32, -33 and -198. (A) Plasmids pBSpliceM, pMT1, pMT2, pMT3, pMT31, pMT32, pMT33 and pMT198. The filled circle represent splice donor and the empty circle represent splice acceptor. Numbers represent nucleotide positions for both splice sites and polyadenylation sites. The mutation in the template pBSpliceM enables the use of the splice site at position 5639, which is normally inhibited by splice site regulators. Blots were probed with an L1 probe. Plasmids pMT1, pMT2, pMT3, pMT31, pMT32, pMT33 and pMT198 all have different deletions in the early region of pBSpliceM. (B) Northern blot representing cytoplasmic RNA extracted from HeLa cells transfected with pMT1, pMT2 or pMT3 in the presence or absence of CMV-E2. (C) Northern blot representing cytoplasmic RNA extracted from HeLa cells transfected with pMT31, pMT32, pMT33 or pMT198 in the presence or absence of CMV-E2.

pMT1, pMT3, pMT31 and pMT33 did not yield any late gene mRNA expression either in the absence or presence of CMV-E2. pMT2, pMT32 and pMT198 did not show any expression in the absence of CMV-E2, but showed weak expression of L1 mRNA in the presence of CMV-E2, figure 10B-C.

3.2.5 Strong expression of late gene mRNA in the presence of E2 despite large deletions
In pMT1SD the whole early region upstream of the splice donor at position 3632 is deleted. pMTDpAE has a deletion that removes the early polyadenylation site. Both of them have the mutation in L1, such as pBSpliceM, which allows usage of the splice site at position 5639.

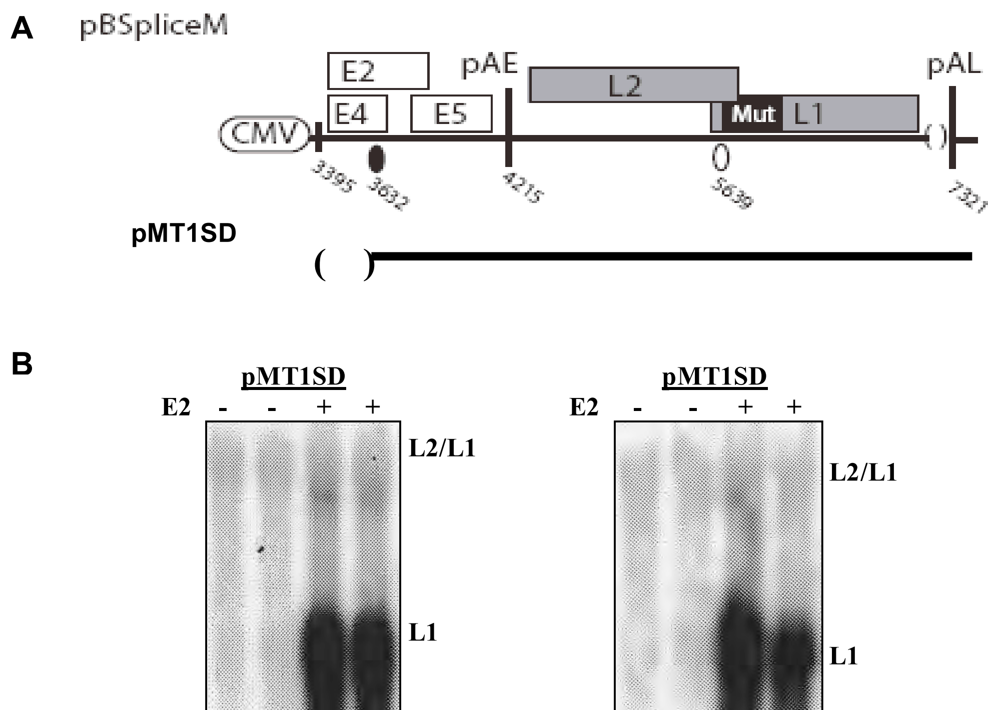
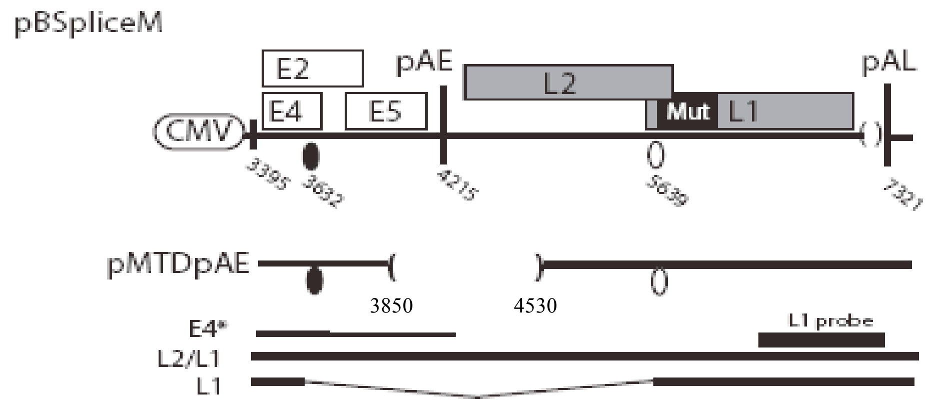


