EPSTEIN-BARR VIRUS EPISOME REPLICATION AND TRANSCRIPTION

AW YONG KOH MENG

NATIONAL UNIVERSITY OF SINGAPORE

2006

EPSTEIN-BARR VIRUS EPISOME REPLICATION AND TRANSCRIPTION

AW YONG KOH MENG
B.Sc. (Hons), NUS

A THESIS SUBMITTED

FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MICROBIOLOGY

NATIONAL UNIVERSITY OF SINGAPORE

Acknowledgements

I would like to take this opportunity to thank my supervisor, Dr Hung Siu Chun for guiding me all this time, even though he is no longer in NUS. I am grateful for all that he has taught me as well as for all the support that he has shown.

I would also like to thank A/Prof. Mary Ng for agreeing to my co-supervisor and Prof. Chan for allowing me to stay in the WHO Immunological Centre to finish my work.

Special thanks to Shyue Wei and Gayathri; you guys have been wonderful.

A big thank you to my wife Yee Sun who has been giving me her unwavering support through this difficult time; you are the best! Not forgetting my family and friends who have all stood behind me supporting me all the way.

Table of contents

List of figures

List of tables

Abbreviations

ACE	Amplification control element
AER	Amplification enhancing region
ARS	Autonomously replicating sequence
bp	Base pairs
CMV	Cytomegalovirus
DHFR	Dihydrofolate reductase
DS	Dyad symmetry
EBNA-1	Epstein-Barr nuclear antigen-1
EBV	Epstein-Barr virus
FR	Family of repeats
kb	Kilobase pairs
LB	Luria Bertani
MCM	Minichromosome maintenance
NTPs	Nucleotide triphosphates
ORC	Origin recognition complex
PCR	Polymerase chain reaction
Pre-RC	Pre-replicative complex
RPA	Replication protein A
SV40	Simian virus 40
TAF	TBP associated factor
TBP	Transcription binding factor
TEC	Transcription elongation complex
UV	Ultra-violet

α	Alpha
β	Beta
γ	Gamma
κ	Kappa
λ	Lambda
ω	Omega

Summary

The relationship between transcription and replication has not been fully understood. In this study I aim to understand more about this relationship by making use of the EBV latent origin of replication *ori*P. *Ori*P is able to initiate replication in the presence of the EBV protein EBNA-1 by recruiting the cellular replication machinery, therefore making it a suitable candidate for this study. Firstly, *ori*P was inserted onto a vector and placed under the transcriptional effect of a promoter at various locations within and without a transcriptional unit. The vectors were transfected into EBNA-1 expressing cells and the replication assayed using Southern blot. In addition, total RNA was also extracted and analyzed using Northern blot. Southern blot results indicated that the presence of a promoter upstream of *ori*P displayed the strongest replication inhibition. Interestingly, Northern blot results indicated that there was a lack of *ori*P containing transcripts both in the presence and absence of EBNA-1, suggesting that *ori*P could have an inhibitory effect on transcription. In an attempt to confirm the inhibitory effect of *ori*P on transcription, *in vitro* transcription was performed, and results obtained were similar to those obtained *in vivo*. There were a few possible explanations for these observations. One of which was that transcription arrest occurred as the transcriptional machinery read through *ori*P. This state of transcriptional arrest would explain for the lack of *ori*P containing transcripts and at the same time; the physical stalling of the machinery along the template could have inhibited replication by preventing the replication initiation machinery from assembling on *ori*P. To test this possibility, we used size exclusion chromatography of *in vitro* transcription reactions to differentiate between arrested transcripts trapped with RNA polymerase and free transcripts. The results were consistent with the dissociation of transcription elongation complex at *ori*P. Thus, the hypothesis of

transcriptional arrest was not supported and the mechanism by which transcription inhibits replication remains uncertain.

Introduction

1. Introduction

The interplay between DNA replication and transcription has long been a focal point of debate between researchers. Various works done by different groups have thrown up different observations and conclusions, each of them seemingly contradicting the other. While some claim that transcription inhibits replication, others propose that transcription is necessary for replication. In this study we attempt to cast this relationship in better light by employing the use of a known viral origin of replication, the *ori*P, from the DNA herpesvirus, Epstein Barr Virus (EBV). But before we talk about the relationship between these two cellular processes, we should first look at them individually and examine what are some of the basic mechanisms that govern and control them.

The bacterial DNA replication system was one of the pioneer models that contributed to our understanding of DNA replication (Jacob *et al.* 1963). In this model, it was proposed that two elements were required for the initiation of DNA replication: a replication initiator protein and a *cis*-acting DNA element. Only when the replication initiator protein binds to the *cis*-acting DNA element can DNA replication initiate and proceed. Further work employing the use of the bacterial chromosome finally elucidated this initiator protein to be a DNA binding protein called DnaA. Multiple binding sites for DnaA could be found on the *cis*-acting DNA element, identified as an origin of replication, *ori*C (Baker and Bell, 1998; Kornberg and Baker, 1992).

So far, DNA replication system in simple eukaryotes such as *Saccharomyces cerevisiae* seems to bear some similar characteristics to the prokaryotic model in the sense that both require an initiator and a *cis*-acting DNA element. Both utilize a DNA

polymerase to synthesize the new strand of DNA from the template strand by adding nucleotides to the 3'-OH end. But that is where the similarities probably end. The complexity and the size of the eukaryotic genome far surpasses that of prokaryotes and the initiator protein consists of different proteins arranged in a complex, called the origin recognition complex (ORC) (Bell, 2002). And the difference is even more obvious in higher order eukaryotes. To solve the problem of having to replicate such a large genome, these eukaryotic organisms are also able to initiate DNA synthesis at multiple sites (Huberman and Riggs, 1968). Another major difference lies in the fact that DNA replication only occurs exclusively in the S phase of these higher order eukaryotic cell cycle as compared to prokaryotes, which occur throughout the bacterial life cycle. This would also mean that DNA replication in eukaryotes is more tightly regulated.

Epstein-Barr Virus (EBV) is a DNA virus that exists as an extrachromosomal episome during its latent stage of infection. Its ability to persist in infected cells latently can be attributed to two viral components, the latent origin of replication *ori*P and the viral nuclear antigen, EBV Nuclear Antigen-1 (EBNA-1). *Ori*P contains multiple EBNA-1 binding sites, and with the help of bound EBNA-1, help recruit cellular proteins necessary for DNA replication. This enables the virus to replicate its genome together with the cell during the S phase, resulting in one copy of each viral replicon being produced per cell cycle (Kieff and Rickinson, 2001). In fact, this particular characteristic of the EBV has been employed in the construction of episomal gene therapy vectors. Episomal gene therapy vectors based on EBV already show great promise in treating disorders like Duchenne Muscular Dystrophy (Tsukamoto *et al.*, 1999). But for such vectors to be successful, efficient replication of the vector and the

expression of the therapeutic gene that is carried on the vector must occur. Therefore, a better understanding of the relationship between replication and transcription of an *ori*P-dependent episome is critical for the optimal design of such vectors. Therefore, the issue addressed in this work is of practical significance besides being an interesting subject in basic molecular biology and virology.

Transcription in prokaryotes occurs in three main steps: initiation, elongation and termination. Initiation first occurs when the RNA polymerase complex binds to the promoter region. Elongation proceeds soon after with the RNA polymerase moving along the DNA template, adding ribonucleotides to the 3'-OH end of the forming RNA transcript and termination occurs when the RNA polymerase meets terminator sequences or when a termination signal protein binds to the RNA polymerase. In prokaryotes, one form of RNA polymerase is apparently responsible for the synthesis of all RNA.

In eukaryotic cells, transcription also requires the three main steps as described for the prokaryotic system. However, it was further described that the initiation stage can be further broken down into another three distinct phases, namely the preinitiation, initiation and promoter clearance (Sims *et al*. 2006). Differences between prokaryotic and eukaryotic transcription systems also include the fact that the eukaryotic system possesses three different RNA polymerases instead of one. They are RNA polymerase I, RNA polymerase II and RNA polymerase III, each of which transcribe a different set of genes. Furthermore, post-transcriptional processing of the RNA transcript such as alternative splicing in eukaryotic cells allow for a larger repertoire of proteins to be synthesized from a single strand of mRNA.

As both transcription and replication utilize the same source of DNA template, there is the inevitable question of whether any conflicts between these two processes could arise. Indeed there have been works describing the physical collision between the RNA polymerase and DNA polymerase in bacterial cells (Brewer, 1988). Haase *et al*. (1994) have also shown that transcription through a known mammalian origin of replication inhibits the ability of the plasmid carrying the origin to replicate. A similar observation can also be extended to another eukaryotic organism, *Saccharomyces cerevisiae*, where transcription through the autonomously replicating sequences (ARS) inhibits the ability the ARS's ability to activate replication (Tanaka *et al*., 1994).

On the other hand, it has been known for some time that regions of chromatin that are transcriptionally active replicate earlier than transcriptionally inactive regions (Stambrook and Flickinger, 1970; Goldman *et al.*, 1984; Taljanidisz *et al*., 1989; Gilbert, 2002). Genome-wide analysis of replication and transcription timing in *Drosophila* have drawn the conclusion that there exits a strong correlation between DNA replication and transcription (Schübeler *et al*., 2002; MacAlpine *et al*., 2004). Analysis studies done on the human genome have drawn very similar conclusions, providing a strong indication that transcription may be essential for replication (Woodfine *et al*. 2004; White *et al*. 2004). In addition, Boucher *et al*. (2004) have shown that transcription was required to ensure the replication and faithful partitioning of plasmids in *Leishmania donavani*.

To aid our study of the relationship between replication and transcription, we decided to utilize the EBV latent origin of replication, *ori*P. In addition to the reason stated

above, *ori*P has also been shown, in the presence of EBNA-1 to recruit cellular replication machinery and this most likely allows the virus to replicate its genome during the latent stage of infection by recruiting the human origin recognition complex to the *ori*P (Chaudhuri *et al*., 2001; Dhar *et al*. 2001; Schepers *et al*. 2001). Therefore it is a suitable candidate for our study. We cloned *ori*P onto a vector containing a SV40 promoter in varying positions from the SV40 and transfected these clones into EBNA-1 expressing cells for a short term replication assay. Southern blot results indicated that vectors with *ori*P immediately downstream of the SV40 promoter showed the most inhibition of replication, regardless of the orientation of the *ori*P. Total RNA was also analyzed to study if transcription through the *ori*P of these clones were affected as well. Interestingly, preliminary Northern blot results also indicated that the clones that exhibited the most replication inhibition also displayed the most inhibition of transcription through the *ori*P.

These preliminary findings seem to indicate that the relationship between replication and transcription is a possibly one of a negative nature. One possible explanation was that for transcription to inhibit replication, the transcription complex would have to prevent the replication complex from assembling on the origin of replication. It could be possible that transcription was arrested at the origin of replication as the transcription complex transverses along the DNA template. This state of arrest would likely cause the complex to be immobilized and prevent the movement of other transcription complexes, resulting in the saturation of transcription complexes on the DNA template and preventing the initiation of replication. Transcriptional arrest would also explain the lack of *ori*P containing transcripts in the *ori*P containing vectors. If this hypothesis is proven to be correct, it could potentially offer a novel

method by which the persistence of episomal gene therapy vectors based on viruses such as the Epstein-Barr virus (EBV) can be regulated.

To determine if inhibition of transcription was due to the physical arrest of the transcription complex on the DNA sequence, we attempted to isolate the complex with the arrested transcript using *in-vitro* transcription and size exclusion chromatography. Preliminary *in-vitro* transcription experiments also indicated that transcription was also inhibited for clones containing *ori*P immediately downstream of the promoter. However, size exclusion chromatography failed to isolate any RNA polymerase-arrested transcript complex, indicating that there is most likely no form of physical arrest of the transcription complex on the DNA template.

The failure to isolate any RNA polymerase-arrested transcript is an indication that termination of transcription rather than arrest most likely occurred as the transcription complex met the *ori*P. If that is the case, our stand that the arrest of the transcription complex along the DNA template prevents the DNA replication complex from recognizing and binding to the *ori*P, thereby inhibiting replication should be re-examined. There are other potential directions that could be explored in the future to help in the further understanding of the interplay between transcription and DNA replication. One of them is chromatin remodeling.

Survey of literature

2. Survey of Literature

2.1 DNA replication

DNA replication is a fundamental process in any living organism. It is also a highly complex procedure with different enzymes being involved at different stages of the process. There are three main stages for replication: initiation, elongation and termination. During initiation of replication, a protein complex first recognizes and binds to a site on the DNA template. The double-stranded parental DNA at that site is then separated into single strands, called the replication fork. Before elongation can occur, priming must first occur by the synthesis of a short RNA primer, a nick in DNA or a small priming protein. Elongation of the daughter DNA strand is carried out by the DNA polymerase bi-directionally and involves the addition of nucleotides to the growing 3'-OH end. This results in a newly synthesized DNA strand being base-paired with the parental strand. This is also called semi-conservative DNA replication. Termination of replication usually occurs when replication is completed. Resolution of the replicated double stranded DNA from one another is required for further partitioning into daughter cells. This section will focus more on the initiation of replication.

2.1.1 Prokaryotic DNA replication

Prokaryotes usually contain only one replicon that exists as a closed circular DNA. One of the most extensively studied replication system in prokaryotes is in *Escherichia coli*, a Gram-negative bacteria.

2.1.1.1 Origin of replication in prokaryotes

Jacob *et al*. (1963) first proposed the idea that for DNA replication in bacteria to proceed, two things are first required: a *trans*-acting initiator protein and a *cis*-acting element at which DNA replication starts. Initiation of replication in *Escherichia coli* occurs at a specific sequence of DNA in the replicon, known as *ori*C. *Ori*C in turn contains multiple binding sites for the initiator protein DnaA, which was arranged as a huge multi-subunit protein complex surrounding *ori*C (Bramhill and Kornberg, 1988; Kornberg and Baker, 1992; Baker and Bell, 1998). The DnaA complex bound to *ori*C serves two purposes in replication: it first helps to unwind the surrounding DNA it is located on and it ultimately recruits helicases to the origin of replication via binding to loading factors. In *Escherichia coli*, the loading factors DnaC are found as a complex with DnaB, with six molecules of DnaC binding with six molecules of DnaB as a multimer. In this form, the bound DnaB is inactive for its helicase activity and only upon the binding of DnaC-DnaB complex to the DnaA, is DnaB released from DnaC in an ATP dependent manner (Wahle *et al*., 1989). DnaB further unwinds the DNA and at the same time activates a primase, DnaG, which synthesizes short RNA primers required for the DNA replication to proceed. A representative diagram of replication initiation for prokaryotes is shown in figure 1 below.

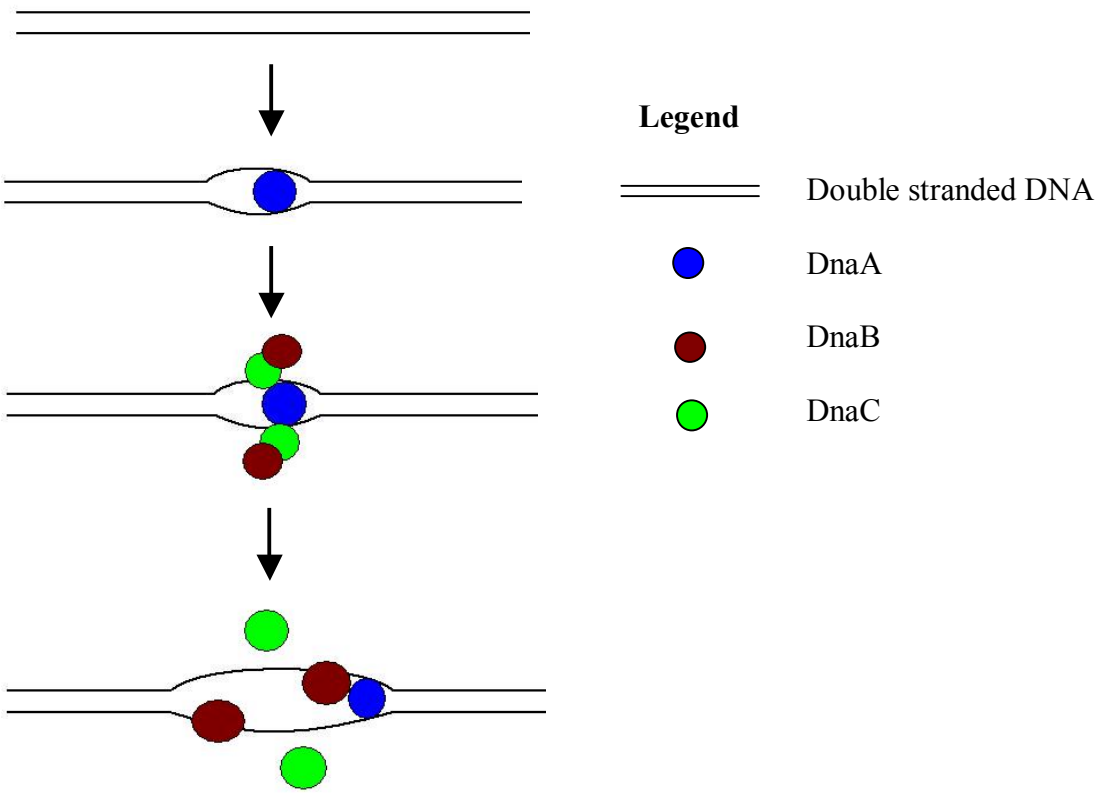


Figure 1. Prokaryotic replication initiation. In prokaryotic replication initiation, DnaA first binds to oriC, resulting in local unwinding of double stranded DNA as symbolized by the bubble. Following that, DnaC and DnaB gets recruited to the origin. DnaC then dissociates from DnaB, leaving DnaB on the origin DNA. DnaB can act as a helicase, resulting in further unwinding of the origin DNA.

2.1.2 Eukaryotic DNA replication

In many ways, the eukaryotic replication is quite similar to the prokaryotes in the sense that they follow the same basic requirements: the presence of *trans*-acting initiator protein and a *cis*-acting DNA element. However, the size and complexity of the eukaryotic genome far surpasses the prokaryotic genome and the eukaryotic system possesses different and more sophisticated ways of regulating DNA replication. The following section will talk more on the origins of DNA replication in several eukaryotic organisms, from simple eukaryotes like *Saccharomyces cerevisiae* to higher eukaryotes like mammalians.

2.1.2.1 Origin of DNA replication in *Saccharomyces cerevisiae*

One of the first origins of replication to be discovered in eukaryotes came from studies done on *Saccharomyces cerevisiae*. These origins of replication like the ones found in prokaryotes, were specific DNA sequences called autonomously replicating sequence (ARS). They were found in the chromosome and were shown to be able to confer to any plasmid the ability to exist extrachromosomally in yeast (Stinchcomb *et al*., 1979). The ARS was later shown to be able to bind to a replication initiator protein called the origin recognition complex (ORC), a multiprotein complex made up of six different subunits, (Bell and Stillman, 1992). And this ORC, together with another protein, Cdc6, help recruit other proteins to form a pre-replicative complex (pre-RC).

Further study into the ARS identified four motifs within the ARS that was responsible for the binding to the ORC and essentially required for replication to occur: A, B1, B2 and B3. The A element contains the ARS consensus sequence (ACS), which was an eleven base pair AT-rich sequence that exists in all known ARS, while the B elements were of varying sizes of ten to fifteen base pairs long that lie 5' of the A element. Seemingly, the A element, while vital for the proper function of the ARS, requires at least one of the B elements to be found in the ARS for the origin to work (Marahrens and Stillman, 1992; Newlon and Theis, 1993; Rao and Stillman, 1995; Rowley *et al.*, 1995).

The ORC is a six protein complex that binds to the ARS in an ATP dependent manner (Bell and Stillman, 1992). Out of the six subunits of the complex, Orc1p-6p, only Orc1p, Orc2p, Orc4p and Orc5p binds directly to the DNA at the AT-rich regions

found in the ACS. The ability of the ORC complex to bind to the DNA requires ATP to be bound to the Orc1p, and although ATP is hydrolyzed by Orc1p, it seems to serve a separate function (Klemm *et al.*, 1997). Although Orc3p and Orc6p do not bind directly to the DNA, both are indispensable. Orc3p may be involved in arranging the other subunits for the proper binding to the origin (Lee and Bell, 1997). While Orc6p does not seem to play a role in affecting the DNA binding specificity of the complex to the DNA, it is still required for DNA replication and cell viability (Li and Herskowitz, 1993).

After the binding of the ORC to the origin, other proteins like Cdc6, Cdt1 and the MCM (minichromosome maintenance) complex are recruited to form the pre-replicative complex. The Cdc6 was identified to be one of the first proteins to be recruited after the association of the ORC with the origin. It is believed to directly bind to the origin as well and in turn helps recruit the MCM complex to the chromatin in an ATP dependent manner, without which replication cannot proceed properly (Tanaka *et al.*, 1997). While the role of Cdt1 in pre-RC formation is not well known, Tanaka and Diffley (2002) have shown that Cdt1 interacts directly with the MCM complex and accumulates in the nucleus during G1 phase, suggesting that the Cdt1 could act as a carrier protein aiding in the localization of the MCM complex to the origin completing the pre-RC formation. The MCM complex is widely believed to be the helicase involved in unwinding double stranded DNA into single strand during DNA replication (Labib and Diffley, 2001). A representative diagram of replication initiation in eukaryotes is shown in figure 2 below.

The purpose of assembling the pre-RC is to ultimately recruit DNA polymerases to

the origin for the purpose of DNA replication. However, for that to happen, additional factors apart from the pre-RC must be present as well. Cdc45 was one such example shown to be needed for the assembly of components necessary for DNA replication. It was found to associate with the MCM complex, replication protein A (RPA) as well as the DNA polymerases themselves, DNA polymerase α and ε (Aparicio *et al*., 1999; Zou and Stillman, 2000). It was suggested that the complex containing the Cdc45, MCM complex and RPA was involved in the unwinding and subsequent assembly of replication forks at the origins.

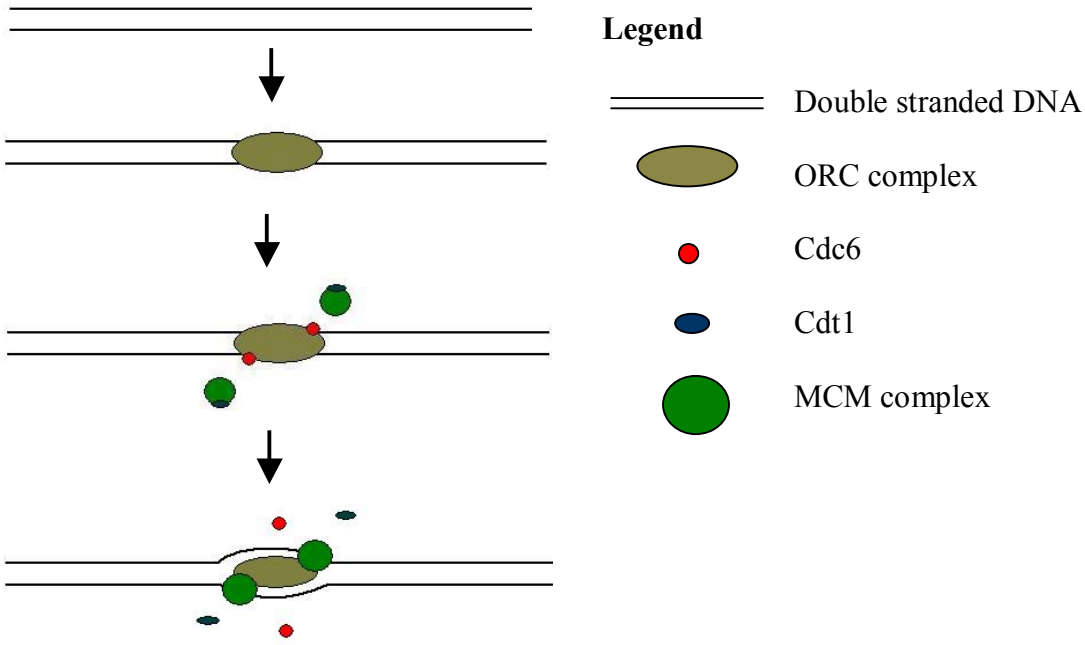


Figure 2. Eukaryotic replication initiation. In eukaryotes, the ORC first binds to DNA, followed by Cdc6. Cdt1 possibly guides the MCM complex to the Cdc6 bound ORC. Hydrolysis of ATP followed by the dissociation of Cdt1 and Cdc6 occurs, leaving the MCM together with the ORC on the DNA. The helicase activity of MCM unwinds the DNA paving the way for the recruitment of other factors such as Cdc45, RPA and DNA polymerases.

2.1.2.2 Origin of DNA replication in *Drosophila melanogaster*

The definition of an origin of replication in *Drosophila melanogaster* is more complicated as compared to *Saccharomyces cerevisiae*. It has been shown that in the

early embryo stage of *Drosophila melanogaster*, initiation of replication is not localized to a specific site, as in *Saccharomyces cerevisiae*. Rather, initiation occurred on multiple sites within the same chromosome, and apparently with little or no regard for specific DNA sequences (Shinomiya and Ina, 1991).

However, during the oogenesis stage of *Drosophila melanogaster*, DNA sequence specific replication initiation is employed for a special function: to produce sufficient levels of eggshell, otherwise known as the chorion. This gene amplification strategy employed by the organism serves as an interesting area of study into DNA replication in metazoans as it also employs similar replication proteins used in replicating the genomic DNA as *Saccharomyces cerevisiae* (Calvi and Spradling, 1999). The *Drosophila* origin recognition complex was identified as a homolog of the *Saccharomyces* ORC (Gossen *et al*., 1995) that could bind to specific DNA sequences, in this case the amplification control element on chromosome 3, ACE3 and amplification enhancing regions, AER-d (Austin *et al.*, 1999). More detailed work identified an 884bp element, oriβ, which overlaps the AER-d as the major site of replication initiation. These two DNA sequences, necessary for gene amplification, were able to induce amplification when inserted into locations other than the chromosomes they were located on (Lu *et al*., 2001). Further mutational studies in oriβ showed that two important elements were required; a 140 base-pair sequence found in the 5' region and a 226 base-pair AT rich sequence in the 3' region that had significant homology to ACE3.

Interestingly, despite the fact that both *Drosophila* and *Saccharomyces* ORC share homology, and that both oriβ and ARS contain AT rich sequences, Zhang and Tower,

2004, showed that the ARS was unable to replace the oriβ in directing amplification. In addition, *in-vitro* studies involving measuring the binding affinity of *Drosophila* ORC with origin and non-origin DNA indicated that the largest difference in binding affinity was only six-fold and it was proposed that this difference was not enough to differentiate between origin and non-origin DNA. On the other hand, the topological state of the DNA appeared to be more crucial, with negatively supercoiled DNA binding 30 fold higher to *Drosophila* ORC when compared to relaxed or linear DNA (Remus *et al.*, 2004). This gives an indication that something else apart from AT rich sequences is needed for ORC binding.

2.1.2.3 Origin of DNA replication in mammalian cells

Similar to *Drosophila,* trying to identify distinct elements of DNA capable of initiating replication in mammalian cells is more complicated as compared to yeast. In 1991, Heinzel *et al*. tried to isolate a human equivalent of the yeast ARS. However, instead of finding a distinct DNA element capable of initiating replication, they found that any piece of large human DNA sequence was sufficient for the autonomous replication in human cells. Since then, more work has been done in trying to identify and understand the mechanisms behind origins of replication in mammalian cells. However, despite the progress made in identifying these sites of replication initiation, a lot of work remains to be done in understanding how they work. Gilbert (2001), suggested that there exist two different types of origins: the first type being a zone of initiation, where replication can be initiated at multiple sites on a large fragment of roughly 10 to 50 kilo-base-pairs of DNA, while the other type include a more localized initiation site consisting of a few kilobase pairs.

An example of the first type of origin is the Chinese hamster ovary dihydrofolate reductase (DHFR) locus. It was early replicating and was found within a highly amplified region of the genome. In addition, this DHFR locus started incorporating nucleotides early in the S phase in three regions, an indication of replication (Milbrandt *et al*., 1981; Heintz and Hamlin, 1982). Further detailed work was done to show that replication initiation was preferably located to two specific loci termed oriβ and oriγ (Anachkova and Hamlin, 1989; Leu and Hamlin, 1989), and that oriβ could initiate replication in both hamster and human cells, even when placed in random locations out of its native state (Altman and Fanning, 2001 and 2004). However, it was also discovered that replication could start from multiple sites within the DHFR locus, although initiation was preferred for the 55kb intergenic region which contained the oriβ and oriγ (Dijkwel and Hamlin, 1995). On top of that, despite oriβ being the preferred choice of replication initiation, deletion studies done with oriβ had no effect on the replication initiation in the DHFR on a whole, indicating that the other origins could initiate replication efficiently even without oriβ (Kalejta *et al*., 1998). Furthermore, the appearance of bubble arcs in almost every restriction fragment of the 55kb intergenic region tested using two-dimensional gel analysis and a PCR-based nascent strand abundance assay on restriction fragments that showed almost all the fragments tested positive for replication initiation corroborated the observation that the region responsible for replication actually contains multiple sites that have varying efficiencies of replication initiation (Dikjwel *et al*., 2002).

An example of the second class of origin as suggested by Gilbert (2001) included the human lamin B2 gene. Lamins are intermediate filaments usually found in the nucleus providing support for the nuclear membrane. It was discovered by Giacca *et al*. (1994)

that the origin of replication for this gene was located at the 3' end and was approximately 500 base-pairs long. Further methods aimed at investigating the proteins bound to the lamin B2 origin were done. UV irradiation using a pulsed laser light source, followed by immunoprecipitation showed that components required for pre-RC formation such as the Cdc6 and MCM proteins were present at the origin (Abdurashidova *et al*., 2004). In addition, the lamin B2 origin was also shown to be able to initiate replication not only in ectopic locations of the chromosome but in hamsters as well (Altman and Fanning, 2004), providing evidence that one, the human lamin B2 origin is a true origin of replication and two, the human DNA replication system may share similar characteristics to other eukaryotic systems. Another replication origin that be classified under this second class of origin was found in the human β-globin gene. It was also shown that like the lamin B2 origin, the β-globin origin could enable replication at ectopic locations in the chromosome (Aladjem *et al*., 1998).

However, despite the success in identifying these origins of replication in mammalian cells, many questions still remain. Why are we still unable to identify other similar origins? Is there or is there not a consensus sequence that can be used to identify these origins as is the case in yeast ARS? A recent work published by Vashee *et al*. (2003) showed that human ORC was able to restore DNA replication in *Xenopus* eggs depleted of *Xenopus* ORC. They also showed that the human ORC did not distinguish between origin and non-origin DNA and could bind to both forms. Not only could they bind to non-origin DNA, they also showed that the human ORC could initiate replication of plasmids containing origins and plasmids that did not have origins.

It would seem that although replication origins have been isolated from eukaryotic systems, specific DNA sequences do not appear to be a prerequisite for the binding of ORC to occur. In order to minimize any ambiguity, it was decided to use a more defined origin of replication for this study, the Epstein-Barr virus, EBV latent origin of replication *ori*P.

2.2 Epstein Barr Virus

In the beginning of 1940's, a British missionary surgeon by the name of Denis Burkitt noted that lymphomas or lymphocyte tumors occurred at a frequent rate in equatorial Africa as compared to the rest of the world. The lymphomas, also known as Burkitt's lymphoma, were also unique in that they were found outside of the lymph nodes. Burkitt then wrote and talked about this unique lymphoma widely, raising the possibility that the cause of this lymphoma could be by an infectious agent (Burkitt and Wright, 1966). After listening to Burkitt speak on his findings, Tony Epstein decided to obtain tumor biopsies and attempt to culture the lymphoma cells, which he was successful in achieving. In 1964, together with Achong and Barr, Epstein managed to identify a herpesvirus in electro micrographs of tumor cells (Epstein *et al*., 1965). In addition, they managed to show that this herpesvirus was different from other members of the family. This virus was unable to replicate in other cell cultures and was non-reactive to antibodies that react with other herpesvirus. It was subsequently called Epstein-Barr virus and became the first virus to play a possible role in causing tumors in humans.

Now, EBV has been classified as a member of the gammaherpesvirus subfamily, which contains potentially tumorigenic herpesviruses (Chang *et al*., 1994). This

subfamily currently includes gamma 1, also known as LCV and gamma 2, also known as RDV, genera. Up to now, EBV is the only LCV known to affect humans and has been implicated as the etiological agent in causing, apart from Burkitt's lymphoma, Hodgkin's disease, some unusual T-cell lymphomas and nasopharyngeal carcinoma. (Klein, 1994; Karimi and Crawford, 1995; Kieff and Rickinson, 2001).

2.2.1 Epstein Barr Virus latent origin of replication *oriP*

In a latent infection with EBV, the 165kb viral genome exists as a piece of circularized extrachromosomal episome that can be maintained autonomously in the proliferating latently infected cells (Lindahl *et al*., 1976 and Nonoyama *et al*., 1972). It was discovered that a 1.7kb region of the viral genome, called *oriP* could mediate replication as well as nuclear retention of the viral episome in the cell. However, the *oriP* can only do so if a single EBV nuclear protein, EBNA-1 is present. Yates *et al*. (1984) have shown that recombinant plasmids containing the *oriP* could be maintained in the presence of EBNA-1, Replication can occur at most once per cell cycle (Yates and Guan, 1991) and the cell cycle machinery seems to be the controlling mechanism (Laskey and Madine, 1996). Reisman *et al*. (1985) have shown that *oriP* consists of two regions that do not touch one another: one being the dyad symmetry (DS) and the other being the family of repeats (FR). Although both contain EBNA-1 binding sites and require EBNA-1 to function, they are structurally distinct from one another and both serve different functions.

