

2009

Comprehensive Analysis of Residues in Conserved Region 3 of the HPV16 E7 protein Involved in Dimerization of E7 and Interaction with pCAF

Katherine Hung

Follow this and additional works at: <https://ir.lib.uwo.ca/digitizedtheses>

Recommended Citation

Hung, Katherine, "Comprehensive Analysis of Residues in Conserved Region 3 of the HPV16 E7 protein Involved in Dimerization of E7 and Interaction with pCAF" (2009). *Digitized Theses*. 3847.
<https://ir.lib.uwo.ca/digitizedtheses/3847>

This Thesis is brought to you for free and open access by the Digitized Special Collections at Scholarship@Western. It has been accepted for inclusion in Digitized Theses by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Comprehensive Analysis of Residues in Conserved Region 3 of the HPV16 E7 protein Involved in Dimerization of E7 and Interaction with pCAF

(Spine title: Analysis of the HPV16 E7-CR3 Dimer)

(Thesis Format: Monograph)

by

Katherine Hung

Graduate Program in Microbiology and Immunology

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Katherine Hung, 2009

ABSTRACT

In this study, I show that the HPV16 E7 protein is able to form homodimers *in vivo* and that the formation of this tertiary structure is primarily dependent on the structural integrity of the CR3 monomer. A collection of E7 mutants targeting residues of the N-terminus, hydrophobic core of the dimer and the solvent exposed residues of CR3 were evaluated for their ability to dimerize in the yeast two-hybrid assay. This new set of mutants was also used to map the interaction between E7 and the pCAF acetyltransferase to the C-terminal region of E7. Binding to pCAF appears to require dimerization of E7 but only one monomer of E7 is involved at any given time. In conclusion, the characterization of these mutants for dimerization and for the interaction with cellular proteins will ultimately pave the way for functional studies to determine the role of dimerization in the HPV life cycle.

Key words: Human papillomavirus, HPV16, E7, CR3, dimer, dimerization, model, site-directed mutagenesis, pCAF acetyltransferase, yeast two-hybrid

ACKNOWLEDGMENTS

First and foremost, I would like to thank my supervisor Dr. Joe Mymryk for his support, help and advice throughout my master's career. His knowledge within and outside his field of research is exceptional and his passion for science and teaching is an inspiration to all young scientists. He has been a remarkable mentor – understanding a student's struggles and encouraging perseverance through it all.

To members of my advisory committee, Dr. Laura Hertel and Dr. Joe Torchia, thank you for all your helpful advice and guidance during my time at Western.

Thanks to all the lab members both past and present. Greg and Ahmed, thanks for teaching me everything I need to know about yeast and keeping the lab a little off beat. Jai thanks for your patience and help with new lab techniques; and especially for sharing your knowledge of what London restaurants have to offer. Peter, thanks for helping me get past my initial rut of failed experiments and for your advice on everything. To Alison, thanks for being a great lab bench buddy and friend. Biljana, the other HPV graduate student, thanks for keeping me alive in an adenovirus lab.

Thanks to my family, for your love and support throughout my education.

To my best friends: Beatrisa, Danielle, Diana, Kim, Preston, Theresa, thank you for all the love, support and happiness you bring into my life and giving me the strength to return to the lab bench everyday with a positive spirit.

TABLE OF CONTENTS

Certificate of Examination / ii	
Abstract/ iii	
Acknowledgements / iv	
Table of Contents / v	
List of Tables / vii	
List of Figures / viii	
List of Abbreviations / x	

CHAPTER 1: INTRODUCTION.....	1
1.1 General Introduction	1
1.2 Introduction to papillomaviruses	1
1.2.1 <i>Classification of papillomaviruses</i>	3
1.2.2 <i>Physical properties of papillomaviruses</i>	4
1.2.3 <i>Overview of the papillomavirus life cycle</i>	5
1.3 Characteristics of the papillomavirus E7 protein	9
1.4 Biological activities and molecular targets of E7	11
1.4.1 <i>Association of E7 with pRb</i>	12
1.4.2 <i>E7 proteins and the p53 tumour suppressor</i>	15
1.4.3 <i>Effects of E7 on other components of the cell cycle machinery</i>	16
1.4.4 <i>Transcriptional regulation by E7</i>	17
1.5 Structural organization of the E7 proteins	20
1.6 Yeast as a model system for studying HPV16 E7	25
1.7 Thesis objectives and hypothesis	26
CHAPTER 2: MATERIALS AND METHODS	27
2.1 <i>Escherichia coli</i> strain, <i>Saccharomyces cerevisiae</i> strains and media.....	27
2.2 Plasmid construction	27
2.3 DNA preparation.....	39
2.3.1 <i>Small scale plasmid DNA preparation</i>	39
2.3.2 <i>Screening potential recombinant plasmids with restriction digestion analysis</i>	39
2.3.3 <i>Large scale plasmid DNA preparation</i>	40
2.4 Yeast transformation	41

2.5	Liquid β -galactosidase assay in yeast	41
2.6	Yeast cell protein extraction and Western blot analysis	42
2.7	Solutions.....	44
CHAPTER 3: RESULTS		46
3.1	Development of a model for HPV16 E7 CR3.....	46
3.2	Application of the CR3 model to identify targets for site-directed mutagenesis 47	
3.2.1	<i>Construction of CR3 mutants</i>	50
3.3	Characterizing the collection of E7 mutations	53
3.3.1	<i>Dimerization properties of all HPV16 E7 mutants</i>	59
3.4	Precise definition of the pCAF binding region within E7	80
CHAPTER 4: DISCUSSION.....		87
4.1	Thesis summary and discussion.....	87
4.1.1	<i>Dimerization of E7 proteins</i>	87
4.1.2	<i>Interaction between pCAF and E7 CR3</i>	95
4.2	Future directions	97
REFERENCES		101
CURRICULUM VITAE.....		112

LIST OF TABLES

Table	Description	
1.1	Overview of HPV Gene Products	6
2.1	List of yeast strains used in this study and the sources from which they were obtained	27
2.2	List of plasmids used in this study and their sources	28
2.3	List of E7 mutants produced in this study and oligonucleotides utilized in PCR	34
3.1	Substitution Scheme for CR3 Mutagenesis	52
3.2	Collection of E7 Mutants	58
3.3	Experimental Design for Testing Dimerization in Yeast 2-Hybrid Assay	60
3.4	Summary of β -galactosidase Activity Relative to Wild-type E7 dimerization	78
3.5	Summary of Dimerization Data for Entire Collection of E7 Mutants	79

LIST OF FIGURES

Figure	Description	
1.1	Schematic representation of the genomic organization of HPV16	6
1.2	The HPV life cycle.....	8
1.3	Location of conserved regions in HPV16 E7 and sequence alignment between conserved regions (CR) 1 and 2 of adenovirus type 5 E1A and regions of HPV16 E7	10
1.4	Relaxation of the G ₁ /S checkpoint by HPV E7 proteins.....	14
1.5	Representative 3-dimensional structures of HPV1 and HPV45 CR3 dimers ..	23
1.6	Multiple sequence alignment of CR3 regions from representative Supergroup A HPVs	24
2.1	Site-directed mutagenesis by the megaprimer method	32
2.2	Site-directed mutagenesis by overlap extension	33
2.3	Maps for yeast expression vectors pBAIT and pJG4-5+	38
3.1	HPV45 and HPV16 E7 CR3 dimers	49
3.2	Molecular surface of the HPV16 E7 CR3 dimer	51
3.3	Predicted molecular surface of the HPV16 E7 CR3 dimer.....	55
3.4	Predicted molecular surface of the HPV16 E7 CR3 monomer.....	57
3.5	Schematic of the yeast two-hybrid system for dimerization assays.....	60
3.6	Controls for the yeast two-hybrid assay.....	62
3.7	Dimer interactions by wild-type and mutant HPV16 E7 proteins	65
3.8	Dimer interactions by HPV16 E7 mutants.....	70
3.9	Equations used to determine HPV16 E7 mutant dimerization relative to wild-type E7	71
3.10	Dimerization ability of selected HPV16 E7 mutants relative to wild-type E7	73

Figure	Description
3.11	Western blots of expression levels of LexA and B42 fusion proteins for dimer defective E7 mutants 77
3.12	Interaction between pCAF and the HPV16 E7 protein..... 83
3.13	Proposed pCAF binding site on HPV16 E7-CR3 86

LIST OF ABBREVIATIONS

Ad	Adenovirus
AD	Activation domain
ADH	Alcohol dehydrogenase promoter
AMP	Ampicillin
bp	Base pairs
BME	β -mercaptoethanol
BPV	Bovine papillomavirus
cdk	Cyclin-dependent kinase
CKI	Cyclin-dependent kinase inhibitor
CR	Conserved region
CKII	Casein kinase II
DBD	DNA binding domain
DHFR	Dihydrofolate reductase
DMEM	Dulbecco's modified Eagle's medium
DNA pol α	DNA polymerase α
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HPV	Human papillomavirus
JMB	Joe Mymryk Bacteria
JMO	Joe Mymryk Oligo
JMY	Joe Mymryk Yeast
kbp	kilobase pair
kDa	kilodalton
LB	Luria Bertani
leu	Leucine
LCR	Long control region
LiAc	Lithium acetate
NMR	Nuclear Magnetic Resonance
NuRD	Nucleosome remodelling and histone deacetylation

ONPG	<i>ortho</i> -Nitrophenyl- β -galactoside
ORF	Open reading frame
PBS	Phosphate buffered saline
PDB	Protein Data Bank
pCAF	p300/CBP-associated factor
PEG	Polyethylene glycol
PVDF	Polyvinylidene fluoride
SC	Synthetic complete
SV	Simian vacuolating virus
TBP	TATA binding protein
TK	Thymidine kinase
Trp	Tryptophan
URR	Upstream Regulatory Region
VLP	Virus-like particle
WT	Wild-type
YEP-D	Yeast extract peptone – dextrose

CHAPTER 1: INTRODUCTION

1.1 General Introduction

The extended family of DNA viruses are as fascinating as they are diverse. The common feature of these viruses is a DNA genome, while specific characteristics, such as size, host species and tissue tropism, vary widely. Viruses are obligate intracellular parasites and often disease causing agents, with a small subset, called tumour viruses, capable of inducing malignant growth in their natural hosts. Many viruses present a significant health, social and economic burden to the human population. Therefore, viruses have become major drivers of research in molecular and cell biology with a two-fold objective. First, viral studies allow us to better understand the mechanisms behind viral infection and production with the goal of preventing infection and developing antiviral therapies. Secondly, viruses redirect host cellular processes for viral production and antagonize host antiviral defence systems. Thus, viruses serve as powerful model systems and tools to discover and elucidate regulatory pathways and cellular processes. Undoubtedly, molecular studies using tumour viruses will also aid in understanding the mechanisms that drive tumour initiation and progression.