Figure 11. Late gene expression from pMT1SD. (A) Plasmids pBSpliceM and pMT1SD. In pMT1SD the whole early region upstream of the splice donor at position 3632 is deleted The filled circles represent splice donor and the empty circles represent splice acceptor. Numbers represent nucleotide positions for both splice sites and polyadenylation sites. This plasmid contains the mutation in L1 that enables the use of the splice site at position 5639. Numbers represent nucleotide positions for both splice sites and polyadenylation sites. Blots were probed with an L1 probe. (B) Northern blot representing cytoplasmic RNA extracted from HeLa cells transfected with pMT1SD in the presence or absence of CMV-E2.

A



B

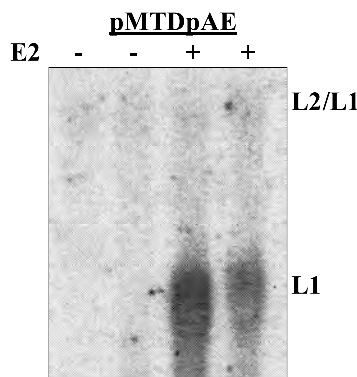


Figure 12. Late gene expression from pMTDpAE. (A) Plasmids pBSpliceM and pMTDpAE. The early poly (A) signal seen in pBSpliceM is deleted in pMTDpAE. The filled circle represent splice donor and the empty circle represent splice acceptor. Numbers represent nucleotide positions for both splice sites and polyadenylation sites. Blots were probed with an L1 probe. (B) Northern blot representing cytoplasmic RNA extracted from HeLa cells transfected with pMTDpAE in the presence or absence of CMV-E2.

Both the pMT1SD and the pMTDpAE plasmids showed strong expression of L1 mRNA in the presence of CMV-E2, but no expression in the absence of CMV-E2, figure 11 and 12.

4 Discussion

Two main questions were to be answered in this project. First, could E2 and ADAR be involved in the induction of late gene expression in HPV-16 and second, with what part of the HPV-16 genome could E2 be interfering?

4.1 E2 and ADAR regulate late gene expression in HPV-16

Based on the results in this study it is clear that both E2 and ADAR individually induce late gene expression in human papillomavirus type 16. What is interesting is that the result is the same for both of them, even though one is a regulatory protein while the other is an RNA editing protein.

4.2 The role of E2

pBSplice and pBSpliceM blots showed that E2 induces late gene expression despite the deletion of all splice sites except for SD3632 and SA5639. Deleting the early poly (A) site and replacing it with a late poly (A) site also resulted in induction of late expression by E2, which could be seen in the results from the p2xpAL blot. Results with pMT22, however, were different. Somberg *et al*. (2008) had found that pMT22 expressed late genes independent of the polypyrimidine tract binding protein (PTB). One would assume, therefore, that the E2 negative blot should show late expression, but that was not the case. The results with pMT1 and pMT3 did not make things more clear. The splice site inhibitors are present in pMT1 and pMT3, but E2 did not have any effect on these plasmids. With pMT2, however, the result was a bit more straight forward. The deletion in this plasmid removes the splice site inhibitors and the late genes were expressed both in the absence and the presence of E2. In the Northern blot representing the plasmid called pMT198, which has a deletion of 198 nucleotides, one would have expected expression of L1 mRNA both in the absence and presence of E2. Somberg *et al.* (2008) had found that this 198 nucleotides long sequence seemed to include splice site inhibitors. Without splice site inhibitors, the late genes should be expressed. However, in this study the blot showed L1 mRNA expression only in the presence of E2. Completely deleting the early polyadenylation site showed the same result, L1 mRNA expression in the presence of E2. pMT1SD has a deletion that reaches to the SD3632 site, and also this plasmid had strong mRNA expression of L1.