Rawlins *et al*. (1985) have shown that the DS region consists of a 120bp region, which contains four EBNA-1 binding sites. Of which, two of the EBNA-1 binding sites are tandem while the other two are arranged in a dyad symmetry. The ultimate

role of the DS region is to initiate the replication of the DNA (Harrison *et al*., 1994). In fact, it has been shown by various groups that the human ORC is loaded onto the DS and they suggested that this is possible through interaction with EBNA-1 (Chaudhuri *et. al*., 2001; Dhar *et. al*., 2001; Schepers *et. al*., 2001).

The FR region is a family of repeats consisting of 21 imperfect repeats, each of 30 basepairs in length; although only 20 of them contain EBNA-1 binding sites (Rawlins *et. al*., 1985). This region plays the role of mediating nuclear retention of the plasmid that involves the physical binding of EBNA-1 to the FR on the plasmid and the chromosomal DNA. This nuclear retention ability ensures that the plasmid is segregated and maintained in the nucleus during mitosis (Mackey *et al*., 1995; Yates *et al*., 2000). Reisman and Sugden (1986) have shown that upon binding to EBNA-1, the FR may also act as an enhancer for transcription. This transcriptional enhancer affects gene expression downstream of the FR. Wysokenski and Yates (1989) have also shown that this enhancer function required at least 6 to 7 copies of the 30bp repeats found in the FR.

2.2.2 EBNA-1 protein

Apart from the *oriP*, EBNA-1, a trans-acting EBV-encoded nuclear antigen, is essential for the long-term persistence and replication of EBV genome in infected cells during latent infection and of any plasmid bearing the *oriP*. As shown in figure 3, EBNA-1 consists of 641 amino acids. Amino acids 33-83 and 328-382 are arginine rich domains. Amino acids 379-386 contain the nuclear localization signal. It was also recently shown that EBNA-1 can function as a transcriptional activator and that amino acids 65 to 89 were necessary to activate transcription (Kennedy and Sugden, 2003).

In addition, Ambinder *et. al*. (1991) as well as Chen *et. al*. (1993) have shown that the DNA binding and dimerization domain are found in amino acids 459-604 while Levitskaya *et al*. (1995) showed that amino acids 90-327 make up a domain of glycine-alanine repeats that render EBNA-1 resistant to degradation by protease and recognition by cytotoxic T cells. In addition, EBNA-1 was found to have induced B-cell neoplasia in transgenic mice in one study (Wilson *et al*., 1996) but not another (Kang *et al.*, 2005).

2.2.2.1 Role of EBNA-1 in the persistence of any plasmids bearing *oriP*

The N-terminus of EBNA-1 was shown to play a critical role in mediating binding of the episome to the chromosome (Hung *et al*, 2001). It was also found that this N-terminus could be replaced by high-mobility group-I amino acids 1-90 or by histone H1-2 to mediate binding to the chromosome and mediate long-term persistence of the episome. The C-terminus of the EBNA-1 binds to the DNA sequences containing binding sites for EBNA-1, such as the DS and the FR element of the *oriP* (Kieff and Rickinson. 2001). This creates a physical association between the EBV and the host chromosome, resulting in the persistence of the virus within the host nucleus.

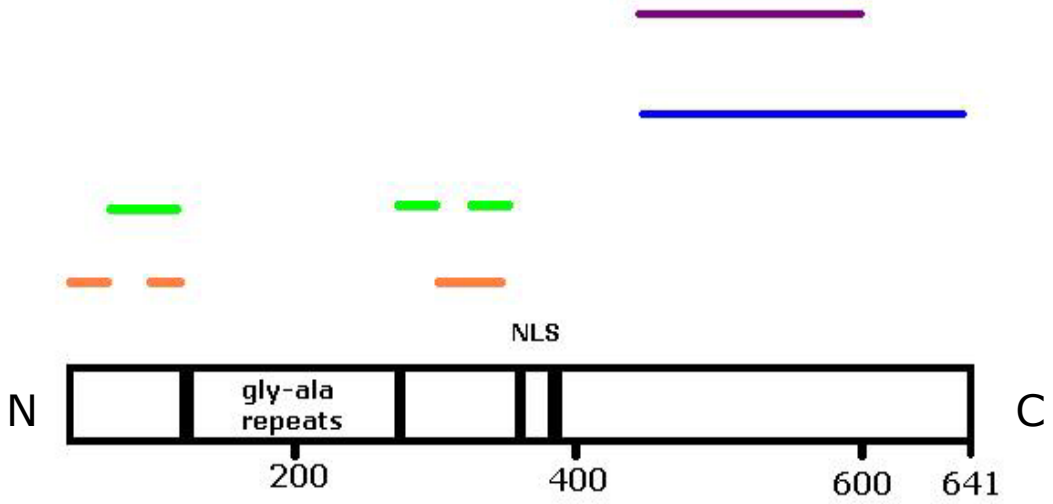


Figure 3. Schematic representation of EBNA-1. EBNA-1 essentially consists of four main components. The orange lines indicate the chromosome-binding domains. These chromosome-binding domains lie in two regions amino acids 32-89 and 328-386, both of which are rich in arginine residues. The nuclear localization signal is located from amino acids 379-386. Both the dimerization (purple) and DNA binding domain (blue) overlap each other within the C terminus.

2.2.2.2 Use of Epstein-Barr Virus based gene therapy vector

There are a few advantages of utilizing a gene therapy vector based on the Epstein-Barr Virus. Being episomal, such vectors hold several advantages over other vectors such as those that employ integration. One advantage of episomal vectors over integrative ones is the fact that there is no need for the negative effects integration might bring about to both the therapeutic gene as well as cellular endogenous genes. Secondly, multiple copies of the episomal vectors can exist in the nucleus, which allow for amplified expression of the therapeutic gene.

A standard vector derived from the Epstein Barr Virus contains the origin of replication *oriP* and the sequence encoding for the *trans*-acting factor EBNA-1. These two sequences are necessary for the retention as well as the replication of the EBV based vector. As seen in figure 4 below, EBV-based vectors can also be used as

shuttle vectors by incorporating a bacterial origin of replication as well as a selection marker for bacteria like ampicillin resistance for easy amplification and manipulation within bacterial cells. However, it was discovered that vectors bearing the *oriP* of EBV could still undergo limited replication even without the presence of EBNA-1, although EBNA-1 is vital for the long-term persistence of plasmids bearing *oriP* (Aiyar *et. al*., 1998). EBV based vectors have been mainly used in primate cells, with reports of failure to replicate in mouse and hamster cells (Yates *et. al*., 1985; Wysokenski and Yates, 1989). This poses a setback for gene therapy applications as testing on mouse models would not be possible. However, even though murine models were not permissive for EBV, it was discovered that when murine cells were transfected with the EBV vector, the marker gene was expressed more intensively than a conventional vector (Tomiyasu *et. al*., 1998). This could be partially due to the transcriptional enhancer effect of the FR and the fault that the vector could not replicate in the murine cells was attributed to the DS.

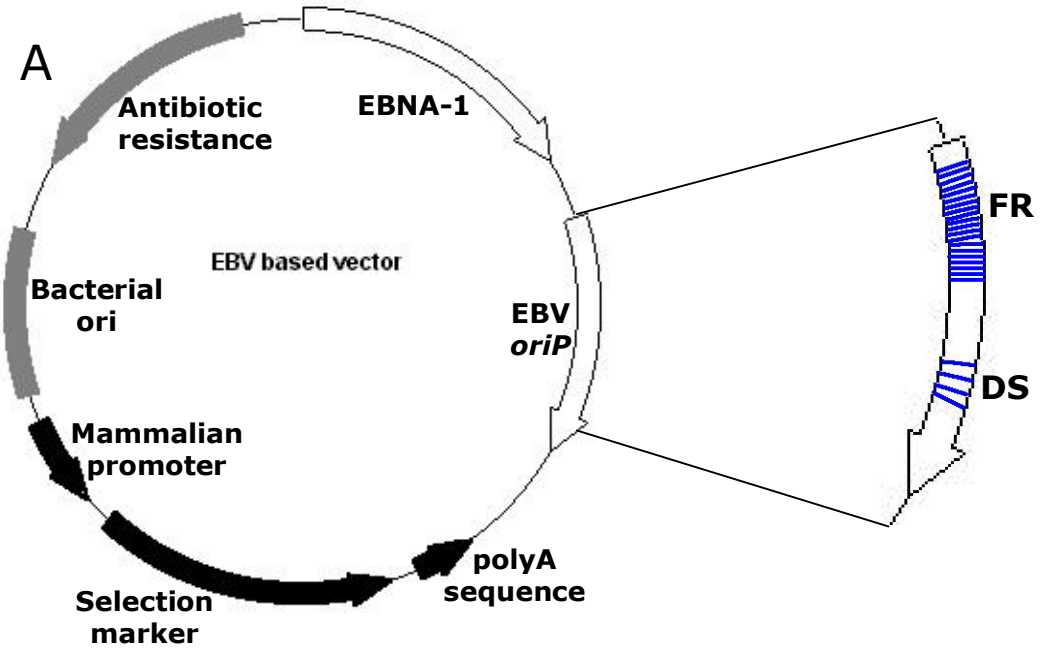


Figure 4. A basic EBV based vector. It consists of the viral sequences *oriP* (white arrow) and ENBA-1 expression cassette (white arrow) as well as a eukaryote selection marker inclusive of promoter and polyA sequence (black arrows). A bacterial origin of replication (grey block) is also included for shuttling between prokaryotic and eukaryotic systems. An antibiotic resistance marker (grey arrow) is also provided for selection within bacteria. The EBV *oriP* consists of two noncontiguous repeats. The dyad symmetry (DS), which has 4 EBNA-1 binding sites denoted by the blue lines and the family of twenty 30bp repeats (FR) which has 20 EBNA-1 binding sites, also denoted by blue lines.

Most of the developments for an efficient episomal vector for gene therapy have come from EBV based vectors, where transient yet high levels of expression of the therapeutic gene is required as in the case of Duchenne Muscular Dystrophy. Research has already shown that intramuscular injection of the EBV based dystrophin expression vector into murine models resulted in a significant enhancement of the expression of the dystrophin as compared to conventional vectors (Tsukamoto *et. al.*, 1999). In addition, the persistence of EBV vectors in humans had been found to be prolonged but not indefinite. This is particularly useful when the goal is to kill off the tumor cells. Already, EBV vectors carrying cytokine genes have been tested on

human B-lymphomas where the gene was shown to be stably expressed over a period of a few weeks (Mücke *et. al*., 1997; Robertson *et. al*., 1996).

The major disadvantage of using EBV based episomal vectors is that it requires the expression of the *trans*-acting viral factor EBNA-1 for stable maintenance. While previous work done by Lutfalla *et. al*. (1989) show that plasmids bearing the EBV *oriP* were able to replicate stably in hepatic cells overexpressing EBNA-1 without interference to the expression of liver-specific proteins, other experiments have showed that EBNA-1 could bind to RNA *in vitro* which meant that EBNA-1 was capable of influencing expression of a gene post-transcriptionally (Snudden *et. al*., 1994). In addition to that, EBNA-1 may induce B-cell lymphoma in transgenic mice (Wilson *et. al*., 1996), thus raising safety concerns regarding the use of this vector for gene therapy.

A possible way to circumvent these potential problems associated with EBV based vectors would be to find a way to regulate the expression of EBNA-1 or the physical replication of the vector.

2.3 Transcription

Transcription is another major process in cells which allows the cell to produce the necessary proteins that are vital for survival. Just like DNA replication, the inability of a cell to undergo transcription often results in fatal consequences. And similar to replication, it mainly involved three main stages: initiation, elongation and termination. This section will deal on the different systems of transcription in prokaryotes and eukaryotes before discussing the possible interplay that could exist

between these two processes.

2.3.1 Prokaryotic transcription

In prokaryotes, transcription is performed by one form of RNA polymerase. The prokaryotic RNA polymerase holenzyme is a multisubunit protein complex consisting of a core enzyme made up of 2 α subunits, 2 β subunits and 1 ω subunit (Zhang *et al*., 1999) and an additional σ factor, which enables the holoenzyme to recognize specific binding sites at the -10 and -35 of the promoter (Dombroski *et al*., 1992).

When the RNA polymerase holoenzyme binds to the promoter, the double stranded DNA is separated into single strands to form a transcription bubble. As the polymerase starts incorporating nucleotides from the +1 site, there exists a chance that abortive initiation, in which transcription is aborted, may occur in the first 17 nucleotides synthesized (Ring *et al*., 1996). When that happens, the RNA polymerase releases the transcript and will start synthesizing a new strand of RNA from the first base.

As the RNA polymerase progresses from initiation to elongation, the σ factor is no longer needed and is generally thought to be released from the core enzyme (Shimamoto *et al*., 1986), although there has been a recent report of a population of RNA polymerases that retain the σ factor through the transition from initiation to elongation (Bar-Nuham and Nudler, 2001). The *Escherichia coli* elongation factor NusA was shown, upon σ factor release, to be capable of interacting with the elongation complex, comprising of the core enzyme, DNA template and the growing RNA transcript, and inducing termination. Upon termination of transcription, the

NusA protein dissociates from the RNA polymerase core enzyme and allows the σ to bind again, thereby allowing transcription to be initiated again (Greenblatt and Li, 1981; Schmidt and Chamberlin, 1984; Gill *et al*., 1991).

During the transcription elongation stage, certain scenarios may occur that may impede the elongation complex. Transcriptional pausing was first discovered *in-vitro* using the *Escherichia coli* RNA polymerase. Despite synchronously initiating transcription, varying lengths of RNA transcripts resulted, demonstrating that elongation was not synchronous. It was attributed to specific pausing sites on the DNA template (Kassavetis and Chamberlin, 1981) and can be suppressed by the presence of transcription factors such as NusG (Burova *et al*., 1995). Transcriptional arrest on the other hand is more severe than pausing in that arrested elongation complex is unable to resume transcript elongation without the addition of accessory factors (Arndt and Chamberlin, 1990). In transcriptional arrest induced by limiting substrate nucleoside triphosphates (NTPs), the RNA polymerase, although active, was unable to proceed with elongation and can be re-activated upon the provision of the missing NTPs. However, for some templates elongation could not be restored demonstrating the potential irreversible effect of transcriptional arrest. Indeed, Komissarova and Kashlev (1997) demonstrated that during transcriptional arrest in *Escherichia coli*, the RNA polymerase could actually disengage from the growing transcript and translocate backward, resulting in the extrusion of the 3' end of the RNA.

2.3.2 Eukaryotic transcription

One of the most well studied eukaryotic transcription systems comes from yeast.

More complex than the prokaryotic system, the eukaryotic transcription consists of three types as compared of one type RNA polymerase in prokaryotic system. RNA polymerase I synthesizes ribosomal RNA, RNA polymerase II synthesizes messenger RNA and RNA polymerase III synthesizes tRNAs and other small RNAs. Although they serve to create different forms of RNA, the subunits of all three polymerases are either identical or homologous (Woychik *et al.*, 1993; Sentenac, 1985). This section shall discuss briefly on some of the basic characteristics of RNA polymerase II initiation and elongation.

2.3.2.1 RNA polymerase II

The RNA polymerase core enzyme is made up of 12 subunits, termed Rpb, 1-12. Crystal structure studies of these 12 subunits have revealed a 10 subunit catalytically active core and a separate 2 subunit heterodimer consisting of Rpb4 and Rpb7 (Armache *et al.*, 2003; Bushnell and Kornberg, 2003). One of the suggested functions of this heterodimer could be to assist the polymerase in interacting with a variety of transcription factors (Armache *et al*., 2003).

Apart from the RNA polymerase, two other components are required for transcription. One of the two components is a 20 subunit protein complex called the Mediator which, like its name suggests, acts as a mediator for transducing signals from transcriptional activators or repressors (Kelleher *et al*., 1990; Kim *et al*., 1994; Gustafsson *et al*., 1998). The other component needed include a set of five additional transcription factors TFIIB, D, E, F and H. TFIID is also a complex made up of a universal TATA-binding protein (TBP) and additional TBP-associated factors (TAFs) (Conaway and

Conaway, 1997; Lee and Young, 1998). Before transcription initiation, the TFIID binds to the promoter region and recruits the mediator complex and the RNA polymerase II as well as other transcriptional factors to form the preinitiation complex. But given that DNA is wrapped around histones in its native state as a nucleosome, how then does the transcription activators and holoenzyme gain access to the sequences to activate transcription? To do that, the nucleosome would have to be remodeled such that the DNA elements are free to interact with the transcription activators and holoenzymes. The SWI (homothallic switching deficient) / SNF (sucrose non-fermenting) protein complex was one such complex capable of remodeling the nucleosome (Brown *et al*., 1996). In the presence of a transcriptional transactivator, such as the human heat shock factor 1 (HSF1), SWI/SNF was recruited to the chromatin, where it may remodel the nucleosome in an ATP dependent manner, and this could possibly result in the opening of the nucleosome, thereby enabling the entry of the transcriptional machinery.

Following the formation of the preinitiation complex, local unwinding of the DNA template occurs resulting in an open complex. This process is dependent on ATP as well as on the transcription factors TFIIE and TFIIH (Holstege *et al*., 1996; Kim *et al*., 2000). Preinitiation then proceeds to initiation, a process marked by the addition of nucleoside triphosphates. Initiation can only move on to elongation after the promoter has been cleared by the RNA polymerase. This early stage of transcription, also known as promoter clearance, is marked by the tendency of the RNA polymerase to slip during the synthesis of the first 23 nucleotides, after which no slippage is detectable (Pal and Luse, 2003), which is not too different from the prokaryotic system of abortive initiation. It was also shown that the RNA polymerase II transcript

elongation complex (TEC) is unstable before the growing RNA transcript-DNA template hybrid reaches 8 nucleotides in length (Kireeva *et al*., 2000). Certain general transcription factors such as TFIIF and TFIIH have been shown to be capable of suppressing the occurrence of abortive initiation during promoter clearance (Dvir *et al*., 1997; Yan *et al*., 1999).

Similar to the prokaryotic system, transcriptional pausing or arrest may also occur during the elongation stage in eukaryotes. One of the first descriptions of transcriptional arrest in eukaryotes came from work on the histone3.3 gene and it was shown that arrest could occur in both the coding and non-coding region. It was also shown that this arrest can be alleviated by the presence of elongation factor TFIIS (Reinberg and Roeder, 1987; Reines *et al*., 1989). Further examination of one of the strongest arrest sites in the histone 3.3 gene revealed that it contained a T-rich region for the non-template strand. It was suggested that the structure of the template could play a role in arrest as such T-rich regions could contain a bend in the DNA double helix (Kerppola and Kane, 1990). TFIIS was found to be able to reactivate arrested polymerase by inducing endonucleolytic cleavage of the nascent transcript near the 3' end, resulting in the creation of a new 3'-OH terminus that is correctly based paired to the DNA template, thus allowing elongation to carry on (Reines, 1992; Reines *et al*., 1992; Izban and Luse, 1992). Apart from TFIIS, there are many other elongation factors such as ELL, TFIIF and Elongin that can interact directly with the polymerase during elongation. However, there exists a second class of elongation factors that while they do not affect the polymerase activity directly, they may affect the progression of elongation through the modification of chromatin. The two main proteins identified so far include the Elongator and FACT. Elongator was first

identified as a six protein subunit complex (Otero *et al*., 1999; Winkler *et al*., 2001) that possessed histone acetyltransferase on one of its subunits Elp3 and could acetylate core and nucleosomal histones (Kim *et al*., 2001) and therefore could possible facilitate elongation. FACT was shown to be able to facilitate RNA polymerase II elongation on chromatin templates *in-vitro*. It was proposed that FACT could promote RNA polymerase II elongation through nucleosomes by binding to and promoting the removal of histones, such as H2A and H2B (Orphanides *et al*., 1999).

2.4 Relationship between DNA replication and transcription

This section examines some of the already existing relationships between DNA replication and transcription and talks about some of the evidence supporting each claim.

2.4.1 Transcription through an origin of replication may inhibit DNA replication

DNA replication and transcription occur throughout the cell cycle in bacteria simultaneously. As they utilize the same DNA template for their purposes, it would seem inevitable that both the replication and transcription machinery would meet one another along the way either traveling in the same direction or head on. And it would seem that collisions between the two would also be unavoidable, given that replication and transcription are both polar. Such a phenomenon was first observed in *Escherichia coli* by French (1992). An inducible origin of replication was placed on either side of a ribosomal RNA operon. It was observed that replication and transcription occurring in opposite directions resulted in a slower replication fork progression. This could be either due to physical collision between the polymerases or it could be due to topological factors. As transcription and replication proceeds, they

generate positive supercoiling in downstream DNA (Liu and Wang, 1987; Peter *et al*., 1998). As they meet head on, the positive supercoil generated by the two could have a negative effect on both. Further work by Mirkin and Mirkin (2005) confirmed the observation that replication was inhibited as it met transcription head on. In addition to the possibility that the positive supercoils generated could have inhibited replication, the authors also suggested that the physical collision between the two machineries could also play a part. At the same time the authors also found that replication elongation traveling in the same direction as transcription did not seem to affect one another. They proposed that in this case, the DNA replication machinery either bypassed or displaced the RNA polymerase from the DNA template as they traveled co-directionally.

When a yeast origin of replication ARS was placed under the effect of an actin promoter, replicative ability was affected negatively (Kipling and Kearsey, 1989; Tanaka *et al*., 1994). It was shown using micrococcal nuclease assay and indirect end-labeling that chromatin structures were not affected and so chromatin remodeling due to the positioning of the ARS near a promoter was ruled out and it was concluded that transcription through the origin of replication altered the activity of the ARS. In addition, transcription into the ARS resulted in an increase in dependence of the MCM complex, an indication that there exists a negative relationship between pre-RC assembly and transcription (Nieduszymski *et al*., 2005). Interestingly, MCM1, which is needed for the initiation of replication in yeast (Chang *et al*., 2004), can also act as a transcriptional co-repressor of MATα2 (Keleher *et al*., 1989).

In *Tetrahymena thermophila*, transcription through the origin of replication rDNA by

placing the origin immediately downstream of a RNA polymerase I promoter resulted in the failure to replicate. Replication could be restored by placing an rRNA transcriptional terminator between the promoter and the origin, or by inducing a mutation to inactivate the promoter, and the authors proposed that transcription through the origin could inhibit replication initiation (Pan *et al*., 1995).

In 1991, Heinzel *et al*. found that any piece of human DNA that was large enough could support autonomous replication of the plasmid it was found on. Using that knowledge, Haase *et al*. (1994) tried to study the effect transcription had on such replicating plasmids. They found out that the ability of plasmids to replicate was inversely correlated to the promoter strength and that similar to what was observed in *Tetrahymena thermophila*, the insertion of a transcription termination sequence downstream of the promoter restored the ability to replicate. The authors also proposed that transcription could inhibit replication by preventing replication initiation.

2.4.2 Transcription factors may affect DNA replication positively

While there has not been much evidence showing that the transcription machinery through an origin could affect replication in a positive way, there have been quite a lot of studies that cast transcription factors playing a positive role in affecting replication. This section will deal on some of the studies done so far.

In the yeast ARS1, a binding site for transcription factor Abf1 exists in the B3 element. The binding of this transcriptional factor had a positive effect on replication and this ability to activate replication was mapped to the C-terminal acidic domain of

the Abf1. In addition, it was discovered that Abf1 could also be replaced by other transcription factors from different organisms such as VP16 from the herpes simplex virus or the p53 tumor suppressor protein from humans, suggesting that conserved mechanisms that could exist between organisms in using transcriptional factors for the activation of replication (Li *et al*., 1998). However, there has been evidence to the contrary. ARS301 is an inactive origin in the chromosome but active when placed on a plasmid. But when the Abf1 transcription factor binding site was introduced near the ARS301, the origin was inactivated (Kohzaki *et al*., 1999).

The *Drosophila* chorion gene amplification system as described in section 2.1.2.2 is one of the most well studied origins in multicellular eukaryotic organisms. As mentioned above, two elements within the origin are vital for replication: the ACE3 and the oriβ. Interestingly, binding sites for Myb, a transcription factor, could be found within the region and was shown to be needed for replication. It was also shown that Myb could interact with ORC subunits and mutations in Myb resulted in reduced or no replication (Beall *et al*., 2002), providing evidence that transcription factors could affect replication in a positive way. In another study, microarray analysis was used on the *Drosophila* genome in a bid to identify replication origins and determine replication timing. It was shown that the ORC localizes to specific regions on the chromosome, many of which actually contain early-activating origins. In addition, these early activating origins contain RNA polymerase II binding sites as well, leading the authors to suggest that transcription activity could have a positive effect on origin activation (MacAlpine *et al*., 2004).

Experiments utilizing human genome microarrays have yielded similar results. By

separating DNA from cells in S phase and DNA from cells in the G1 phase, the authors were able to compare the difference in DNA copy number using genomic array hybridization. They found that there was a positive correlation between replication timing and GC content, gene density as well as transcriptional activity (Woodfine *et al*., 2004). The fact that origins of replication were found near gene loci, such as the lamin B2 gene and the β-globin gene is a strong indication that transcription could have a positive effect on replication rather than a negative one. Further proof came from work done on the origin found in the β-globin gene locus in which a locus control region (LCR) that contained binding sites for transcription factors was found to be vital for replication initiation (Aladjem *et al*., 1995).

A clue of how transcription factor may activate replication comes from Hu *et al*., (1999). In that study, the authors fused the transcriptional activator BRCA1 (breast cancer protein 1) to the DNA binding domain of the GAL4. Using this construct they showed that the GAL4-BRCA-1 resulted in a significant increase of plasmid stability. In addition, they showed that the presence of BRCA1 could remodel the chromatin. They proposed that one way in which transcription factors may affect replication positively may be by increasing chromatin accessibility to replication initiation proteins.

Materials and Methods

3.1 Polymerase Chain Reaction (PCR) amplification of EBV latent origin of replication *ori*P

PCR was used to amplify the EBV latent origin of replication *oriP*. Primers were designed using the DNA sequence analysis software VectorNTI (InforMax. Inc. Maryland, USA). They were also analysed by the same software for any primer dimer formations that may potentially disrupt the PCR reaction. And they were synthesized by MWG-Biotech AG (Germany). The forward primer was designed in such a way that an *Aat* II restriction enzyme site was incorporated at its 5' end while the reverse primer was designed with an *Eco* RI restriction enzyme site at its 3' end. The sequences of the forward and reverse primers are provided below.

*oriP*for- GCCCTGACGTCTCACATTGGTCTGTACCTCCACACT

*oriP*rev- CCTCCTGGAATTCTATCATTAAACGGC

3.1.1 PCR reaction setup

Hotstartaq DNA polymerase (Qiagen) was used for PCR. This polymerase used was unique in that it was totally inactive at room temperature and required a pre-PCR cycle step of heating at 95°C for 15 minutes for the enzyme to be activated. This pre-PCR cycle heating step was useful in ensuring target specificity and the yield of the PCR product. In addition, any primer dimers that could have resulted from complementary base pairing between the added primers would also be separated and extension of the primers was prevented. The 10 mM dNTP solution was supplied by New England Biolabs. A final PCR reaction volume of 100 µl for amplifying the *oriP* fragment was performed to obtain sufficient amount of PCR product for cloning. The reaction mix for the PCR reaction is shown in table 1. B95-8 whole cell DNA (250

ng), purified according to section 3.4.2 was used as the template and the PCR reaction was carried out using a thermal cycler (Applied Biosystems GeneAmp PCR System 2400) with the parameters as shown in table 2. A total of 35 PCR cycles were run with an initial step of pre-PCR cycle heating at 95°C and a final holding step of 4°C. Cycle steps are indicated in bold.

Reagents	**Final concentration**
Template DNA	Specified in text
PCR Buffer	1X
Forward primer (100μM)	0.8 μM
Reverse primer (100μM)	0.8 μM
dATP, dCTP, dGTP, dTTP	500 μM each
Polymerase	0.025 u/μl
H_2O	q.s as specified in text

Table 1. PCR reaction mix for amplifying *ori*P

Step	**Temperature**	**Time**
Hold	95°C	15 minutes
Denaturation	**94°C**	**45 seconds**
Annealing	**Specified in text**	**45 seconds**
Extension	**72°C**	**1 minute**
Final extension	72°C	10 minutes
Hold	4°C	∞

Table 2. PCR reaction conditions.

3.1.2 Gel purification of PCR products

After amplification by PCR, the products were run on a 0.8 % agarose gel (appendix)

and the DNA band that had the closest fragment size to the expected size of *oriP* was purified using Qiagen Gel purification Kit (Qiagen) and eluted using 50 µl of 10 mM Tris pH8.0 solution. After gel purification, the amount of PCR end product was estimated by measuring the absorbance of the PCR product at 260 nm using a spectrophotometer.

3.1.3 Non-gel based purification of enzyme reactions

Alternatively, instead of using gel purification, an easier way would be to utilize the Qiagen QiaQuick PCR purification kit (Qiagen) where the product does not need to be run on ethidium bromide containing agarose gel, thereby reducing any contamination with agarose or ethidium that may interfere with downstream steps.

3.2 Restriction enzyme digestion

Restriction enzyme digestion is a vital step in the cloning done throughout this entire study. All the restriction enzymes utilized in this study are from New England Biolabs and the buffers used are as recommended by the manufacturer. To achieve as complete a digestion by the enzymes as possible, the maximum units of enzymes are usually added, except for the reactions involving known enzymes with known non-specific activity when used at high concentrations. In those cases, only 1 unit of enzyme per µg of plasmid was applied. Reaction conditions were usually at 37ºC for at least 5 hours to ensure complete digestion.

For double restriction enzyme digests, certain enzymes share buffers which they are both active in. In cases where neither enzymes share a common buffer, sequential restriction enzyme digestion was conducted. The plasmid would first be digested with

one enzyme and then the buffer would be changed to an appropriate one using the Qiagen QiaQuick PCR purification procedure (section 3.1.3) before digestion with the second enzyme.

The tables of all the restriction enzyme reactions can be found in appendix.

3.3 Filling in of 5' overhang and removal of 3' overhang of restriction-digested plasmid DNA

Filling in of 5' overhangs and removal of 3' overhangs was carried out using the T4 DNA polymerase. The purified digested plasmid DNA (50 µl) was mixed together with the polymerase, BSA and dNTPs as shown in table 3 to a final reaction volume of 60 µl. The reaction was incubated at 12ºC for 1 hour before heat inactivating the enzyme by treating the reaction at 75ºC for 10 minutes.

Reagent	Final concentration
DNA sample	Specified in text
BSA(10mg/ml)	50 µg/ml
10mM dNTP	100 µM
10X T4 polymerase buffer	1X
T4 polymerase	0.1 u/µl
H_2O	q.s as specified in text

Table 3. T4 DNA polymerase reaction mix

3.4 Ligation of the insert and plasmid vector DNA

Ligation was another vital reaction used in the cloning steps. The insert DNA was ligated to the plasmid vector DNA using T4 ligase (New England Biolabs). The proportion of vector to insert was kept at a 1:3 molar ratio. The reaction was set up as

described in table 4 below and incubated at 16ºC overnight using a thermal cycler.

Reagent	Final concentration
Vector	Specified in text
Insert	Specified in text
Ligase	40 c.e.u/µl
10X Ligase buffer	1X
H_2O	q.s as specified in text

Table 4. Ligation reaction mix c.e.u= cohesive end units

3.5 Manipulation of *Escherichia coli* DH10B strain

3.5.1 Preparation of electrocompetent DH10B cells

A single colony from a streak plate of DH10B was picked and used to inoculate 100 ml of LB broth. This starter culture was incubated at 37ºC with shaking at 200 rpm overnight in a shaker. (Innova 4300 incubator shaker by New Brunswick Scientific) The overnight starter culture was then used to further inoculate 1 L of LB broth. The 1 L of culture was then placed into a shaker and incubated at 37ºC with shaking at 200 rpm. O.D readings at 600 nm were taken every 30 minutes and incubation terminated when a suitable O.D reading of about 0.5A to 0.7A was observed. The cells were incubated on ice for 15-30 minutes before being centrifuged at 1500 *g* (Kubota 2000) for 15 minutes at 4ºC. After centrifugation, the supernatant was carefully decanted away and the cell pellet re-suspended in 1 L of sterile 10 % ice-cold glycerol. After re-suspension, the cells were centrifuged again at 1500 *g* for 15 minutes at 4ºC. The supernatant was again carefully discarded after centrifugation and the cells re-suspended in 20 ml of sterile 10 % ice-cold glycerol. The re-suspended cells were centrifuged one last time at 1500 *g* for 15 minutes at 4ºC and finally re-suspended in 2

ml of sterile 10 % ice-cold glycerol. 50 µl aliquots were dispensed into sterile 1.5 ml micro-centrifuge tubes for a single use tube of electro-competent DH10B cells and kept at a -80ºC freezer for long term storage.

3.5.2 Electro-transformation of electrocompetent DH10B cells

1 ng of supercoiled pcDNA3.1+ plasmid vector (Invitrogen) was used to test the efficiency of the electro-competent DH10B cells. The plasmid was first diluted to a concentration of 1 ng/µl using 10 mM Tris, pH 8.0. One tube of electro-competent cells was then allowed to thaw on ice before adding 1 µl of plasmid solution. This was mixed well by gently pipetting the cell suspension up and down. The cells were dispensed into a pre-chilled 1 mm cuvette (BioRad) and electro-transformed at 1800 V, 25 µF, 200 Ω in an electro-porater (BioRad). 450 µl of SOC broth was added immediately after electro-transformation and the reconstituted cells were left at 37ºC for 1 hour with shaking. The cells were plated onto LB agar plates containing 50 µg/ml of ampicillin and incubated overnight at 37ºC.