1.2 Introduction to papillomaviruses

Papillomaviruses comprise a group of small DNA tumour viruses that have a marked tropism for epithelial cells. They are highly species specific and infection typically induces the formation of benign lesions of the skin (wart or papilloma) or mucous membranes (condylomas). The viral nature of human warts was first demonstrated in the early 1900s when cell-free filtrates from lesions were shown to transmit the disease (Ciuffo, 1907). The first papillomavirus was identified in 1933 – by Richard Shope, who recognized the virus as the etiologic agent responsible for cutaneous papillomas in wild cottontail rabbits (Shope & Hurst, 1933). Despite this early start, efforts to study papillomaviruses in conventional cell culture was hampered during the 1950s and 1960s, when it was demonstrated that their replication is associated with the

differentiation process of the infected epithelium (Moore *et al.*, 1959; Noyes, 1959; Noyes & Mellors, 1957). In addition, papillomaviruses are not lytic viruses and so no plaque assay exists to study this virus. It was not until the late 1970s, that the advent of molecular cloning enabled the first papillomavirus genome to be cloned using bacteria. Application of these molecular techniques provided standardized reagents to study the biological and biochemical properties of papillomaviruses (Orth *et al.*, 1977). The availability of viral sequences eventually made the identification and molecular cloning of new viral genotypes possible. A noteworthy advance came with the discovery that several human papillomavirus (HPV) genotypes exist. Despite the finding that HPV infection was associated with the development of cutaneous warts (Gissmann *et al.*, 1977; Orth *et al.*, 1977), this virus was treated with little medical concern, because it was thought to cause lesions that had little or no potential to progress to malignancy. It was not until research focused on genital-mucosal HPVs, that the medical importance of this virus was heightened. The observation that the histological appearance of cervical dysplasia resembled that of viral papillomas (Meisels & Fortin, 1976), combined with several lines of evidence primarily from research conducted by Dr. Harald zur Hausen and colleagues, led to the proposal in the early 1980s, that HPV infection might be involved in cervical carcinoma (Meisels & Morin, 1981; zur Hausen *et al.*, 1981). These significant medical findings led to a renewed interest in HPV research. Many advances continued in this field, and by the 1990s, it was widely recognized that HPV is closely associated with certain human cancers, including anogenital cancers, head and neck cancers, and skin cancers in immunosuppressed individuals; but most notably as the etiologic agent for cervical carcinoma, which is the second most common gynaecological malignancy in the world. Dr. Harald zur Hausen's work in this field and initial hypothesis that HPV plays an important role in the cause of cervical cancer was recognized in 2008, when he received the Nobel Prize in Physiology and Medicine. Since the introduction of a cytology based test, the Papanicolaou test (Pap smear), the number of deaths caused by cervical cancer have seen a dramatic reduction by up to 90% in some populations where regular screening is implemented (Hans-Olov, A., 1994). Identifying HPV as a major risk factor for anogenital cancers has led to additional methods of preventive screening for HPV-induced lesions, including the detection of

HPV DNA in precancerous lesions. Although the study of animal papillomaviruses continues to shed new light on this family of viruses, the shift in focus towards HPV research in recent years has contributed greatly towards our understanding of viral tumourigenesis.

1.2.1 *Classification of papillomaviruses*

Papillomaviruses are widespread in nature, but thus far, have only been identified in vertebrates. This is a highly species-specific family of viruses and there has been no evidence of papillomaviruses from one species causing a productive infection in another. Papillomaviruses form a family of their own, the *Papillomaviridae*, which is further subclassified. The broadest phylogenetic branch – is a “genus” (formerly called “supergroups”), and is formed by papillomaviruses that are distantly related to one another and each designated by a letter of the Greek alphabet. A genus is comprised of several “species” that group together separate “types.” Classification of new viral types occurs when the nucleotide sequences of the L1 gene, which encodes the major capsid protein, differs from corresponding sequences of known papillomavirus genotypes by at least 10%. To date, more than 100 different types of HPVs have been fully genetically characterized (de Villiers *et al.*, 2004). Despite a similar genomic organization, different HPVs are tropic for epithelia at distinct anatomical locations. Approximately 30 different HPVs preferentially infect the anogenital and oral mucosa (Supergroup A, also known as Alpha papillomaviruses) and are termed “low-risk” or “high-risk” according to the propensity of the lesions that they cause to progress to malignancy. Low-risk viruses (e.g. types 1, 6 and 11) cause benign mucosal warts that usually regress with time. In contrast, infection with high-risk HPVs (e.g. types 16, 18, 31, 33 and 45) can cause epithelial lesions that often progress to invasive carcinomas. Most HPV infections are cleared or suppressed by cell-mediated immunity within 1-2 years of exposure (Stanley, 2006) and the rate of carcinogenic progression is relatively low. However, persistent infection with the same high-risk HPV type is the major risk factor for invasive carcinoma, accounting for virtually all cervical carcinomas and approximately 20% of head and neck cancers (Gillison *et al.*, 2000; zur Hausen, 1996). Amongst the high-risk genotypes, HPV16 is the most often associated with cancer, and is responsible for

approximately half of all cases of cervical cancer (Munoz *et al.*, 2003). The causal role that HPV16 plays in a variety of human cancers has made this high-risk genotype the most studied and well documented amongst all HPVs. HPV16 is also the genotype of focus in this thesis.

1.2.2 *Physical properties of papillomaviruses*

Papillomaviruses are small, nonenveloped, DNA viruses with an icosahedral shell that ranges from 52 to 55 nm in diameter. Virion particles contain a single, double-stranded circular genome, approximately 8000 base pairs (bp) in size. Within the capsid, viral DNA is found associated with cellular histones to form a chromatin-like complex (Favre *et al.*, 1975; Pfister *et al.*, 1977). The capsid is composed of 72 pentameric capsomers of the major capsid protein (L1) arranged on a T=7 surface lattice (Baker *et al.*, 1991). The internal minor protein (L2) makes up 20% of the remaining icosahedral shell and associates with only a subset of the L1 pentamers. Fine structural analysis by cryo-electron microscopy has also revealed that the capsomeres can arrange themselves in two distinct positions: one capable of making contact with six neighbouring pentamers (hexavalent) and another making contact with five neighbouring pentamers (pentavalent) (Baker *et al.*, 1991). L1 is capable of self-assembling into virus-like particles (VLPs) (Chen *et al.*, 2000). Expression of L1 alone or in combination with L2 using mammalian expression systems can produce VLPs that are virtually indistinguishable from authentic virions (Hagensee *et al.*, 1993; Kirnbauer *et al.*, 1992). This property of L1 has been utilized in the first generation of prophylactic vaccines against human papillomaviruses. Using VLPs is an ideal vaccine strategy, because not only are they noninfectious and nononcogenic, but they have also been shown to be capable of generating high titers of type-specific neutralizing antibodies (Kirnbauer *et al.*, 1992). Gardasil®, manufactured by Merck & Co., is currently the only prophylactic HPV vaccine available in North America. This vaccine is designed to prevent infection with four of the most common HPV genotypes: HPV types 6 and 11, which are responsible for 90% of genital warts cases; and the highly oncogenic types 16 and 18, which account for 70% of cervical cancer cases.

The genomic organization of all types of papillomaviruses is remarkably similar. All open reading frames (ORF) are located on the same strand of viral DNA (Figure 1.1). The HPV16 genome encodes 8 ORFs designated as either early (E) or late (L) ORF and are each followed by a polyadenylation site. The early region encodes non-structural, viral regulatory proteins that are expressed in non-productively infected cells and transformed cells. The late region encodes two capsid proteins that are expressed only in productively infected cells. There exists a region in the genome, approximately 1 kilobase pair (kbp) in size, which contains no ORFs. Termed the upstream regulatory region (URR) or alternatively, the long control region (LCR), this portion of the genome contains the origin of replication and a variety of cis elements which regulate DNA replication and gene expression, including the origin of viral DNA replication and transcription factor binding sites.

During transcription, only one strand serves as a template as the ORFs are transcribed into polycistronic mRNA. The mechanisms by which these mRNAs are ultimately translated are still poorly understood. Termination-reinitiation, leaky scanning, and shunting, have all been previously suggested as potential mechanisms (Remm *et al.*, 1999; Smotkin *et al.*, 1989; Stacey *et al.*, 1995; Tan *et al.*, 1994). However, translation of the E6 and E7 oncoproteins appears to occur predominantly through a leaky scanning mechanism (Stacey *et al.*, 2000). This method is commonly used by many viruses to express polycistronic RNAs. Ribosomes bind at the 5' end of mRNA transcripts and scan in a linear fashion for favourable AUG start sites. In this way, translation can occur at one or more initiation codons and this allows multiple viral proteins to be synthesized from a single mRNA.

1.2.3 Overview of the papillomavirus life cycle

Papillomaviruses are rather unique in that their productive life cycle is tightly linked to the differentiation program of the infected tissue (Figure 1.2). As previously stated, these viruses have a strong tropism for cells of the epithelium. Initial infection requires access of infectious particles to actively dividing cells in the epithelium. This is thought to occur through microlesions in the stratified epithelium that exposes cells within the basal layer. The cell surface receptor that allows for initial attachment of the

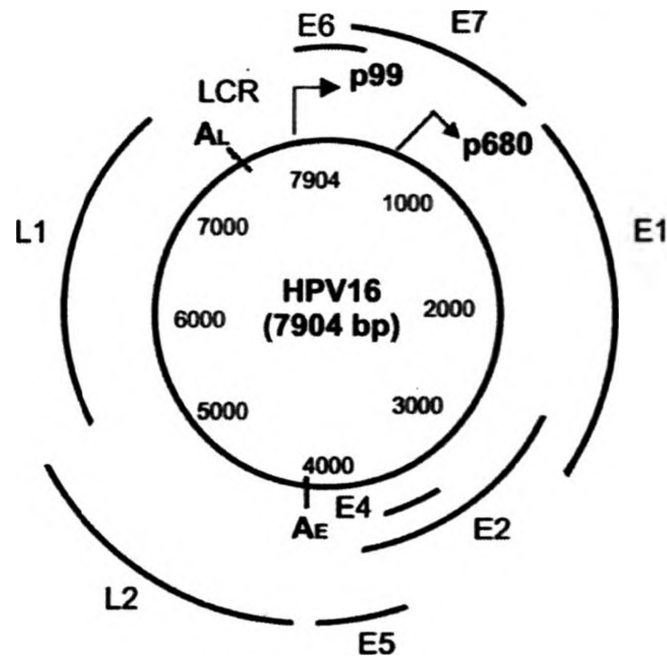


Table 1 – Overview of HPV gene products

Gene product	Description
E1	DNA replication and control of gene expression, helicase function
E2	DNA replication and regulation of gene transcription
E4	Viral assembly and release; interaction with cytoskeleton proteins
E5	Membrane signalling protein; growth stimulation by interaction with growth factor receptors
E6	Transformation; disruption of cell cycle; p53-degradation; anti-apoptotic effect; induction of genomic instability
E7	Transformation; disruption of cell cycle; interaction with pocket proteins; transactivation of E2F-dependent promoters; induction of genomic instability
L1	Major capsid protein
L2	Minor capsid protein; involvement in nuclear transport of viral DNA

(E) early gene; (L) late gene ; LCR: long control region; p99: early promoter; p680: late promoter; A_E: early polyadenylation site; A_L: late polyadenylation site

Figure 1.1. Schematic representation of the genomic organization of HPV16.