It is hard to make a conclusion based on these results. E2 does not seem to work in the same way as does PTB. It is important to remember, however, that when a persistent infection is established, the virus can integrate its genome into the host genome, whereupon E2 is often disrupted. Overexpression of E2 clearly affected late gene expression from different plasmids. One explanation could be that E2 actually works somewhat like PTB, by downregulating splice site inhibitors, but not as efficiently. This would explain the lack of late gene expression in the case of pMT1 and pMT3 in the presence of E2.

Many things can affect the outcome of a Northern blot, and it takes a while to get significant results. RNA is very easily degraded and in the somewhat long procedure of extracting the cytoplasmic RNA and blotting it, the risk of contamination is always present.

4.3 Future perspective

This study shows that E2 and ADAR induce late gene expression, but not how they do it. It would be interesting to investigate further how they affect the genome of HPV-16. The function of ADAR was not at all investigated and a start would be to test ADAR on the same plasmids that were used for E2. However, since they are completely different proteins this

might not at all tell us more about the function of ADAR, but it would be a start. Also for E2, new experiments have to be carried out since the current ones have yielded somewhat contradictory information. It would be of interest to see if E2 from low risk strains of HPV has the same effect as E2 from HPV-16. For this study, the two plasmids constructed in this project, HPV-16 E2 and HPV-5 E2 could be used.

5 Materials and methods

5.1 Cell line, bacterial strain and plasmids

HeLa is a cell line isolated from cervical cancer patient Henrietta Lacks in 1951 which doubles in only 24 hours under appropriate conditions. These cells carry the oncogenes E6 and E7 that allow uninhibited cell growth (Kiessling *et al* 1993). HeLa cells were grown on 10 cm Ø plates (Sarstedt) in DMEM + GlutaMAX (+ 4.5 g/L glucose-pyruvate. GIBCO) + 10% foetal bovine serum (Gibco BRL) and 1% penicillin + streptomycin at 37º C and 7 % CO_2. *E. coli* DH5α was used for plasmid transformation. DH5α is a strain that by the company is made chemically competent to take up DNA efficiently. Strain specific mutations in the *endA1* and *hsdR17* genes prevent degradation of foreign DNA by intracellular endonucleases. Also, mutation of the *recA* gene nullifies homologous recombination which causes slow growth, but preserves the integrity of plasmid DNA (Taylor *et al* 1993). All the plasmids in this project are presented in table 1.

5.2 Plasmids used or constructed in this project

Plasmids were generated in two different ways, table 2. One way was to amplify a fragment of interest with PCR and then insert it into a vector. The second way was to simply digest a plasmid, and thereby delete a part of the genome, and then re-ligate the ends.

For plasmid constructs pMT31 and pMT198, PCR fragments were amplified from pBEL and introduced into vector plasmid, pMT2. Vector and fragment were cleaved with XbaI and MluI for pMT31 and with ApaI and MluI for pMT198. For plasmid constructs HPV-5 E2 and HPV-16 E2 PCR fragments were amplified from HPV-5 and pFL-16 respectively, and inserted into plasmid vector PC16L1. Vector and fragment were cleaved with BssHII and XhoI. For the generation of pMTpAEmut, a PCR fragment was amplified from pBEL and inserted into pMT32, both cleaved with ApaI and MluI.

Ligations were performed for one hour at room temperature by combining PCR fragment with vector at a ratio of 5:2 and then adding 1U of T4 DNA ligase (Fermentas) and 1xT4 DNA ligase buffer to the mixture.

Plasmid pMTDpAE was constructed by digesting pMT3 with restriction enzymes ApaI and MluI and pMT1SD was constructed by digesting pT1SD with SalI and BssHII, table 2. Overhangs of restriction product were then "filled in" using 1U Klenow polymerase (Fermentas) in 1x reaction buffer, and the ends were then ligated by adding 1U of T4 DNA ligase (Fermentas) and 1xT4 DNA ligase buffer. The ligation mixture was incubated at room temperature for 1-2 h.

Table 1. Plasmids used (for transfections) or constructed (underlined) in this project.