3.5.3 Preparation of small amount of plasmid

Individual colonies that grew overnight on the LB + ampicillin agar plate were picked using a sterile pipette tip. Each colony was suspended in 3 ml of LB broth with 50 µg/ml of ampicillin (appendix). 10 µl of the resuspended bacteria was streaked on a fresh LB + ampicillin agar plate using a sterile loop to obtain a purity plate of the transformants. The LB + ampicillin broth inoculated with individual colonies were then incubated at 37°C overnight with shaking at 220 rpm until a confluent bacterial culture was obtained as a preliminary culture for plasmid analysis. The bacterial cells

were harvested from this culture by centrifuging at 1500 *g* using a centrifuge (Beckman) for 10 minutes and the plasmid extracted using Qiagen plasmid miniprep kit (Qiagen).

3.5.4 Preparation of higher amount of plasmid

Upon confirming the identity of the plasmid, a larger amount of the plasmid was obtained. A single colony was picked from the purity plate to inoculate 150 ml of LB broth + 50 µg/ml ampicillin and incubated at 37°C with shaking at 220 rpm overnight. 800 µl of the overnight culture was mixed with 70 µl of DMSO and frozen at -80°C for long-term storage. The remaining culture was harvested by centrifugation at 1500 *g* for 15 minutes and the plasmid was extracted using HiSpeed Plasmid Maxi Kit (Qiagen) to obtain a final yield of approximately 1 mg of plasmid DNA. Further analysis using the appropriate restriction enzymes (section 3.2) was conducted on the plasmid isolated. The plasmid was then placed in a refrigerator at -20°C for storage.

3.6 Maintenance and manipulation of B95-8 or BJAB cell line

The B95-8 cell line was grown in 20 ml of R10 media (appendix) incubated at 37°C with 5 % CO_2 and passaged every time the culture reached a cell density of 10^6 cells/ml. No trypsin was added, as both B95-8 and BJAB are suspension cell lines although B95-8 does exhibit weak adherent characteristics. By using a sterile pipette, R10 media was used to wash any cells that remained adhered to the culture flask even after gentle shaking and the re-suspended cells were diluted twenty-fold using fresh R10 buffer to a final cell density of approximately 2 X 10^5 cells/ml.

3.6.1 Cell count using a hemocytometer

Cells were well suspended by pipetting the culture up and down several times using a sterile pipette. 50 µl of suspended cells were diluted 2 X by mixing with 50 µl of trypan blue and the mixture was loaded onto a hemocytometer counting chamber (American Optics) and viewed under a phase contrast light microscope (Olympus IX51). Cells found within the four large squares at each of the corners of the counting chamber were counted and a pacer was used to help keep track of the count of viable cells, which appeared as bright spots under the microscope field. The cell density of the culture was calculated as follows:

$$\frac{\text{Total cell count in large squares in counting chamber}}{4} \text{ X } 2 \text{ X } 10^4 = \text{cell density of cell culture (cells/ml)}$$

3.6.2 Extraction of genomic DNA from B95-8

A total of 180 ml of B95-8 culture was grown to confluence. Upon reaching confluent growth, the cells were harvested first by using a sterile pipette to re-suspend all cells. The re-suspended cells were collected in 50 ml sterile propylene tubes and centrifuged at 300 *g* for 10 minutes. The genomic DNA from the resulting cell pellet was extracted from the harvested cells using the Qiagen Genomic DNA kit. 500 µg of genomic DNA obtained from B95-8 cell line was dissolved in sterile TE buffer to a final concentration of 0.5 µg/µl. 1 µl of the final DNA prep was taken for analysis with gel electrophoresis using a 0.4 % agarose gel in TAE buffer (appendix) with an applied voltage of 50 V for 3 hours.

3.6.3 Transfection of B lymphocyte cell lines B95-8 and BJAB

The B95-8 cell line was chosen as the cell line to be transfected because it actively expresses the viral protein EBNA-1 and can be transfected more efficiently than other EBV positive cell lines. B95-8 is also a permissive cell line that results in a small number of cells entering the lytic cycle. BJAB was chosen as it did not express any form of viral antigen like EBNA-1 and would not support the replication of *oriP* containing plasmids. For both cell lines, cells were grown until mid-log phase, which has an approximate cell density of 0.5 X 10^6 cells/ml, before being harvested for transfection by electroporation. 20 ml or 10^7 of these mid-log phase cells were used for each transfection cuvette. The 20 ml cells were pipetted into sterile 50 ml propylene tubes and harvested by centrifugation at 300 *g* for 6 minutes. 400 µl of R10 media was used to resuspend the cell pellet before transferring to the cuvette. The volume of DNA added, regardless of the concentration of DNA did not exceed 20 µl as the buffer in which the DNA was dissolved could affect the overall salt concentration in the cuvette and resulted in a less efficient transfection. Electroporation of the cells was done using BioRad GenePulser II with a selected voltage and the capacitance set at 950 µF. After electroporation, the cells were quickly re-constituted in 20 ml of fresh R10 media and incubated at 37ºC with 5% CO_2.

3.6.3.1 Optimization of transfection efficiency

In the optimization of the transfection parameters, a GFP expressing plasmid pTracer (Invitrogen) was transfected into B95-8 and BJAB cells using different electroporator voltage settings. The transfected cells were incubated for 1 day at 37°C with 5% CO_2. 1.5 ml of the recovered cells was collected in a 1.5 ml microcentrifuge tube and harvested by centrifuging at 1800 rpm for 6 minutes. The supernatant was removed

and the cell pellet was re-suspended in 50 µl of non-sterile 1 X PBS. 1 ml of the re-suspended cells was placed onto an immunofluorescence slide and viewed under an immunofluoroscence microscope. The voltage with which the transfected cells exhibited the most intense fluorescence with the least amount of cell death was determined to be the most optimal voltage for transfection.

3.7 Southern and Northern blot analysis

50 µg of plasmid DNA was transfected into B95-8 and BJAB cells and the plasmids rescued after 72 hours and analyzed by southern blot to determine the proportion of plasmid that could replicate.

3.7.1 Transfection of B95-8 and BJAB cells

Transfection of B95-8 and BJAB cells was carried out as described in section 3.6.3 using 50 µg of plasmid. The parameters used to transfect B95-8 and BJAB cells were 250 V, 950 µF. After electroporation, the cells were quickly re-constituted in 20 ml of fresh R10 media and incubated at 37 ºC with 5 % CO_2 for 72 hours.

3.7.2 Isolation of whole cell RNA and DNA from transfected samples

After 72 hours, the transfected cells were harvested by centrifuging at 1800 rpm for 5 minutes. The supernatant was carefully removed without disturbing the cell pellet by using a sterile filtered pipette tip. Whole cell RNA and DNA were then extracted using the RNA and DNA isolation kit (Qiagen). As RNA is sensitive to RNase, gloves were worn throughout the entire procedure and all apparatus were cleaned with RNase AWAY (Amersham). The RNA obtained was redissolved in 5 µl of RNase free water

and then quantified by measuring the absorbance at 260 nm. An additional step of treating the RNA with RNase free DNase I (Qiagen) was necessary to remove any contaminating DNA. The DNase I reaction was set up as described in table 5 below to a final reaction volume of 6 µl and incubated at 37 ºC for 1 hour. The treated RNA was then stored at -20ºC and saved for further Northern Blot analysis. The whole cell DNA isolated from the kit was re-suspended in 18 µl of 10 mM Tris.Cl, pH 8.0 and further digested with both *Dpn* I and another suitable restriction enzyme (to linearize the plasmid) in a double enzyme digestion reaction.

Reagent	Final concentration
RNA	Specified in text
10X NEBuffer 2	1X
RNase free DNase I	0.3 µl

Table 5. RNase free DNase I digestion mix

3.7.3 Preparation of labeled probes

Three types of labeled probes were made using the neomycin phosphotransferase gene fragment from pcDNA3.1+, the EBV latent origin of replication *oriP* DNA and GAPDH cDNA. The neomycin resistance gene fragment was excised from 10 µg of pcDNA3.1+ plasmid DNA using the restriction enzymes *Stu* I and *Sal* I. The reaction was set up as described (appendix) to a final volume of 50 µl and incubated at 37 ºC for 5 hours. The neomycin resistance gene fragment was purified using gel extraction as described in section 3.1.2. *OriP* DNA was obtained as described in section 4.1.1.3 using *Sac* II and *Eco* RI (appendix). GAPDH cDNA was obtained by first performing reverse transcription on total RNA and then amplifying out the product using 2 sets of primers as described below.

3.7.3.1 Reverse transcription-PCR

Firstly, total RNA was isolated from BJAB cells as described in section 3.7.2. RT-PCR was carried out using QIAGEN One-step RT-PCR kit (Qiagen). The reaction was set up as suggested by the manufacturer and can be seen in table 6 below to a final volume of 50 µl. The thermal parameters used were also suggested by the manufacturer and can be seen in table 7 below. A total of 35 cycles were run. The RT-PCR products were analysed using agarose gel electrophoresis.

Reagent	Final concentration
5X Qiagen One step RT-PCR buffer	1 X
10 mM dNTPs	400 µM
Forward primer	0.6 µM
Reverse primer	0.6 µM
Qiagen One step RT-PCR enzyme mix	0.2 u/µl
RNA template	30 ng/µl
RNase free water	q.s as specified in text

Table 6. RT-PCR reaction mix

Step		Temperature	Time
Reverse transcription		50 ºC	30 minutes
Initial PCR activation step		95 °C	15 minutes
3 step cycling	**Denaturation**	**94 °C**	**45 seconds**
	Annealing	**55 ºC**	**45 seconds**
	Extension	**72 °C**	**1 minute**
Final extension		72 °C	10 minutes
Hold		4 °C	∞

Table 7. One Step RT-PCR thermal cycling conditions

3.7.3.2 Purification of cDNA products

Purification of GAPDH cDNA products was performed as described in section 3.1.2 using gel purification. The eluate was stored at -20ºC before downstream reactions.

3.7.3.3 Labeling reaction

The creation of labeled probes was done using the *Gene Images* Random Prime Labelling Module (Amersham Biosciences). The template DNA was first diluted to 25 ng/µl using 10 mM Tris.Cl, pH 8.0 to a final volume of 100 µl. The nucleotide mix, primers and water supplied by the kit were allowed to thaw on ice and the enzyme left in the -20ºC freezer. The DNA template was then denatured by heating in boiling water bath for 5 minutes before chilling on ice. The labeling reaction was set up as shown in table 8 below. The enzyme was added last to the reaction and gently mixed by pipetting up and down using filtered pipette tips. The reaction was incubated at 37 ºC for 1 hour and stopped by adding EDTA to a final concentration of 20 mM. The labeled probes were then stored in the dark at -20ºC.

Reagent	Final concentration
DNA template	50 ng
Nucleotide mix	10 µl
Primers	5 µl
Enzyme solution (5u/µl)	1 µl
Water	q.s to 50 µl

Table 8. Reaction mix for labeling reaction

3.7.4 Southern Blot

The restriction enzymes treated DNA was loaded onto a 0.8 % agarose gel and

subjected to electrophoresis at 100 V for 80 minutes. This was to allow for sufficient separation of the DNA fragments. The gel was then viewed under a UV trans-illuminator and any redundant gel removed. A picture was also quickly taken for future reference. The DNA in the gel was then depurinated by washing the gel in 250 mM hydrochloric acid for 10 minutes with gentle agitation. After depurination, the DNA was denatured using denaturation buffer (appendix) for 25 minutes with gentle agitation before neutralizing with neutralization buffer (appendix). Adequate buffer to cover the gel was added for these three steps and a rinse step with distilled water was included in between washes. The DNA was then transferred onto a marked nylon membrane (Amersham Biosciences) by capillary action. To do this, the gel was first placed face down onto 3 mm Whatman paper acting as a wick in a reservoir containing 10 X SSC buffer. The nylon membrane which was approximately larger than the gel by 1 cm in both width and length was positioned on top of the gel with the marked surface facing up. Six pieces of 3 mm Whatman paper cut to the size of approximately 1 cm smaller than that of the nylon membrane in both width and length was placed on top of the nylon membrane. Lastly, stacks of paper towels cut to about 0.5 cm smaller than the 3 mm Whatman paper and stacked up to a height of approximately 5 cm was placed on the six pieces of 3 mm Whatman paper. A weight was then placed on top to provide pressure and capillary transfer was allowed to proceed overnight.

The next day, the capillary apparatus was disassembled and the DNA was fixed onto the nylon membrane by cross-linking. Cross-linking was done by subjecting the membrane (marked surface facing down) to UV (VilberLourmat BLX-254, 0.120 joules). The hybridization buffer (appendix) was thawed and pre-heated to 60ºC using

a hybridization oven. The amount of hybridization buffer used was about 0.125 ml per cm^2 of membrane. After cross-linking, the blot was wetted using 5X SSC buffer and transferred to the hybridisation buffer for a pre-hybridisation step of approximately 4-5 hours. The labeled probed were thawed and denatured by heating in a boiling water bath for 5 minutes before chilling in ice. The amount of probe added was dependent on the volume of the hybridization buffer, with the concentration of probe added roughly about 10 ng per ml of hybridisation buffer used. Care was also taken to make sure that the probes were not added directly onto the membrane. Hybridisation was allowed to take place in the oven at 60ºC overnight. After the hybridization step, the membrane was subjected to a stringency wash using 1 X SSC, 0.1 % SDS at 60ºC for 5 minutes with a total of 3 washes. Fresh buffer was used for every wash. A second stringency wash using 0.5 X SSC, 0.1 % SDS was also conducted for 5 minutes with a total of 3 washes at 60ºC. Fresh buffer was also used for every wash. After the stringency washing step, the membrane was placed in blocking buffer and incubated at room temperature with shaking for 1 hour, before transferring the membrane to the antibody binding solution. Approximately 0.75 ml to 1.0 ml of blocking buffer per cm^2 of membrane was used and approximately 0.3 ml of antibody binding buffer per cm^2 of membrane was used. The antibody was diluted 5000 fold in the antibody binding solution. Subsequently, the membrane was washed with 0.3 % (v/v) Tween 20 in buffer A (appendix) for a total of 3 washes each 10 minutes long, using fresh wash buffer in between washes. Excess wash buffer was removed by touching the corner of the membrane onto a clean piece of cling wrap and placed on a flat clean plastic tray. The amount of detection reagent added was approximately 40 µl per cm^2 of membrane and left at room temperature for 2 to 5 minutes. Excess detection reagent was drained off by again touching the corner of the membrane onto a clean

piece of cling wrap and transferred into a clean hybridization bag and heat sealed. The bag containing the membrane was then exposed to film (Fujifilm) in a dark room for an appropriate amount of time and developed using the developer found in CRC (Kodak)

3.7.5 Northern Blot

As RNA is susceptible to degradation to RNase which can be found on un-clean surfaces or hands, gloves were worn throughout the entire procedure and all apparatus used were cleaned with RNaseZap (Amersham). RNase-free 1.5 ml and 2.0 ml microcentrifuge tubes (Axygen) were used and only RNase free water or milliQ grade water was used to prepare buffers and reactions. A 0.8 % FA gel containing 4% formaldehyde (appendix) was cast and placed in running buffer (appendix). The RNA samples were then mixed with sample loading dye (appendix) before heated at 60ºC for 15 minutes to remove any secondary RNA structures. After heating, the samples were loaded into the gel and subjected to electrophoresis at 80 V until the bromophenol blue dye had ran to a suitable distance in the gel. Lanes containing the RNA samples were excised using a clean surgical blade and washed with excess RNase free water with gentle agitation for 30 minutes to dilute out the formaldehyde. The gel was then washed in Alkaline Buffer A (appendix) for 30 minutes with gentle agitation before neutralizing with Neutralization buffer A (appendix) for 30 minutes with gentle agitation. All the wash steps were done in room temperature. Finally, the gel was soaked in 10 X SSC buffer for 5 minutes. This step was repeated once using fresh 10 X SSC buffer. To transfer the RNA from to gel to a nylon membrane, the gel was first placed face down onto 3 mm Whatman paper acting as a wick in a reservoir containing 10 X SSC buffer. The nylon membrane which was approximately larger

than the gel by 1 cm in both width and length was positioned on top of the gel with the marked surface facing up. Six pieces of 3 mm Whatman paper cut to the size of approximately 1 cm smaller than that of the nylon membrane in both width and length was placed on top of the nylon membrane. Lastly, stacks of paper towels cut to about 0.5 cm smaller than the 3 mm Whatman paper and stacked up to a height of approximately 5 cm was placed on the six pieces of 3 mm Whatman paper. A weight was then placed on top to provide pressure and capillary transfer was allowed to proceed overnight.

The next day, the capillary apparatus was disassembled and the RNA was fixed onto the nylon membrane by cross-linking. Cross-linking was done by subjecting the membrane (marked surface facing down) to UV. The hybridization buffer (appendix) was thawed and pre-heated to 65ºC using a hybridization oven. The amount of hybridization buffer used was about 0.125 ml per cm^2 of membrane. After cross-linking, the blot was wetted using 5 X SSC buffer and transferred to the hybridisation buffer for a pre-hybridisation step of approximately 4-5 hours. The labeled probes were thawed and denatured by heating in a boiling water bath for 5 minutes before chilling in ice. The amount of probe added was dependent on the volume of the hybridization buffer, with the concentration of probe added roughly about 10 ng per ml of hybridisation buffer used. Care was also taken to make sure that the probes were not added directly onto the membrane. Hybridisation was allowed to take place in the oven at 65ºC overnight. After the hybridization step, the membrane was subjected to a stringency wash using 1 X SSC, 0.1 % SDS at 60ºC for 5 minutes with a total of 3 washes. Fresh buffer was used for every wash. A second stringency wash using 0.1 X SSC, 0.1 % SDS was also conducted for 5 minutes with a total of 3 washes at 65ºC.

Fresh buffer was also used for every wash. After the stringency washing step, the membrane was placed in blocking buffer and incubated at room temperature with shaking for 1 hour, before transferring the membrane to the antibody binding solution. Approximately 0.75 ml to 1.0 ml of blocking buffer per cm^2 of membrane was used and approximately 0.3 ml of antibody binding buffer per cm^2 of membrane was used. The antibody was diluted 5000 fold in the antibody binding solution. Subsequently, the membrane was washed with 0.3 % (v/v) Tween 20 in buffer A (appendix) for a total of 3 washes each 10 minutes long, using fresh wash buffer in between washes. Excess wash buffer was removed by touching the corner of the membrane onto a clean piece of cling wrap and placed on a flat clean plastic tray. The amount of detection reagent added was approximately 40 µl per cm^2 of membrane and left at room temperature for 2 to 5 minutes. Excess detection reagent was drained off by again touching the corner of the membrane onto a clean piece of cling wrap and transferred into a clean hybridization bag and heat sealed. The bag containing the membrane was then exposed to film (Fujifilm) in a dark room for an appropriate amount of time and developed using the developer found in CRC (Kodak).

3.7.6 Reprobing of membrane

Prior to reprobing, the membrane is soaked in 5 X SSC for 1-2 minutes. After soaking, the membrane is added to boiling 0.1 % (w/v) SDS for 10 minutes with shaking, using approximately 5 ml of SDS per cm^2 of membrane. The procedure is performed a total of three times, each time using fresh 0.1 % (w/v) SDS. After washing, proceed from pre-hybridisation onwards.

3.8 *In-vitro* transcription

In-vitro transcription was done using Promega's HelaScribe® Nuclear Extract *in vitro* Transcription System. The reaction was set up as recommended in the manufacturer's protocol, as shown in table 9 below. rNTPs were added last with a preincubation period of 30 minutes at 30ºC consisting of just the template, $MgCl_2$, nuclear extract, transcription buffer and RNase free water. This was to allow the transcription complex to first bind onto the promoters. After 30 minutes of pre-incubation, rNTPs were added and incubated at 30ºC for a specified time.

Reagent	Final concentration
Hela Nuclear Extract 1X transcription buffer	7.4 µl
Nuclear extract	3.6 µl
$MgCl_2$	8 µM
Template DNA	24 ng/µl
RNase free water	q.s 25 µl
rNTPs	as specified in text

Table 9. Reaction mix for *in-vitro* transcription

After the reaction, any RNA transcripts were isolated according to the manufacturer's protocol and treated with RNase free DNase I (table 5) to remove any plasmid template.

3.8.1 *In-vitro* transcription optimization experiments

Two optimization experiments were first conducted to help identify the optimal conditions for *in-vitro* transcription. A time course study as well as rNTP concentration optimization experiment was carried out, using pEGFP-lacZ as the

template DNA. For both experiments, 600 ng of template was added per 25 µl reaction volume.

3.8.1.1 Preparation of template DNA

pEGFP-lacZ digested with the suitable restriction enzymes for use as template. 10 µg of plasmid DNA was digested in a final reaction volume is 50 µl. The digested plasmid was then purified using the Qiagen QiaQuick PCR purification procedure (section 3.1.3) and placed in -20ºC freezer for long term storage.

3.8.1.2 Time course study

The *in-vitro* transcription reaction was set up as described above in section 3.8 and table 8. Upon addition and mixing of rNTPs, an initial aliquot was taken at 30 seconds. The reaction was incubated at 30ºC. Aliquots were taken at 5, 10, 15, 20, 25, 30, 45, 60, 75 and 90 minutes and transferred to separate tubes containing stop buffer (appendix) to terminate the reaction. Northern blot analysis was done as described in section 3.5.5.

3.8.1.3 rNTP concentration optimization experiment

The *in-vitro* transcription reaction was set up as described above in section 3.8 and table 8. The concentrations of rNTPs used were 10 µM, 25 µM, 50 µM, 100 µM, 200 µM and 400 µM. The reaction was allowed to proceed at 30ºC for 20 minutes before the addition of an equal volume of 40 µM EDTA to terminate the reaction. RNA isolated as described in the manufacturer's protocol.

3.9 Size exclusion chromatography

To determine if the RNA generated from in vitro transcription is free or in large molecular complexes, size exclusion chromatography was performed using Sephacryl-S1000 (Amersham). The column was packed into 1 ml sterile pipettes (Falcon). A 1 ml pipette tip was used as the reservoir. Elution was by gravity and the fractions collected manually into nuclease free 1.5 ml microcentrifuge tubes.

3.9.1 Packing and calibration of the column

The Sephacryl-S1000 beads were first re-suspended in elution buffer (appendix). Before casting the column, the cotton stopper in the 1 ml sterile pipette was first re-positioned to the tip of the pipette using a vacuum pump. The tip of the pipette was further shortened so as to reduce the dead space under the column. This was to minimize band broadening, which happens when the eluate undergoes non-laminar flow within this dead space. RNaseZAP ® (Ambion) was first applied to the openings of the column to remove any possible RNase contamination. A schematic diagram of the column can be seen in figure 5 below.

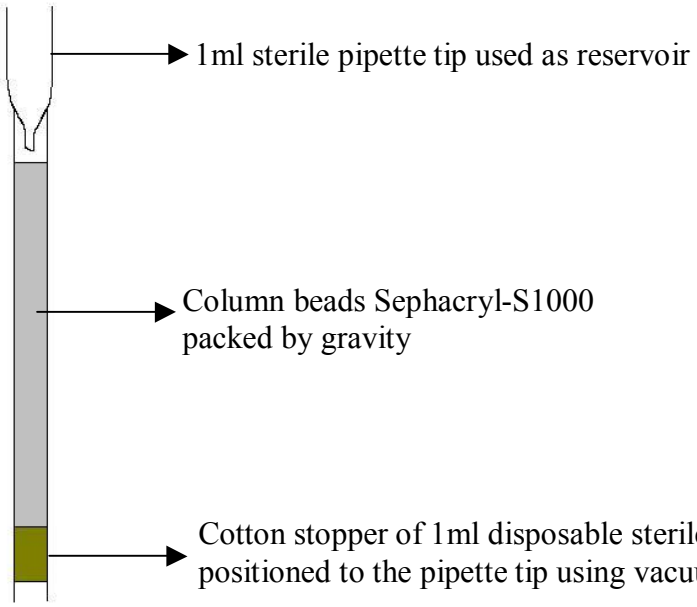


Fig. 5. Schematic diagram of column used for size exclusion chromatography
After packing, the column was first calibrated before it could be used for size exclusion chromatography. 2 X the column length of elution buffer was allowed to flow through the column by gravity before the sample was loaded.

3.9.2 Loading of sample onto column

As the meniscus of the liquid phase reached the top surface of the column, the sample was quickly but carefully loaded. The sample was then allowed to flow into the column before more elution buffer was added.

3.9.2.1 Collection of fractions and nucleic acid precipitation

Fraction collection only started on the 21st drop. The fractions were collected into nuclease free 1.5 ml microcentrifuge tubes containing 0.3 M NaAc and 20 µg of yeast tRNA. After collection, the fractions were subjected to phenol:chloroform treatment to remove any contaminating proteins. 1 X final fraction volume of phenol:chloroform was added to each fraction and vortexed for 1 minute before centrifuging at 14,000 rpm for 5 minutes. After centrifugation, the aqueous layer was carefully removed and placed in a new nuclease-free 1.5 ml microcentrifuge tube. 2.5 X the aqueous layer volume of absolute ethanol was added and the fractions were then

stored at -70ºC for at least 15 minutes, before being centrifuged at 14,000 rpm for 15 minutes at 4ºC. After centrifugation, the supernatant was carefully removed by aspiration. 300 µl of 70% ethanol was added as a washing step to remove any phenol:chloroform carryover. After adding 300 µl of 70% ethanol, the fractions were again centrifuged at 14,000 rpm for 15 minutes before the supernatant removed by aspiration, with care taken to ensure that the RNA pellet was not accidentally removed at the same time. The precipitated nucleic acid pellet was allowed to air dry for 10 minutes before being resuspended in 5µl of 1 X NEB Buffer 2. The resuspended RNA pellet was further treated with RNase-free DNase I as described in section 3.7.2. 1 µl of the treated was then used for Northern Blot analysis as described in 3.7.5.

Results

4.1 Replication of EBV *ori*P-containing plasmids in EBNA-1-expressing cells

During the latent infection stage of the Epstein-Barr Virus (EBV), only two components of the virus are needed for replication. The latent origin of replication *ori*P, together with the Epstein-Barr virus nuclear antigen (EBNA-1) have been shown to be capable of mediating the replication and long-term persistence of the virus genome as an episome in the host cell (Yates *et al*., 1984), making sure that the episome replicates synchronously once per cell cycle and is properly partitioned into daughter cells. In this study, I first placed *ori*P in varying locations and orientations in the pcDNA3.1+ plasmid (Invitrogen) and made selective deletions of various promoters from the plasmid to understand the effect of transcription on replication in vivo. The useful plasmids constructed and their distinctive features are listed in table 10 below. The structural elements of these plasmids and their parent plasmid pcDNA3.1+ are depicted in figures 6 to 10.

Name	Distinctive features
p-oriP-S	*ori*P outside transcription unit
p-S-oriP	*ori*P inside transcription unit: FR more proximal to the promoter than DS
p-S-oriP.1	*ori*P inside transcription unit: DS more proximal to the promoter than FR
p-oriP-ΔS	Without known transcription promoter (SV40-promoter deleted)

Table 10. Plasmids designed for in vivo study of effect of transcription on *ori*P-dependent replication. A transcription unit is defined as the region from a transcription enhancer/promoter to the downstream polyadenylation signal. The promoter for the transcription unit in the first three plasmids listed is the SV40 early promoter. The plasmids listed do not contain the CMV immediate early promoter, which is present in their parent plasmid pcDNA3.1+ but removed during their construction (see below).

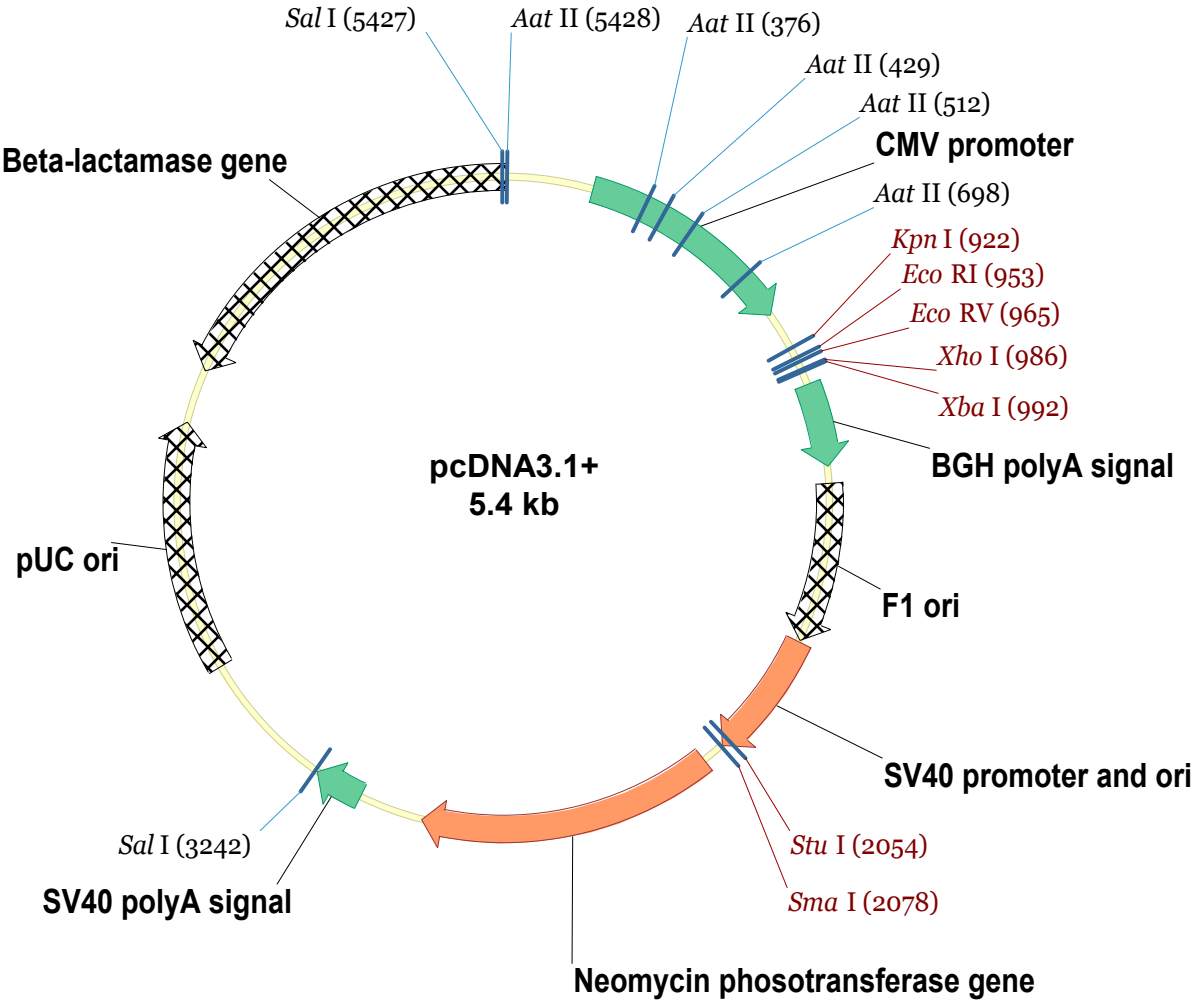


Figure 6. Map of plasmid pcDNA3.1+. This plasmid contains a CMV immediate early promoter that is used for expressing the gene of interest, and SV40 early promoter that drives the expression of the neomycin phosotransferase gene for selection in transfected mammalian cells. Important restriction enzyme sites are shown and their positions on the plasmid in parenthesis.