The 7.9kb genome of HPV16 can be divided into three functional regions: the early region (E), which encodes six early proteins E1 to E7; the late region (L), encoding the viral capsid proteins (L1 and L2); and a long control region (LCR). Adapted from Crosbie and Kitchener, Clinical Science (2006) 110, (543–552).

virus to the host cell is still unknown, although a few studies have suggested heparin sulphate as a candidate (Giroglou *et al.*, 2001; Joyce *et al.*, 1999). Following infection and uncoating, the virus migrates to the nucleus where it maintains itself as a low copy episome at approximately 10-200 copies per cell. As basal cells divide, the viral genome replicates co-ordinately with the host genome and is divided evenly among each daughter cell. The basal layer of the epithelium is comprised of undifferentiated progenitor cells that are used to continually replenish the outer layers of the epithelium as they are shed. Following replication, one of the daughter cells migrates away from the basal layer and begins a program of differentiation. Under normal conditions, uninfected cells withdraw from the cell cycle and undergo terminal differentiation as they migrate upward. With the exception of the viral E1 and E2 proteins, papillomaviruses lack the necessary enzymes to replicate their DNA and must rely on the host DNA replication machinery for a productive life cycle. Therefore, the virus must maintain an environment that is conducive for viral production. For this reason, virally-infected cells re-enter the cell cycle, using a process primarily mediated by the E6 and E7 proteins. As infected cells differentiate in the suprabasal layers of the epithelium, the viral genome undergoes a high-level of amplification that coincides with activation of the late promoter. It is only in the terminally differentiated layers of the upper epithelium, that viral DNA is packaged and new virions are assembled. Papillomaviruses are non-lytic viruses and, therefore, release their viral progeny into the extracellular environment as epithelial cells are shed from the surface.

The differentiation-dependent viral life cycle of papillomaviruses initially made propagation of this virus in the laboratory impossible, significantly hampering papillomavirus research efforts. However, in recent years, many of these technical difficulties have been overcome. Initial approaches utilized nude mice with xenografts from HPV-infected epithelial tissue (Kreider *et al.*, 1987). Although this method allows for the synthesis of HPV virions, it remains technically challenging. Much more success has been achieved through the use of organotypic “raft” cultures. Initially developed by Asselineau *et al.*, this *in vitro* system is designed to mimic epithelial differentiation (Asselineau & Prunieras, 1984). Most recently, a transient-transfection-based system was developed to encapsidate the papillomaviral genome, independently of viral DNA

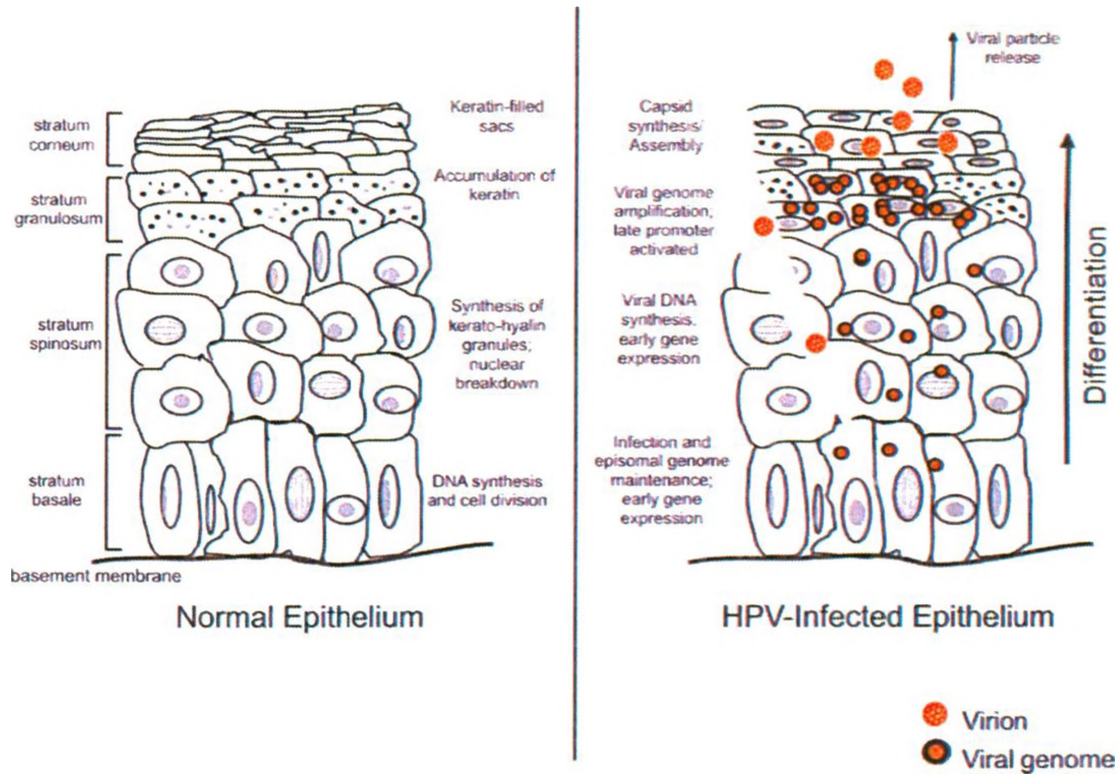


Figure 1.2. The HPV life cycle. Following infection (in this case, through a cut in the epithelium), the viral genome is maintained at the basal layer of the epithelium, where HPV infection is established. Early proteins are expressed at low levels for genome maintenance and cell proliferation. As the basal cells divide and migrate away from the basement membrane, they differentiate and withdraw from the cell cycle. However, differentiation and cell cycle arrest are effectively uncoupled in HPV-infected cells. This allows the viral life cycle to go through successive stages of genome amplification, virus assembly, and virus release, with a concomitant shift in expression patterns from early to late genes. As infected cells move toward the apical surface, dead epithelial cells are shed, releasing infectious virions into the extracellular environment.

replication and epithelial differentiation (Pyeon *et al.*, 2005). Although this system is able to produce higher yields of infectious virus than traditional methods, this technique is relatively new, and organotypic raft cultures remain the method of choice for most investigators.

1.3 Characteristics of the papillomavirus E7 protein

The E7 gene is expressed early after infection, and its transcription is upregulated as the host epithelial cell undergoes differentiation. In these cells, E7 together with the other major viral oncoprotein, E6, are tasked with reprogramming the cellular environment to make it permissive for viral replication. The E6 protein functions primarily by binding to p53, targeting it for degradation and abrogating its ability to induce cell cycle arrest or apoptosis (Hubbert *et al.*, 1992; Huibregtse *et al.*, 1991; Scheffner *et al.*, 1990).

The E7 protein is the topic of this thesis and the remainder of the introduction will focus on the current state of our knowledge of its properties, functions and structure. The E7 protein of HPV16 is a small, acidic polypeptide of 98 amino acids. The amino terminus of HPV16 E7 has both sequence and functional homology to a small portion of conserved region (CR)1 and to the entire CR2 of adenovirus (Ad) E1A and related sequences in the simian vacuolating virus 40 large tumour antigen (SV40 T-Ag). Based on amino acid sequence homology within E7 proteins, E7 can be separated into three conserved regions in an analogous fashion to AdE1A (Figure 1.3). CR1 spans residues 2-15, CR2 spans residues 16-37 and CR3 spans residues 58-94. The CR2 homology domain includes a Leu-X-Cys-X-Glu (LXCXE, where X is any amino acid) motif, which is the canonical binding site for the retinoblastoma susceptibility locus product (pRb) and the related pocket proteins, p107 and p130. Immediately adjacent to this motif is a consensus phosphorylation site for casein kinase II (CKII). The carboxyl terminus of HPV16 E7 contains a zinc-binding domain that is composed of two Cys-X-X-Cys motifs separated by 29 amino acids (C₅₈-X-X-C₆₁-X₂₉-C₉₁-X-X-C₉₄) (Patrick *et al.*, 1992). This region of E7 is vital in maintaining the structural integrity of the protein and also

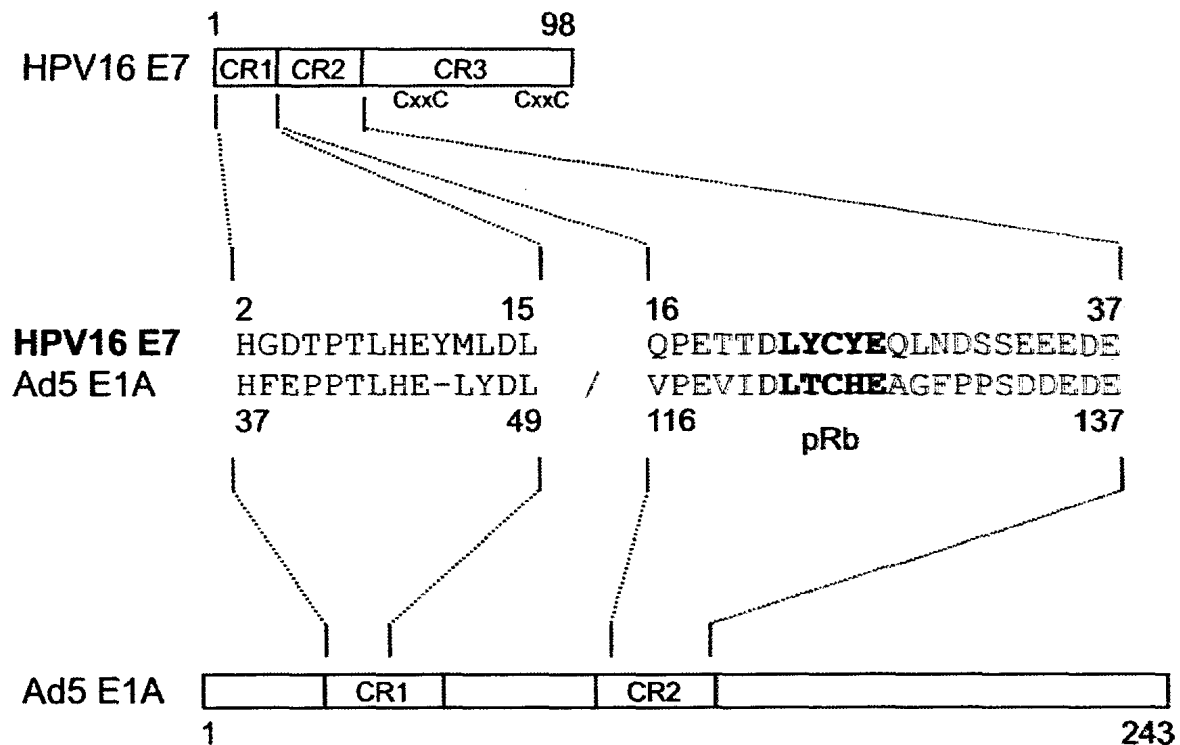


Figure 1.3. Location of conserved regions in HPV16 E7 and sequence alignment between conserved regions (CR) 1 and 2 of adenovirus type 5 E1A and regions of HPV16 E7. The positions of CR1, CR2 and CR3 in HPV16 E7 and human adenovirus type 5 (Ad5) E1A and the depicted regions of amino acid similarity are indicated. The pRb binding regions within CR2 of both proteins are indicated in bold, and shares the consensus sequence LXCXE. The amino terminal region of the E7 proteins share sequence similarity with portions of CR1 and the entire CR2 AdE1A. Although the sequences in the carboxy terminal half of E7 are well conserved among the HPVs, this region possesses no substantial amino acid sequence similarity to any part of E1A with the exception of the repeated Cys-X-X-Cys motifs, which are also found in CR3 of E1A.

functions as a dimerization domain (Clemens *et al.*, 1995; Liu *et al.*, 2006; McIntyre *et al.*, 1993; Ohlenschlager *et al.*, 2006).