Name	**Properties**	**Source**
pl0806	Empty vector DNA for negative control	Collier *et al* 2002
CMV-E2	A plasmid encoding E2 under the control of a CMV promoter, used for all E2 positive transfections in this project	With thanks to David Pin for providing us with this plasmid
ADAR	A plasmid encoding ADAR 1	-
pBEL	E6, E7, p97 and p670 replaced with a CMV promoter	Zhao *et al.* 2004
pBELM	Properties of pBEL, but contains a mutation in gene L1 that allows for usage of the splice site SA5639	Zhao *et al.* 2004
pBSplice	All splice sites except for SD3632 and SA5639 are deleted	Zhao *et al.* 2004
pBSpliceM	Properties of pBSplice, but with the mutation in L1	Zhao *et al.* 2004
p2xpAL	Has the early poly (A) signal replaced with the late poly (A) signal	Somberg *et al.* 2008
pMT22	The whole early region is deleted, including SD3632	Somberg *et al.* 2008
pMT1	Properties of pBSpliceM, with a deletion between nt 3395-4212	Somberg *et al.* 2008
pMT2	Properties of pBSpliceM, with a deletion between nt 3653-4212	Somberg *et al.* 2008
pMT3	Properties of pBSpliceM, with a deletion between nt 3850-4212	Somberg *et al.* 2008
pMT31	Properties of pBSpliceM, with a deletion between nt 3700-4212	This investigation
pMT32	Properties of pBSpliceM, with a deletion between nt 3750-4212	Unpublished data, Somberg
pMT33	Properties of pBSpliceM, with a deletion between nt 3800-4212	Unpublished data, Somberg
pMT198	Properties of pBSpliceM, with a deletion between nt 3652-3850	This investigation
pMT1SD	Properties of pBSpliceM, but has the whole region until SD3632 deleted	This investigation
pMTDpAE	Has the early poly (A) signal deleted	This investigation
HPV-5 E2	Expresses HPV-5 E2 in a vector plasmid called PC16L1	This investigation
HPV-16 E2	Expresses HPV-16 E2 in a vector plasmid called PC16L1	This investigation
pMTpAEmut	Has a mutation in the early poly (A) site	This investigation
PC16L1	Vector for HPV-5 E2 and HPV-16 E2	Collier *et al* 2002

5.3 Polymerase chain reaction

5.3.1 Amplifying a fragment for plasmid constructions

For a few of the new plasmids to be constructed, an insert was made by PCR amplification. The PCR reaction was made in a final volume of 100 µl (3.75 mM $MgCl_2$ (Fermentas), 1 µM dNTP´s (Amershan Biosciences, Piscataway, NJ, US), 500 nM forward and reverse primers (table 3), 1xTaq-buffer (+KCl, -$MgCl_2$) (Fermentas) and 5U Taq DNA polymerase (Fermentas)). Cycling conditions are shown in table 2.

Table 2. Overview of plasmid constructions.

Construct name	Vector	Primers used for amplification	PCR template	PCR program	Restriction enzymes
pMT31	pMT2	3455s ,M23A	pBEL	30 cycles of (94ºC 1 min, 55ºC 1 min, 72ºC 2 min)	MluI, XbaI
pMT198	pMT2	M14S, M17A	pBEL	29 cycles of (94ºC 1 min , 50ºC 1 min, 72ºC 2.5 min)	ApaI, MluI
HPV-5 E2	PC16L1	5E2S, 5E2A	HPV-5	30 cycles of (94ºC min, 50ºC 1 min, 72ºC 2.5 min)	BssHII, XhoI
HPV-16 E2	PC16L1	16E2S, 16E2A	pFL-16	30 cycles of (94ºC 1 min, 50º 1 min, 72ºC 2 min)	BssHII, XhoI
pMTpAEmut	pMT32	K3s, M17A	pBEL	29 cycles of (94ºC 1 min, 50ºC 1 min , 72ºC 2.5 min)	ApaI, MluI
pMTDpAE	pMT3	-	-		ApaI, MluI
pMT1SD	pT1SD	-	-		SalI, BssHII (PauI)

Table 3. Primers and primer sequences

Primer name	Sequence
3455s	GCGCGCTCTAGACCAACCACCCCGCCGCGACC
M23A	GACGCGTCACTGCAGTATACAATGTCAAATGCTT
M14S	GGACGCGTATGACAAATCYYGATACTGCAT
M17A	CACTAAAGAAACTATAGAAGGATCAG
5E2S	GGCGCGCACCATGGAGAATCTCAGCGAGCGTTTCA
5E2A	CCTCGAGTTAAAGACTGTCCAGGTTGCCATAG
16E2S	GGCGCGCACCATGGAGACTCTTTGCCAACGTTTAA
16E2A	CCTCGAGTCATATAGACATAAATCCAGTAGAC
K3s	ACGCGTGGTACCCTGTTATTACTTAACAATGCG

5.3.2 Reverse transcription polymerase chain reaction

A sample with the final volume of 12 µl was prepared (200 ng RNA, 100 ng Random Hexamer (GE Healthcare) and 1 µl dNTP´s (Amershan Biosciences, Piscataway, NJ, US) and incubated at 65º C for 5 min. After incubation, RT (reverse transcriptase) and buffers were added to a final volume of 19 µl (1 µl RT (Invitrogen), 2 µl 0.1M DTT and 4 µl 5 x First strand buffer). This was then incubated at 37º C for 50 min and then at 70º C for 15 min, followed by 30 cycles of 94º C 2 min, 55º C 1 min, 72º C 1,5 min.