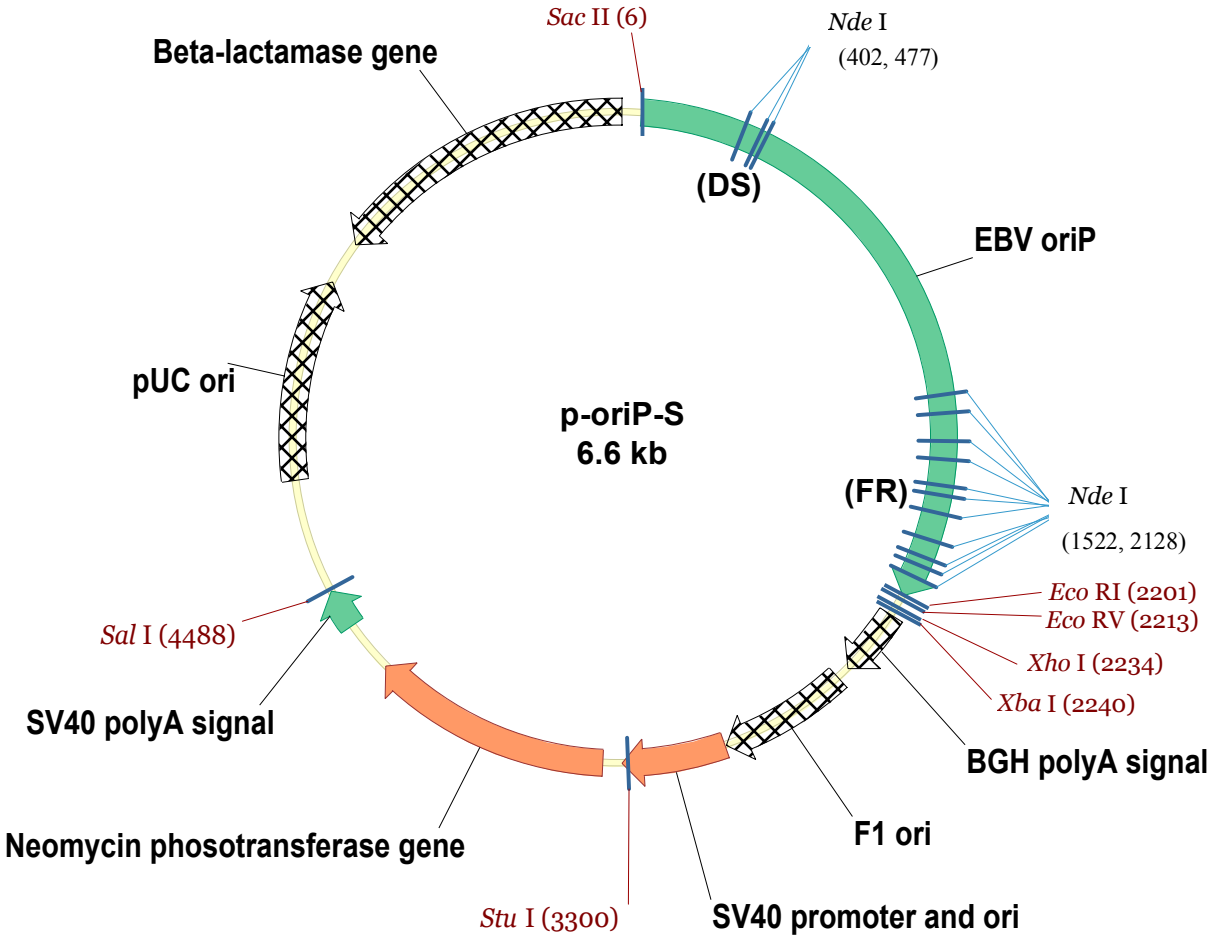


Figure 7. Map of plasmid p-oriP-S. In this construct, the CMV promoter was removed from pcDNA3.1+. *Ori*P was inserted outside of the SV40-promoter-driven transcription unit and approximately 2 kb downstream of the SV40 polyadenylation signal. Important restriction enzyme sites are shown and their positions on the plasmid in parenthesis. For *Nde* I sites, only the first and the last sites are represented.

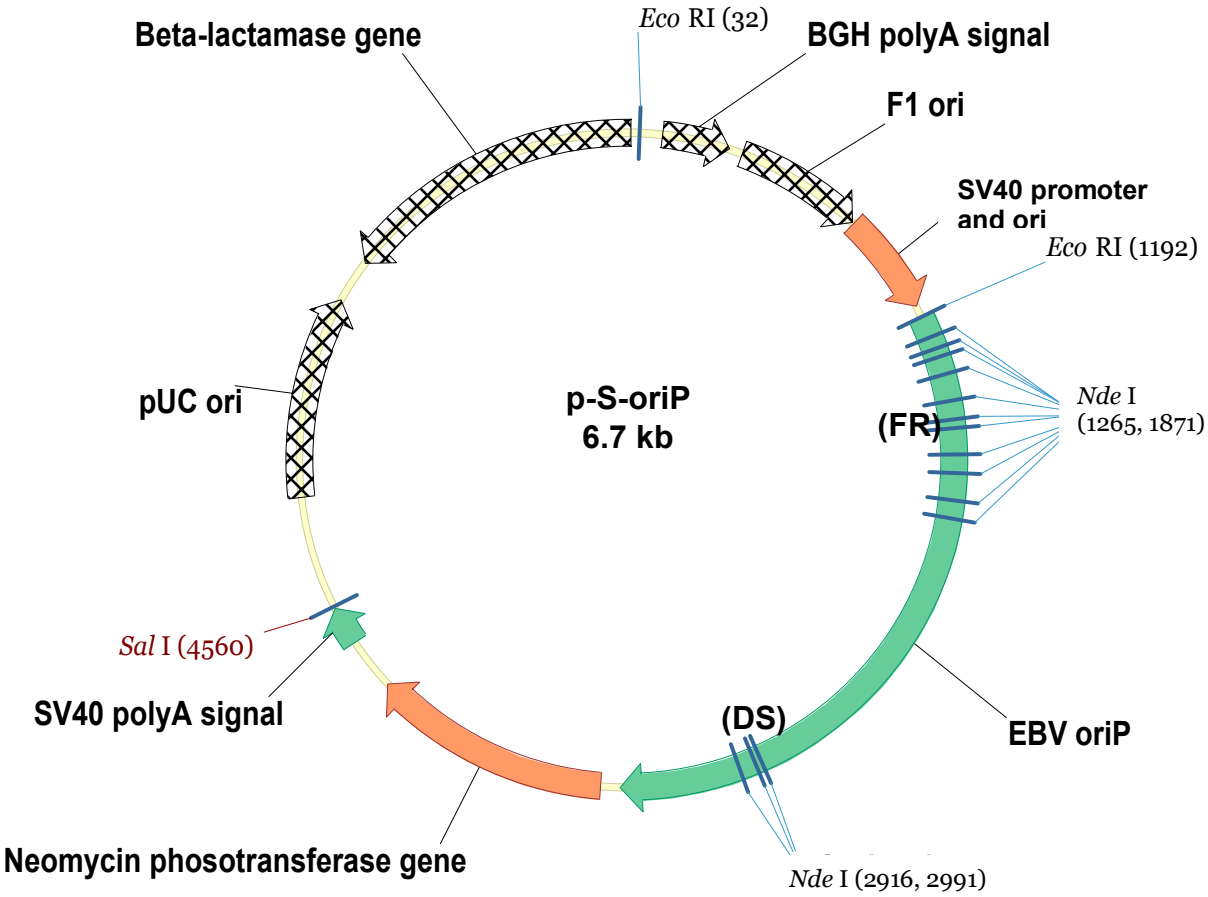


Figure 8. Map of plasmid p-S-oriP. In this construct, the CMV promoter was also removed from pcDNA3.1+ whereas the *ori*P was inserted into the SV40-promoter-driven transcription unit. The FR is closer to the SV40 promoter; located about 73 base pairs downstream while the DS is located approximately 1.9 kb further downstream. Important restriction enzyme sites are shown and their positions on the plasmid in parenthesis. For *Nde* I sites, only the position of the first and last sites are shown.

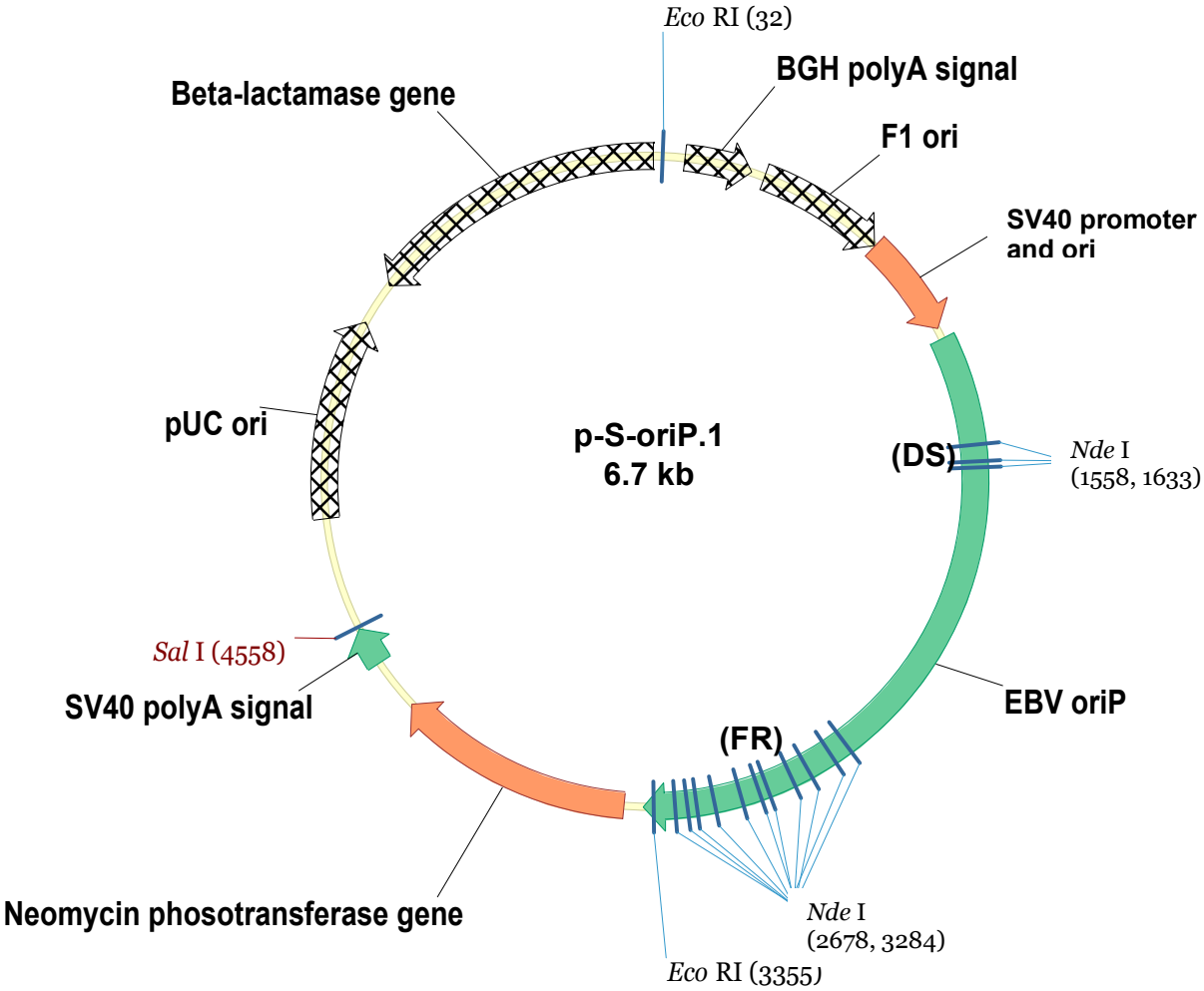


Figure 9. Map of plasmid p-S-oriP.1. This construct is similar to p-S-oriP, except that the orientation of *ori*P is different. In this case, the DS is closer to the SV40 promoter, located approximately 400 base pairs downstream, while the FR is approximately 1.5 kb further from the promoter. Important restriction enzyme sites are shown and their positions on the plasmid in parenthesis. For *Nde* I sites, only the position of the first and last sites are shown.

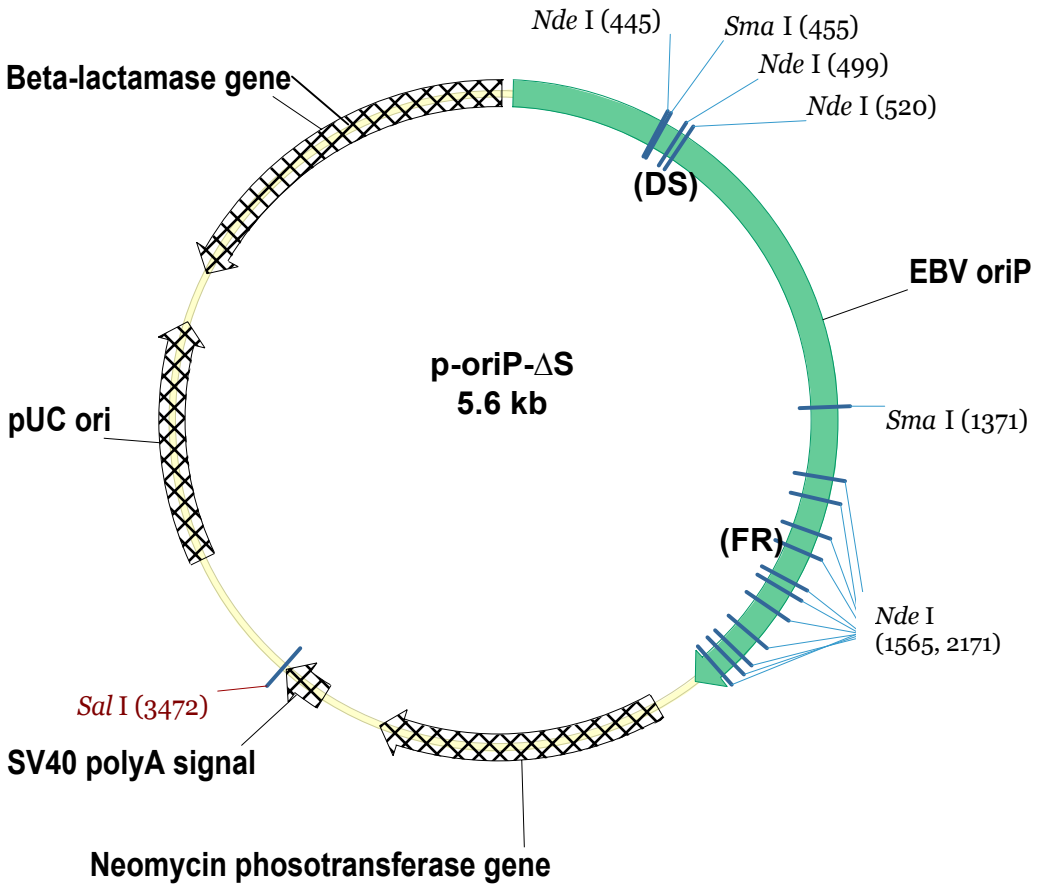


Figure 10. Map of plasmid p-oriP-ΔS. In this construct, both the CMV and SV40 promoters were removed, thus eliminating any form of possible eukaryotic transcription on this vector. Important restriction enzyme sites are shown and their positions on the plasmid in parenthesis. For *Nde* I sites, only the position of the first and last sites are shown.

4.1.1 Construction of *ori*P-containing plasmids

4.1.1.1 Amplification of *ori*P DNA

A 2.4 kb EBV genomic DNA containing the latent origin of replication *ori*P (figure 11) was amplified from a known EBV containing cell line B95-8. Whole cell genomic DNA was isolated as described in section 3.6.2 and PCR amplification was performed on the genomic DNA as described in section 3.1.1. After PCR, 2 µl of the PCR product was loaded onto a 0.8 % agarose gel for electrophoresis at 100 V for 1 hour. After electrophoresis, the gel was viewed under UV light and a picture of the PCR

products taken. The gel photograph of the *ori*P PCR product is shown in figure 12 below.

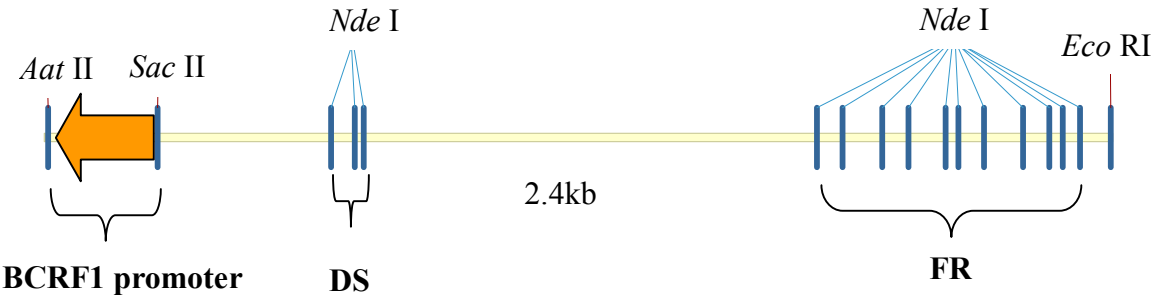


Figure 11. Map of *ori*P-containing EBV genomic fragment to be PCR-amplified and cloned in this study. Some important restriction enzyme sites are shown. They were located based on EBV B95-8 strain genomic sequence (GENBANK Accession Number V01555). The promoter for EBV ORF BCRF1 is present in this fragment as indicated.

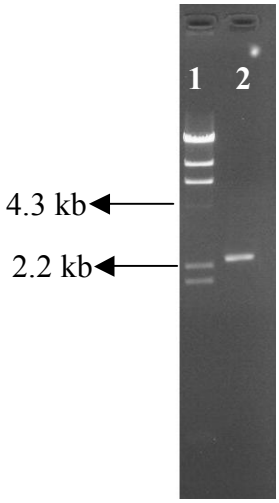


Figure 12. PCR-amplified *ori*P-containing EBV genomic fragment. (A): Lane 1 is λ*Hin*dIII DNA ladder; lane 2 is the PCR-amplified *ori*P-containing EBV genomic fragment.

As can be seen from the above figure, a PCR product of the expected size was obtained. Restriction analysis using *Nde* I was done and confirmed the identity of the PCR product and this result can also be seen in all later restriction enzyme analyses of all *ori*P containing plasmids using *Nde* I. Sequencing of the *ori*P fragment was performed but while I was able to confirm the presence of *ori*P sequences, I was unable to obtain clean sequencing results because of the presence of numerous repeat

sequences of EBNA-1 binding sites. In addition, the original aim of the project is to study the effect of transcription on *ori*P's ability to allow replication in the presence of EBNA-1. What was really needed was a functional *ori*P that is able to initiate replication and it was not necessary to confirm every base pair of the *ori*P. The *Nde* I digestion serves as a way of confirming the presence of *ori*P in each vector construct. As will be demonstrated below, the presence of *ori*P on plasmid vector clearly allowed replication, indicating a functional *ori*P.

4.1.1.2 pcDNA3.1+

pcDNA3.1+ was digested with restriction enzymes *Stu* I and *Sal* I for downstream reactions. After digestion, the fragments were analyzed using 0.8 % agarose gel electrophoresis. A photo of the gel can be seen in figure 13 below.

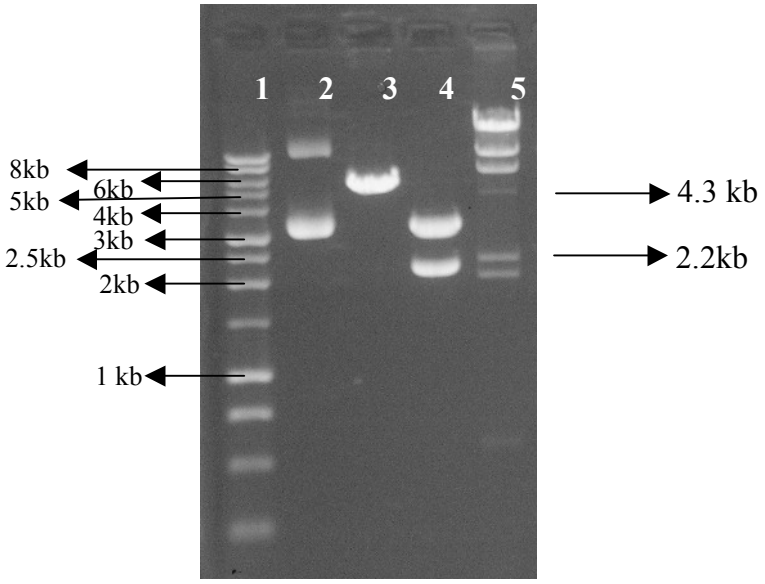


Figure 13. Restriction analysis of plasmid pcDNA3.1+. Lane 1 is 1kb DNA ladder (Promega) Lane 2 is undigested pcDNA3.1+ plasmid; Lane 3 is pcDNA3.1 digested with *Stu* I; Lane 4 is pcDNA3.1+ digested with *Sal* I; Lane 5 is λ *Hin*dIII.

As expected from the plasmid map (figure 6), *Stu* I cleaves pcDNA3.1+ at one site resulting in a linear band corresponding to 5.4 kb in size. *Sal* I cleaves pcDNA3.1+ at two sites, resulting in a larger 3.2 kb band and a smaller 2.2 kb band.

4.1.1.3 p-oriP-S

To construct the initial plasmid containing *ori*P, the 2.4 kb *ori*P-containing EBV genomic fragment (section 4.1.1.1) and the pcDNA3.1+ vector were digested with restriction enzymes *Aat* II and *Eco* RI. After digestion, the products were purified, ligated and transformed into electrocompetent DH10B *Escherichia coli* cells. Plasmid DNA was purified from a few transformants and analyzed by *Nde* I digestion. An isolate with the restriction pattern expected of the cross-ligation product was kept for further construction described below.

The initial *ori*P-containing plasmid described above contained the transcription promoter for the EBV ORF BCRF1 (figure 11). Although this promoter requires activation by a viral transcription activator, which is not normally expressed in the cell line B95-8 that I used in this study, it is better to remove this promoter to eliminate the possibility of undesired transcription initiated from it. The BCRF1 was included in the original PCR product as it was difficult to design PCR primers within the exact *ori*P region. It was therefore an easier strategy to include the BCRF1 ORF and remove it later by restriction enzymes. Sequence analysis showed that this promoter is closely flanked by the unique *Aat* II and *Sac* II sites in the initial *ori*P-containing plasmid. To construct p-oriP-S, the desired *ori*P-containing plasmid without the BCRF1 promoter, the initial *ori*P-containing plasmid DNA was first digested with *Aat* II and *Sac* II. The large fragment resulted from the digestion was purified, subjected to T4 polymerase blunt end repair, purified again and finally self-ligated before transforming into DH10B electrocompetent *Escherichia coli* cells.

Plasmid DNA was purified from a few transformants and subjected to restriction enzyme analysis.

Single digests using the restriction enzymes *Nde* I, *Aat* II, *Stu* I and *Sal* I of a positive isolate are shown in figure 14 below. As expected from the plasmid map (figure 7), *Nde* I digestion yielded small fragments that are visualized as a smear near the bottom of lane 2, indicative of the EBNA-1 binding sites on the *ori*P, the 1.1 kb fragment between DS and FR and a larger 4.9 kb band from the vector. Due to the T4 polymerase blunt end repair, the *Aat* II site was destroyed after ligation with the repaired *Sac* II site and the plasmid was not cleaved by *Aat* II. As expected, both *Sal* I and *Stu* I cleaved the plasmid once, resulting in a 6.6 kb band.

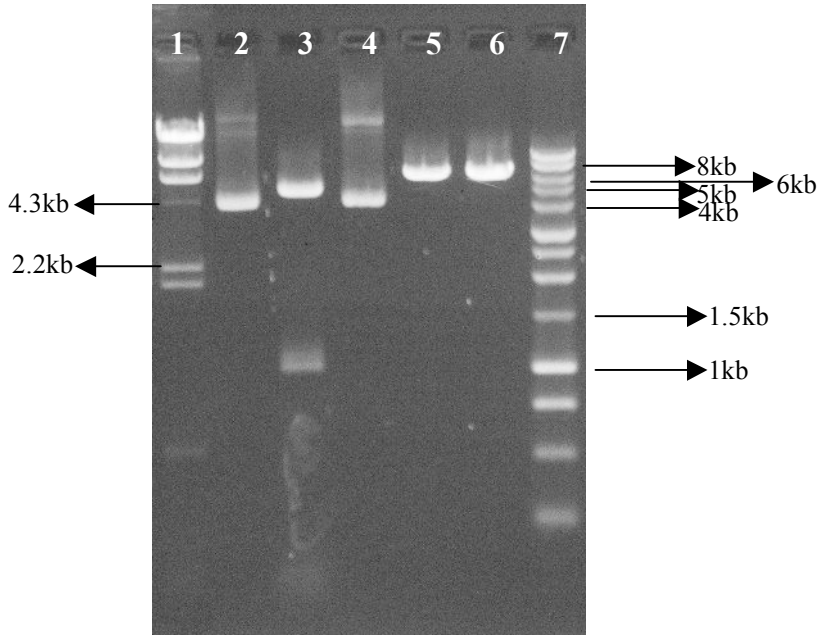


Figure 14. Restriction analysis of plasmid p-oriP-S. Lanes 1 and 7 are λ *Hin*dIII and 1 kb DNA ladder respectively. Lane 2 is undigested p-oriP-S; Lane 3 is p-oriP-S digested with *Nde* I; Lane 4 is p-oriP-S digested with *Aat* II; Lane 5 is p-oriP-S digested with *Stu* I; Lane 6 is p-oriP-S digested with *Sal* I

4.1.1.4 p-S-oriP and p-S-oriP.1

To construct p-S-oriP and p-S-oriP.1, the precursor plasmid pcDNA3.1+ΔC, i.e. pcDNA3.1+ with CMV immediate early promoter removed, was first constructed.

The CMV promoter of pcDNA3.1+ was removed by digesting the plasmid with *Aat* II and *Kpn* I. The largest product of the reaction was purified and the incompatible ends treated with T4 polymerase. After that, the reaction product was purified again, followed by self-ligation and transformation. Plasmid DNA purified from positive transformants that showed resistance to *Aat* II and *Kpn* I and other expected restriction patterns were directly used for the construction of both p-S-oriP and p-S-oriP.1. The *ori*P-containing DNA fragment was extracted from the initial *ori*P-containing plasmid DNA used to construct p-oriP-S (section 4.1.1.3) with restriction enzymes *Sac* II and *Xba* I and further treated with T4 polymerase to blunt the ends. pcDNA3.1+ΔC DNA was linearized using *Sma* I, a blunt end cutter that cleaves just downstream of the SV40 promoter. The blunt-end *ori*P-containing fragment was then ligated to the linearized pcDNA3.1+ΔC and the ligation product was used to transformed DH10B cells. Since these two blunt-end reactants can be ligated in two opposite relative orientations, the two desired constructs, p-S-oriP and p-S-oriP.1, could be obtained from a single reaction. Plasmid DNA was purified from a few transformants and analyzed using the restriction enzymes *Nde* I, *Eco* RI and *Sal* I. Agarose gel electorphoresis was performed to analyze the restriction fragments. Figure 15 shows the results of the restriction analysis of one positive isolates for each plasmid.

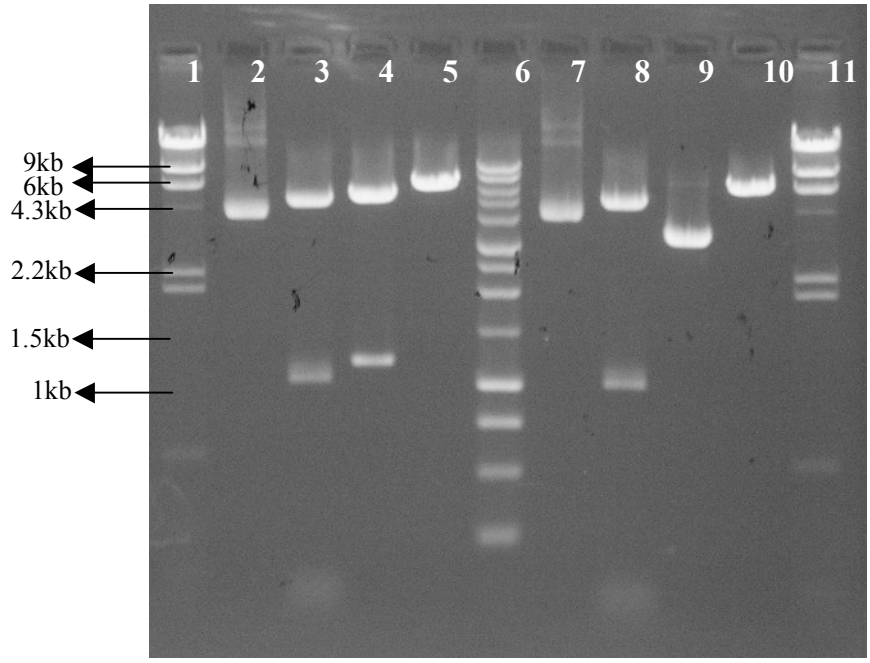


Figure 15. Restriction analysis of plasmids p-S-oriP and p-S-oriP.1. Both plasmids were digested with the restriction enzymes *Nde* I, *Eco* RI and *Sal* I. Lanes 1 and 11 are λ *Hin*dIII DNA ladders. Lane 6 is 1kb ladder. Lane 2: undigested p-S-oriP; Lane 3: p-S-oriP digested with *Nde* I; Lane 4: p-S-oriP digested with *Eco* RI; Lane 5: p-S-oriP digested with *Sal* I; Lane 7: undigested p-S-oriP.1; Lane 8: p-S-oriP digested with *Nde* I; Lane 9: p-S-oriP.1 digested with *Eco* RI; Lane 10: p-S-oriP digested with *Sal* I.

As can be seen from the gel photograph and plasmid maps (figures 8 and 9), upon digestion with *Nde* I, both plasmids yield small molecular weight bands that appear as a smear due to the multiple *Nde* I restriction enzyme sites on both plasmids, indicative of the presence of *ori*P. *Eco* RI digestion of p-S-oriP yielded two bands, a larger 5.6 kb band and a smaller 1.1 kb band due to the orientation of *ori*P. It can be seen that for p-S-oriP, the FR is closer to the promoter while the DS is further away. From the gel photograph, it would appear that *Eco* RI digestion of p-S-oriP.1 yielded only a single band corresponding to 3.3 kb in length. In fact, there are two equal molecular weight bands, each 3.3 kb long due to the position of the *Eco* RI sites on p-S-oriP.1 as evidenced from the plasmid map (figure 9), which indicates that the DS is closer to the SV40 promoter in p-S-oriP.1. *Sal* I yielded a single band of 6.7 kb in length for

both plasmids.

4.1.1.5 p-oriP-ΔS

To construct p-oriP-ΔS, the precursor plasmid pcDNA3.1+ΔCS, i.e. pcDNA3.1+ with both CMV immediate early and SV40 immediate promoters removed, was first constructed. To construct pcDNA3.1+ΔCS, the SV40 early promoter-containing *Eco* RV-*Sma* I fragment was removed from pcDNA3.1+ΔC (described in section 4.1.1.4). The construction of p-oriP-ΔS from pcDNA3.1+ΔCS was done in the same way as the construction of p-S-oriP and p-S-oriP.1 from pcDNA3.1+ΔC described in 4.1.1.4. The results of the restriction enzyme analysis of a positive isolate using *Nde* I, *Sma* I, *Sal* I and *Stu* I are shown in figure 16 below.

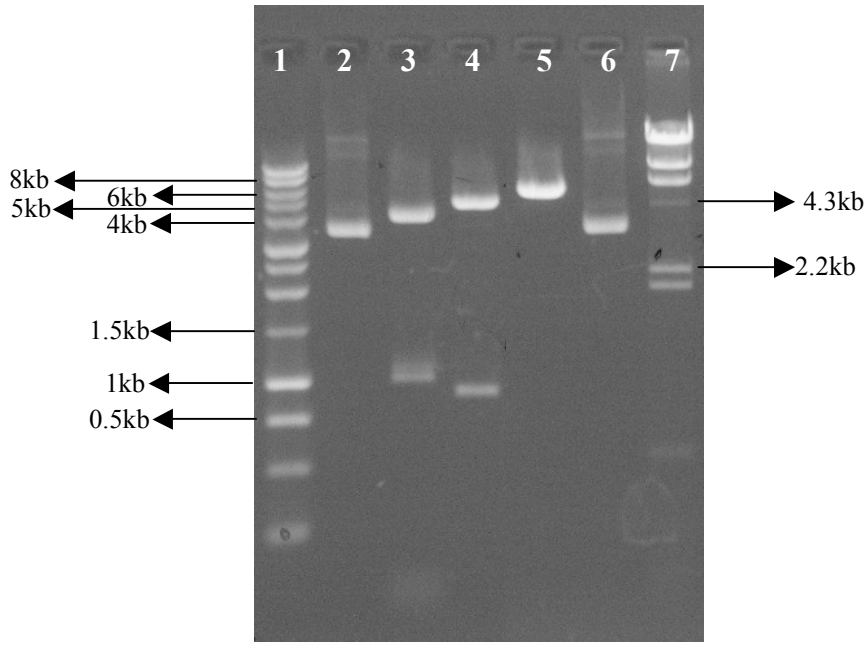


Figure 16. Restriction analysis of plasmid p-oriP-ΔS. *Nde* I, *Sma* I, *Sal* I and *Stu* I were used to digest the plasmid. Lane 1: 1kb DNA ladder; Lane 2: undigested p-oriP-ΔS; Lane 3: p-oriP-ΔS digested with *Nde* I; Lane 4: p-oriP-ΔS digested with *Sma* I; Lane 5: p-oriP-ΔS digested with *Sal* I; Lane 6: p-oriP-ΔS digested with *Stu* I; Lane 7: λ *Hin*dIII DNA ladder.

As can be seen from the above figure, *Nde* I digestion of p-oriP-ΔS yielded small fragments from the multiple *Nde* I enzyme sites, a 1 kb band and a larger 3.8 kb fragment, which corresponds to the expected digestion pattern from the plasmid map (figure 10). *Sma* I digestion yielded two fragments: a 900 bp and a larger 4.7 kb band. As there is only one *Sal* I site on the plasmids, *Sal* I digestion gave an expected 5.6 kb fragment. As the SV40 promoter was removed, the plasmid is now expectedly resistant to cleavage by *Stu* I. The orientation of *ori*P in this clone is not expected to matter as there is not promoter upstream of the *ori*P and therefore only one clone was used for subsequent experiments. For better comparisons, I chose the one with the relative orientation of different components the same as that in p-oriP-S and that in p-S-oriP.1 too.

4.1.2 Replication of p-oriP-S in EBNA1-expressing cells

Replication of transfected plasmid DNA in mammalian cells can be assayed conveniently based on the change in DNA methylation pattern. Plasmid DNA synthesized by a *Escherichia coli* strain expressing DAM methylase is methylated by the enzyme at the adenosine residue in the specific DNA sequence 5'-GATC-3'. Mammalian cells do not possess DAM methylase and therefore the DNA synthesized by them is not methylated this way. The restriction enzyme *Dpn* I has the same sequence specificity as DAM methylase and it cleaves DNA only if the adenosine residue in its recognition site is methylated. To assay for plasmid DNA replication in mammalian cells, plasmid DNA carrying *Dpn* I recognition sites is prepared from DAM-positive bacterial cells and transfected into mammalian cells. Plasmid DNA is then recovered from the transfected cells after a suitable time of growth and digested with *Dpn* I. DNA replicated in mammalian cells is identified by its resistance to *Dpn* I.