E7 is a predominantly nuclear protein, although cytoplasmic pools of E7 have also been detected *in vivo* (Smotkin & Wettstein, 1987). It lacks a prototypical nuclear localization sequence and enters the nucleus through a novel Ran-dependent pathway (Angeline *et al.*, 2003). The low steady-state levels of E7 in the cell can be attributed to the relatively short half life of this protein of less than 2 hours (Smotkin & Wettstein, 1987). E7 is also targeted for degradation by the ubiquitin-dependent pathway (Reinstein *et al.*, 2000). Interestingly, HPV16 E7 is known to aberrantly migrate on SDS polyacrylamide gels at an apparent molecular weight size of 18-20 kDa as opposed to its predicted size of approximately 11 kDa. This has been found to be caused primarily by the acidic nature of the amino terminal CR1 domain of HPV16 E7 and is not shared by HPV6 E7 proteins (Heck *et al.*, 1992; Munger *et al.*, 2001).

1.4 Biological activities and molecular targets of E7

The E7 protein of HPV16 has been studied the most extensively as it was the first high-risk HPV oncogene to be discovered (Kanda *et al.*, 1988; Phelps *et al.*, 1988; Vousden *et al.*, 1988; Yutsudo *et al.*, 1988). E7 of high-risk HPVs possesses several properties characteristic of an oncogene. E7 alone can transform cultured rodent fibroblasts (Bedell *et al.*, 1989; Phelps *et al.*, 1988) and bypass serum starvation-induced growth arrest (Pei *et al.*, 1998). In cooperation with an activated *ras* oncogene, E7 can transform primary rodent fibroblasts (Phelps *et al.*, 1988), and in the presence of E6, it can efficiently immortalize primary human keratinocytes (Halbert *et al.*, 1991). Low-risk E7 proteins possessed a marked decrease in transforming and immortalizing activities. The amino terminal domain of E7 has been shown to be largely responsible for the differing transforming abilities between the high-risk viruses and low risk viruses.

In low-grade HPV-induced lesions, the HPV genome is maintained episomally. However, during the malignant progression of these high-risk HPV-associated lesions, the viral genome is often found integrated into the host chromosomes. This is a key event in HPV-induced carcinogenesis, and is accompanied by termination of the viral life cycle and loss of portions of the HPV genome. Viral tumourigenesis can therefore be

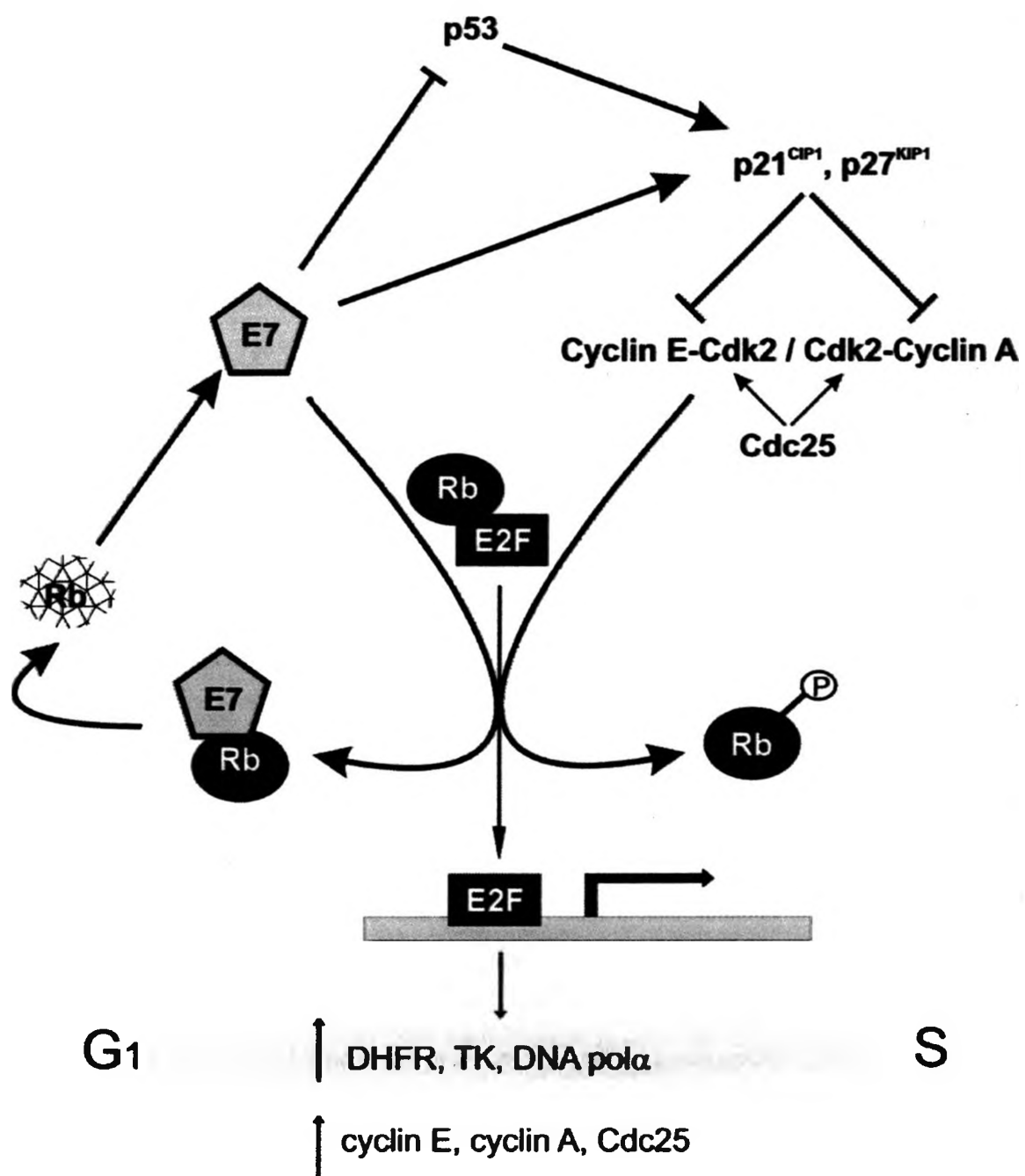
regarded as a by-product of virus propagation. If integration disrupts the region encoding the viral E2 transcriptional regulator, loss of the repressive action of E2 results in the deregulated overexpression of the E6 and E7 oncoproteins (Choo *et al.*, 1987). In fact, both E6 and E7 are consistently expressed in malignant tissue and are required for the induction and maintenance of the transformed phenotype, as it has been shown that re-expression of E2 represses E6 and E7 transcription in addition to causing rapid growth arrest and senescence in cervical cancer cell lines (Goodwin & DiMaio, 2000; Goodwin *et al.*, 2000).

1.4.1 Association of E7 with pRb

In normal epithelial cells, the differentiation program is tightly linked to withdrawal from the cell cycle. The HPV E7 protein is tasked with the critical role of altering the cellular environment into a DNA replication competent state that would support viral replication. The E7 protein of HPV16 lacks any intrinsic enzymatic activity, and therefore, it primarily exerts its functions by interacting with cell regulatory proteins (Figure 1.4).

The binding of E7 to pRb is the most studied interaction between this viral protein and a cellular target. The main function of pRb is to act as a negative regulator of the cell cycle. pRb can regulate the activity of the E2F family of transcription factors. Typically, the association of pRb with E2F blocks the transition through the G₁/S checkpoint of the cell cycle. In HPV-infected cells, E7 binds to hypophosphorylated pRb and related pocket proteins, p107 and p130, primarily through the LXCXE motif in CR2 (Dyson *et al.*, 1992; Munger *et al.*, 1989). Several residues within CR3 appear to provide a second low-affinity pRb binding site (Liu *et al.*, 2006; Patrick *et al.*, 1994) and facilitate the displacement of E2F (Helt & Galloway, 2001; Huang *et al.*, 1993; Wu *et al.*, 1993). In contrast to AdE1A, amino terminal sequences of E7 do not appear to enhance the interaction between E7 and Rb. The initial school of thought was that E7, like AdE1A and SV40 T-Ag, would stoichiometrically interact with pRb and the related pocket proteins, sequestering them away from E2Fs and the gene promoters they regulate. However, E7 proteins are expressed at much lower levels than their cellular target proteins, suggesting that this strategy would not be favourable. Instead, it is now known

Figure 1.4. Relaxation of the G₁/S checkpoint by HPV E7 proteins. High-risk E7 proteins destabilize pRb, thus abrogating the pRb/E2F transcriptional repressor complex. This causes an increased expression of E2F responsive cell-cycle regulators, including Cdc25, cyclin E and cyclin A, and replication enzymes such as dihydrofolate reductase (DHFR), thymidine kinase (TK) and DNA polymerase α (pol α). HPV16 E7 proteins can also associate with and abrogate the inhibition of cdk2 activity by interacting with the cyclin-dependent kinase inhibitors, p21^{CIP1} and p27^{KIP1}. These activities contribute to the ability of HPV to create and/or maintain an environment in the differentiating host cell that is conducive for viral replication.



that E7 can circumvent this problem by reducing the steady-state levels of pRb through proteasome-mediated proteolysis (Berezutskaya & Bagchi, 1997; Boyer *et al.*, 1996; Giarre *et al.*, 2001), independently of E7 degradation (Gonzalez *et al.*, 2001). In this fashion, a low level of E7 protein expression can impair the entire spectrum of biological activities of the pocket proteins, with the further advantage of aberrantly activating E2F transcription. Binding of E7 to pRb is necessary for degradation of pRb, but additional sequences outside the pocket protein-binding domain are known to be involved in destabilization of pRb. Evidence for this is seen with transformation-impaired E7 mutations within the N-terminal region and CKII phosphorylation site that can efficiently associate with pRb but are defective for pRb degradation (Edmonds & Vousden, 1989; Gonzalez *et al.*, 2001; Helt & Galloway, 2001; Jones *et al.*, 1997; Phelps *et al.*, 1992). Furthermore, low-risk HPV E7 proteins can also efficiently bind to pRb, but are unable to target pRb for degradation (Giarre *et al.*, 2001; Gonzalez *et al.*, 2001). Above all, it has been shown that E7's oncogenic potential correlates more closely with the ability to induce pRb proteolysis than merely with its pRb-binding potential (Giarre *et al.*, 2001; Gonzalez *et al.*, 2001).