5.4 Gel electrophoresis

5.4.1 Polyacrylamide gel

The RT-PCR products was run on a 5% acrylamide gels (5.7 ml PAGE 29:1 (40 %) (Saveen Werner AB), 3.6 ml ddH_2O, 4.5 ml 5 X TBE (for a final volume of 5 L: 270 g Tris, 137.5 g boric acid and 100 ml 0.5M EDTA), 450 µl 10 % APS (adenosine phosphosulfate) (Merck) and 30 µl TEMED (Amershan Biosciences)).

5.4.2 Agarose gel

- DNA

When constructing new plasmids, both the digested vector and the PCR fragment were run on a gel. 1.4 µl EtBr (1 % EtBr in H_2O) was added to 30 ml agarose gel (1 or 2 % agarose depending on band size) and poured into a tray with a comb and let set. The gel was run at 130 V in 5 X TAE (TRIS-acetate-EDTA) buffer (for a final volume of 2 L of 50 X TAE: 484 g TRIS, 114.2 ml HAc and 200 ml 0.5 M EDTA pH 8 (pH adjusted with NaOH pellets)). A high range or a low range marker (Fermentas) was used to determine the size of the different bands. The band was cut out from the gel. The product was extracted with NucleoSpin™ Extract II kit (Macherey-Nagel).

- RNA

The gel equipment was carefully washed, sprayed with 70 % EtOH and allowed to air dry. 1.4 µl EtBr (1 % EtBr in H_2O) was added to 30 ml agarose gel (1 % agarose) and poured into a tray with a comb and let set. 1 µl RNA sample was mixed with 4 µl dH_2O and 3 µl loading dye (for a final volume of 25 ml: 3.75 g Ficoll 400, 0.06 g bromophenolblue, 0.06 g xylene cyanol). The gel was run at 130 V for 20 min in 5 X TAE buffer (for a final volume of 2 L: 484 g TRIS, 114.2 ml HAc and 200 ml 0.5 M EDTA pH 8 (pH adjusted with NaOH pellets)).

5.5 Transfection

5.5.1 Cell culture

Cells were split one day prior to transfection with plasmids. The medium was removed and the cells were washed twice with 1 X PBS (phosphate buffered saline) (10 x stock solution with a final volume of 2 L = 160 g NaCl, 4 g KCl, 4.8 g KH_2PO_4, 35.6 g Na_2HPO_4 x $2H_2O$). 1 ml of trypsin (Gibco) was added to detach the cells from the plate and the plate was incubated 5-10 min, at 37º C. During incubation, new plates were prepared containing 6 ml of medium (10 cm Ø plates, Sarstedt), or 2 ml of medium (2 cm Ø plates, Sarstedt). 10 ml medium was added to the incubated plate and the detached cells were thoroughly resuspended. To each new plate containing medium, 1 ml of cell suspension was added. The plates were then incubated over night at 37º C.

5.5.2 Transfection with plasmids

The confluency of the cells split the previous day was checked using a light microscope. In order to get efficient transfection, 50-60 % confluence of cells was required.

The DNA-concentration of each plasmid was measured by OD 260. 1 µg of plasmid was used for each transfection. In the same tubes as the plasmids of interest either 1 µg positive control (plasmids CMV-E2 or ADAR) or 1 µg negative control (empty vector DNA from pl0806) were added for co-transfection. A mastermix of serum free medium, 200 µl for each transfection (1 X DMEM + GlutaMAX, GIBCO) and Fugene, 3 µl Fugene/µg plasmid DNA, i.e. 6 µl Fugene for each transfection (Fugene 6 Transfection Reagent, Roche) was prepared and incubated at

room temperature for 5 min. 200 µl of this mastermix was then added to each sample and this was incubated at room temperature 15-30 min. Each sample was dropwise added to a plate with confluent cells and this was then incubated at 37º C and 7 % CO_2 for 20-24 h.