All the plasmids used in this study contain multiple *Dpn* I recognition sites.

Once the p-oriP-S plasmid was constructed, its ability to replicate in EBNA1-expressing mammalian cells was analyzed, with its parent plasmid pcDNA3.1+ as a negative control. A time course was done to understand the kinetics of its replication. Thus, 20 µg of p-oriP-S and pcDNA3.1+ were transfected separately into EBNA-1 expressing B95-8 cells as described in section 3.6.3 using 250 V and 950 µF. The transfected cells were allowed to grow at 37 ºC, 5 % CO_2. For each transfection, plasmid DNA was harvested from 5 X 10^5 cells at 48, 72 and 96 hours post-transfection as described in section 3.7.2. The recovered plasmid DNA was then linearized using an appropriate restriction enzyme. An aliquot of it was further digested by *Dpn* I. The so-treated plasmid DNA samples were subjected to Southern blot analysis using neomycin phosotransferase gene specific probes as described in section 3.7.4. Various amounts of *Eco*R I linearized pcDNA3.1+ plasmid were included as standards in the analysis. The results can be seen in figure 17 below.

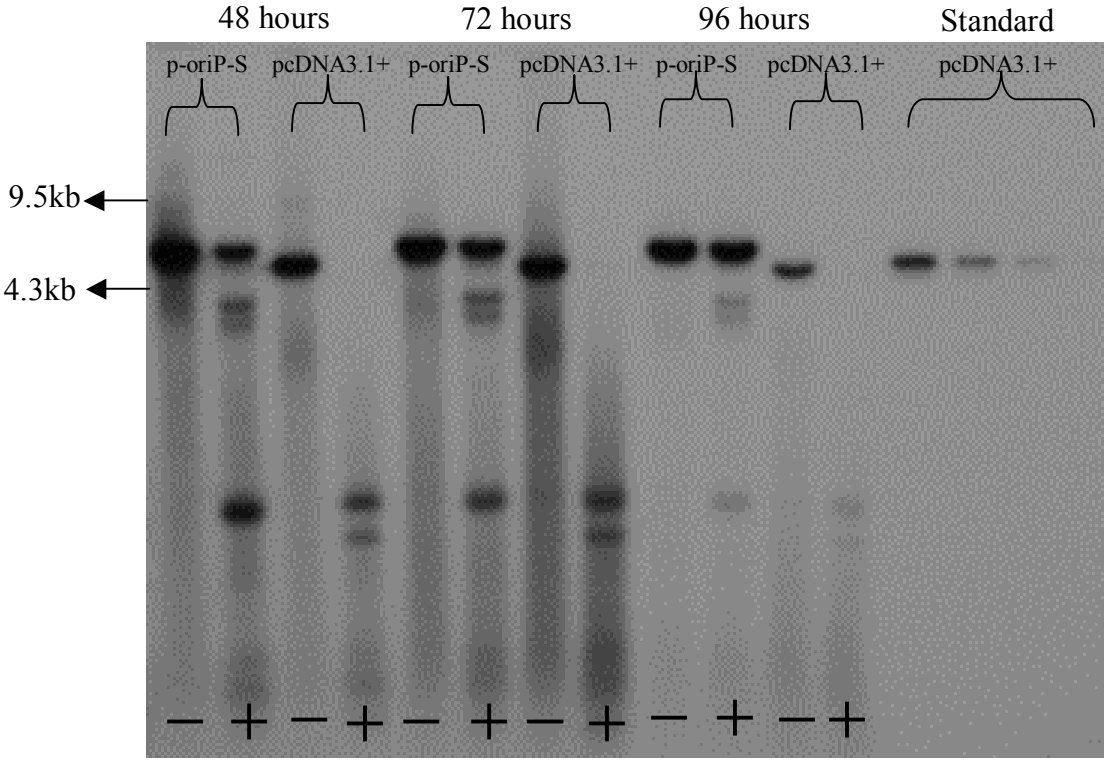


Figure 17. Kinetics of replication of p-oriP-S in EBNA-1 expressing cells. Southern blot analysis of plasmid DNA recovered from transfected cells. (+): *Dpn* I treated; (-): Non-*Dpn* I treated. Lanes 1-2: p-oriP-S recovered after 48 hours; Lanes 3-4: pcDNA3.1+ recovered after 48 hours; Lanes 5-6: p-oriP-S recovered after 72 hours; Lanes 7-8: pcDNA3.1+ recovered after 72 hours; Lanes 9-10: p-oriP-S recovered after 96 hours; Lanes 11-12: pcDNA3.1+ recovered after 96 hours; Lane 13: 1 ng of linearised pcDNA3.1+; Lane 14 300 pg of linearized pcDNA3.1+; Lane 15: 100 pg of linearized pcDNA3.1+; Lane 16: 30 pg of linearized pcDNA3.1+. For lanes 1 to 12, each lane contains plasmid DNA recovered from the same number of transfected cells.

Replicated plasmid DNA in the *Dpn* I-digested sample should have a size of the full-length plasmid because of its resistance to *Dpn* I digestion, whereas the input plasmid DNA should be digested to smaller fragments. As expected, there was no full-length pcDNA3+ left after *Dpn* I digestion (lanes 4, 8 & 12), showing that pcDNA3+ did not replicate at all throughout the 96 hours in B95-8 cells. On the other hand, some full-length p-oriP-S remained after *Dpn* I digestion (lanes 2, 6 & 10) showing that p-oriP-S had replicated in B95-8 cells, as expected. Since the *ori*P-containing EBV genomic sequence is the only additional sequence in p-oriP-S relative to pcDNA3.1+, replication of p-oriP-S is most probably dependent on *ori*P. To estimate roughly the

percentage of replicated p-oriP-S plasmid DNA in each time point, the intensities of the full-length plasmid DNA signals from *Dpn* I-digested and undigested samples were compared by naked eye with reference to the signal intensities of the pcDNA3.1+ DNA standards on X-ray films exposed for various times to the chemilluminescently probed blot. In this estimation, roughly 10% of the plasmids recovered from the transfected B95-8 cells were replicated at 48 hours post-transfection (lanes 1 & 2), while approximately 30% of plasmids were replicated at 72 hours post-transfection (lanes 5 & 6). Almost 100 % the plasmids were replicated at 96 hours post-transfection (lanes 9 & 10). Since 72 hours post-transfection appeared to be the midway through the accumulation of the replicated plasmid DNA in the transfected cells, this time point was chosen to be the time for plasmid recovery from the transfected cells in the subsequent analyses of the replication efficiency of different plasmid constructs. Since real-time PCR analysis is more quantitative, I originally attempted to use it in conjunction with DpnI digestion to measure the amount of replicated DNA. This attempt failed because of two reasons. First, I was unable to design suitable primers amplifying for a region in *ori*P due to the highly repetitive nature of *ori*P as well as the lack of *Dpn* I sites within amplifiable regions. Second, when I used the bacterial backbone of the plasmids as the amplicon, the background amplification was unacceptably high due to the contamination with bacterial plasmid from the laboratory environment.

4.1.3 Replication of *ori*P-containing plasmids is negatively influenced by the presence of transcription promoter in the replicon

To study the effect of transcription on *ori*P-dependent DNA replication, 20 μg of the each following plasmids, p-oriP-S, p-S-oriP, p-S-oriP.1 and p-oriP-ΔS, was

transfected separately into EBNA-1 expressing B95-8 cells as described in section 3.6.3 using 250 V and 950µF. The cells were allowed to grow at 37ºC, 5% CO_2 for 72 hours. Plasmid DNA was then recovered from the transfected cells as mentioned in section 3.7.2. An aliquot of the recovered plasmid DNA was treated with *Dpn* I. Southern blot analysis was performed on the *Dpn* I-treated and untreated plasmid DNA and the results can be seen in figure 18 below.

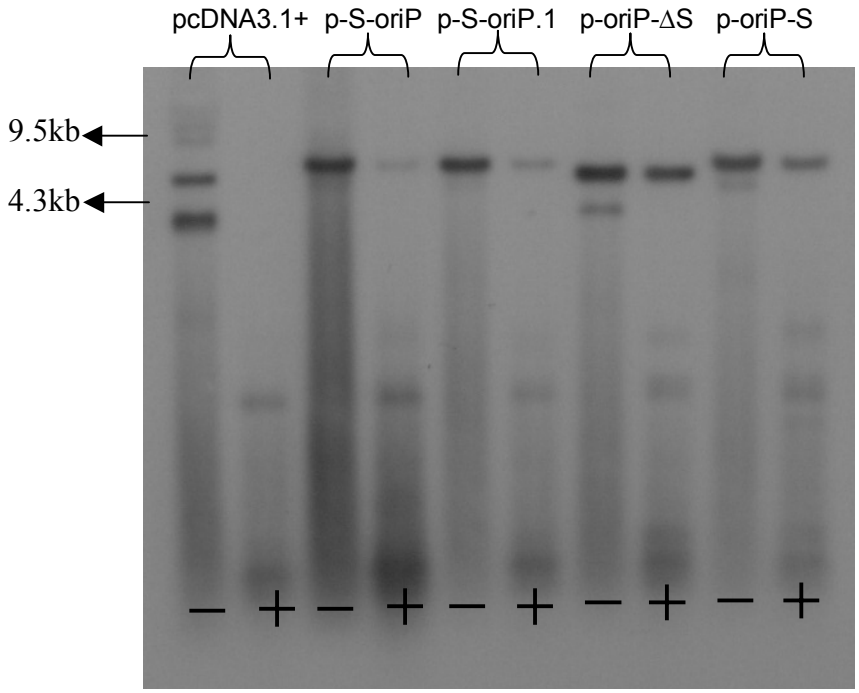


Figure 18. Replication efficiencies of pcDNA3.1+-derived *ori*P-containing plasmids. Southern blot analysis of plasmid DNA recovered from transfected cells. (+): *Dpn* I treated (-): Non-*Dpn* I treated. Lanes 1 and 2 contain rescued pcDNA3.1+ vector; lanes 3 and 4 contain p-S-oriP; lanes 5 and 6 contain p-S-oriP.1; lanes 7 and 8 contain p-oriP-ΔS; lanes 9 and 10 contain p-oriP-S. All lanes contain plasmid DNA recovered from the same number of transfected cells.

As can be seen from the Southern blot analysis (figure 18), 72 hours after transfection in EBNA-1 expressing cell, parental vector pcDNA3.1+ exhibited no signs of replication (lane 2) while all *ori*P-containing plasmids replicated to certain extent (lanes 4, 6, 8 and 10). These results suggest that the plasmid replication observed is dependent on *ori*P. P-oriP-ΔS does not contain any known transcription promoter. Its

ability to replicate shows that *ori*P-dependent DNA replication does not require concomitant transcription anywhere in the replicon. In fact, p-oriP-ΔS exhibited a higher level of replication (roughly 60% of total plasmid replicated) than the other *ori*P-containing plasmids which contain the strong SV40 early promoter (compare lane 8 with lanes 4, 6 and 10). These results suggest that the presence of a promoter on the plasmid negatively influences the ability of *ori*P to induce replication. Among the plasmids containing the SV40 early promoter, p-oriP-S consistently showed higher level of replication (approximately 30 % of total plasmid replicated) than p-S-oriP and p-S-oriP.1, while the latter two showed similar levels of replication (approximately 10% of total plasmid replicated). In p-oriP-S, *ori*P is located outside of the only transcription unit on the plasmid and is more than 2-kb downstream of the relevant polyadenylation signal. It is likely that most of the transcription complexes have fallen off the DNA template before reaching *ori*P or been destabilized when they reach *ori*P. On the other hand, in p-S-oriP and p-S-oriP.1, *ori*P is located within the transcription unit and expected to experience heavy traffic of transcription complexes. The reduction in the level of replication in these two plasmids compared to that of p-oriP-S suggests that transcriptional activity at *ori*P is inhibitory to plasmid replication.

The experiment was repeated once. The results obtained were basically the same.

4.2 Transcription through *ori*P is inhibited *in vivo*

To analyze plasmid transcription in the transfected cells, total RNA was isolated from them (section 3.7.2) and northern blot analysis was performed on the isolated RNA (section 3.7.5). The analysis was done with neomycin phosphotransferase gene sequence-specific probes and the results can be seen in figure 19A below.

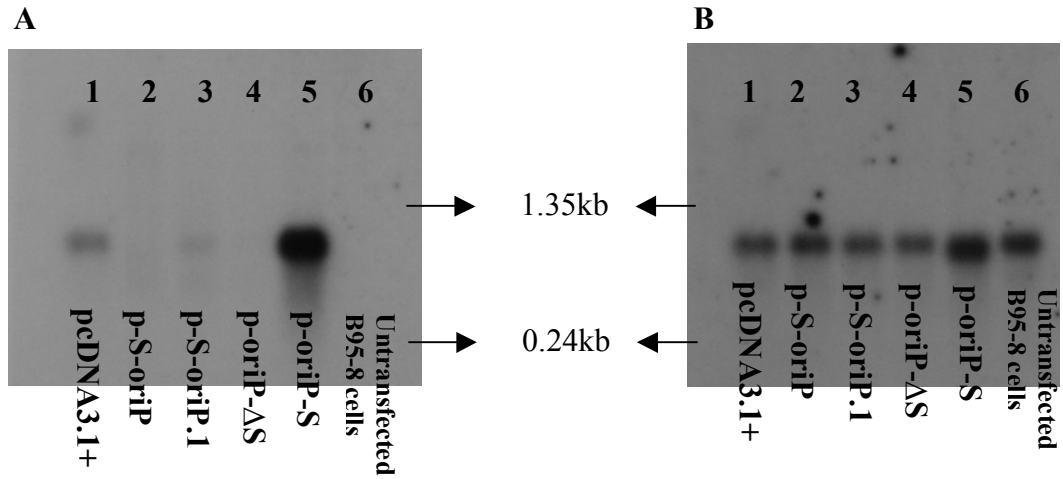


Figure 19. Transcription of pcDNA3.1+-derived *ori*P-containing plasmids in B95-8 cells. Northern blot analysis of RNA isolated from plasmid transfected cells. (A): Northern blot probed with neomycin phosotransferase gene specific probes. (B): Blot is stripped of probes and re-probed with GAPDH specific probes. Lane 1: pcDNA3.1+; Lane 2: p-S-oriP; Lane 3: p-S-oriP.1; Lane 4: p-oriP-ΔS; Lane 5: p-oriP-S; Lane 6: untransfected B95-8 cells.

Plasmids p-S-oriP and p-S-oriP.1 have *ori*P inserted upstream of the transcriptional start site of the neomycin resistance gene. Lane 6 is the negative control containing total RNA from untransfected cells. As expected, no hybridization signal can be observed. According to the sequence of pcDNA3.1+, the neomycin phosphotransferase gene transcript produced from this plasmid should be about 1 kb in size. As expected, an RNA signal of this size was obtained from the cells transfected with pcDNA3.1+ (lane 1). An RNA signal of the same size but much higher intensity was obtained from the cells transfected with p-oriP-S (lane 5). This observation can be explained by the fact that FR element in *ori*P is an EBNA1-dependent transcriptional enhancer (Reisman and Sudgen, 1986). Since FR occurs immediately upstream of the SV40 early promoter in p-oriP-S, the transcription of the neomycin phosphotransferase gene is enhanced and this lead to the observed intense RNA signal obtained from p-oriP-S-transfected cells. No neomycin phosphotransferase gene RNA signal was obtained from cells transfected with p-oriP-

ΔS (lane 4). This is expected because there is no promoter on this plasmid to drive transcription. In p-S-oriP and p-S-oriP.1, a 2.1-kb *ori*P-containing EBV sequence is inserted between the SV40 early promoter and the neomycin phosphotransferase gene sequence (section 4.1.1.3). There is no known transcription promoter, transcript splicing or polyadenylation signal in this sequence insert. Because of this, the neomycin phosphotransferase gene sequence-containing transcript produced from these two plasmids should be about 3.1 kb in size. Lanes 2 and 3 contain total RNA from cells transfected with p-S-oriP and p-S-oriP.1 respectively. Unexpectedly, there was very little or no signal corresponding to RNA of this size on these lanes. There is a faint band of approximately 1 kb observed in lane 3. The origin of this RNA is unknown. Given that the probes used are highly specific for the neomycin transcript (as evidenced by the negative control in lane 6), it could probably mean that there is possibly a weak cryptic transcriptional promoter occurring in the DS-distal end of the FR region. The only way a 3.1kb band can be seen in lanes 2 and 3 is if transcription proceeded without disruption through the *ori*P and into the neomycin resistance gene. Any transcripts that terminated within the *ori*P itself would not generate any signal. Figure 21 also presents a summary of the constructs used. To check if the lack of expected RNA signal on these two lanes could be caused by unintended RNA degradation during the experimental process, the blot was stripped of the neomycin phosphotransferase gene-specific probes as described in section 3.7.6 and re-probed with the housekeeping cell gene GAPDH gene-specific probes as described in section 3.7.5. The results are showed in figure 19B. As shown, the intensities of RNA signals from all samples are similar, except the one from p-oriPS-transfected cells which is somewhat higher. The apparently stronger GAPDH RNA signal from p-oriP-S-transfected cells could be due to incomplete stripping of neomycin phosphotransferase

gene probes. The neomycin phosphotransferase and GAPDH gene transcripts obtained from this plasmid are by coincidence of very similar sizes and thus should occur on very nearby locations on the Northern blot. The particularly abundant neomycin phosphotransferase gene probes hybridized to the corresponding transcripts from p-oriP-S-transfected cells (figure 19A, lane 5), if not completely removed, could add to the GAPDH signal from the re-probed blot. In any case, GAPDH probing shows that RNA isolated from p-SoriP- and pSoriP.1-transfected cells did not suffer random degradation. Southern blot analysis shows that the amounts of plasmid DNA in p-S-oriP- and p-S-oriP.1-transfected cells were not much different from the amounts of plasmid DNA in the cells transfected by other plasmids (compare figure 18 lanes 3 and 5 with lanes 1, 7 and 9). Therefore, a possible explanation for the particular lack of neomycin phosphotransferase gene sequence-containing RNA in p-S-oriP- and p-S-oriP.1-transfected cells is that the transcription on these two plasmids was somehow inhibited and it is the presence of *ori*P sequence within the transcription unit that causes the inhibition.

4.2.1 Transcription through *ori*P *in vivo* is inhibited even in the absence of EBNA-1

There are two likely mechanisms leading to the observed inhibition of transcription of p-S-oriP and p-S-oriP.1 in B95-8 cells. First, transcription elongation is blocked by the tight binding of EBNA1 protein its cognate sites in *ori*P. Second, transcription elongation could be inhibited by the head-on replication initiated at *ori*P. In other words, replication and transcription occurring simultaneously on the same DNA template can be mutually inhibitory. Since both of these possible transcription

inhibitory mechanisms require EBNA1 protein, they can be tested by comparing the efficiencies of transcription in EBNA1-expressing and non-expressing cells. Thus, the same plasmids were transfected into BJAB cells, which do not express EBNA1. Total RNA was isolated and northern blot performed using neomycin specific probes. The results are shown in figure 20 below.

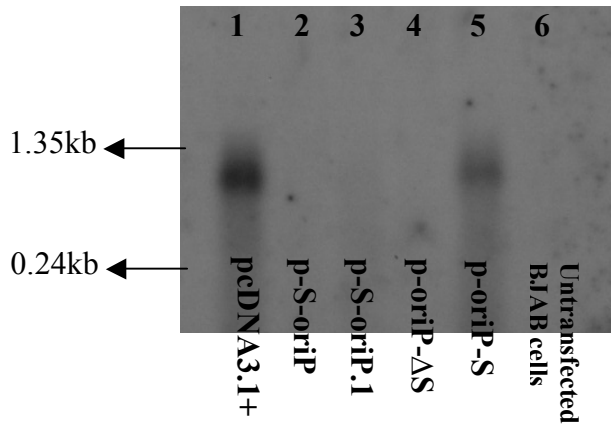


Figure 20. Transcription of pcDNA3.1+-derived *ori*P-containing plasmids in BJAB cells. Northern blot analysis of RNA isolated from plasmid transfected cells using neomycin phosotranasferase gene specific probes. Lane 1: pcDNA3.1+; Lane 2: p-S-oriP; Lane 3: p-S-oriP.1; Lane 4: p-oriP-ΔS; Lane 5: p-oriP-S; Lane 6: untransfected B95-8 cells.

Lane 6 is the negative control containing total RNA from untransfected BJAB cells. As expected, no signal was generated. Lane 1 contains RNA from BJAB cells transfected with the positive control pcDNA3.1+ parental vector. As expected, a 1 kb band identical to that obtained from pcDNA3.1+-transfected B95-8 cells (figure 19A, lane 1) can be seen. An RNA signal of the same size was also obtained from p-oriP-S-transfected BJAB cells (lane 5). Recall that the level of transcription on p-oriP-S was much higher than that on pcDNA3.1+ in B95-8 cells because of the EBNA1-dependent enhancer effect of *ori*P (compare lanes 1 and 5 of figure 19A). Here in BJAB cells which do not express EBNA1, the level of transcription on p-oriP-S was not elevated relative to that on pcDNA3.1+ (compare lanes 1 and 5 of figure 20). As p-oriP-ΔS contained no transcriptional promoter, no transcripts were observed (lane

4). The main focus of this experiment was the transcription of p-S-oriP and p-S-oriP.1. As explained under section 4.2, the neomycin phosphotransferase gene sequence-containing transcript produced from these two plasmids should be about 3.1 kb in size. As can be seen from lanes 2 and 3, very little or no such transcript was produced in BJAB cells transfected with these two plasmids. These results are similar to those obtained from B95-8 cells (figure 19A, lanes 2 and 3). Therefore, the inhibition of transcription in these constructs is independent of EBNA1 protein, and the two mechanisms of transcription inhibition suggested above are invalidated. The DNA sequence of *ori*P itself is the likely cause of transcription inhibition.

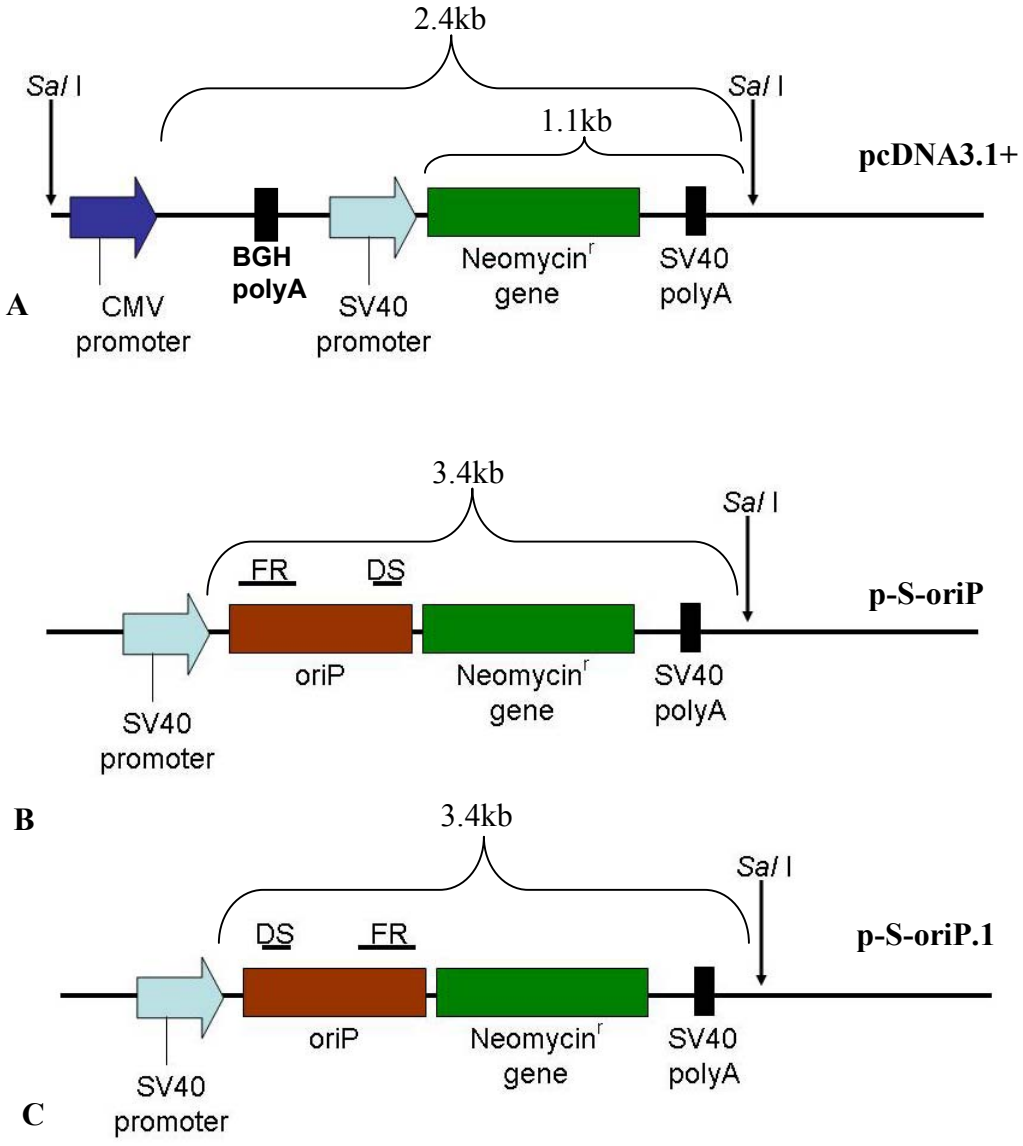
An interesting scenario that can explain both the inhibition of *ori*P-dependent replication by transcription as well as the inhibition of transcription by *ori*P in p-S-oriP and p-S-oriP.1 is that *ori*P blocks the translocation of transcription elongation complexes without dissociating them. In other words, *ori*P induces transcription arrest. Thus, transcription elongation cannot proceed to the neomycin phosphotransferase gene sequence occurring downstream of *ori*P in these plasmids. At the same time, the transcription complexes arrested at *ori*P blocks the initiation and/or elongation of replication of these plasmids. The subsequent part of this thesis work was devoted to the testing of this possible scenario through in vitro studies.

4.3 *In vitro* transcription of *ori*P-containing template

4.3.1 *In vitro* transcription of pcDNA3.1+, p-oriP-S, p-S-oriP, p-S-oriP.1 and p-oriP-ΔS

I first tried to see if inhibition of transcription by *ori*P was reproducible *in-vitro*. As

indicated in section 3.8, *in vitro* transcription was done using Promega's HelaScribe® Nuclear Extract. In the initial attempts, I carried out in vitro transcriptions using the same plasmids employed in the *in vivo* study. Since polyadenylation machinery was not expected to be functional in this in vitro transcription system, the transcription templates had to be linear in order to give full-length run-off transcripts of fixed sizes. Therefore, the plasmid DNA was first subjected to *Sal* I restriction enzyme digestion. The resulting transcription templates as well as their corresponding run-off transcript sizes are represented in figure 21 below.



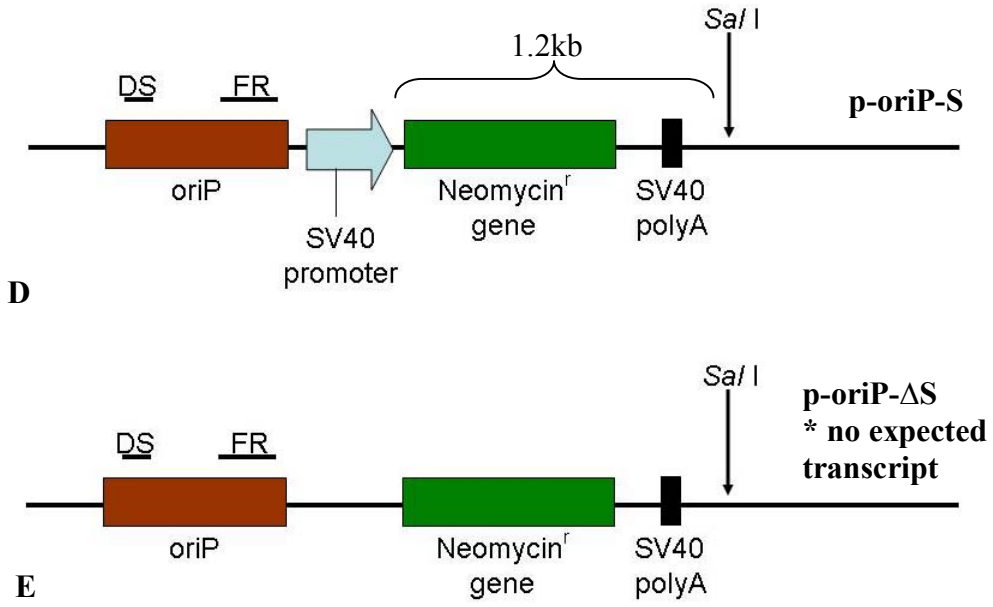


Figure 21. *In vitro* transcription templates and expected transcripts from pcDNA3.1+-derived *ori*P-containing plasmids. *Sal* I sites in each plasmid are indicated by arrows. The sizes of expected run-off transcript(s) from each *Sal* I-linearized transcription template are indicated by braces. Plasmids are not drawn according to size.

The *Sal* I digestion products are shown in figures 13 to 16. They were purified using the QiaQuick PCR purification kit before use in the *in vitro* transcription reactions. The transcription reactions were done as described in section 3.8, with rNTPs used at 400 μM and incubation done at 30 ºC for 90 minutes. Transcripts were extracted as described in section 3.8 and analyzed by Northern blot analysis using neomycin phosotransferase gene specific probes. The results of this analysis can be seen in figure 22 below.

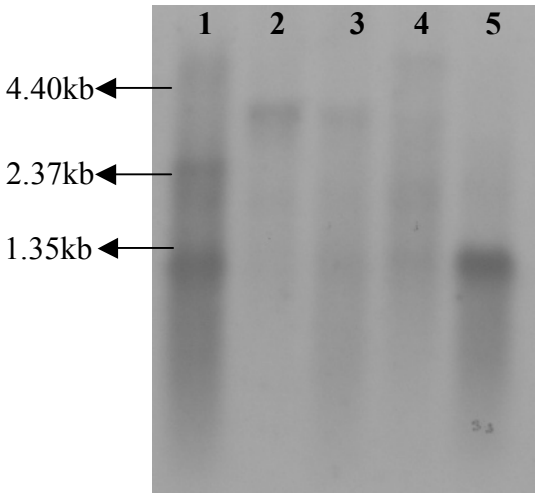


Figure 22. Transcription of pcDNA3.1+-derived *ori*P-containing plasmids in vitro. Northern blot analysis of transcription products using neomycin phosotransferase gene specific probes. Lane 1: parent vector pcDNA3.1+; Lane 2: p-S-oriP; Lane 3: p-S-oriP.1; Lane 4: p-oriP-ΔS; Lane 5: p-oriP-S

Lane 1 contains the positive control pcDNA3.1+ parent vector. From figure 22, it can be observed that there are two very distinct bands in lane 1 upon probing with neomycin phosotransferase gene specific probes. This is expected as *Sal* I digest of pcDNA3.1+ yielded a template containing two promoters, the CMV and SV40, as can be seen from figure 21. The larger transcript is the run-off transcript from the CMV promoter which is roughly estimated to be about 2.5kb and the smaller transcript is the run-off from the SV40 promoter, which is estimated to be 1.2kb long. A similar 1.2 kb transcript was also observed for p-oriP-S (lane 5). As there is no CMV promoter on p-oriP-S, only one run-off transcript is produced and that is the neomycin phosotransferase gene driven by the SV40 promoter. p-oriP-ΔS contains no promoters and no discrete products were observed (lane 4). The expected 3.4 kb transcripts of p-S-oriP and p-S-oriP.1 (lanes 2 and 3 respectively) did appear but at much lower levels than the transcripts from pcDNA3.1+ and p-oriP-S, indicating that transcription on the former two templates was much less efficient. Thus, the inhibition of transcription by *ori*P observed in vivo was reproducible in vitro.

4.4 Construction of templates for analysis of transcription arrest *in vitro*

In the studies described above, inhibition of transcription by *ori*P was observed in the plasmids p-S-oriP and p-S-oriP.1, which have *ori*P occurring right downstream of the SV40 promoter and upstream of neomycin phosphotransferase gene sequence. In the Northern blot analysis, the transcripts were detected by probing for the neomycin phophotransferase gene sequence. Should transcription arrest occur at *ori*P, the transcripts resulted from this event would not contain the probed sequence and thus could not be detected. Probing for *ori*P sequence may not be able to provide sensitive detection of the arrested transcripts because the detectable region in the arrested transcripts could be very short, depending on the actual arrest site(s) within *ori*P. Nevertheless, the possibility of transcription arrest could be investigated only if sensitive detection of the arrested transcripts is achieved. To achieve this, a new plasmid construct would have to be created that has a sufficiently long sequence between the promoter and *ori*P, which is to be probed for in Northern blot analysis. That would also most likely enable us to determine the region within *ori*P that is responsible for inhibition of transcription.