The initial analysis of the E7-pRb interaction led to the assumption that the ability of E7 to disrupt pRb/E2F complexes was linked to HPV tumorigenicity. As mentioned above, it was subsequently shown that the ability of E7 to destabilize pRb and the pocket proteins correlated more tightly with oncogenic potential and cellular transformation. However, this association has been somewhat weakened by the observation that destabilization of the pocket proteins is not sufficient to overcome cell cycle arrest (Helt & Galloway, 2001) or efficiently transform some cell lines (Banks *et al.*, 1990). These observations highlight the importance of additional cellular targets of E7 involved in stimulating cell cycle progression.

1.4.2 *E7 proteins and the p53 tumour suppressor*

Normal cells expressing a single oncogene are predisposed to apoptosis, particularly under conditions of growth factor deprivation. This “tropic sentinel pathway” represents a cellular defence mechanism to eliminate potential neoplastic cells and is p53 dependent (Eichten *et al.*, 2002; Evan & Vousden, 2001). Subversion of this

intrinsic apoptotic response is an important aspect of carcinogenic progression and occurs in HPV infected cells (Eichten *et al.*, 2002; Hanahan & Weinberg, 2000). Steady-state levels of p53 are increased in E7-expressing cells (Demers *et al.*, 1994), suggesting that E7 may perturb p53 degradation (Jones *et al.*, 1997). The mechanism underlying E7-mediated p53 stabilization appears to be independent of p19^{ARF}, an E2F-regulated inhibitor of mdm2-mediated p53 degradation (Seavey *et al.*, 1999). Typically, stabilization of p53 is accompanied by an increase in transcription of p53 targets. However, the increased levels of p53 are transcriptionally “inert” (Eichten *et al.*, 2004); and the augmented steady-state levels of p53 targets, such as p21, appears to be the consequence of protein stabilization (Jones *et al.*, 1999; Noya *et al.*, 2001). It has been suggested that this phenomenon may be due to a disturbance in transcriptional activity of p53 in E7-expressing cells. In fact, it has been shown that E7 interferes with the activation of p53-responsive constructs in transient transfection assays (Massimi *et al.*, 1997). Nevertheless, it has been demonstrated that normal mdm2-mediated p53 turnover does not occur in cervical cancer cell lines and that the rapid turnover of p53 in these cell lines is primarily a result of the action of the E6/E6-associated protein ubiquitin ligase complex (Hengstermann *et al.*, 2001).

1.4.3 *Effects of E7 on other components of the cell cycle machinery*

Although replication of the HPV genome critically depends on the ability of the infected differentiating keratinocyte to support DNA synthesis, certain aspects of the viral life cycle, particularly the early to late switch and the synthesis of capsid proteins require the milieu of a differentiated host cell. For this reason, it is important that E7 does not reverse or alter the differentiated phenotype, but rather, uncouple differentiation from proliferation. There are a number of mechanisms by which E7 attains this goal (Figure 1.4).

E7 associates with and regulates the expression of several cell cycle regulators during the G₁ phase of the cell cycle encouraging transition through the S-phase to support viral replication. As described in the preceding section, inactivation of pocket proteins by E7 increases the expression of a number of E2F responsive genes. Most notably, HPV16 E7 causes an increase in the transcription of two cyclin dependent kinase

(cdk) activators, cyclin E and cyclin A (Martin *et al.*, 1998; Zerfass *et al.*, 1995), and of the Cdc25A phosphatase (Katich *et al.*, 2001). E7 can also directly associate with cyclinA/Cdk2 and cyclinE/Cdk2 complexes (Nguyen & Munger, 2008). HPV16 E7 is also capable of abrogating the inhibition of cdk2 activity by interacting with cdk inhibitors (CKI), p21^{Cip1} and p27^{Kip1} (Funk *et al.*, 1997; Jian *et al.*, 1998; Zerfass-Thome *et al.*, 1996). Levels of p21 are also increased due to protein stabilization by E7 (Jian *et al.*, 1998). Interestingly, cdk2 remains active in E7-expressing cells regardless of the increase in p21. These inhibitors are known to function in cell cycle arrest in terminally differentiating keratinocytes (Alani *et al.*, 1998; Di Cunto *et al.*, 1998; Missero *et al.*, 1996). Thus, the ability of E7 to interfere with the ability of these CKIs to block cell cycle progression is critical to the ability of HPV to uncouple differentiation from DNA synthesis. The culmination of these activities resulting from HPV16 E7 expression is the inappropriate phosphorylation of substrates that allows cells to transverse the G₁/S boundary. Consequently, differentiating cells are induced to enter and sustain a replication-competent cellular milieu.

1.4.4 Transcriptional regulation by E7

The E7 protein can act as a transcriptional regulator and can influence E2F-induced transcription through a number of means that are independent of pRb inactivation. E2F proteins play an important role in regulating the transcription of several genes that are rate-limiting for S-phase progression and cellular DNA replication, including cell cycle regulators (e.g. cyclins E and A), DNA polymerase α , and enzymes involved in nucleotide biosynthesis (e.g. dihydrofolate reductase and thymidine kinase) (Munger *et al.*, 2001).

The organization of eukaryotic chromatin has a major impact on gene expression. The exposed amino terminal tails of histones are available to undergo a number of reversible post-translational modifications. These are site-specific changes that locally modulate the configuration of chromatin and include acetylation, phosphorylation, methylation and ubiquitination (Li, 2002). The E7 proteins are able to associate directly and indirectly with the two families of histone-modifying enzymes: histone acetyltransferases (HATs) and histone deacetylases (HDACs). Site-specific histone

acetylation alters the charge of histone tails affecting their ability to interact with the negatively-charged DNA. While acetylation of specific lysine residues neutralizes the positively-charged histones, deacetylases favour the formation of compact, repressive chromatin (Li, 2002). In this way, the ability of E7 to interact with these enzymes affects the accessibility and expression of specific chromosomal loci. HPV16 E7 is able to indirectly associate with HDAC-1 via a direct interaction with Mi2 β of the NuRD (nucleosome remodelling and histone deacetylation) chromatin remodelling complex (Brehm *et al.*, 1999). This is thought to occur in a pRb-independent manner. pRb and the related pocket proteins, p107 and p130 are also capable of associating with HDAC-1 through sequences that are distinct from the E7 binding domain (Brehm *et al.*, 1998; Dick *et al.*, 2000; Luo *et al.*, 1998). The formation of pRb/E2F/HDAC complexes represses E2F responsive promoters and transcription of S-phase specific genes. In HPV-expressing cells, it is not known whether E7 competes with pRb for HDAC-1 or if E7/pRb/E2F/HDAC inhibitory complexes are formed. In either event, the consequence of the interaction between E7 and HDACs increases the levels of E2F-mediated transcription in differentiating keratinocytes (Longworth & Laimins, 2004). HPV16 E7 is also able to interact with various HATs including p300, CBP, and p300/CBP-associated factor (pCAF) (Avvakumov *et al.*, 2003; Bernat *et al.*, 2003; Huang & McCance, 2002). In support of this observation, the acetylation of histone H3 was observed to be increased at E2F-responsive promoters in HPV16-expressing human foreskin keratinocytes (Zhang *et al.*, 2004). The ability of E7 to bind components of both the acetyltransferase and deacetylase complexes may be difficult to reconcile. However, it has been suggested that these interactions may be regulated by post-translational modifications to E7 or promoter-specific interactions (Avvakumov *et al.*, 2003).

Another method by which E7 is able to enhance transcriptional activation is to be recruited to promoters by interacting with transcription factors or components of the transcriptional apparatus. E7 interacts with several members of the AP-1 family of transcription factors, including c-Jun, c-Fos, JunB and JunD, and activate transcription of their responsive promoters (Antinore *et al.*, 1996). E7 also binds the TATA-binding protein (TBP) (Massimi *et al.*, 1997) and several TBP-associated factors (TAFs), including TBP-associated factor-110 (Mazzarelli *et al.*, 1995). Furthermore, HPV16 E7

targets multiple members of the E2F transcription factor family, including both the transcriptional activator, E2F1 and repressor, E2F6 (Hwang *et al.*, 2002; McLaughlin-Drubin *et al.*, 2008). Interaction with E2F1 can activate E2F1-driven transcription independently of pRb and drive quiescent cultured fibroblasts through the G₁/S cell cycle checkpoint. In contrast, E2F6 is normally activated in a negative feedback loop during the S-phase to slow down or exit progression through the cell cycle (Cartwright *et al.*, 1998; Gaubatz *et al.*, 1998; Trimarchi *et al.*, 1998). E2F6 is therefore an attractive target for E7 to deregulate. Blockade of E2F6-mediated transcriptional repression would presumably extend the S-phase competence of differentiating cells (McLaughlin-Drubin *et al.*, 2008). These specific interactions with E7 can have profound implications on the induction of malignancy, as several AP-1 proteins and E2F1 have been implicated as potent transforming agents (Adams & Kaelin, 1996).

The activities of E7 described above complement one another and primarily act during G₁ to allow aberrant transition through the cell cycle. Together, this co-ordinated deregulation of the complex control over the host cell cycle creates and maintains an environment that is conducive for viral replication in differentiated keratinocytes.

1.5 Structural organization of the E7 proteins

Undoubtedly, studying the structure of the HPV E7 proteins can provide great insight into the mechanisms behind E7's broad specificity of cellular targets. Having the structure of E7 is an invaluable tool that can significantly aid in planning mutational studies that address E7 functions and interactions. These studies may also help explain the functional differences between the high-risk and low-risk E7 proteins, and the malignant progression of lesions induced by high-risk HPVs. Ultimately, this information may suggest new antiviral strategies to interfere with the action of E7.

Recently, a solution structure of the high-risk HPV45 E7 protein was solved using Nuclear Magnetic Resonance (NMR) spectroscopy (Ohlenschlager *et al.*, 2006) (Figure 1.5A). Experiments using full-length HPV45 E7 and a construct representing its CR3 region revealed that the N-terminus (amino acids, 1-54) of this oncoprotein is unstructured and flexible in solution, whereas the C-terminus (residues 55-106) folds autonomously into a well-structured zinc-binding domain. Interestingly, results obtained from this NMR analysis are consistent with the intrinsic disorder of high-risk E7 proteins predicted by bioinformatic analysis (Uversky *et al.*, 2006) and with a previously reported secondary structure of E7 proteins derived by consensus averaging and spectroscopic studies (Ullman *et al.*, 1996).