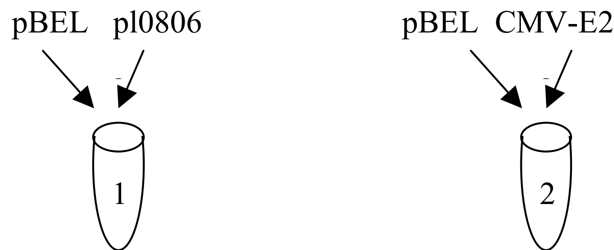


Figure 13. An example of transfection in the absence or presence of E2. Eppendorf tube number 1 shows a transfection with pBEL in the absence of E2. pBEL is instead co-transfected with the empty vector DNA from pl0806. Eppendorf tube number 2 shows a transfection in the presence of E2, where pBEL is co-transfected with the E2 encoding plasmid CMV-E2.

5.6 Extraction of cytoplasmic ribonucleic acid

Plates with transfected cells were kept on ice. The medium was discarded and the plates were washed twice with 2 ml 1 X PBS each. To disrupt the cell membrane, 600 µl of ISO-B and Np40 (Nonidet P40) were added (for a final volume of 50 ml ISOB: 0.5 ml Tris pH 9 (adjusted pH with HCl), 1.5 ml NaCl and 0.075 ml $MgCl_2$) (65 µl of 10 % Np40 /ml ISOB). The plates were then left on ice 5-10 min with occasional shaking. While the plates were incubating on ice, new, sterile eppendorf tubes were prepared. The cells were collected using a cell scraper (Sarstedt). The cells were then pipetted into the previously prepared sterile eppendorf tubes. The cell scraper was rinsed in 1 X PBS in between collection of cells from each plate. The cell suspensions were kept on ice and then centrifuged at 4º C, 16 000 x g, for 3 min. New sterile tubes were prepared containing 150 µl 5 X RPS (0.5M Tris, 50mM EDTA, 2.5 % SDS, ddH_2O). After centrifugation, the supernatant was transferred into these tubes. The tubes were vortexed and left on ice. 650 µl phenol (AppliChem, Germany)-chloroform (Merck), ratio of 1:1, was added to each tube to remove possible cytoplasmic proteins. The tubes were then vortexed thoroughly and centrifuged at 4º C, 16 000 x g, for 3 min. The upper RNA containing phase was then transferred to new tubes containing 650 µl phenol-chloroform. This was vortexed thoroughly and centrifuged at 4º C, 16 000 x g, for 3 min. This step was then repeated once. After the last centrifugation the RNA phase was transferred to new tubes containing 830 µl iso-propanol (2-Propanol, Merck) and 30 µl 5M NaCl (Merck). The tubes were vortexed and stored in freezer, at -20º C, for 30 min or over night.

Samples from freezer were centrifuged at 4º C, 16 000 x g, for 30 min. A master mix of buffer containing 90 µl dH_2O, 9 µl DNase-buffer (for a final volume of 10 ml of 10 X DNase buffer: 5 ml Tris, 200 µl EDTA, 1 ml $MgCl_2$ and 100 µl DTT) and 1 µl DNase I (FPL Cpture™, 10.000 u/ml. GE-Healthcare) for each sample was prepared. After centrifugation, the supernatant was discarded and the pellet was re-suspended in 100 µl of the master mix and incubated at 37º C for 15 min. 100 µl phenol-chloroform was added to each sample that was then thoroughly vortexed and centrifuged at 4º C, 16 000 x g, for 3 min. The upper phase was then transferred to new tubes containing 10 µl 3M NaAc and 250 µl 99.7 % EtOH (Etax Aa, Solveco). Samples were then stored in freezer, at -20º C, for at least 20 min or over night. After incubation the samples were centrifuged at 4º C, 16 000 x g, for 20 min and supernatant was removed. The pellet was allowed to air dry and then, dependent on size, re-suspended in 10-20

µl RNAse-free H_2O. The re-suspension was kept on ice from this step or stored at -20º C until use.

5.7 Northern Blot analysis

5.7.1 Northern blots

All DNAse-treated samples were thawed on ice. One litre of 1 X MOPS buffer (3-[N-Morpholino]propanesulfonic Acid) (10 x MOPS buffer = 0.4M MOPS ($(C_4H_8NO)(CH_2)_3SO_3H$), 0.1M NaAc, 0.01M EDTA, UBS Corporation, Cleveland, Ohio, USA) was prepared. A 1.2 % agarose gel was made that included 1 X MOPS buffer and 2.2 mM formaldehyde (Merck). The gel was pre-run for 5 minutes in 1 X MOPS buffer. RNA-concentration of the samples was measured by OD_{260} and 5 µg of RNA from each sample, was put into tubes containing 4.5 µl formaldehyde (Merck), 12.5 µl formamide (50 % formamide, Merck) and 2.5 µl 10 X MOPS buffer. The tubes were incubated at 60º C, 15 min. 5 µl of loading dye was added and the gel was run at 85 V for 3-4 h. To transfer the RNA to a nitrocellulose membrane (Schleicher & Schuell Bioscience), the gel was put onto paper soaked in 20 X SSC (saline sodium citrate) (140 g NaCl, 70 g Na_3-citrate and 800 ml ddH_2O). The membrane was placed onto gel and covered with three pieces of paper soaked in 20 X SSC (saline sodium citrate). Paper towels and a heavy weight were put on top and this was incubated at room temperature over night to let the RNA-product transfer to the membrane.