Thus, the plasmid p-E-oriP was constructed by inserting an *ori*P-containing 2.2-kb EBV genomic sequence into the vector pEGFP-C1 downstream of the enhanced green fluorescent protein (EGFP) coding sequence in the CMV immediate early promoter-driven transcription unit. If transcription of this plasmid is arrested at anywhere within *ori*P, the resultant transcript will contain the EGFP sequence and thus be detectable in Northern blot analysis using EGFP sequence-specific probes. To accompany this plasmid, a plasmid (p-E-lacZ) was constructed by the insertion of a 2.2-kb partial lacZ

ORF into the corresponding site in p-EGFP-C1. The relevant regions of these plasmids and their parent plasmid pEGFP-C1 are depicted in figures 23 to 25.

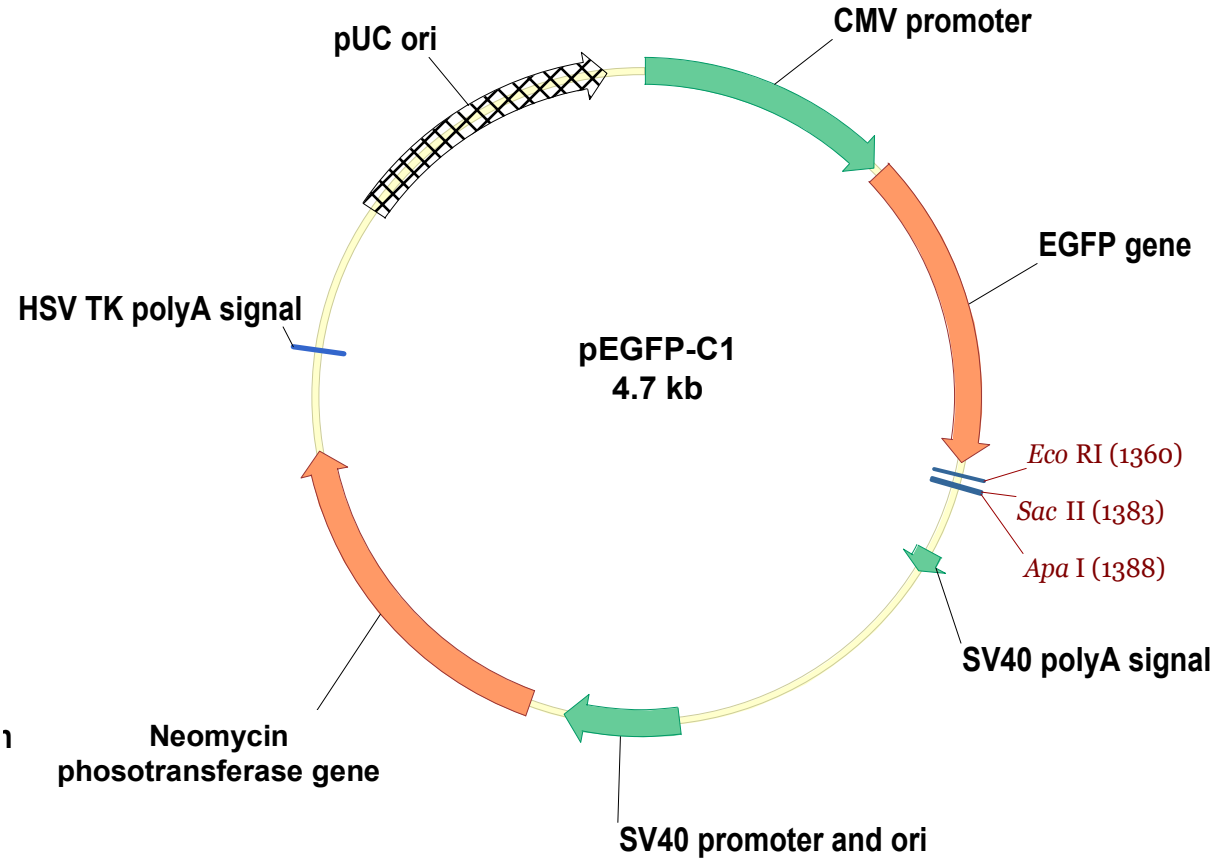


Figure 23. Map of pEGFP-C1. This vector contains two promoters, the CMV immediate early and the SV40 early promoter. The EGFP ORF is driven by the CMV promoter and the SV40 polyA signal is located downstream. The neomycin phosotransferase gene is driven by the SV40 promoter for selection purposes in transfected mammalian cells. Several important restriction enzyme sites used for cloning are indicated.

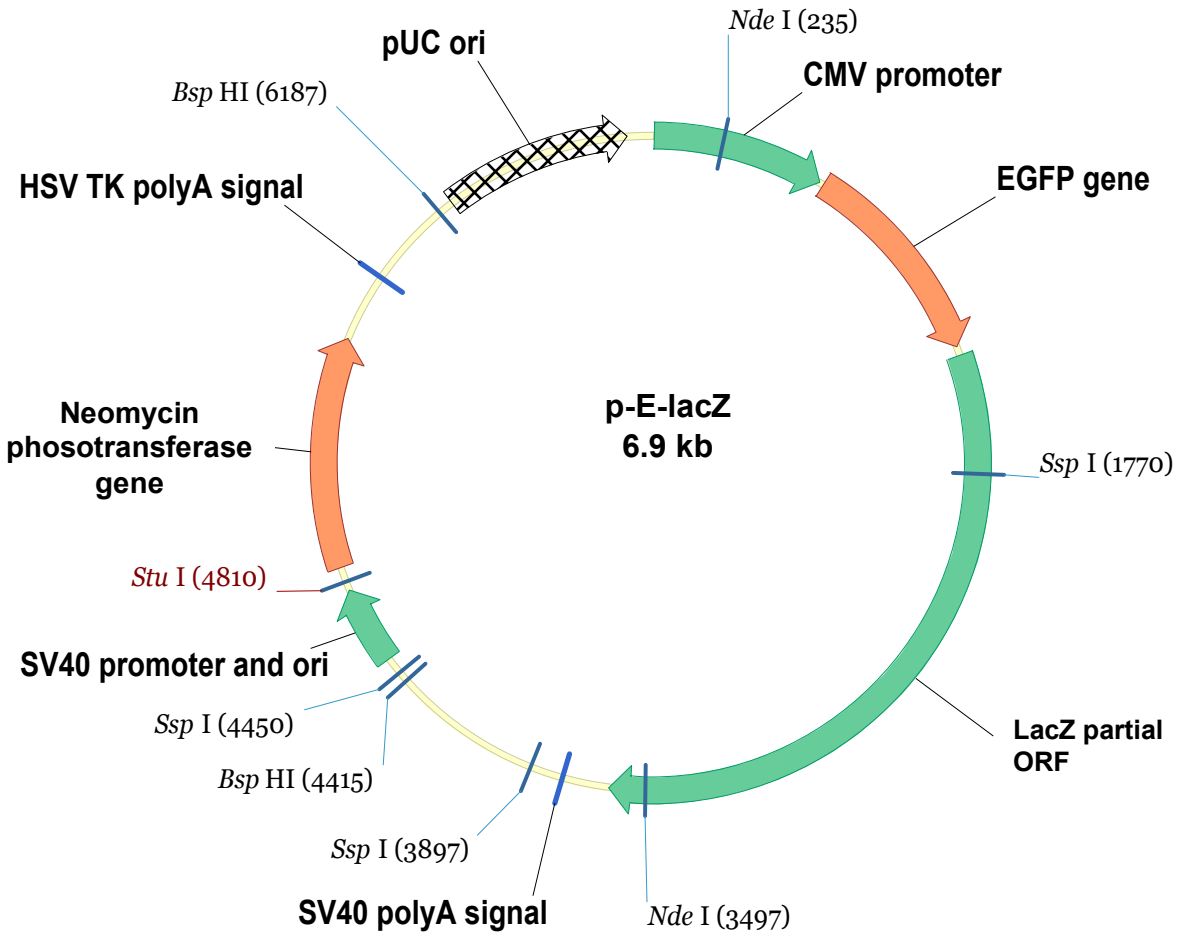


Figure 24. Map of plasmid p-E-lacZ. This plasmid was constructed by the insertion of a 2.2-kb partial lacZ-coding sequence into pEGFP-C1, between EGFP-coding sequence and SV40 polyadenylation signal. Important restriction enzyme sites are shown and their positions on the plasmid in parenthesis.

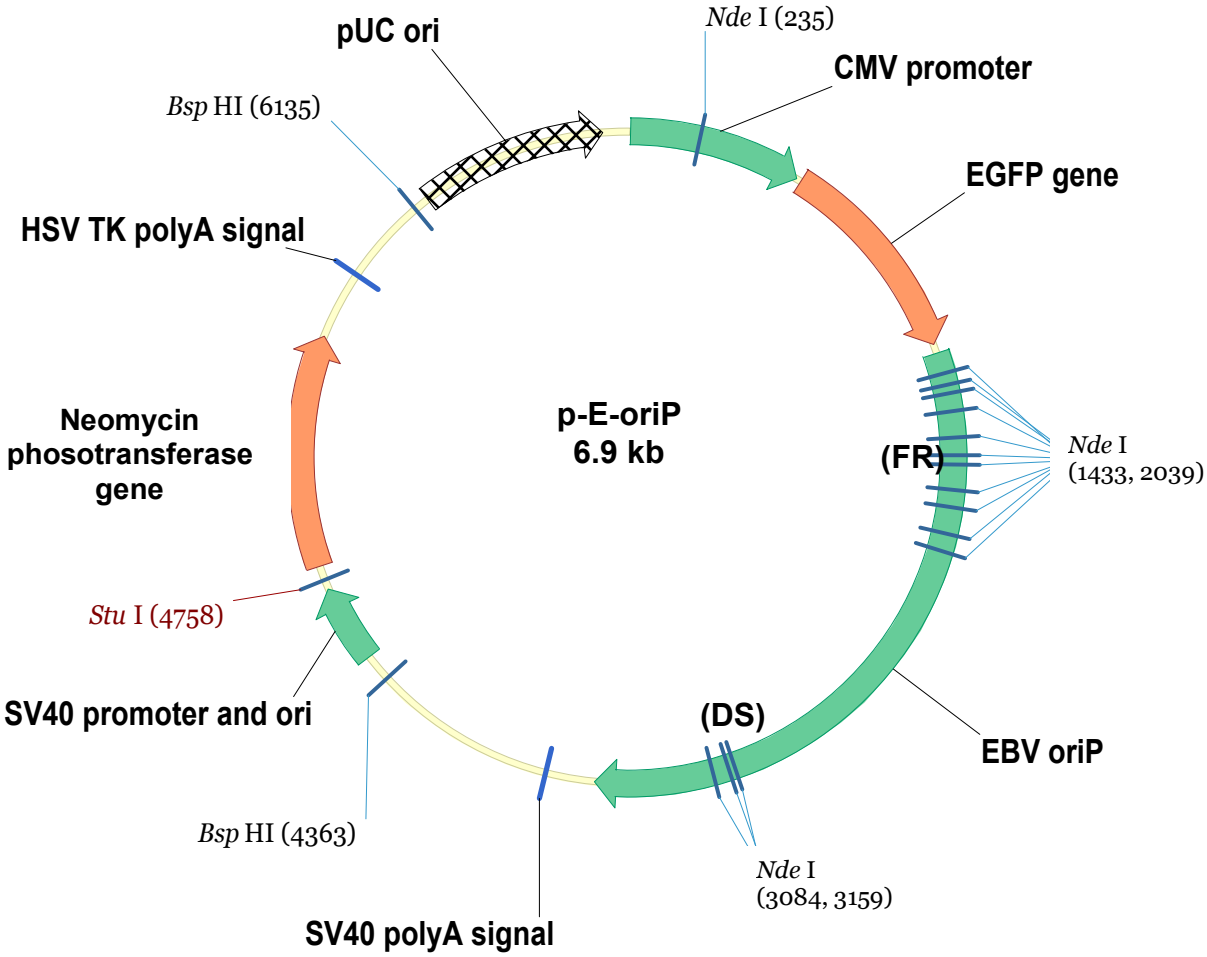


Figure 25. Map of plasmid p-E-oriP. This plasmid was constructed by the insertion of an *ori*P-containing 2.2-kb EBV genomic sequence into pEGFP-C1, between EGFP-coding sequence and SV40 polyadenylation signal. Important restriction enzyme sites are shown and their positions on the plasmid in parenthesis. For *Nde* I sites, only the first and last position are shown.

4.4.1 p-E-lacZ

For the construction of p-E-lacZ, pEGFP-C1 (Clontech) was digested using the restriction enzymes *Eco* RI and *Apa* I. The lacZ insert was excised from pTracer-CMV/Bsd/lacZ (Invitrogen) using the enzymes *Cla* I and *Apa* I. Both vector and insert were purified using the QiaQuick PCR purification kit before ligation. After ligation, the products were again purified and subjected to T4 DNA polymerase blunt end repair. After repair, the products underwent a final round of purification and ligated before transformation into DH10B electrocompetent *Escherichia coli*.

Plasmids extracted from several successful transformants were analyzed using the restriction enzyme *Nde* I. After enzyme digestion, the fragments were analyzed using gel electrophoresis. In addition, p-E-lacZ was also treated with the following enzymes to generate different restriction fragments: *Stu* I and *Bsp* HI double digest; *Stu* I and *Ssp* I double digest. After enzyme digestion, agarose gel electrophoresis analysis was performed. Figure 26 shows the gel photograph of one positive isolate.

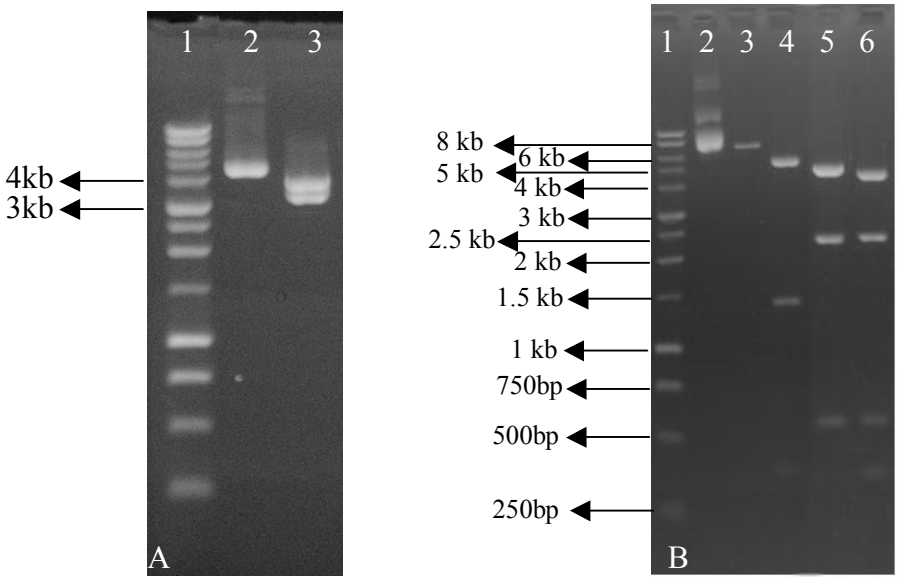


Figure 26. Gel photos of p-E-lacZ restriction enzyme analysis. (A) Lane 1: 1kb DNA ladder; lane 2: undigested p-E-lacZ; lane 3: p-E-lacZ digested with *Nde* I. (**B**) lane 1: 1kb DNA ladder; lane 2: undigested p-E-lacZ; lane 3: p-E-lacZ digested with *Stu* I; lane 4: p-E-lacZ digested with *Stu* I and *Bsp* HI; lane 5: p-E-lacZ digested with *Ssp* I; lane 6: p-E-lacZ digested with *Stu* I and *Ssp* I

From figure 26A lane 3, it can be seen that an extra *Nde* I restriction enzyme site was introduced by the lacZ insert, resulting in two closely migrating fragments, one 3.3 kb in size and a second fragment 3.6 kb in size. *Stu* I cleaved p-E-lacZ at a single site and generated a fragment of 6.9 kb as evidenced in figure 26B lane 3. As expected from the plasmid map (figure 26), a double digestion with *Stu* I and *Bsp* HI yielded a small 400 bp fragment, a 1.3 kb fragment and a large 4.2 kb fragment as seen in figure 26B, lane 4. *Stu* I and *Ssp* I double digest also yielded the expected larger 3.8 kb, 2.1 kb

fragments as well as the smaller 500 bp and 360 bp fragments (figure 26B, lane 6).

4.4.2 p-E-oriP

For the construction of p-E-oriP, pEGFP-C1 was digested with *Eco* RI and *Sac* II. The *oriP* insert was also isolated from the original *ori*P containing plasmid as described in section 4.1.1.3, using the *Eco* RI and *Sac* II. After the reaction, both vector and insert were first purified before ligation. The ligated products were transformed into electrocompetent DH10B *Escherichia coli*. Plasmids isolated from a few successful transformants were analyzed using *Nde* I restriction enzyme. In addition, the plasmid was also subjected to a double enzyme digest using *Stu* I and *Bsp* HI. The gel photograph of one positive isolate can be seen in figure 27 below.

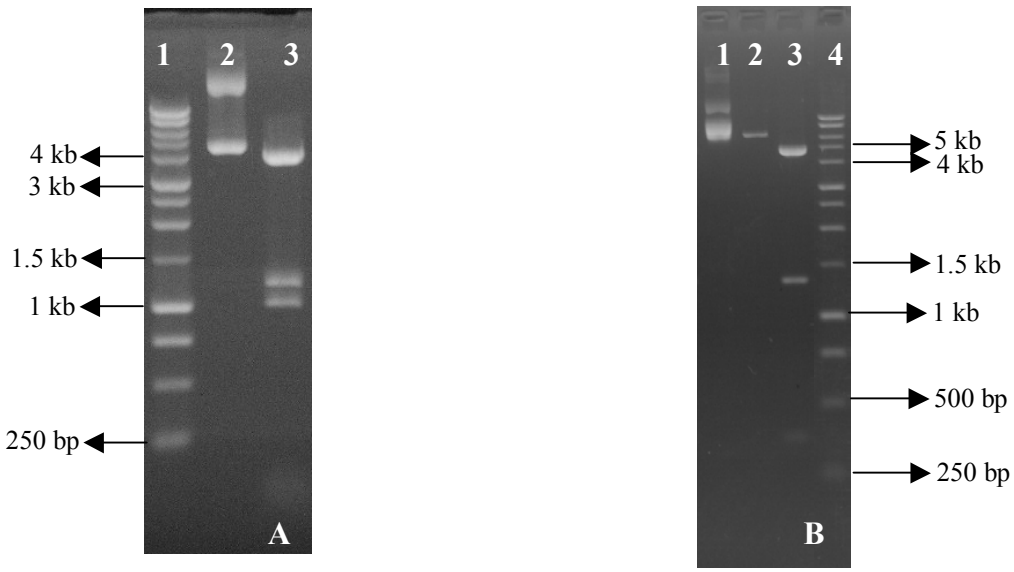


Figure 27. Gel photo p-E-oriP restriction enzyme analysis. (**A**) Lane 1: 1kb DNA ladder; lane 2: undigested p-E-oriP; lane 3: p-E-oriP digested with *Nde* I. (**B**) Lane 1: undigested p-E-oriP; lane 2: p-E-oriP digested with *Stu* I; lane 3: p-E-oriP digested with *Stu* I and *Bsp* HI; lane 4: 1kb DNA ladder.

As can be seen from figure 27A and the plasmid map (figure 25), digestion of p-E-oriP with *Nde* I yields small molecular weight DNA that appear as a smear at the bottom of Figure 27A lane 3 due to the multiple *Nde* I site, as well as the appearance

of the expected 1.1kb fragment found between the *Nde* I repeats of DS and FR, an indication of the presence of *ori*P. In addition, the presence of an additional *Nde* I site within the CMV promoter resulted in an additional 1.2kb fragment that could also be observed in the same lane. The digestion of p-E-oriP with *Stu* I and *Bsp* HI also yielded the expected bands: a larger 5.1kb band, a smaller 1.4kb and 390bp band (figure 27B, lane 3).

4.5 *In vitro* transcription of p-E-lacZ and p-E-oriP

To prepare the template for in vitro transcription, p-E-oriP was subjected to double enzyme digestion with *Stu* I and *Bsp* HI as described in section 4.4.2. p-E-lacZ was double-digested with *Stu* I plus *Ssp* I and *Stu* I plus *Bsp* HI to generate two templates, the unimpeded transcription of which will give a 1.1-kb and 4.4-kb transcripts respectively. The resulting transcription template of each plasmid as well as expected transcript size is represented in figure 28 below.

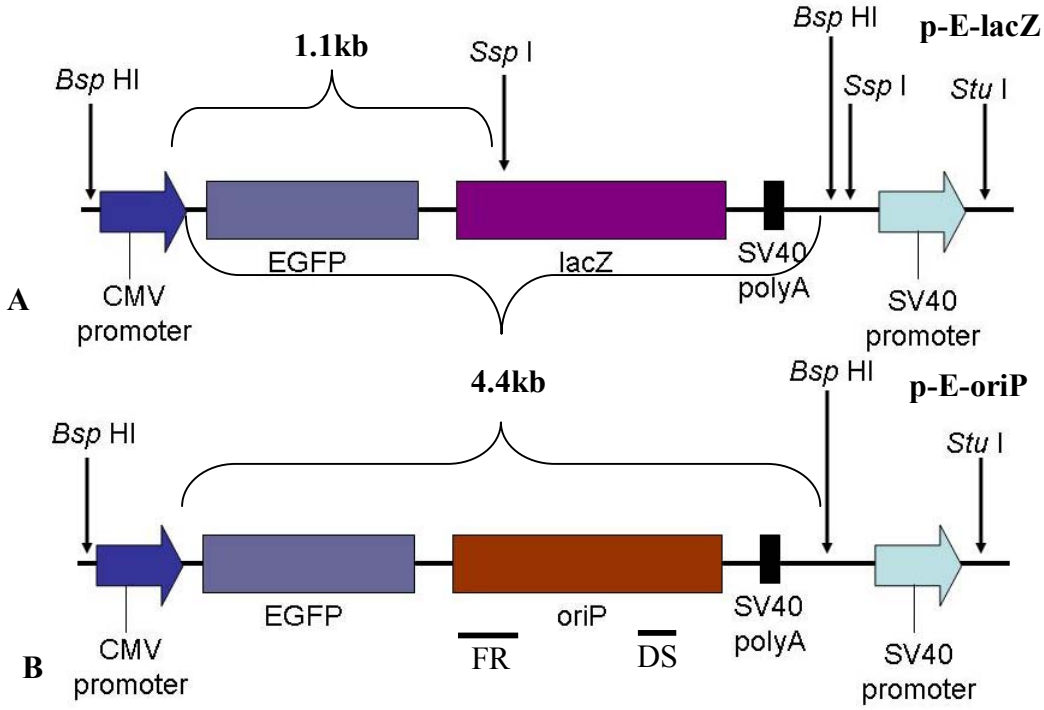


Figure 28. Linear maps of plasmids. Linearized plasmids digested with restriction enzymes (indicated by arrows). Expected run-off transcript sizes are indicated by braces. Plasmids not drawn according to size.
A: p-E-lacZ digested with *Stu* I and *Bsp* HI; expected run-off transcript: 4.4kb. If digested with *Stu* I and *Ssp* I, expected size of run–off transcript to be 1.1kb.
B: p-E-oriP digested with *Stu* I and *Bsp* HI; expected run-off transcript: 4.4kb
The EGFP ORF will help to generate a transcript at least 700 bp in length.

After digestion, the reaction was purified and 300ng of each digested plasmid was used as template for *in vitro* transcription. Northern blot analysis of the *in vitro* transcription was done and the results can be seen in figure 29 below.

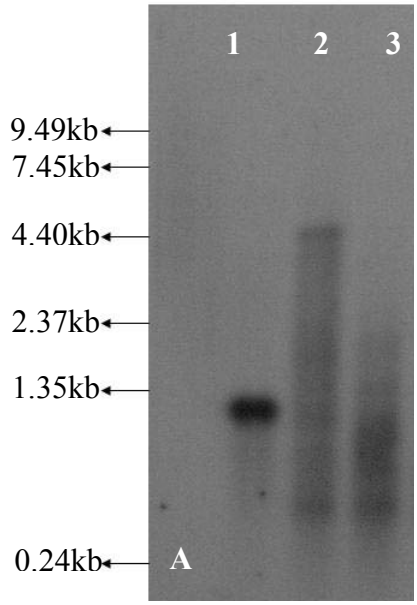


Figure 29. Northern blot analysis of *in vitro* transcription. The transcription templates were p-E-lacZ digested with *Stu* I and *Ssp* I. Lane 1: p-E-lacZ digested with *Stu* I and *Bsp* HI; lane 2: p-E-oriP digested with *Stu* I and *Bsp* HI; lane 3: Probes specific to the EGFP gene was used.

Lane 1 is the *in vitro* transcripts obtained from using p-E-lacZ digested with *Stu* I and *Ssp* I as a template. As expected, *in vitro* transcription yielded a small molecular weight run-off transcript with an estimated size of 1.1kb, while p-E-lacZ digested with *Stu* I and *Bsp* HI (lane 2) yielded an estimated 4.4kb run-off transcript, also as expected. This shows that the *in vitro* transcription kit employed is capable of transcribing long templates. If transcriptional arrest did not occur at *ori*P, p-E-oriP digested with *Stu* I and *Bsp* HI (lane 3) would be expected to yield a run-off transcript similar in size to lane 2. However, it can be seen that whatever transcripts that were present in lane 3 were not full length transcript. A few conclusions can be drawn from the above data. Firstly, the presence of *ori*P clearly poses an inhibitory effect on transcription *in vitro*. Taken together with the inhibition of transcription of p-S-oriP and p-S-oriP.1 (section 4.3.1), these results show that the inhibitory effect of *ori*P is independent of the promoter. Secondly, the appearance of partial length products for p-E-oriP shows that it is transcription elongation, rather than initiation being inhibited.

The majority of the partial length products were roughly about 1 kb. It would also seem that transcription of p-E-oriP, after 1 kb was drastically reduced. Interestingly, the FR of the *ori*P is located immediately after the EGFP ORF in this construct. This would suggest that that transcription could be arrested at the FR region of the *ori*P.

The data presented thus far show *ori*P blocks transcription elongation. There are two types of transcription elongation block. If the transcription elongation complexes dissociate under this block, the event is known as transcription termination. Alternatively, if the blocked elongation complexes remain intact, the event is known as transcription arrest (Wiest *et. al.*, 1992). A terminated transcript, being free, is in a much smaller molecular framework than an arrested transcript, which is still associated with the DNA template, RNA polymerase and other transcription elongation protein factors. Therefore, transcription termination and arrest can be distinguished by size exclusion chromatographic analysis of the transcription products. Terminated transcripts will elute from a gel filtration column much later than arrested transcripts. As mentioned in section 4.2.1, the inhibition of both replication and transcription by the presence of *ori*P in a transcription unit can be fully explained if *ori*P induces transcription arrest. Therefore, I intended to carry out gel filtration chromatographic analysis of the *ori*P-blocked transcription products to test if *ori*P actually induces transcription arrest. In order to achieve this goal, terminated and arrested transcript controls were first needed to calibrate the gel filtration column. For the terminated transcript control, purified RNA or the run-off transcript from any reaction could be used. However, the arrested transcript control would be harder to obtain because there has not been any DNA sequence known to induce transcription

arrest. To obtain that, I attempted to arrest transcription by providing insufficient amounts of substrates, the rNTPs.

4.5.1 Transcription arrest under shortage of substrates

In the attempt to achieve transcription arrest under substrate shortage, in vitro transcription reactions were done using *Stu* I and *Ssp* I digested p-E-lacZ as the templates and different concentrations of rNTPs. The results can be seen in figure 30 below.

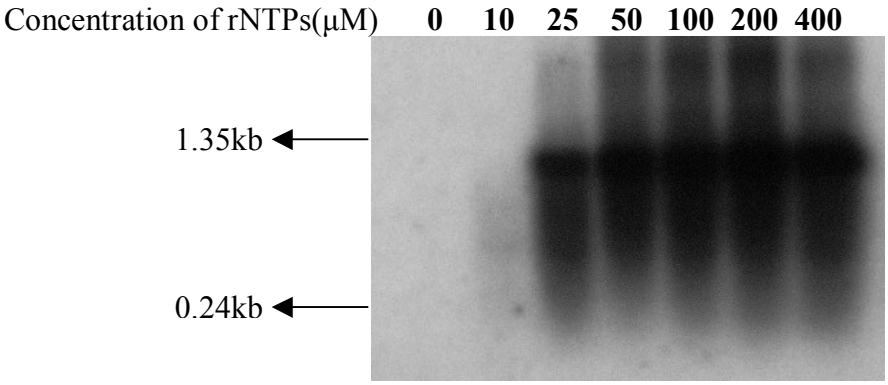


Figure 30. Transcription arrest under shortage of substrates. *In vitro* transcription was performed using *Stu* I and *Ssp* I digested p-E-lacZ as the templates and different concentrations of rNTPs. Time of reaction incubation is 20 minutes.

From the figure, at 0 μM rNTPs, there was no full length transcripts produced as expected. The 1.1-kb full-length transcript can be seen from 25 μM onwards, although the intensity of the band started to reach a maximum at 50 μM. This shows that the concentrations of rNTPs, 400 μM, regularly used in transcription reactions in vitro are vastly excessive. Reaction at 10 μM of rNTPs gave products that were smaller than the full length transcript. These small transcripts were most probably from the elongation complexes arrested at various sites due to shortage of rNTPs. The amounts of these arrested transcripts were much lower than those from reactions with higher

concentrations of rNTPs. This indicates that transcription initiation in vitro is much less efficient at these low substrate concentrations. Using 400µM of rNTPs would give us maximum transcriptional reaction rate, which may affect any arrest by "forcing" the polymerase to read through the arrest site. Therefore, the idea of optimizing the rNTPs is to obtain the minimal amount of rNTPs that would give a full length transcript. On the other hand, using too little a concentration of rNTP could "induce" arrest. From the results of the rNTP optimization assay, it can be seen that 50µM rNTPs was sufficient. This, in conjunction with work done by Freund and McGuire (1986), which characterized human term placental RNA polymerase II and found that the Km for rNTP to range from 45µM to 62µM, justified the use of 50µM rNTP as a workable concentration for *in vitro* transcription reactions in this study. In addition, due to the fact that I had limited time I decided not to optimize any further.

4.6 Size exclusion chromatography able to separate DNA/RNA according to size

Home-made gel filtration columns were prepared, as described in section 3.9.1, to analyze the in vitro transcription products of p-E-oriP-derived template to determine if *ori*P induces transcription arrest. Before subjecting transcription products to gel filtration chromatography analysis, preliminary studies on the columns were first done to see if the columns cast this way were generally RNase-free and able to separate nucleic acid molecules according to their sizes. Thus, DNA and RNA markers were loaded onto the column and eluted into fractions. 10µl of each fraction were then subjected to agarose gel electrophoresis analysis. After electrophoresis, the gel was viewed under UV light and a picture taken. The results are shown in figure 31 and 32 below.

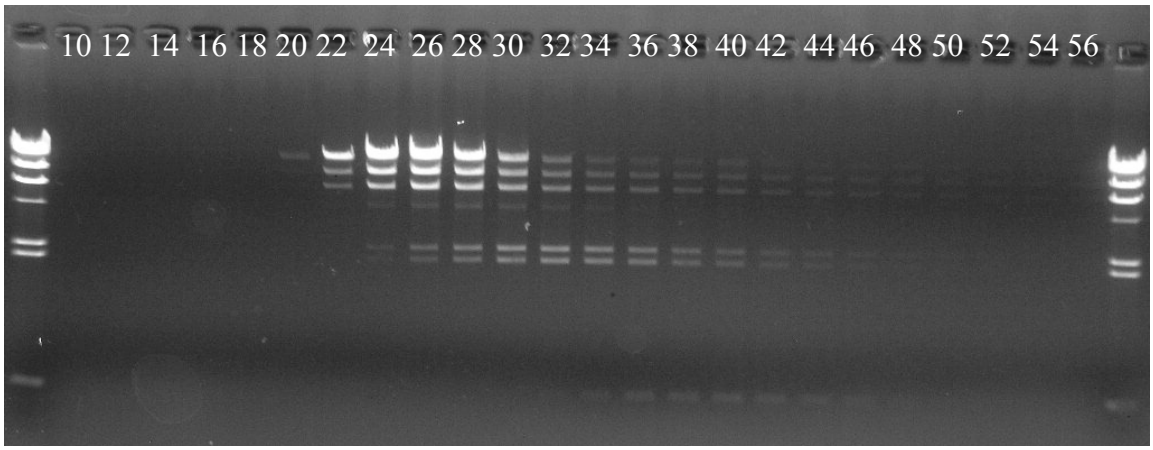


Figure 31. Gel analysis of size exclusion chromatography using λ *Hin*d III DNA ladder. Only fractions 10 and beyond were collected.

As can be seen in the figure 31, the higher molecular weight 23kb band started to elute out from fraction 20, while the 2 kb band eluted out from fractions 24 onwards. The 500bp band eventually eluted out from fraction 34.