No protein structures have been reported for the CR1 and CR2 portions of HPV E7, with the exception of a nine residue peptide in CR2 containing the LxCxE motif. This short peptide was co-crystallized with pRb and was shown to bind to a shallow groove of pRb in a highly extended β -strand conformation. The N-terminal region of E7 represents an example of another functional intrinsically disordered region that undergoes a localized conformational change upon interaction with its biological target (Dyson & Wright, 2005). This is not surprising given the fact that intrinsic protein disorder is very common in cancer-associated proteins, with approximately 79% of cancer-associated and 66% of cell-signalling proteins containing predicted regions of disorder of 30 residues or longer (Iakoucheva *et al.*, 2002). The dynamic nature of the N-terminus may aid HPV E7 proteins in fulfilling some of their biological functions.

The solution structure of the C-terminal domain of HPV45 E7 is virtually identical to an X-ray crystallography structure of a CR3 construct derived from the low-

risk HPV 1 E7 protein (Figure, 1.5; (Liu *et al.*, 2006)). Comparison of the HPV1 and HPV45 dimeric structures of E7 and the surface-exposed residues for the two proteins, did not reveal a scheme that could help decipher the properties between the low-risk and high-risk E7 proteins. The E7-CR3 domain forms a well-structured zinc-binding domain with a unique $\beta 1\beta 2\alpha 1\beta 3\alpha 2$ topology containing a zinc-binding fold that is not found in zinc-binding proteins unrelated to E7. This region contains a C4-type zinc finger which co-ordinates one molecule of zinc (II) ion. The distance between the two Cys-X-X-Cys motifs appeared to be too large to form a classical zinc finger structure that was first proposed for the *Xenopus* transcription factor IIIA (TFIIIA) (Lee *et al.*, 1989; Miller *et al.*, 1985).

Both structural studies demonstrate that E7 assembles as a roughly globular, obligate zinc-dependent dimer. Although the co-ordinated zinc ions are not directly involved in dimer formation, they are important for maintaining the folded state of the monomeric CR3 as a prerequisite for dimerization. Indeed, the removal of zinc ions causes unfolding of CR3. E7 forms a stable dimer involving the $\alpha 1$ helices of each monomer and β -sheet interactions between the $\beta 2$ and $\beta 3$ strands of opposing strands. Dimerization leads to the formation of a contiguous hydrophobic core that is further stabilized by a subset of residues that form intersubunit contacts. The formation of these dimeric complexes does not appear to be dependent on intermolecular disulfide bonding (McIntyre *et al.*, 1993). A sequence alignment of the CR3 domains from E7 proteins previously predicted that the E7 dimer is maintained in large part by a hydrophobic core, based on the observations that 7 hydrophobic residues in CR3 exhibit more than 90% conservation, and 5 more exhibit 70% conservation (Ullman *et al.*, 1996). Furthermore, the high degree of sequence conservation within the CR3 region suggests that these structural features are likely to be conserved among E7 proteins, with the most highly conserved residues within CR3 representing structural components of the monomer core and dimer interface (Figure 1.6).

This structure of E7 forces a re-evaluation of much of the existing literature, which addresses E7 function and its interaction with cellular proteins using mutants of these highly conserved residues. A number of mutational studies have mapped the interaction of cellular targets of E7 to CR3, using mutations that target the hydrophobic

core, including p21^{CIP1}, p27^{KIP1}, TBP, TAF110, Mi2 β and pCAF (Avvakumov *et al.*, 2003; Brehm *et al.*, 1999; Funk *et al.*, 1997; Jian *et al.*, 1998; Massimi *et al.*, 1997; Mazzearelli *et al.*, 1995; Zerfass-Thome *et al.*, 1996). The loss of interaction and/or function observed in those cases can possibly be attributed to gross structural effects upon E7 or to the loss of dimerization rather than to impairment of specific functions. This issue has consistently been raised with mutations in the coordinating cysteines that destroy protein stability (Clemens *et al.*, 1995; Phelps *et al.*, 1992; Watanabe *et al.*, 1990) and impair transformation and transactivation functions (Edmonds & Vousden, 1989; McIntyre *et al.*, 1993; Phelps *et al.*, 1992; Storey *et al.*, 1990). However, to date, it has not been established whether dimerization is necessary for the biological activities of E7. There are several possible states in which E7 can exist in a cell. The predominant functional form of E7 can either be a dimer or monomer, or both forms can be active *in vivo*. In addition, the concentration levels of the E7 protein in the cell may or may not affect this outcome by altering the steady-state equilibrium between the monomeric and dimeric forms. To explore this question, a thorough structure-function analysis will need to be performed using a comprehensive collection of mutants in CR3.

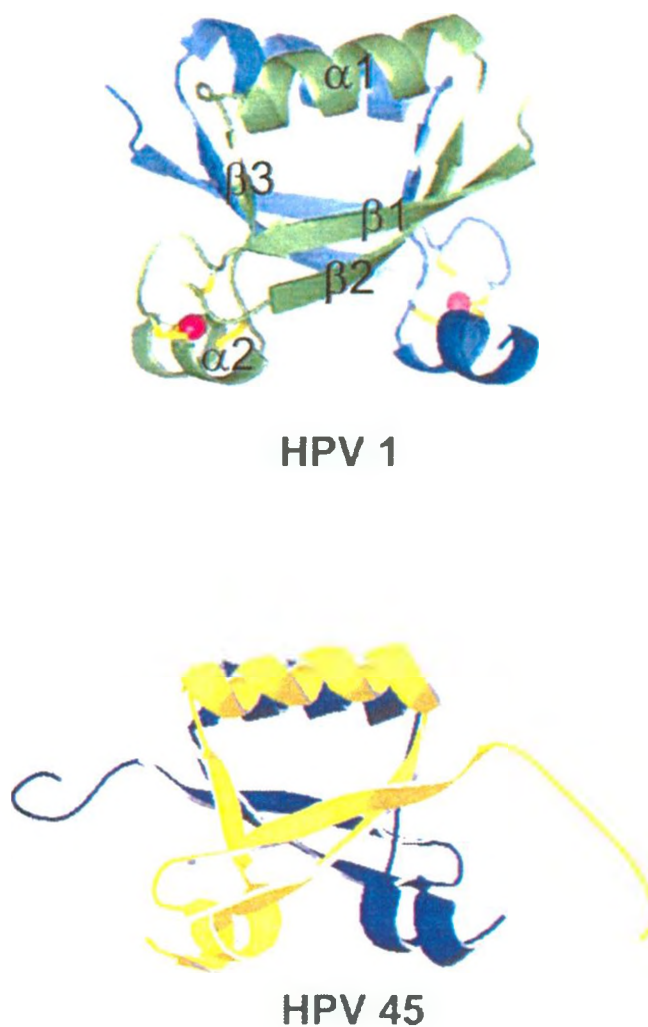


Figure 1.5. Representative 3-dimensional structures of the HPV1 and HPV45 CR3 dimers. A structural schematic of the HPV1 (residues 44-93) and HPV45 E7 (residues 55-106) CR3 dimer. CR1 and CR2 are not shown as they are intrinsically disordered. Structures of monomers are coloured either in green and blue, or blue and yellow. Bound zinc ions are coloured in red (balls) and cysteine ligands in yellow (sticks). (Structures from Liu, X. et al. J. Biol. Chem. 2006;281:578-586; and Ochenschlager, O. et al. Oncogene. 2006; 25: 5953-5959).

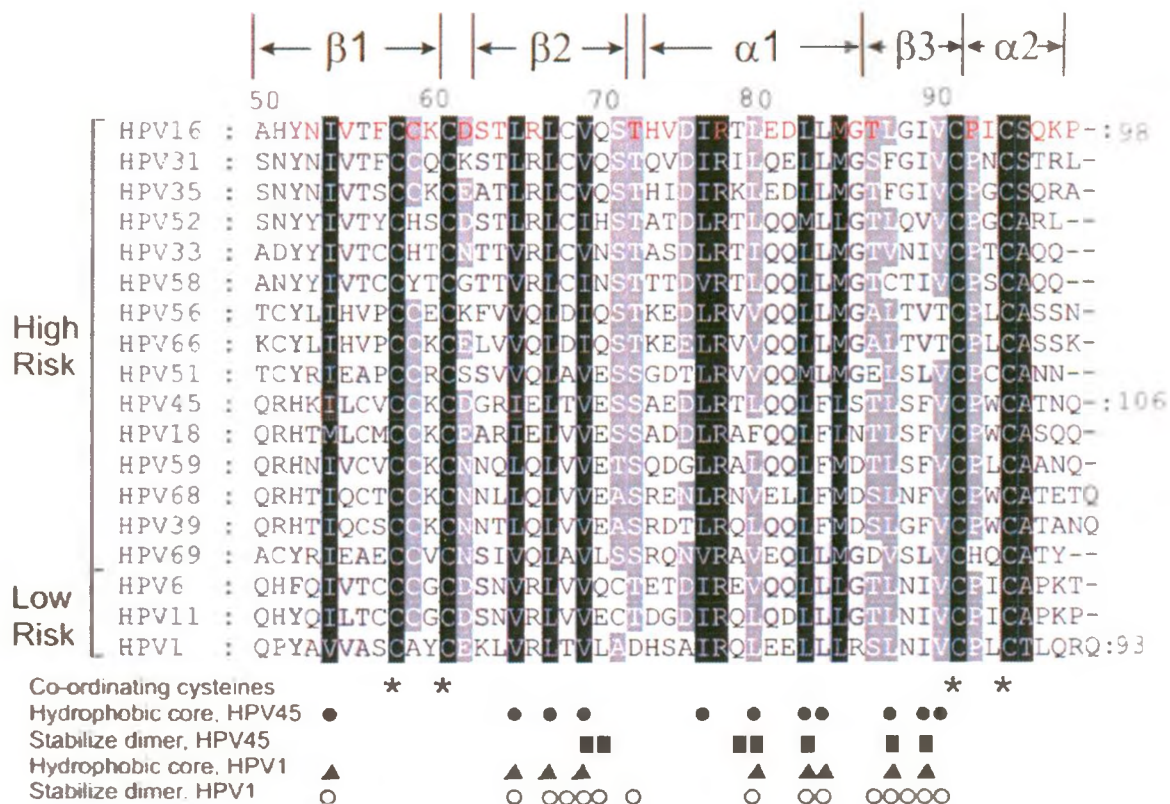


Figure 1.6. Multiple sequence alignment of CR3 regions from representative Supergroup A HPVs. A) Perfectly conserved residues are highlighted in black with progressively less conserved residues in dark to light gray. Numbering scheme corresponds to that of HPV16 E7 and starts at amino acid 49. Secondary structure elements determined from the crystal structure and solution structure of HPV1 and HPV45 E7 are indicated at the top. Residues that are involved in zinc co-ordination, contribute to the hydrophobic core of each monomer, or stabilize the dimer, are indicated below the alignment.