After the blot, the membrane was baked at 80º C in complete vacuum for 1.5 h or more. The membrane was then stored at -20º C, until use.

5.7.2 Hybridization

The membrane was soaked in 6 X SSC (saline sodium citrate) and placed in a thermo tube, with the RNA staining facing inwards in ~3 ml hybridization buffer (5 x SSC, 5 x Denhardt´s (100 x Denhardt´s with a final volume of 50 ml: 1 g Ficoll, 1 g polyvinylpyrrolidone, 1 g BSA (bovine serum albumin)) , 10 % dextransulphate (Merck), 50 % formamide (Merck), 0.1 % Na-pyrophosphate ($Na_4P_2O_4$ x 10 H_2O) and 1 % SDS (sodium dodecyl sulphate) (Merck) and about 3 ml of this was then added to each membrane (1 ml/10 cm^2). This was incubated for 1 h or more at 42º C while rotating.

5.7.3 Radiolabelling of probe

A mixture of 4 µl DNA template, 6 µl ddH_2O and 2.5 µl decamer (10 x Decamers, DECAprime TM II, Ambion) was prepared on ice and then incubated at 95º C for 5 min. The decamer reaction was then put on dry ice and 5 µl of (-) dCTP-reaction buffer (DECAprime tm II, Ambion) and 1.5 µl ddH_2O was added. In the radioactivity lab, 1 µl Klenow DNA-polymerase and 5 µl isotope [α^{32} P] dCTP (Amersham Biosciences) was added to the reaction and incubated at 37º C for 8 min. 1 µl EDTA and 25 µl ddH_2O was also added. A MicroSpin TM G-50 Column (GE Healthcare) was centrifuged prior to use for 1 min at 3400 x g. The 50 µl probe was then added to the centre of the column, which was centrifuged for an additional 1 min at 3400 x g. Before usage, the probe was incubated for 10 min at 95º C, and then put on ice.
The probe was added to the thermo tube containing the membrane and the hybridization buffer. This was then incubated over night at 42º C while rotating.

Upon concluded hybridization the membrane was at first washed twice with 2 X SSC, 0.1 % SDS for 5 min at room temp. After this it was washed 2 X 5 min with 0.2 X SSC 0.1 % SDS at room temp, followed by 2 X 15 min with 0.2 X SSC 0.1 % SDS and 2 X 15 min with 0.1 X SSC 0.1 % SDS at 42º C

When the radioactivity was low throughout the membrane except for where the probe had bound, the membrane was placed in a cassette and covered with an autoradiography film. The film was then, after a few hours, developed.

6 Acknowledgements
I would like to thank Stefan Schwartz and Monika Somberg for excellent supervision during this project work. I would also like to thank Anna Tranell, Sofia Lindström, Samir Abdurahman, Fiona Cassidy, Xioaze Li and the group members of Göran Akusjärvi for nice collaboration. Further I would like to thank my wonderful family and relatives for all the support and care I have received all these years. At last but not least I would like to thank Daniel Hagey for always being there when needed and also my sweet friends, whom made this education to what it was.

7 References

Adams, M., Jasani, B. & Fiander, A. 2007. Human papilloma virus (HPV) prophylactic vaccination: Challenges for public health and implications for screening. *Vaccine* 25:3007-3013.

Ault, K.A. 2006. Epidemiology and natural history of human papillomavirus infections in the female genital tract. *Infectious diseases in obstetrics and gynaecology* Suppl:40470.

Blumenthal, P.D. & Gaffikin, L. 2005. Cervical cancer prevention: Making programmes more appropriate and pragmatic. *Journal of the American medical association* 294:2225-2228.

Collier, B., Öberg, D., Zhao, X. & Schwartz, S.M. 2002. Specific inactivation of inhibitory sequences in 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. *Journal of virology* 76:2739-2752.

Doorbar, J. 2005. The papillomavirus life cycle. *Journal of clinical virology* 32 Suppl 1:S7-15.

Doorbar, J. 2006. Molecular biology of human papillomavirus infection and cervical cancer. *Clinical science* 110:525-41.