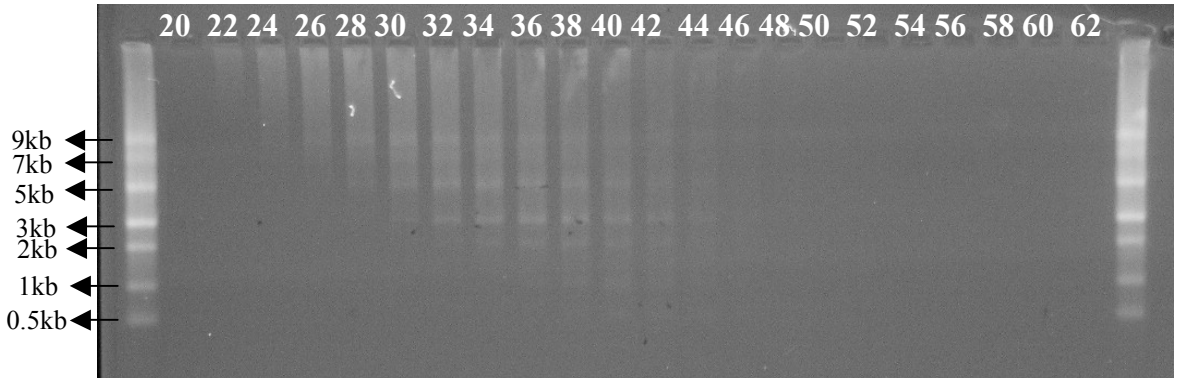


Figure 32. Gel analysis of size exclusion chromatography of RNA ladder (New England Biolabs). Fractions were collected from fractions 20 onwards. Size of RNA ladder is indicated by arrows in kilo-base-pairs.

As for the separation of RNA, it can be seen from figure 32 that the RNA molecules started to elute out from fractions 22 onwards. The 9kb band started eluting out from fraction 26 onwards, while the 5kb band started eluting out from fraction 28. The 2kb ladder started eluting out from 32 and finally the 0.5kb ladder elutes out at fraction 40.

From the two experiments described above, it can be concluded that the home-made gel filtration columns can separate nucleic acids according to size and be RNase-free. They are thus suitable for the analysis of transcription products.

4.6.1 Unimpeded transcription

With the success in size-separating purified DNA and RNA using the gel filtration columns, the gel filtration columns was tested further on the products of unimpeded or run-off transcription. The run-off transcription products were first prepared by *in vitro* transcription using *Stu* I- and *Ssp* I-digested p-E-lacZ as the templates and 50µM rNTPs, as described in section 4.5.1. An aliquot was taken after 20 minutes for input and the rest loaded onto a column and fractions collected. Northern blot analysis was done and the results can be seen in figure 33 below.

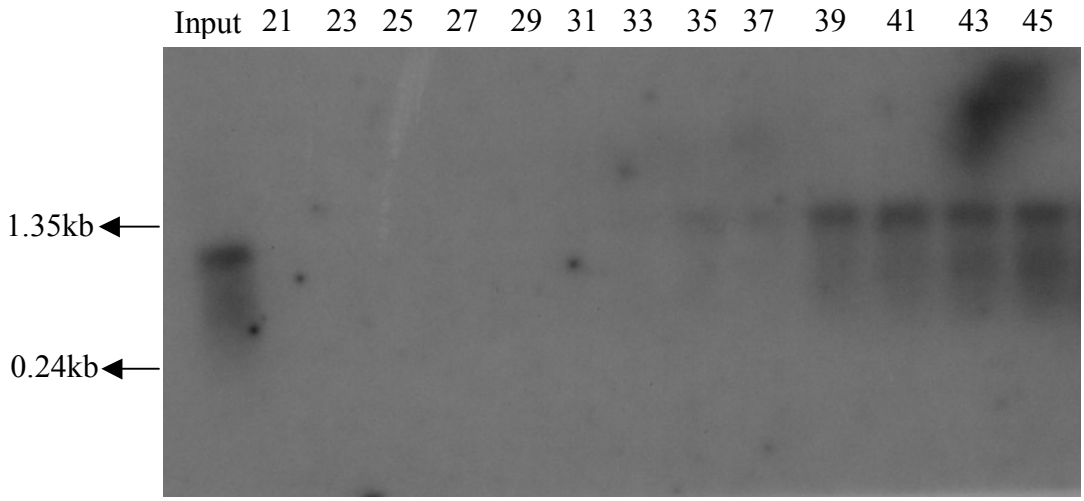


Figure 33. Northern blot analysis of size exclusion chromatography of *in vitro* transcription. (A): size exclusion of p-E-lacZ (digested with *Stu* I and *Ssp* I). The first lane is input run-off transcripts. Only fractions 21 and later are analyzed. Each lane contains 2 fractions.

As expected, the transcripts produced in this reaction were the 1.1-kb full-length transcripts (input lane) indicating that transcription elongation had proceeded to the end of the linear template. These transcripts eluted out from the gel filtration column

from fractions 35 onward (figure 33). This elution pattern was very similar to that of the 1-kb RNA marker (figure 32). In fact, run-off transcripts are thought to be free from the association with any other molecules. Therefore, these results show that the gel filtration column is functional in separating components from a complex sample, such as the in vitro transcription reaction. The input RNA may appear to be smaller in size than the eluted RNA but that was due to a misalignment of the film and the camera resulting in a slanted image.

4.6.2 Column is capable of excluding artificially induced arrested RNA polymerase

Another chromatographic run was done to determine how arrested transcription complexes eluted from the gel filtration column. To do that, arrested transcription complexes were prepared by *in vitro* transcription using *Stu* I- and *Ssp* I-digested p-E-lacZ and 10μM rNTPs, as described in section 4.5.1. An aliquot of the reaction was then taken as input control and the rest of the reaction was loaded directly onto a column and fractions collected from 20 onwards. The fractions were analyzed by Northern blot and the results can be seen in figure 34 below.

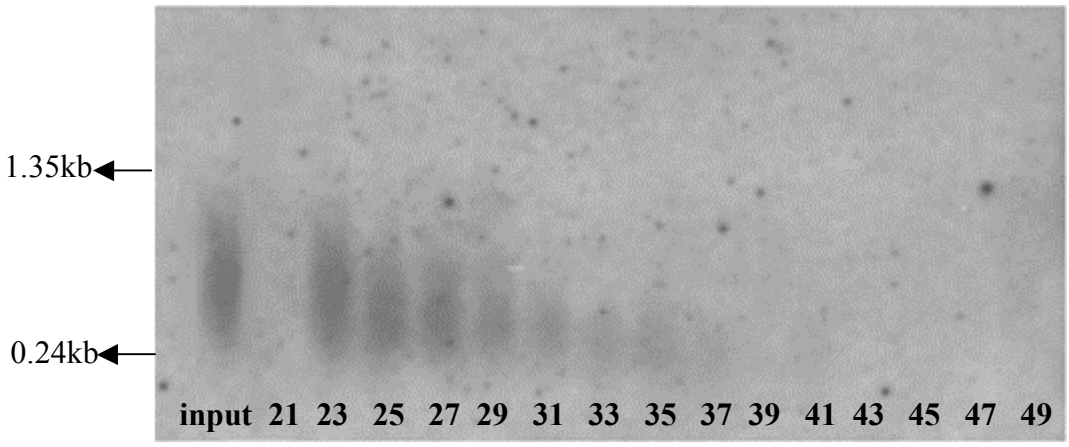


Figure 34. Exclusion chromatography of artificially arrested *in vitro* transcription. The first lane contains input transcripts. The numbers at the bottom of each lane represents the fraction number. 2 fractions were collected as one sample.

As expected, the transcripts produced in the reaction were shorter than the 1.1-kb full-length transcript indicating that transcription arrest occurred (input lane). These transcripts started eluting out from the column as early as fraction 23 (figure 34), much earlier than the similar size RNA species the RNA ladder (figure 32). This elution pattern is expected of arrested transcripts, which remain associated with RNA polymerase, DNA template and probably other transcription factors. This result also provides evidence for the capability of the gel filtration column to distinguish free and complex-bound transcripts, setting the stage for the analysis of nature of *ori*P-dependent transcription block.

4.6.3 *Ori*P induces transcriptional termination, rather than arrest, in pEGFP-oriP

In vitro transcription was performed using *Stu* I- and *Bsp* HI-digested pEGFP-oriP as the template and 50μM rNTPs. An aliquot of the reaction was then taken as input control and the rest of the reaction was loaded directly onto a column and fractions collected from 20 onwards. The fractions were analyzed by Northern blot and the

results can be seen in figure 35 below.

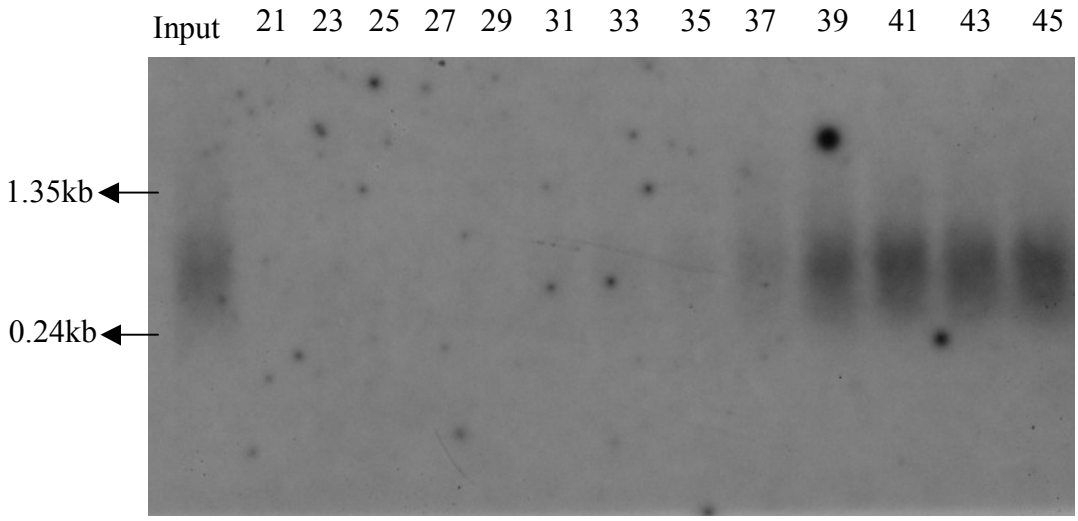


Figure 35. Northern blot analysis of size exclusion chromatography of *in vitro* transcription. Size exclusion of pEGFP-oriP (digested with *Stu* I and *Bsp* HI). The first lane is input run-off transcripts. Only fractions 21 and later are analyzed. Each lane contains 2 fractions.

Consistent with the results shown in section 4.5, transcripts of lengths between 0.3-1 kb, rather than the 4.4-kb full-length transcripts, were obtained (input lane), indicating that transcription elongation was blocked within *ori*P. These transcripts eluted from the gel filtration column mainly from fraction 39 onwards. The elution pattern was similar to that of the run-off transcripts (figure 33) or purified RNA of similar sizes (figure 32), and distinct from that of transcripts in arrested transcription complexes (figure 34).

The appearance of free rather than arrested transcripts points to termination rather than arrest, as being the probable cause of transcription inhibition by *ori*P. This result does not support the possibility that inhibition of *ori*P-dependent replication is caused by the arrest of transcription complexes at *ori*P. Thus, how replication is inhibited when *ori*P is located within a transcription unit remains unexplained.

Discussion

5. Discussion

5.1 Inhibition of *ori*P replication function is dependent on the presence of promoter

From figure 18, it is clear that the presence of the SV40 early promoter in a plasmid bearing *ori*P inhibits replication. The *ori*P function was severely impaired upon placing it, in both orientations, immediately downstream of this strong and constitutive promoter. Inhibition of replication was still present but to a lesser degree when *ori*P was placed further downstream of the promoter. Several conclusions can be drawn from these observations. Firstly, inhibition of replication is independent of the orientation of the *ori*P within the transcriptional unit. Secondly, the further away the *ori*P is from the direction of transcription, the less inhibition of replication there is, indicating that the progress of the transcriptional machinery through the *ori*P most probably played a major role in reducing the function of *ori*P. This is in agreement with some of the observations in previous works done using other known replication systems (Tanaka *et al*., 1994; Haase *et al*., 1994; Pan *et al*., 1995).

5.2 Transcription through *ori*P was inhibited

From the Northern blot analysis of total RNA isolated from the transfected cells (section 4.2), it was clear that the transcriptional elongation complex was somehow prevented from reading through the *ori*P. This observation was further corroborated by a similar finding using *in vitro* transcription of the same templates (section 4.3). *In vitro* transcription of templates containing *ori*P immediately downstream of the SV40 promoter was highly inefficient. There are two possibilities that could have resulted in

such an observation: transcription termination or transcriptional arrest. While transcription termination would not be able to offer much explanation about any possible interplay between transcription and replication, transcription arrest on the other hand, if it really occurred, could potentially help explain the observations obtained in section 4.2 as to why transcription could inhibit replication, and to such a strong extent.

Based on the preliminary results obtained, transcriptional arrest along *ori*P seemed like a probable explanation. As arrest is irreversible, it results in complexes trapped on the DNA template, preventing the progression of other elongation complexes. This could represent a spatial obstruction that prevents replication preinitiation complexes from assembling at the *ori*P, thus inhibiting replication. The fact that transcriptional arrest occurred not only *in vivo* (both in the presence and absence of EBNA-1) but *in vitro* as well indicated that the DNA sequence was the main culprit in inducing arrest. This was not surprising as *ori*P is an A-T rich region, containing many repeat sequences, especially near the FR (family of repeats) region. And drawing parallels with work done by Kerppola and Kane (1990); one of the strongest known arrest site in the histone 3.3 gene contains a T-rich region. It was therefore highly be possible that *ori*P could also contain transcription arrest sites.

However, it was not easy to use the plasmids p-S-oriP and p-S-oriP.1 to determine exactly where within *ori*P transcription got inhibited because the *ori*P was too close to the promoter and any arrested transcripts would be too short for detection. In order to solve this problem, elongation of transcription would have to be allowed to proceed for a certain distance on the template before encountering arrest. This would enable

the easy detection of any arrested transcripts. Two new plasmids were constructed: one involving the insertion of *ori*P downstream of EGFP gene (p-E-oriP; section 4.4.1) and the other involving the insertion of a lacZ ORF downstream of the EGFP as a positive control (p-E-lacZ; section 4.4.2). These plasmids were digested with *Ssp* I or *Bsp* HI to generate different size templates used for *in vitro* transcription to test for transcriptional arrest. An additional restriction enzyme digest with *Stu* I is needed to remove the SV40 promoter from the template. By doing so, transcriptional elongation cannot proceed from the SV40 promoter and no contaminating transcripts will be produced. True enough, transcriptional inhibition still seemingly occurred in p-E-oriP, as evidenced in figure 27. From the range of the molecular size of the smear, the majority of the inhibition of elongation seemed to occur at the FR region, with a minority of inhibition occurring at the DS region.

One of the potential problems that could surface was the possibility that the concentration of rNTPs utilized during *in vitro* transcription. The recommended concentration of 400μM was for the purpose to ensure maximum reaction rate. However, it was unnecessary to ensure maximum reaction rate for this study. In fact, there was a worry that the high reaction rate resulting from using 400μM rNTPs could affect arrest in an unknown way or by "forcing" the polymerase to read through the arrest site. Therefore, the optimization experiment of minimum rNTPs needed for elongation to proceed to completion was performed. From the results of the rNTP optimization assay, it can be seen that 50μM rNTPs was sufficient. This, in conjunction with work done by Freund and McGuire (1986), which characterized human term placental RNA polymerase II and found that the Km for rNTP to range from 45μM to 62μM, justified the use of 50μM rNTP for *in vitro* transcription

reactions in this study.

One of the other problems was the need for a positive control to show that size exclusion chromatography was capable of isolating arrested transcripts cum RNA polymerase. From the optimization of rNTPs study, it would seem no full length transcripts was observed using 10μM of rNTPs, most likely due to the slow reaction rate. This low reaction rate, coupled with the need for the transcription complex to clear the promoter before the elongation complex is stable, could result in more transcription complexes undergoing abortive initiation. This also probably accounts for the low intensity of the smear observed for 10μM rNTP lane in figure 30. The preincubation step without the addition of rNTPs was necessary for ensuring synchronized transcription elongation.

5.3 Something else other than transcriptional arrest causes replication inhibition

As figures 31 and 32 showed, the column used in this study was capable of differentiating between arrested RNA transcripts and free RNA transcripts. Although there exist the problem of overlapping peaks but that does not present a major obstacle to this study as the interest was not in isolating pure free or arrested transcripts.

Unfortunately, the results obtained from section 4.6.3 seemed to indicate that the partial length transcripts isolated from *in vitro* transcription of p-E-oriP (digested with *Stu* I and *Bsp* HI) using 50μM of rNTPs were free rather than arrested transcripts. Strong evidence for this stems from the fact that the elution profile was similar to the elution profile of run-off transcripts isolated from *in vitro* transcription of p-E-lacZ

(digested with *Stu* I and *Ssp* I). What this probably indicates is that the transcriptional termination rather than transcriptional arrest probably occurred at *ori*P and that the hypothesis that arrest of the transcription elongation complex at *ori*P resulted in the obstruction of replication initiation had to be reexamined. But the clear inhibition of replication by the presence of a promoter nearby still meant that transcription could play an inhibitory role in replication either directly or indirectly.

5.4 Passage of transcription machinery could prevent replication initiation

It has been shown that the arrest of transcription on *ori*P that could have resulted in prevention of replication initiation does not occur in this system; rather, it was transcription termination that most likely occurred. However, transcription termination itself does not explain how it may affect replication so negatively. One more plausible explanation would be that the progression of transcriptional elongation complex along the template could inhibit replication and physical arrest of the machinery along the template was not needed. This was shown elegantly by Haase *et al*., 1994 and Pan *et al*., 1995. Inserting transcriptional termination sequences between the origin of replication and the promoter, they prevented the elongation complex from reading through the origin and thus reduced the level of replication inhibition. It is possible that this might be the case in this study, and the observation that transcription through *ori*P was inhibited may not be related to the interplay between transcription and replication at all.

How can transcription elongation inhibit replication then? Two possible explanations exist: First, the physical collision between the transcription and replication machineries along the same template as they met head-on could have accounted for

the inhibition. This was shown in the prokaryotic system by Mirkin and Mirkin (2005). In that study, the authors showed that the replication fork progression stalled when it is within a DNA region which is being transcribed. They also showed that although transcription and replication proceeding in the same direction had no effect on one another, stalling of the replication fork occurred only if the DNA polymerase were to meet the RNA polymerase head on during elongation.

In addition to the mere physical collision between transcription and replication, the positive supercoils generated by both transcription and replication (Liu and Wang, 1987; Peter *et al*., 1998) as they met head-on could help further worsen the inhibitory effect of transcription on replication. This build-up of positive supercoiling, (termed "knotting" by Olavarrieta *et. al*., 2002) was an effect brought about by the head-on collision of transcription and replication and had disastrous results. A similar explanation could be provided for this study: similar to what was observed in the prokaryotic system, as the elongation complexes of both transcription and replication meet head on, inhibition of replication was severely impaired either due to physical collision and/or due to the positive supercoils generated.

If head-on physical collision between the two machineries was the case, it would still mean that replication elongation in the same direction as transcription would be allowed to proceed to the end. And since DNA replication is bidirectional, it would mean that at least one molecule of daughter DNA would still be produced as compared to two, and replication would be 50 % of what it could be per cell cycle. From the results obtained in section 4.1.3, an estimated 10 % of pSV40-oriP and pSV40-oriP.1 constructs were replicated after 72 hours. Assuming that the transfected

cells had undergone at least two to three rounds of replication (1 cell cycle for every 24 hours) before the plasmids were extracted, the expected amount of replicated plasmid would be from the 10 % to 25 % range, which would also explain for the observed levels of replication in this study. Given that the expected and observed percentages of replicated plasmids were roughly the same; the possibility that physical collision between transcription and replication resulted in inhibition of replication should be included in future work.

The second possible explanation could be that as transcriptional elongation complex progresses through the *ori*P, it prevents the preinitiation complex from forming, as was suggested by Nieduszynski *et al*., 2005, thus resulting in an overall inhibition of replication. The authors showed that transcription into known origins (ARS1 and ARS121) increased their dependence for Mcm2 to 7. This observation indicated that transcription elongation could also inhibit replication by preventing the formation of pre-replication complex formation. Further work would be needed to be done to show if this is really the case.

5.5 Chromatin remodeling could also affect replication

As mentioned in section 2.3.2.1, to initiate transcription, the nucleosome must be first remodeled such that access to the promoter is granted to the RNA polymerase holoenzyme. Certain transcription factors such as SWI/SNF, BRCA1 and even elongation factors such as Elongator and FACT, have been found to contain chromatin remodeling properties. Since replication also requires an initiation stage in which replication machinery needs to obtain access to origins wrapped in histones, it could be possible that chromatin remodeling plays a vital role in replication as well.

And in this case, transcription factors could well play a double role in not only transcription but replication as well.

However, most of the works reviewed so far seem to indicate that transcriptional factors play a positive role in replication. As described in section 2.3.2, the Abf1 binding domain in the ARS1 was found to play a positive effect on replication and the C-terminal acidic domain can be replaced with transcription factors from other organisms (Li *et al*., 1998). In fact, Abf1 has been known to be involved in the repositioning of nucleosomes near the origin of replication and has been also shown to be able to recruit Esa1, which is an essential histone acetyltransferase (Reid *et al*., 2000; Lascaris *et al*., 2000). It could be possible that Abf1 facilitates replication by remodeling the chromatin structure near the origin to be more accessible to the replication initiation complex.

However, an exception to this claim can be found in another known yeast replication origin ARS301. In its normal state in the chromosome, the ARS301 was described as an inactive origin of replication, incapable of initiating replication. When ARS301 was taken out of its native context and inserted into a plasmid, the ARS301 becomes active and is able to induce replication of the plasmid. This ability could be negated by inserting Abf1 transcription binding sites near the ARS301 region on the plasmid (Kohzaki *et al*., 1999).

Combining the observations in both works, there is no clear answer whether a transcription factor like Abf1 plays a positive or negative role in replication. It could be possible that it can play both roles depending on where it binds to relative to the

origin. Take ARS301 for example: it could be that in the native chromosome state, the chromatin structure near and on the ARS301 region was arranged in a way such that replication initiation was prevented from occurring. But, when the ARS301 was taken out of the chromosomal environment and inserted onto a plasmid, the chromatin structure surrounding or within the ARS301 could have potentially changed to a form more conducive for replication to occur. As Abf1 possesses the ability to recruit chromatin remodeling enzymes like Esa1, the insertion of Abf1 transcription factor binding sites could have the potential to revert the chromatin structure of the plasmid back to a similar state found in the chromosome, or potentially change it to a structure inhibitory to replication.

Similarly, the inhibition of replication observed in this study could have originated from the presence of transcription factors rather than the passage of the elongation complex through the *ori*P. This is highly likely as both promoters used to drive transcription through *ori*P are strong promoters. The CMV immediate-early promoter used in this study has been shown to contain binding sites for transcription factors, such as NF (nuclear factor)-κB, ATF (activating transcription factor)/ CREB (cAMP response element binding protein), SP1 and AP1 (activator protein) to name a few (Meier *et al*., 2002). Similarly, the SV40 early promoter is also known to contain binding sites for SP1 and AP1 as well (Lee *at al*., 1987). It could be possible that upon binding to the respective sites found on the CMV or SV40 promoter, these transcription factors could have influenced the chromatin structure by recruiting remodeling enzymes. And under the circumstances of this study, the chromatin could have been remodeled to a configuration that prevented the initiation of replication.

As described in section 2.2.1, *ori*P consists of two separable and distinct regions, DS and FR. The DS has been shown to be the origin of replication where the ORC assembles through interaction with the EBNA-1 protein, whereas the FR has been described as playing a role of a transcriptional enhancer as well as ensuring that replicated episomes are properly segregated to daughter cells during mitosis. However, it was shown that in certain established EBNA-1 expressing cell lines, the presence of FR in the *ori*P could in fact inhibit the ability of the *ori*P containing plasmid to replicate (Leight and Sugden, 2001). This observation can be related to the both positive and negative effects insertion of transcription factor Abf1 binding sites can have on the ability of different ARS to act as an origin of replication. Being described as a transcriptional enhancer element, it would be expected to find certain binding sites for transcription factors within the FR. Indeed, it was shown recently that transcription factors Oct1 and Oct2 can also bind to FR and activate transcription without the presence of EBNA-1 (Almqvist *et al*., 2005). It should be noted that there may exist other hitherto undiscovered transcription factors that can also bind to the FR.

It could therefore be possible that the presence of a CMV or SV40 promoter upstream of the *ori*P could have interfered with the chromatin structure of the *ori*P, resulting in a configuration that was inhibitory to replication. Or the presence of more than one enhancer element (one from the CMV promoter or SV40 promoter; the second from the FR element of the *ori*P) could have resulted in a recruitment of different chromatin remodeling enzymes to each enhancer element. That could potentially result in a competition between the remodeling enzymes to shape the chromatin. Although the insertion of *ori*P upstream of the SV40 promoter did not yield the same

level of inhibition, it could also mean that the position of the transcription factors binding sites may be vital in determining the chromatin structure. More work would be needed to ascertain this theory.

5.6 Future directions

As described in sections 5.4 and 5.5 above, there are two main other possible explanations as to how transcription can inhibit replication. The first possibility is that transcriptional arrest is not needed and the passage of the elongation complex through the *ori*P would itself be sufficient in inhibiting replication. qTranscriptional elongation can inhibit replication through the physical collision and/or positive supercoils generated when they meet head on; or elongation could prevent the assembly of pre-replication complexes, thereby inhibiting replication. One way to test these two possibilities would be to first insert transcriptional termination sequences between the promoter and the *ori*P in p-S-oriP and p-S-oriP.1 and study the effects these termination sequences on the level of replication inhibition. It would also be interesting to see if the distance of *ori*P from the termination sequences would have any added effect on alleviating the inhibition of replication. This could be done by inserting the *ori*P at different sites downstream of the termination sequences. Southern and Northern blot analysis to would have to be performed to see if inhibition of replication would still occur with the introduction of termination sites as well as to see if the termination sites would be sufficient to terminate transcription *in vivo* and *in vitro*.

The second possibility would be that chromatin remodeling could have a role in inhibiting replication. Although it is not certain as to how chromatin remodeling

enzymes can be recruited to the promoter or *ori*P, the presence of transcription factor binding sites within both remain a plausible explanation. Regardless of whether the binding of transcription factors to its cognate sequences result in the recruitment of chromatin remodeling enzymes; or if chromatin remodeling enzymes are directly recruited through hitherto unknown sequences, one of the first things to do would be to examine the nucleosome arrangement of the plasmids p-S-oriP, p-S-oriP.1, p-oriP-S and p-oriP-ΔS.

One of the ways to determine nucleosome arrangement on the plasmids would be to utilize non-sequence specific nucleases such as P1 nuclease or micrococcal nuclease that are able to cleave nucleosome free DNA sequences (Chu *et al*., 1990; Telford and Stewart, 1989). The plasmids would have to be first transfected into EBNA-1 expressing cells and incubated at 37ºC at 5% CO_2 after transfection for 48 hours. This is to allow for any replication of DNA and nucleosome arrangement to take place. After 48 hours, the cell membrane would have to be lysed in a manner that the nuclei are left intact. P1 or micrococcal nuclease would then be added to the nuclei and incubated at a suitable temperature. After nuclease treatment, the transfected plasmids would be extracted and analyzed by Southern blot and probed the appropriate probes to determine where the nucleases cleaved the plasmid. One of the problems facing this experiment would be how to ensure that the nucleases can pass through the nuclear pore and into the nucleus to act on the plasmids.

It would be also important to determine whether the inhibitory effect of the promoter can be pinpointed to a more specific sequence. From there, it can also be determined whether these sequences could contain binding sites for any other protein factors,

transcription or otherwise that could potentially play a role in chromatin remodeling. This can be done by breaking the promoter into smaller fragments and inserting them upstream of the *ori*P and studying the inhibitory effect any of them might have on replication. It would also be interesting to see if the replacement of the promoter with additional FR elements could also remodel chromatin and inhibit replication.

5.7 Conclusions

In this study, we have managed to show that replication is inhibited by the presence of a promoter upstream of the *ori*P. In addition, transcription through the *ori*P was also shown to be inhibited. Arrest of the elongation complex along *ori*P resulting in the physical obstruction of the replication machinery from assembling on the origin was thought to have occurred. This offered a novel explanation of the interplay between transcription and replication. However, it was shown through size exclusion chromatography that transcription was most likely terminated and the inhibition of transcription through *ori*P was most likely not related to the inhibition of replication.

Two other possible explanations of the interplay between transcription and replication in this study exist. The first explanation is that the passage of the elongation complex through the *ori*P is sufficient in preventing the assembly of the replication initiation complex. The second explanation is that transcription factor binding sites could have recruited transcription factors that either have the ability to remodel chromatin or are able to further recruit remodeling enzymes. The existence of transcription factor binding sites on both the promoter and the *ori*P also suggest the possibility that competition between these two elements for the correct chromatin structure could exist and that could have potentially contributed to the replication inhibition observed

in this system. Further work would need to be done in order to understand more about the true interplay between transcription and replication.

References

References

Abdurashidova G, Danailov MB, Ochem A, Triolo G, Djeliova V, Radulescu S, Vindigni A, Riva S, Falaschi A. 2003. Localization of proteins bound to a replication origin of human DNA along the cell cycle. EMBO J. 22(16): 4294-4303

Aiyar A, Tyree C, Sugden B. 1998. The plasmid replicon of EBV consists of multiple cis-acting elements that facilitate DNA synthesis by the cell and a viral maintenance element. EMBO J. 17: 6394-6403

Aladjem MI, Groudine M, Brody LL, Dieken ES, Fournier RE, Wahl GM, Epner EM. 1995. Participation of the human β-globin locus control region in initiation of DNA replication. Science. 270: 815-819

Aladjem MI, Rodewald LW, Kolman JL, Wahl GM. 1998. Genetic dissection of a mammalian replicator in the human β-globin locus. Science. 281: 1005-1009

Altman AL, Fanning E. 2001. The Chinese hamster dihydrofolate reductase replication origin beta is active at multiple ectopic chromosomal locations and requires specific DNA sequences for activity. Mol Cell Biol. 21: 1098-1110

Altman AL, Fanning E. 2004. Defined sequence modules and an architectural element cooperate to promote initiation at an ectopic mammalian chromosomal replication origin. Mol Cell Biol. 24: 4138-4150

Ambinder RF, Mullen MA, Chang YN, Hayward GS, Hayward SD. 1991. Functional domains of Epstein-Barr virus nuclear antigen EBNA-1. J Virol. 65: 1466-1478

Anachkova B, Hamlin JL. 1989. Replication in the amplified dihydrofolate reductase domain in CHO cells may initiate at two distinct sites, one of which is a repetitive sequence element. Mol Cell Biol. 9: 532-540

Aparicio OM, Stout AM, Bell SP. 1999. Differential assembly of Cdc45p and DNA polymerases at early and late origins of DNA replication. Proc Natl Acad Sci USA. 96: 9130-9135

Armache KJ, Kettenberger H, Cramer P. 2003. Architechure of initiation-competent 12-subunit RNA polymerase II. Proc Natl Acad Sci USA. 100: 6964-6968

Arndt KM, Chamberlin MJ. 1990. RNA chain elongation by Escherichia coli RNA polymerase. Factors affecting the stability of elongating ternary complexes. J Mol Biol. 213: 79-108

Austin RJ, Orr-Weaver TL, Bell SP. 1999. Drosophila ORC specifically binds to ACE3, an origin of DNA replication control element. Genes Dev. 13: 2639-2649

Baker TA, Bell SP. 1998. Polymerases and the replisome: machines within machines. Cell 92: 295-305

Bar-Nuham G, Nudler E. 2001. Isolation and characterization of σ^{70}-retaining

transcription elongation complexes from Escherichia coli. Cell. 106: 443-451

Beall EL, Manak JR, Zhou S, Bell M, Lipsick JS, Botchan MR. 2002. Role for a Drosophila Myb-containing protein complex in site-specific DNA replication. Nature. 420: 833-837

Bell SP. 2002. The origin recognition complex: from simple origins to complex functions. Genes Dev 16: 659-672

Bell SP, Stillman B. 1992. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. Nature. 357: 128-134.