1.6 Yeast as a model system for studying HPV16 E7

A large majority of the experiments reported in this thesis utilize the budding yeast *Saccharomyces cerevisiae*. This simple eukaryotic organism has been widely used as a model system for studying many aspects of mammalian cell function. It was the first eukaryotic genome to be completely sequenced (Goffeau *et al.*, 1996) and a significant number of yeast genes have mammalian homologues (Botstein *et al.*, 1997). In addition to the conservation of many genes across species, yeast are particularly suitable for the study of many cellular pathways. The understanding of many cellular mechanisms and signalling pathways in mammalian cells began with studies in yeast, including the genes involved in regulating the cell cycle and its checkpoint controls (Botstein *et al.*, 1997; Hartwell *et al.*, 1970). Yeast are often utilized in genetics and cell biology, largely because they are easy to manipulate, and are cheaper and faster to work with than mammalian cells. These characteristics also make them ideal for large-scale screens and analyses. Unlike more complex eukaryotes, this unicellular organism has a short generation time (doubling time 1.5-2 hours at 30°C) and can be grown on defined media which gives the investigator complete control over environmental parameters. Yeast have both a stable haploid and diploid state, which makes them particularly favourable for genetic analysis. The ability to grow as a haploid allows recessive mutations to be easily manifested and isolated, while complementation tests can be performed in the diploid state. The ease of genetic manipulation in yeast has allowed for the identification of new genes and for the functional dissection of previously identified genes (Forsburg, 2001) and a large variety of protocols for genetic analysis in yeast are available (Adams *et al.*, 1998). *S. cerevisiae* provides a particularly useful, genetically amenable system for the study of HPV because the full-length HPV16 genome can replicate as an episome when linked in *cis* to a selected yeast marker (Angeletti *et al.*, 2002). Furthermore, yeast, which lack pRb or orthologous proteins allow one to focus on the secondary functions and interactions of E7 without the competing effects of pRb binding and its consequences.

1.7 Thesis objectives and hypothesis

The overall objective of this thesis was to construct a series of mutations in CR3 of the HPV16 E7 protein and to comprehensively analyze these mutants to determine whether E7 exists as a functional dimer. The characterization of these mutants described here has laid the groundwork for a more extensive analysis of CR3, which will address the role of dimerization in E7 function. At this time, this structure/function analysis represents the first detailed study conducted on this region of the HPV16 E7 protein.

The hypothesis underlying the specific aims outlined below is that the HPV16 E7 protein can exist as an obligate zinc-dependent dimer *in vivo*; and that CR3 residues that are not involved in maintaining this tertiary structure, contribute to the ability of E7 to interact with cellular proteins.

Objectives:

- 1) Systematically mutate all solvent-exposed residues of CR3 in HPV16 E7
- 2) Assess the dimerization properties of these CR3 mutants as well as of additional mutants previously constructed by others.
- 3) Utilize these mutants to map the E7 surface that binds to the pCAF acetyltransferase.

CHAPTER 2: MATERIALS AND METHODS

2.1 *Escherichia coli* strain, *Saccharomyces cerevisiae* strains and media

The DH5 α strains of *Escherichia coli* (F⁻, ϕ 80d/*lacZ* Δ M15, Δ (*lacZYA-argF*), U169, *deoR*, *recA1*, *hsdR17*(r_k⁻, m_k⁺), *phoA*, *supE44*, Λ mbd⁺, *thi-1*, *gyrA96m relA1*) used in this study was provided by Dr. Gabriel DiMattia (The University of Western Ontario). At all stages of plasmid construction and replication, DH5 α bacteria were grown in Luria Bertani Broth (LB), with or without selective antibiotics.

The yeast strains utilized in this study and their sources are listed in Table 2.1. The *Saccharomyces cerevisiae* L-40 strain was used for the dimerization studies while the wild-type *w303-1a* strain was used to map the binding surface of pCAF on E7. Yeast culture medium was prepared using standard techniques (Adams *et al.*, 1998).

Table 2.1 List of yeast strains used in this study and the sources from which they were obtained

Strain (JMY #)	Genotype	Source
<i>w303 1A</i> (JMY 33)	<i>MATa leu2-3, 112; his3-11,15; trp1-1; ura3-1; ade2-1, can1-100</i>	Dr. C. Boone (U. Toronto)
L-40 (JMY 109)	<i>MATa his3Δ200; trp1-90; leu2-3, 112; ade2; lys2-801am</i> <i>LYS2: (lexAop) 4-HIS3 URA3: (lexAop)8-lacZ Gal4</i>	Dr. D. Mangroo (U. Guelph)

*JMY = Joe Mymryk Yeast database

2.2 Plasmid construction

The plasmids utilized in this study are listed in Table 2.2. Enzymes utilized in this study were purchased from New England Biolabs (Mississauga, ON, CA) and

Invitrogen (Carlsbad, CA, USA). Reactions were performed in accordance to the supplier's recommendations.

Table 2.2 List of plasmids used in this study and their sources

Gene	Parent Vector	JMB #	Source
None	pJG4-5+	405	Joe Mymryk
None	pBAIT	440	Joe Mymryk
HPV 16 E7 (wild-type)	pBAIT	2915	Joe Mymryk
HPV 16 E7 H2P	pBAIT	3108	Dr. N. Avvakumov
HPV 16 E7 Δ PTLHE 6-10	pBAIT	3107	Dr. N. Avvakumov
HPV 16 E7 M12V	pBAIT	3121	Dr. N. Avvakumov
HPV 16 E7 Δ DLYC 21-24	pBAIT	3106	Dr. N. Avvakumov
HPV 16 E7 S31G/S32G	pBAIT	3105	Dr. N. Avvakumov
HPV 16 E7 E46Q/D48N	pBAIT	3104	Dr. N. Avvakumov
HPV 16 E7 Y52A	pBAIT	2939	This study
HPV 16 E7 N53D	pBAIT	2940	This study
HPV 16 E7 V55T	pBAIT	2941	This study
HPV 16 E7 F57A	pBAIT	2925	This study
HPV 16 E7 C58G/C91G	pBAIT	3103	Dr. N. Avvakumov
HPV 16 E7 HPV 16 E7 C59S	pBAIT	2924	This study
HPV 16 E7 K60E	pBAIT	2923	This study
HPV 16 E7 D62K	pBAIT	2922	This study
HPV 16 E7 S63D	pBAIT	2904	Sean Jewell
HPV 16 E7 T64D	pBAIT	2942	This study
HPV 16 E7 L65A	pBAIT	3109	Dr. N. Avvakumov
HPV 16 E7 R66E	pBAIT	2903	This study
HPV 16 E7 L67R	pBAIT	3110	Dr. N. Avvakumov
HPV 16 E7 V69A	pBAIT	3116	Dr. N. Avvakumov
HPV 16 E7 S71I	pBAIT	3122	Dr. N. Avvakumov
HPV 16 E7 T72D	pBAIT	2907	Sean Jewell

HPV 16 E7 H73E	pBAIT	2902	Sean Jewell
HPV 16 E7 V74T	pBAIT	2906	Sean Jewell
HPV 16 E7 I76A	pBAIT	3117	Dr. N. Avvakumov
HPV 16 E7 R77E	pBAIT	2905	Sean Jewell
HPV 16 E7 L79A	pBAIT	3118	Dr. N. Avvakumov
HPV 16 E7 E80K/D81K	pBAIT	2918	This study
HPV 16 E7 L82A/L83A/M84A/G85D	pBAIT	3123	Dr. N. Avvakumov
HPV 16 E7 M84S	pBAIT	2882	Sean Jewell
HPV 16 E7 G85A	pBAIT	2917	Dr. Joe Mymryk
HPV 16 E7 T86D	pBAIT	2880	Sean Jewell
HPV 16 E7 L87A	pBAIT	3119	Dr. N. Avvakumov
HPV 16 E7 I89A/V90A	pBAIT	3120	Dr. N. Avvakumov
HPV 16 E7 C91G	pBAIT	3055	This study
HPV 16 E7 P92A	pBAIT	2916	Dr. Joe Mymryk
HPV 16 E7 I93T	pBAIT	2878	Sean Jewell
HPV 16 E7 Q96E/K97E/P98A	pBAIT	2877	Sean Jewell
HPV 16 E7 (wild-type)	pJG4-5+	893	Dr. N. Avvakumov
HPV 16 E7 H2P	pJG4-5+	1496	Dr. N. Avvakumov
HPV 16 E7 Δ PTLHE 6-10	pJG4-5+	323	Dr. N. Avvakumov
HPV 16 E7 M12V	pJG4-5+	522	Dr. N. Avvakumov
HPV 16 E7 Δ DLYC 21-24	pJG4-5+	324	Dr. N. Avvakumov
HPV 16 E7 S31G/S32G	pJG4-5+	325	Dr. N. Avvakumov
HPV 16 E7 E46Q/D48N	pJG4-5+	3124	This study
HPV 16 E7 Y52A	pJG4-5+	3099	This study
HPV 16 E7 N53D	pJG4-5+	3100	This study
HPV 16 E7 V55T	pJG4-5+	3101	This study
HPV 16 E7 F57A	pJG4-5+	3083	This study
HPV 16 E7 C58G/C91G	pJG4-5+	455	Dr. N. Avvakumov
HPV 16 E7 HPV 16 E7 C59S	pJG4-5+	3082	This study
HPV 16 E7 K60E	pJG4-5+	3081	This study

HPV 16 E7 D62K	pJG4-5+	3080	This study
HPV 16 E7 S63D	pJG4-5+	3072	This study
HPV 16 E7 T64D	pJG4-5+	3102	This study
HPV 16 E7 L65A	pJG4-5+	1203	Dr. N. Avvakumov
HPV 16 E7 R66E	pJG4-5+	3096	This study
HPV 16 E7 L67R	pJG4-5+	556	Dr. N. Avvakumov
HPV 16 E7 V69A	pJG4-5+	1202	Dr. N. Avvakumov
HPV 16 E7 S71I	pJG4-5+	3125	This study
HPV 16 E7 T72D	pJG4-5+	3075	This study
HPV 16 E7 H73E	pJG4-5+	3071	This study
HPV 16 E7 V74T	pJG4-5+	3074	This study
HPV 16 E7 I76A	pJG4-5+	1201	Dr. N. Avvakumov
HPV 16 E7 R77E	pJG4-5+	3073	This study
HPV 16 E7 L79A	pJG4-5+	1200	Dr. N. Avvakumov
HPV 16 E7 E80K/D81K	pJG4-5+	3079	This study
HPV 16 E7 L82A/L83A/M84A/G85D	pJG4-5+	3126	This study
HPV 16 E7 M84S	pJG4-5+	3070	This study
HPV 16 E7 G85A	pJG4-5+	3078	This study
HPV 16 E7 T86D	pJG4-5+	3069	Sean Jewell
HPV 16 E7 L87A	pJG4-5+	1198	Dr. N. Avvakumov
HPV 16 E7 I89A/V90A	pJG4-5+	1197	Dr. N. Avvakumov
HPV 16 E7 C91G	pJG4-5+	454	Dr. N. Avvakumov
HPV 16 E7 P92A	pJG4-5+	3077	This study
HPV 16 E7 I93T	pJG4-5+	3098	This study
HPV 16 E7 Q96E/K97E/P98A	pJG4-5+	3097	This study
pCAF (residues 310-732)	pBAIT	732	Dr. N. Avvakumov
8opLexA-LacZ	pSH18-34	313	Dr. M. Shuen

*JMB = Joe Mymryk Bacteria database

Two methods of site-directed mutagenesis were used to mutate the predicted solvent exposed residues in CR3 of full-length HPV16 E7. The megaprimer method of site-directed mutagenesis involves two rounds of PCR and uses three oligonucleotide primers (Figure 2.1) (Kammann *et al.*, 1989). One of the primers is mutagenic, and the other two are forward and reverse primers that lie upstream and downstream from the binding site for the mutagenic primer. The mutagenic primer and one of the flanking primers (oriented in the opposite direction) are used in the first PCR reaction to generate and amplify a mutated fragment of DNA. This mutagenic fragment of DNA, called the megaprimer, is used in the second PCR reaction alongside both the forward and reverse flanking primers to amplify a longer region of the template DNA. The second method of creating site-specific mutations involves a step of overlap extension (Figure 2.2) (Ho *et al.*, 1989). In two separate PCR reactions, the wild-type E7 gene is amplified with the mutation of interest. This method requires two sets of forward and reverse primers. One set will contain the mutation of interest and the other set will have sequences complementary to the ends of the initial fragments. After the first round of PCR, the overlapping DNA fragments will each contain the mutation of interest in the overlapping region. The products from the two different reactions are then mixed, denatured, annealed and extended to form full-length mutant DNA. The mutant DNA is then amplified using primers that bind to the extreme ends of the two initial fragments.