Gammoh, N., Sterlinko Grm, H., Massimi, P. & Banks, L. 2006. Regulation of human papillomavirus type 16 E7 activity through direct protein interaction with the E2 transcriptional activator. *Journal of virology* 80:1787-1797.

Gu, R. Zhang, Z., Carmichael, G.G. 2006. How a small DNA virus uses dsRNA but not RNAi to regulate its lifecycle. *Cold spring harbor laboratory press* 71:293-9.

Hernandez-Ramon, E.E., Burns, J.E., Zhang, W., Walker, HF., Allen, S., Antson, A.A., Maitland, N.J. 2008. Dimerization of the human papillomavirus type 16 E2 N terminus results in DNA looping within the upstream regulatory region. *Journal of virology* 82:4853-61.

Javier F. Cáceres and Alberto R. Kornblihtt. 2002. Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends in genetics* 4:186-193.

Kiessling, U., Kube, D., Janda, J., Bergmann, S., Johannsmeyer, K.D., Hauff, P., Arnoldt, W. & Waehlte, H. 1993. Properties of two epithelial cells lines derived from HPV-associated cervical and vulvar lesions. *European journal of cancer* 29A:1746-53.

Lodish, H., Berk, A., Matsudaira, P., Kaiser,. C.A., Krieger, M., Scott, M.K., Zipursky, S.L., Darnell, J. 2004. *Molecular cell biology* 5:th edition. W. H. Freeman and company, New York.

Milligan, S.G., Veerapraditsin, T., Ahamat, B., Mole, S. and Graham, S.V. 2007. Analysis of novel human papillomavirus type 16 late mRNAs in differentiated W12 cervical epithelial cells. *Virology* 360**:**172–181.

Motoyama, S., Ladines-Llave, C.A., Luis Villanueva, S., Maruo, T. 2004. The role of human papilloma virus in the molecular biology of cervical carcinogenesis. *The Kobe journal of medical sciences* 50:9-19.

Murray, P.R., Rosenthal, K.S. & Pfaller, M.A. 2005. *Medical microbiology* 5:th edition. Elsevier Mosby, Philadelphia.

Oberg, D., Fay, J., Lambkin, H. & Schwartz, S. 2005. A downstream polyadenylation element in human papillomavirus type 16 L2 encodes multiple GGG motifs and interacts with hnRNP H. *Journal of virology* 79: 9254-69.

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

Schiffman, M., Castle, P.E., Jeronimo, J., Rodriguez, A.C. & Wacholder, S. 2007. Human papillomavirus and cervical cancer. *Lancet* 370:890-907.

Somberg, M., Zhao, X., Fröhlich, M., Evander, M., Schwartz, S. 2008. Polypyrimidine tract binding protein induces human papillomavirus type 16 late gene expression by interfering with splicing inhibitory elements at the major late 5' splice site, SD3632. *Journal of virology* 82:3665-78.

Stone, K.M., Karem, K.L., Sternberg, M.R., McQuillan, G.M., Poon, A.D., Unger, E.R., Reeves, W.C. 2002. Seroprevalence of human papillomavirus type 16 infection in the United States. *The journal of infectious diseases* 186:1396-402.

Taylor, R.G., Walker, D.C. & McInnes, R.R. 1993. E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic acids research* 21:1677-1678.

Terhune, S.S., Milcarek, C. & Laimins, L.A. 1999. Regulation of human papillomavirus type 31 polyadenylation during the differentiation-dependent life cycle. *Journal of virology* 73:7185-92.

Tindle, R.W. 2002. Immune evasion in human papillomavirus-associated cervical cancer. *Nature reviews cancer* 2:1-7.

White, P.W., Titolo, S., Brault, K., Thauvette, L., Pelletier, A., Welchner, E., Bourgon, L., Doyon, L., Ogilvie, W.W., Yoakim, C., Cordingley M.G. & Archambault, J. 2003. Inhibition of human papillomavirus DNA replication by small molecule antagonists of the E1-E2 protein interaction. *The journal of biological chemistry* 278:26765-26772.

Wilson, V.G., West, M., Woytek, K. & Rangasamy, D. 2002. Papillomavirus E1 proteins: form, function, and features. *Virus genes* 24:275-90.

Zhao, X. & Schwartz, S. 2008. Inhibition of HPV-16 L1 expression from L1 cDNAs correlates with the presence of hnRNP A1 binding sites in the L1 coding region. *Virus genes* 36:45-53.

Zheng, Z-M. & Baker, C.C. 2006. Papillomavirus genome structure, expression, and post-transcriptional regulation. *Frontiers in bioscience* 11:2286-2302.