Boucher N, McNicoll F, Laverdière M, Rochette A, Chou MN, Papadoloulou B. 2004. The ribosomal RNA gene promoter and adjacent cis-acting DNA sequences govern plasmid DNA partitioning and stable inheritance in the parasitic protozoan Leishmania. Nucleic Acids Res. 32(9): 2925-2936

Bramhill D, Kornberg A. 1988. Duplex opening by DnaA protein at novel sequences in initiation of replication at the origin of E.coli chromosome. Cell. 52: 743-755

Brewer, BJ. 1988. When polymerases collide: replication and transcriptional organization of the E.coli chromosome. Cell 53: 679-686

Brown SA, Imbalzano AN, Kingston RE. 1996. Activator-dependent regulation of transcriptional pausing on nucleosomal templates. Genes Dev. 10: 1479-1490

Burkitt D, Wright D. 1966. Geographical and tribal distribution of the African lymphoma in Uganda. Br Med I. 5487: 569-573

Burova E, Hung SC, Sagitov V, Stitt BL, Gottesman ME. 1995. Escherichia coli NusG protein stimulates transcription elongation rates in vivo and in vitro. J Bacteriol. 177: 1388-1392

Bushnell DA, Kornberg RD. 2003. Complete, 12-subunit RNA polymerase II at the 4.1-Å resolution: implications for the initiation of transcription. Proc Natl Acad Sci USA. 100: 6969-6973

Calvi BR, Spalding AC. 1999. Chorion gene amplification in Drosophila: a model for metazoan origins of DNA replication and S-phase control. Methods 18: 407-417

Chang VK, Donato JJ, Chan CS, Tye BK. 2004. MCM1 promotes replication initiation by binding specific elements at replication origins. Mol Cell Biol. 24 (14): 6514-6524

Chang Y, Cesarman E, Pessin MS, Lee F, Culpepper J, Knowles DM, Moore PS. 1994. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. Science 266(5192): 1865-1869

Chaudhuri B, Xu H, Todorov I, Dutta A, Yates J. 2001. Human DNA replication intiation factors, ORC and MCM associate with oriP of Epstein-Barr Virus. Proc Natl Acad Sci U.S.A 98: 10085-10089

Chen MR, Middeldorp JM, Hayward SD. 1993. Separation of the complex DNA binding domain of EBNA-1 into DNA recognition and dimerization subdomains of novel structure. J Virol. 67: 4875-4885

Chu Y, Huang T, Hsu MT. 1990. P1 nuclease defines a subpopulation of active SV40 chromatin-a new nuclease hypersensitivity assay. Nucleic Acids Res. 18: 3705-3711

Conaway RC, Conaway JW. 1997. General transcription factors for RNA polymerase II. Prog Nucleic Acid Res Mol Biol. 56: 327-346

Dhar SK, Yoshida K, Machida Y, Kaira P, Chaudhuri B, Wohlschlegel JA, Leffak M, Yates J, Dutta A. 2001. Replication from oriP of Epstein-Barr Virus requires human ORC and is inhibited by geminin. Cell 106: 287-296

Dijkwel PA, Wang S, Hamlin JL. 2002. Initiation sites are distributed at frequent intervals in the Chinese hamster dihydrofolate reductase origin of replication but are used at very different efficiencies. Mol Cell Biol. 22: 3053-3065

Dijkwel PA, Hamlin JL. 1995 The Chinese hamster dihydrofolate reductase origin consists of multiple potential nascent-strand start sites. Mol Cell Biol. 15: 3023-3031

Dombroski AJ, Walter WA, Record MT Jr, Siegele DA, Gross CA. 1992. Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA.

Dvir A, Conaway RC, Conaway JW. 1997. A role for TFIIH in controlling the activity of early RNA polymerase II elongation complexes. Proc Natl Acad Sci USA. 94: 9006-9010

Epstein M, Achong B, Barr Y. 1965. Morphological and biological studies on a virus in cultured lymphoblasts from Burkitt's lymphoma. J Exp Med. 121: 761-770

French S. 1992. Consequence of replication fork progression through transcription units in vivo. Science. 258: 1362-1365

Freund E, McGuire PM. 1986. Characterization of RNA polymerase type II from human term placenta. J Cell Physiol. 127: 432-438

Giacca M, Zentilin L, Norio P, Diviacco S, Dimitrova D, Contreas G, Biamonti G, Perini G, Weighardt F, Riva S, Falaschi A. 1994. Fine mapping of a replication origin of human DNA. Proc Natl Acad Sci. 91: 7119-7123

Gilbert D. 2002. Replication timing and transcriptional control: Beyond cause and effect. Curr Opin Cell Biol. 14: 377-383

Gilbert DM. 2001. Making sense of eukaryotic DNA replication origins. Science. 294: 96-100

Gill SC, Weitzel SE, von Hippel PH. 1991. Escherichia coli σ-70 and NusA proteins. I. Binding interactions with core RNA polymerase in solution and within the transcription

complex. J Mol Biol. 220: 307-324

Goldman M, Holmquist G, Gray M, Caston L, Nag A. 1984. Replication timing of genes and middle repetitive sequences. Science 224: 686-692

Gossen M, Pak DTS, Hansen SK, Achary JK, Botchan MR. 1995. A Drosophila homolog of the yeast origin recognition complex. Science. 270: 1674-1677

Greenblatt J, Li J. 1981. Interaction of the σ factor and the nusA gene protein of E.coli with RNA polymerase in the initiation-termination cycle of transcription. Cell. 24: 421-428

Gustafsson CM, Myers LC, Beve J, Spahr H, Lui M, Erdjument-Bromage H, Tempst P, Kornberg RD. 1998. Identification of new mediator subunits in the RNA polymerase II holoenzyme from Sacchaormyces cerevisiae. J Biol Chem. 273: 30851-30854

Haase SB, Heinzel SS, Calos MP. 1994. Transcription inhibits the replication of autonomously replicating plasmids in human cells. Mol Cell Biol. 14(4): 2516-2524

Harrison S, Fisenne K, Hearing J. 1994. Sequence requirements of the Epstein-Barr virus latent origin of replication. J Virol. 68: 1913-1925

Heinzel SS, Krysan PJ, Calos MP. 1991. Autonomous DNA replication in human cells is affected by the size and source of the DNA. Mol Cell Biol. 11(4): 2263-2272

Heintz NH, Hamlin JL. 1982. An amplified chromosomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific restriction fragments. Proc Natl Acad Sci USA. 79: 4083-4087

Holstege FC, van der Vliet PC, Timmers HT. 1996. Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcriptional factors IIE and IIH. EMBO J. 15: 1666-1677

Hu YF, Hao ZL, Li R. 1999. Chromatin remodeling and activation of chromosomal DNA replication by an acidic transcriptional activation domain from BRCA1. Genes Dev. 13: 637-642

Huberman JA, Riggs, AD. 1968. On the mechanism of DNA replication in mammalian chromosomes. J. Mol. Biol. 32: 327-341

Hung SC, Kang MS, Kieff E. 2001. Maintenance of Epstein-Barr virus (EBV) oriP-based episomes requires EBV-encoded nuclear antigen-1 chromosome-binding domains, which can be replaced by high mobility group-I or histone H1. Proc. Natl. Acad. Sci U.S.A. 98: 1865-1870

Izban MG, Luse DS. 1992. The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5' direction in the presence of elongation factor SII. Genes Dev. 6: 1342-1356

Jacob F, Brenner S, Cuzin F. 1963. On the regulation of DNA synthesis in bacteria: the

hypothesis of the replicon. Cold Spring Harbor Symp Quant Biol 28: 329-348

Kalejta RF, Li X, Mesner LD, Dijkwel PA, Lin HB, Hamlin JL. 1998. Distal sequences, but not oriβ/OBR-1, are essential for initiation of DNA replication in the Chinese hamster DHFR origin. Mol Cells 2: 797-806

Kang MS, Lu HX, Yasui T, Sharpe A, Warren H, Cahir-McFarland E, Bronson R, Hung SC, Kieff E. 2005. Epstein-Barr virus nuclear antigen 1 does not induce lymphoma in transgenic FVB mice. Proc Natl Acad Sci USA. 102: 820-825

Karimi L, Crawford H. 1995. Epstein-Barr virus: Mechanisms of oncogenesis. In: DNA tumor viruses: Oncogenic mechanisms. Barbanti-Brodano G (eds) 347-373 Plenum Press, NY

Kassavetis GA, Chamberlin MJ. 1981. Pausing and termination of transcription within the early region of bacteriophage T7 DNA in vitro. J Biol Chem 256: 2777-2786

Keleher CA, Passmore S, Johnson AD. 1989. Yeast repressor α2 binds to its operator cooperatively with yeast protein MCM1. Mol Cell Biol. 9: 5228-5230

Kelleker RJ 3rd, Flanagan PM, Kornberg RD. 1990. A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. Cell. 61: 1209-1215

Kerppola TK, Kane CM. 1990. Analysis of the signals for transcription termination by purified RNA polymerase II. Biochemisty. 29: 269-278

Kieff E, Rickinson AB, 2001. Epstein-Barr virus and its replication. In: Knipe DM, Howley P (eds.) Chapter 34 pp. 1249-1284 Fundamental Virology Lippincott, Williams and Wilkins, Philadelphia

Kim YJ, Bjorklund S, Li Y, Sayre MH, Kornberg RD. 1994. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell. 77: 599-608

Kim TK, Ebright RH, Reinberg D. 2000. Mechanism of ATP-dependent promoter melting by transcription factor IIH. Science. 288: 1418-1422

Kim JH, Lane WS, Reinberg D. 2002. Human Elongator facilitates RNA polymerase II transcription through chromatin. Proc Natl Acad Sci USA. 99: 1241-1246

Kireeva ML, Komissarova N, Waugh DS, Kashlev M. 2000. The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex. J Biol Chem. 275: 6530-6536

Klein G. 1994. Epstein-Barr virus strategy in normal and neoplastic B cells. Cell 77: 791-793

Klemm RD, Austin RJ, Bell SP. 1997. Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. Cell. 88: 493-502

Kohzaki H, Ito Y, Murakami Y. 1999. Context dependent modulation of the replication activity of Saccharomyces cerevisiae autonomously replicating sequences by transcription factors. Mol Cell Biol. 19: 7428-7435

Komissarova N, Kashlev M. 1997. Transcriptional arrest: Escherichia coli RNA polymerase translocates backwards, leaving the 3' end of the RNA intact and extruded. Proc Natl Acad Sci USA. 94: 1755-1760s

Kornberg A, Baker TA. 1992. DNA replication. New York: Freeman. 526

Labib K, Diffley JF. 2001. Is the Mcm2-7 complex the eukaryotic DNA replication fork helicase? Curr Opin Genet Dev 11:64-70

Lascaris RF, Groot E, Hoen PB, Mager WH, Planta RJ. 2000. Different roles for abf1p and a T-rich promoter element in nucleosome organization of the yeast RPS28A gene. Nucleic Acids Res. 28: 1390-1396

Laskey R, Madine M. 1996. Roles of nuclear structure in DNA replication In: DePamphilis ML (eds) 119-130 DNA replication in eukaryotes. Cold Spring Harbor Laboratory Press, Plainview, NY

Lee D, Bell SP. 1997. Architecture of the yeast origin recognition complex bound to origins of DNA replication. Mol Cell Biol. 17: 7159-7168

Lee TI, Young RA. 1998. Regulation of gene expression by TBP-associated proteins. Genes Dev. 12(10): 1398-1408

Lee W, Haslinger A, Karin M, Tjian R.1987. Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. Nature. 325: 368-372

Leight ER, Sugden B. 2001. The cis-acting family of repeats can inhibit as well as stimulate establishment of an oriP replicon. J Virol. 75: 10709-10720

Leu TH, Hamlin JL. 1989. High-resolution mapping of replication fork movement through the amplified dihydrofolate reductase domain in CHO cells by in-gel renaturation analysis. Mol Cell Biol 9: 523-531

Levitskaya J, Coram M, Levitsky V, Imreh S, Steigerwald-Mullen PM, Klein G, Kurilla, MG, Masucci MG. 1995. Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen-1. Nature 375: 685 – 688

Li JJ, Herskowitz I. 1993. Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. Science. 262: 1870-1874

Lindahl T, Adams A, Bjursell G, Bornkamm GW, Kaschka-Dierich C, Jehn U. 1976. Covalently closed circular duplex DNA of Epstein-Barr virus in a human lymphoid cell line. J Mol Biol. 102(3): 511-530

Liu LF, Wang JC. 1987. Supercoiling of the DNA template during transcription. Proc

Natl Acad Sci USA. 84: 7024-7027

Lu L, Zhang H, Tower J. 2001. Functionally distinct, sequence specific replicator and origin elements are required for Drosophila chorion gene amplification. Genes Dev. 15: 134-146

Lutfalla G, Armbruster L, Dequin S, Bertolotti R. 1989. Construction of an EBNA-producing line of well-differentiated human hepatoma cells and of appropriate Epstein-Barr virus-based shuttle vectors. Gene. 76: 27-39

MacAlpine D, Rodríguez, Bell SP. 2004. Coordination of replication and transcription along a Drosophila chromosome. Genes Dev 18: 3094-3105

Mackey D, Middleton T, Sugden B. 1995. Multiple regions within EBNA-1 can link DNAs. J Virol. 69(10): 6199-6208

Marahrens Y, Stillman B. 1992. A yeast chromosomal origin of DNA replication defined by multiple functional elements. Science. 255: 817-823

Meier JL, Keller MJ, McCoy JJ. 2002. Requirement of multiple cis-acting elements in the human cytomegalovirus major immediate-early distal enhancer for viral gene expression and replication. J Virol. 76: 313-326

Milbrandt JD, Heintz NH, White WC, Rothman SM, Hamlin JL. 1981. Methotrexate-resistant Chinese hamster ovary cells have amplified a 135-kilobase-pair region that includes the dihydrofolate reductase gene. Proc Natl Acad Sci USA. 78: 6043-6047

Mucke S, Polack A, Pawlita M, Zehnpfennig D, Massoudi N, Bohlen H, Doerfler W, Bornkamm G, Diehl V, Wolf J. 1997. Suitability of Epstein-Barr virus-based episomal vectors for expression of cytokine genes in human lymphoma cells. Gene Ther. 4(2): 82-92

Newlon CS, Theis JF. 1993. The structure and function of yeast ARS elements. Curr Opin Genet Dev. 3: 753-758

Nieduszynski CA, Blow JJ, Donaldson AD. 2005. The requirement of yeast replication origins for pre-replication complex proteins is modulated by transcription. Nucleic Acids Res. 33: 2410-2420

Nonoyama M, Pagano JS. 1972. Separation of Epstein-Barr virus DNA from large chromosomal DNA in non-virus producing cells. Nat New Biol. 9:238(84):169-171

Olavarrieta L, Hernández P, Krimer DB, Schvartzman JB. 2002. DNA knotting caused by head-on collision of transcription and replication. J Mol Biol. 322: 1-6

Orphanides G, Wu WH, Lane WS, Hampsey M, Reinberg D. 1999. The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. Nature. 400: 284-288

Otero G, Fellows J, Li Y, de Bizemont T, Dirac AM, Gustafsson CM, Erdjument-

Bromage H, Tempst P, Svejstrup JQ. 1999. Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation. Mol Cell. 3:109-118

Pal M, Luse DS. 2003. The initiation-elongation transition: Lateral mobility of RNA in RNA polymerase II complexes is greatly reduced at +8/+9 and absent by +23. Proc Natl Acad Sci USA. 100: 5700-5705

Pan WJ, Gallagher RC, Blackburn EH. 1995. Replication of an rRNA gene origin plasmid in the Tetrahymena thermophila macronucleus is prevented by transcription through the origin from an RNA polymerase I promoter. Mol Cell Biol. 15: 3372-3381

Peter BJ, Ullsperger JC, Hiasa H, Marians KJ, Cozzarelli NR. 1998. The structure of supercoiled intermediates in DNA replication. Cell. 94: 819-827

Rao H, Stillman B. 1995. The origin recognition complex interacts with a bipartite DNA binding site within yeast replicators. Proc Natl Acad Sci 92: 2224-2228

Rawlins DR, Milman G, Hayward SD, Hayward GS. 1985. Sequence-specific DNA binding of the Epstein-Barr virus antigen (EBNA-1) to clustered sites in the plasmid maintenance region. Cell. 42(3): 859-868

Reid JL, Iyer VR, Brown PO, Struhl K. 2000. Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. Mol Cell. 6: 1297-1307

Reinberg D, Roeder RG. 1987. Factors involved in specific transcription by mammalian RNA polymerase II. Transcription factor IIS stimulates elongation of RNA chains. J Biol Chem. 262: 3331-3337

Reines D. 1992. Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II. J Biol Chem. 267: 3785-3800

Reines D, Chamberlin MJ, Kane CM. 1989. Transcription elongation factor SII (TFIIS) enables RNA polymerase II to elongate through a block to transcription in a human gene in vitro. J Biol Chem. 264: 10799-10809

Reines D, Ghanouni P, Li QQ, Mote J Jr. 1992. The RNA polymerase II elongation complex. Factor-dependent transcription elongation involves nascent RNA cleavage. J Biol Chem. 267: 15516-15522

Reisman D, Sugden B. 1986. Trans activation of an Epstein-Barr viral transcriptional enhancer by the Epstein-Barr virus nuclear antigen 1. Mol Cell Biol. 6(11): 3838-3846

Reisman D, Yates JL, Sugden B. 1985. A putative origin of replication of plasmids derived from Epstein-Barr virus is composed of two cis-acting components. Mol Cell Biol. 5(8): 1822-1832

Remus D, Beall EL, Botchan MR. 2004. DNA topology, not DNA sequence is a critical determinant for Drosophila ORC-DNA binding. EMBO J. 23(4): 897-907

Ring BZ, Yarnell WS, Roberts JW. 1996. Function of E. coli RNA Polymerase σ Factor-σ^{70} in Promoter-Proximal Pausing. Cell. 86: 485-493

Ritzi M, Tillack K, Gerhardt J, Ott E, Humme S, Kremmer E, Hammerschmidt W, Schepers A. 2003. Complex protein-DNA dynamics at the latent origin of DNA replication of Epstein-Barr virus. J Cell Sci. 116: 3971-3984

Robertson ES, Ooka T, Kieff ED. 1996. Epstein-Barr virus vectors for gene delivery to B lymphocytes. Proc Natl Acad Sci U S A. 93(21): 11334-40

.

Rowley A, Tada S, Cocker JH, Harwood J, Diffley JF. 1995. Initiation complex assemblt at budding yeast replication origins begins with the recognition of a bipartite sequence by limiting amounts of the initiator, ORC. EMBO J. 14: 2631-3641

Schepers A, Ritzi M, Bousset K, Kremmer E, Yates J, Harwood J, Diffley JFX, Hammerschmidt W. 2001. Human origin recognition complex binds to the region of the latent origin of DNA replication of Epstein-Barr Virus. EMBO J. 20: 4588-4602

Schmidt MC, Chamberlin MJ. 1984. Amplification and isolation of Escherichia coli nusA protein and studies of its effects on in-vitro RNA chain elongation. Biochemistry. 23: 97-203

Shimamoto N, Kamigochi T, Utiyama H. 1986. Release of the σ subunit of Escherichia coli DNA-dependent RNA polymerase depends mainly on time elapsed after the start of initiation, not on length of product RNA. J Biol Chem. 261: 11859-11865

Shinomiya T, Ina S. 1991. Analysis of chromosomal replicons in early embryos of Drosophila melanogaster by two-dimensional gel electrophoresis. Nucleic Acids Res. 19 (14): 3935-3941

Schübeler D, Scalzo D, Kooperberg C, Steensel B, Delrow J, Groudine M. 2002. Genome-wide DNA replication profile for Drosophilia melanogaster: a link between transcription and replication timing. Nature Gen. 32: 438-442

Sentenac A. 1985. Eukaryotic RNA polymerases. CRC Crit Rev Biochem. 18: 31-90
Sims RJ, 3rd, Belotserkovskaya R, Reinberg D. 2006. Elongation by RNA polymerase II: the short and long of it. Genes & Dev 18: 2437-2468

Snudden DK, Hearing J, Smith PR, Grasser FA, Griffin BE. 1994. EBNA-1, the major nuclear antigen of Epstein-Barr virus, resembles 'RGG' RNA binding proteins. EMBO J. 1994 13(20): 4840-7

Stambrook P, Flickinger R. 1970. Changes in chromosomal DNA replication patterns in developing frog embryos. J Exp Zool. 174: 101-113

Stinchcomb DT, Struhl K, Davis RW. 1979. Isolation and characterisation of of a yeast chromosomal replicator. Nature. 282: 39-43

Taljanidisz J, Popowski J, Sarkar N. 1989. Temporal order of gene replication in Chinese hamster ovary cells. Mol Cell Biol. 9: 2881-2889

Tanaka S, Diffley JF. 2002. Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. Nat Cell Biol. 4: 198-207

Tanaka S, Halter D, Livingstone-Zatchej M, Reszel B, Thoma F. 1994. Transcription through the yeast origin of replication ARS1 ends at the ABFI binding site and affects the extrachromosomal maintenance of minichromosomes. Nucleic Acids Res. 22(19): 3904-3910

Tanaka T, Knapp D, Nasmyth K. 1997. Loading of an Mcm protein onto DNA replication origins is regulated by Cdc6p and CDKs. Cell.90: 649-660

Telford DJ, Stewart BW. 1989. Micrococcal nuclease: its specificity and use for chromatin analysis. Int J Biochem. 21: 127-137

Tomiyasu K, Satoh E, Oda Y, Nishizaki K, Kondo M, Imanishi J, Mazda O. 1998. Gene transfer in vitro and in vivo with Epstein-Barr virus-based episomal vector results in markedly high transient expression in rodent cells. Biochem Biophys Res Commun. 253: 733-738

Tsukamoto H, Wells D, Brown S, Serpente P, Strong P, Drew J, Inui K, Okada S, Dickson G. 1999. Enhanced expression of recombinant dystrophin following intramuscular injection of Epstein-Barr virus (EBV)-based mini-chromosome vectors in mdx mice. Gene Ther. 6(7): 1331-5

Wahle E, Lasken RS, Kornberg A. 1989. The dnaB-dnaC replication protein complex of Escherichia coli. II. Role of the complex in mobilizing dnaB functions. J Biol Chem. 264(5): 2469-2475.

White EJ, Emanuelsson O, Scalzo D, Royce T, Kosak S, Oakeley EJ, Weissman S, Gerstein M, Goudine M, Synder M, Schübeler D. DNA replication-timing analysis of human chromosome 22 at high resolution and different development states. Proc. Natl. Acad. Sci U.S.A. 101(51): 17771-17776

Wiest DK, Wang D, Hawley DK. 1992. Mechanistic studies of transcription arrest at the adenovirus major late attenuation site. Comparison of purified RNA polymerase II and washed elongation complexes. J Biol Chem. 267: 7733-7744

Wilson JB, Bell JL, Levine AJ. 1996. Expression of Epstein-Barr virus nuclear antigen-1 induces cell neoplasia in transgenic mice. EMBOJ 15 (12): 3117-3126

Winkler GS, Petrakis TG, Ethelberg S, Tokunaga M, Erdjument-Bromage H, Tempst P, Svejstrup JQ. 2001. RNA polymerase II elongator holoenzyme is composed of two discrete subcomplexes. J Biol Chem. 276: 32743-32749

Woodfine K, Fiegler H, Beare DM, Collins JE, McCann OT, Young BD, Debernadi S, Mott R, Dunham I, Carter NP. 2004. Replication timing of the human genome. Hum Mol Gen 13(2): 191-202

Woychik NA, McKune K, Lane WS, Young RA. 1993. Yeast RNA polymerase II subunit

RPB11 is related to a subunit shared by RNA polymerase I and III. Gene Expr 3(1): 77-82

Wysokenski DA, Yates JL. 1989. Multiple EBNA-1 binding sites are required to form an EBNA-1 dependent enhancer and to activate a minimal replicative origin within oriP of Epstein-Barr virus. J Virol. 63(6): 2657-2666

Yan Q, Moreland RJ, Conaway JW, Conaway RC. 1999. Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. J Biol Chem. 274: 35668-35675

Yates JL, Camiolo SM, Bashaw JM. 2000. The minimal replicator of Epstein-Barr virus oriP. J Virol. 74: 4512-4522

Yates JL, Guan N. 1991. Epstein-Barr virus-derived plasmids replicate only once per cell cycle and are not amplified after entry into cells. J Virol 1: 483-488

Yates JL, Warren N, Reisman D, Sugden B. 1984. A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc Natl Acad Sci USA. 81(12): 3806-3810

Yates JL, Warren N, Sugden B. 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature. 13: 812-815

Zhang H, Tower J. 2004. Sequence requirements for function of the Drosophila chorion gene locus ACE3 replicator and ori-β origin elements. Development. 131: 2089-2099

Zhang G, Campbell EA, Minakhin L, Ricter C, Severinov K, Darst SA. 1999. Crystal structure of Thermus aquaticus RNA polymerase at 3.3Å resolution. Cell. 98: 811-824

Zou L, Stillman B. 2000. Assembly of a complex containing Cdc45p, replication protein A and Mcm2p at replication origins controlled by S-phase cyclin dependent kinases and Cdc7p-Dbf4p kinase. Mol Cell Biol. 20 (9): 3086-3096

Appendices

Appendix

Restriction enzyme reaction set-up

Reagent	Final concentration
DNA	Specified in text
10X NEBuffer 4	1X
Aat II	1 u/µl
H_2O	q.s as specified in text

***Aat* II single enzyme digestion reaction mix**

Reagent	Final concentration
DNA	Specified in text
10 X NEBuffer 4	1 X
BSA (10 mg/ml)	100 µg/µl
Aat II	1 u/µl
Eco RI	0.5 u/µl
H_2O	q.s as specified in text

***Aat* II and *Eco* RI double enzyme digestion reaction mix**

Reagent	Final concentration
DNA	Specified in text
10 X NEBuffer 4	1 X
Aat II	1 u/µl
Sac II	1 u/µl
H_2O	q.s as specified in text

***Aat* II and *Sac* II double enzyme digestion reaction mix**

Reagent	Final concentration
DNA	Specified in text
10 X NEBuffer 4	1 X
BSA (10 mg/ml)	100 µg/µl
Aat II	1 u/µl
Kpn I	0.5 u/µl
H_2O	q.s as specified in text

***Aat* II and *Kpn* I double digestion reaction mix**

Reagent	Final concentration
DNA	Specified in text
10 X NEBuffer 3	1 X
BSA (10 mg/ml)	100 µg/µl
Apa I	0.5 u/µl
Cla I	0.5 u/µl
H_2O	q.s as specified in text

***Cla* I and *Apa* I double enzyme digestion reaction mix**

Reagent	Final concentration
DNA	Specified in text
10 X NEBuffer 3	1X
BSA (10 mg/ml)	100 µg/µl
Apa I	0.5 u/µl
Eco RI	0.5 u/µl
H_2O	q.s as specified in text

***Eco* RI and *Apa* I double enzyme digestion reaction mix**

Reagent	Final concentration
DNA	Specified in text
10 X NEBuffer 4	1 X
Dpn I	2 u/µl
H_2O	as specified in text

***Dpn* I single enzyme digestion reaction mix**

Reagent	Final concentration
DNA	Specified in text
10 X *Eco* RI buffer	1 X
Eco RI	0.5 u/µl
H_2O	q.s as described in text

***Eco* RI single enzyme digestion reaction mix**

Reagent	Final concentration
DNA	Specified in text
10X NEBuffer 3	1X

BSA (10mg/ml)	100µg/µl
Eco RV	1 u/µl
Sma I	1 u/µl
H_2O	q.s as specified in text

***Eco* RV single enzyme digestion reaction mix**

Reagent	**Final concentration**
DNA	Specified in text
10X NEBuffer 4	1X
Nde I	1u/µl
H_2O	q.s as specified in text

***Nde* I single enzyme digestion reaction mix**

Reagent	**Final concentration**
DNA	Specified in text
10X NEBuffer 4	1X
BSA (10mg/ml)	100µg/µl
Sac II	1u/µl
Eco RI	0.5u/µl
H_2O	q.s as specified in text

***Sac* II and *Eco* RI double enzyme digestion reaction mix**

Reagent	**Final concentration**
DNA	Specified in text
10X NEBuffer 4	1X
BSA (10 mg/ml)	100 µg/ml
Sac II	1 u/µl
Xba I	1 u/µl
H_2O	as specified in text

***Sac* II and *Xba* I double enzyme digestion reaction mix**

Reagent	Final concentration
DNA	Specified in text
10 X NEBuffer 3	1 X
BSA (10 mg/ml)	100 µg/ml
Sal I	1 u/µl

H_2O	Specified in text

***Sal* I single enzyme digestion reaction mix**

Reagent	**Final concentration**
DNA	Specified in text
10X NEBuffer 4	1X
Sma I	2 u/µl
H_2O	q.s as specified in text

***Sma* I single enzyme digestion reaction mix**

Reagent	**Final concentration**
DNA	Specified in text
10X NEBuffer 4	1X
Stu I	1u/µl
H_2O	q.s as specified in text

***Stu* I single enzyme digestion reaction mix**

Reagent	**Final concentration**
DNA	Specified in text
10X NEBuffer 1	1X
Stu I	0.5 u/µl
Bsp HI	0.5 u/µl
H_2O	q.s as specified in text

***Stu* I and *Bsp* HI double enzyme digestion reaction mix**

Reagent	**Final concentration**
DNA	Specified in text
10X NEBuffer 3	1X
Stu I	0.5 u/µl
Sal I	1 u/µl
H_2O	as specified in text

***Stu* I and *Sal* I double enzyme digestion reaction mix**

Reagent	**Final concentration**
DNA	Specified in text

10X NEBuffer 2	1X
Stu I	0.5u/µl
Ssp I	0.5u/µl
H_2O	q.s as specified in text

***Stu* I and *Ssp* I double enzyme digestion**

***In vitro* transcription reagents**

Elution Buffer
15 mM K-HEPES (pH7.9)
50 mM KCl
1 mM EDTA
1 mM DTT
0.2 % Triton X-100

Stop Buffer
0.3 M Tris-HCl, pH 7.4
0.3 M NaAc
0.5 % (w/v) SDS
2 mM EDTA
3 µg/ml tRNA

Southern and Northern blot Reagents

Alkaline Buffer A
50 mM NaOH
100 mM NaCl

Buffer A
100 mM Tris-HCl
300 mM NaCl

Denaturation Buffer
1.5 M NaCl
0.5 M NaOH

Hybridization Buffer
5 X SSC
20 fold dilution of liquid block provided
0.1 % (w/v) SDS
5 % (w/v) Dextran Sulphate
5 X MOPS Buffer
0.1 M MOPS
40 mM NaAc
5 mM EDTA

Neutralization Buffer
1.5 M NaCl

0.5 M Tris-HCl, pH 7.5 (NUMI)

Neutralization Buffer A
100 mM Tris-HCl

Running buffer; Northern blot
To make 100 ml of Running Buffer, add the following
20 ml 5 X MOPS Buffer
12.5 ml 37 % Formaldehyde
67.5 ml RNase free water

Sample loading dye; Northern blot(10 µl volume)
5 µl Ambion Sample Buffer (Ambion)
2 µl 5 X MOPS buffer
1.6 µl 37 % formaldehyde
0.4 % Glycerol
1 µl RNA sample

20 X SSC
0.3 M Na_3citrate
3 M NaCl
RNase free water to be used for Northern blot

NEBuffer 1
10mM Bis Tris Propane HCl
10mM $MgCl_2$
1mM dithiothreitol
pH 7.0

NEBuffer 2
50mM NaCl
10mM Tris-HCl
10mM $MgCl_2$
1mM dithiothreitol
pH 7.9

NEBuffer 3
100mM NaCl
50mM Tris-HCl
10mM $MgCl_2$
1mM dithiothreitol
pH 7.9

NEBuffer 4
50mM potassium acetate
20mM Tris-acetate
10mM magnesium acetate
1mM dithiothreitol
pH 7.9

NEBuffer *Eco*R I
50mM NaCl
100mM Tris-HCl
10mM $MgCl_2$
0.025% Triton X-100
pH 7.5

T4 DNA polymerase buffer
50mM NaCl
10mM Tris-HCl
10mM $MgCl_2$
1mM dithiothreitol
pH 7.9

T4 DNA ligase buffer
50mM Tris-HCl
10mM $MgCl_2$
10mM dithiothreitol
25μg/μl bovine serum albumin
pH 7.5

RPMI media
To obtain 10 litres of R10 media, add the following:
1 X 10 litre RPMI powder dissolved in 5 litres of nanopure water
35.7g Hepes
3g L-glutamine
1.1g pyruvic acid sodium salt
10g glucose
20g $NaHCO_3$
Adjust to pH 7.2 by adding 1N NaOH or 1N HCl
Add another 5 litres of nanopure water and sterilize by membrane filtration

R10 media
Add the following to 500ml of RPMI to obtain R10 media,
5ml penicillin and streptomycin (10,000u/ml each) from Gibco
50ml FCS

LB agar
For 1 litre of LB agar, add the following to 1 litre of RO water:
10g tryptone
5g yeast extract
5g NaCl
10g agar base
Autoclave at 121°C for 15 minutes, cool to 50°C before pouring into steril petri dish
Store agar plate at 4°C

LB + ampicillin agar
For 1 litre of LB + ampicillin agar, add the following to 1 litre of RO water:

10g tryptone
5g yeast extract
5g NaCl
10g agar base
Autoclave at 121°C for 15 minutes, allow to cool to 50°C before adding ampicillin
Pour into sterile petri dish
Store agar plates at 4°C

LB broth
For 1 litre of LB broth, add the following to 1 litre of RO water:
10g tryptone
5g yeast extract
5g NaCl
Autoclave at 121°C for 15 minutes cool to room temperature
Store at room temperature

LB + ampicillin broth
For 1 litre of LB + ampicillin broth, add the following to 1 litre of water:
10g tryptone
5g yeast extract
5g NaCl
Autoclave at 121°C for 15 minutes, allow to cool to 50°C before adding ampicillin
Store at 4°C

SOC medium
For 1 litre of SOC medium, add the following to 1 litre of water:
20g tryptone
5g yeast extract
0.5g NaCl
2.5mM KCl
Adjust to pH 7.0 with 5M NaOH
10mM $MgCl_2$
Autoclave at 121°C for 15 minutes
20mM filter-sterilzed glucose solution
Store at room temperature

TE buffer
10 mM Tris-Cl, pH8.0
1 mM EDTA

1.0 % agarose gel
To 100 ml of TBE (NUMI), add the following,
1.0 g agarose
Microwave to dissolve agarose
Allow to cool for 5 minutes before adding 5µl of ethidium bromide

0.8 % agarose gel
To 100 ml of TBE (NUMI), add the following:
0.8 g agarose
Microwave to dissolve agarose

Allow to cool for 5 minutes before adding 5µl of ethidium bromide

1.0 % Formaldehyde gel
To make 100 ml of gel, add the following
20 ml of 5 X MOPS buffer
67.5 ml of RNase free water
1.0 g of agarose
Microwave to dissolve agarose
Allow to cool for 5 minutes before adding the following
12.5 ml of 37 % formaldehyde