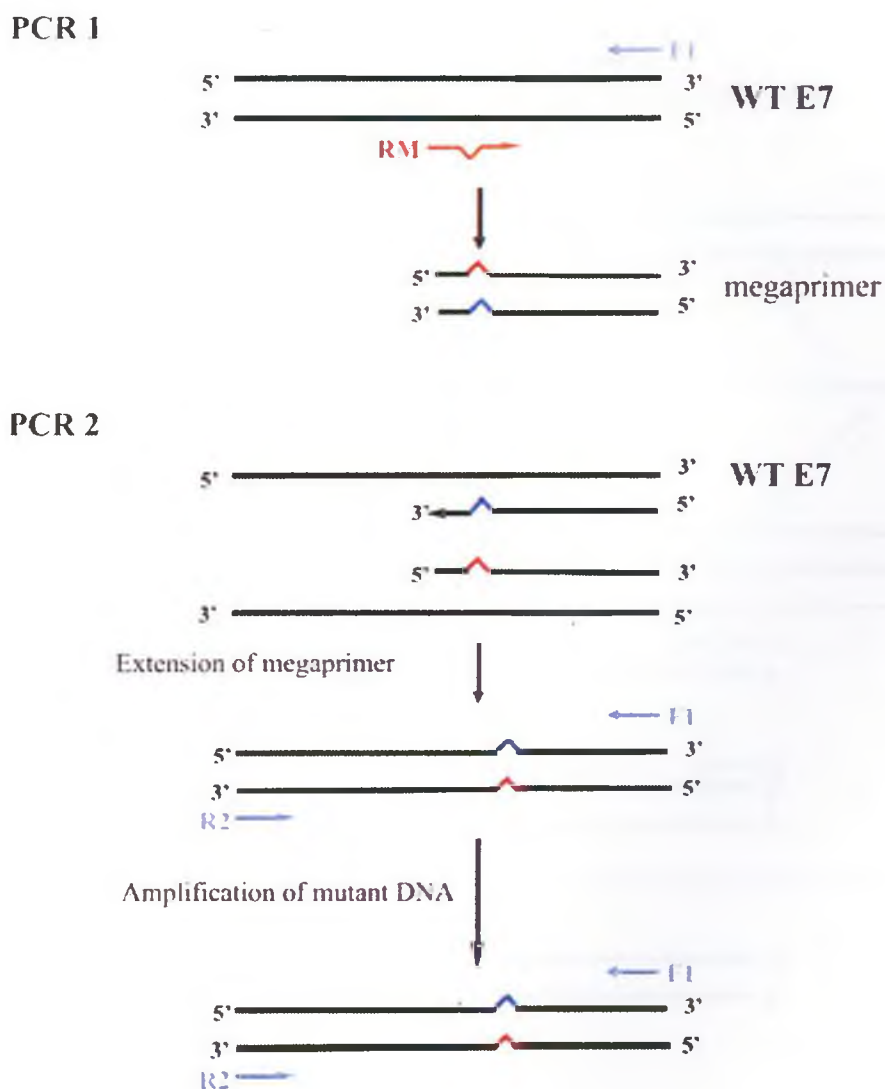


Figure 2.1. Site-directed mutagenesis by the megaprimer method. The mutation of interest is highlighted in blue on one strand and red on the complementary strand. Three different primers are used: 1) A mutagenic primer containing the mutation of interest (RM); and 2) two wild-type primers with binding sites at the ends of the E7 sequence. This approach involves two initial rounds of PCR. In the first reaction, a megaprimer is created. This megaprimer is then used in the second round of PCR, alongside the wild-type primers to create a mutation in the full-length sequence of E7. In subsequent cycles, there is a preferential amplification of the mutant DNA using the forward and reverse primers.

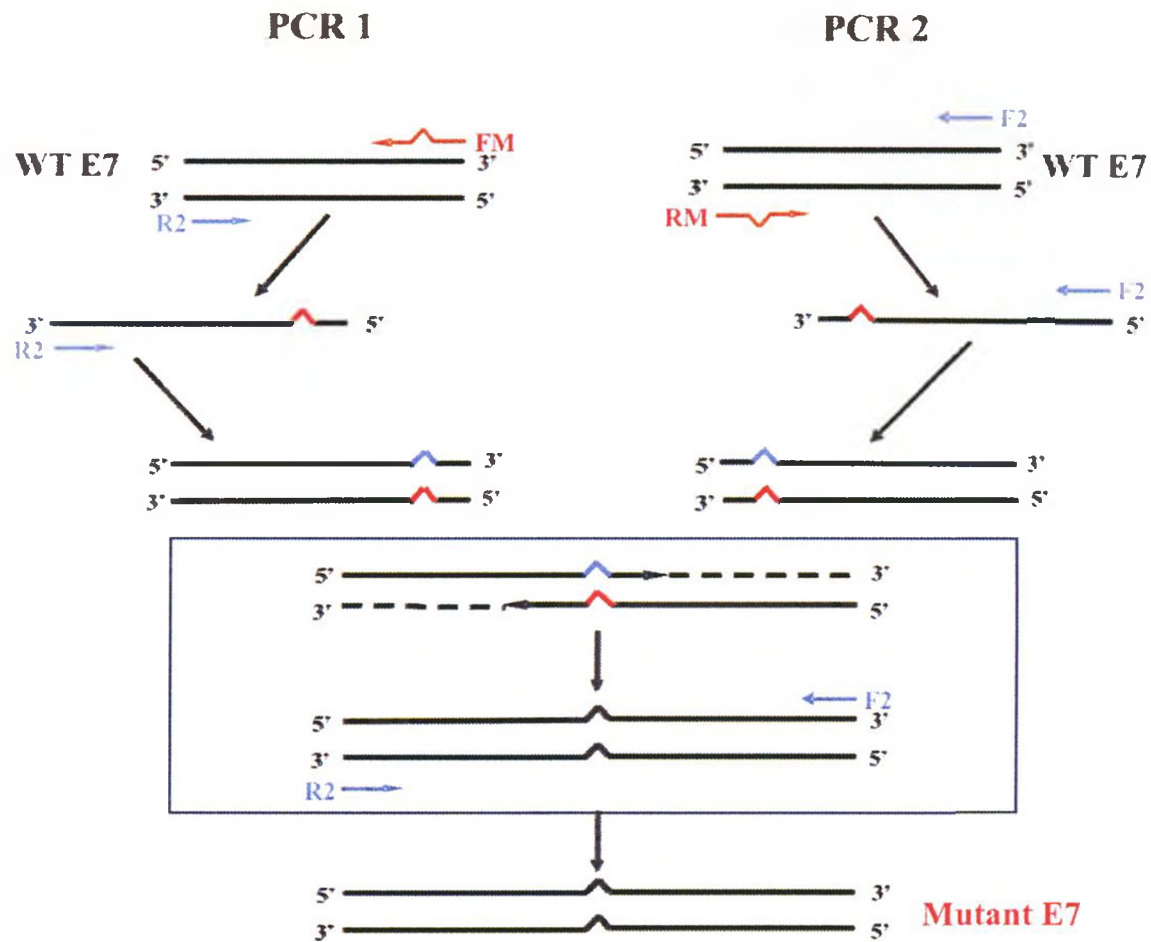


Figure 2.2. Site-directed mutagenesis by overlap extension. The mutation of interest is highlighted in blue on one strand and red on the complementary strand. Two sets of primers are used: 1) mutagenic primers containing the mutation of interest (FM, RM); and 2) wild-type primers that bind to the ends of each E7 sequence (F2, R2). This approach involves two initial rounds of PCR to produce two overlapping DNA fragments, each containing the mutation of interest in the overlapping region. The mutant is then amplified in a final PCR reaction using wild-type primers that bind to the extreme ends of the two initial PCR templates.

The panel of solvent exposed CR3 mutants were generated with flanking *EcoRI* and *SalI* restriction sites and inserted into the multiple cloning sties (MCS) of pBAIT (Figure 2.3), which contains a *LEU2* selectable marker. Proteins expressed from this vector are fused to a LexA DNA binding domain (DBD). Table 2.3 illustrates the oligonucleotides used in the construction of the E7 CR3 mutants in this study. All PCR products were sequenced at the York University Core Molecular Biology and DNA Sequencing Facility (Toronto, Ontario). For yeast two-hybrid assays, all E7 mutants were excised using *EcoRI* and *SalI* restriction enzymes and were inserted into the same sites of pJG4-5+, which contains the *TRP1* selectable marker. Proteins expressed from this vector are fused to the synthetic B42 transcriptional activation domain (AD).

Table 2.3 List of E7 mutants produced in this study and oligonucleotides utilized in PCR

Construct in pBAIT (JMB#)	Oligos (5' – 3')
	Forward (F, JMO #)
	Reverse (R, JMO #)
HPV16 E7 (wild-type)	GTCGAATTCATGCATGGAGATACACCTACATTGC (F, JMO 680) CGTGTCGACTTATGGTTTTTGTGAGAACAGATGGGG (R, JMO 713)
HPV16 Y52A	GAACCGGACAGAGCCCATGCCAATATTGTAACCTTTTGTTG C (F, JMO 722) GCAACAAAAGGTTACAATATTGGCATGGGCTCTGTCCGGTT C (R, JMO 702)
HPV16 N53D	CCGGACAGAGCCCATTACGATATTGTAACCTTTTGTTGCAAG (F, JMO 719) CTTGCAACAAAAGGTTACAATATCGTAATGGGCTCTGTCCG G (R, JMO 701)

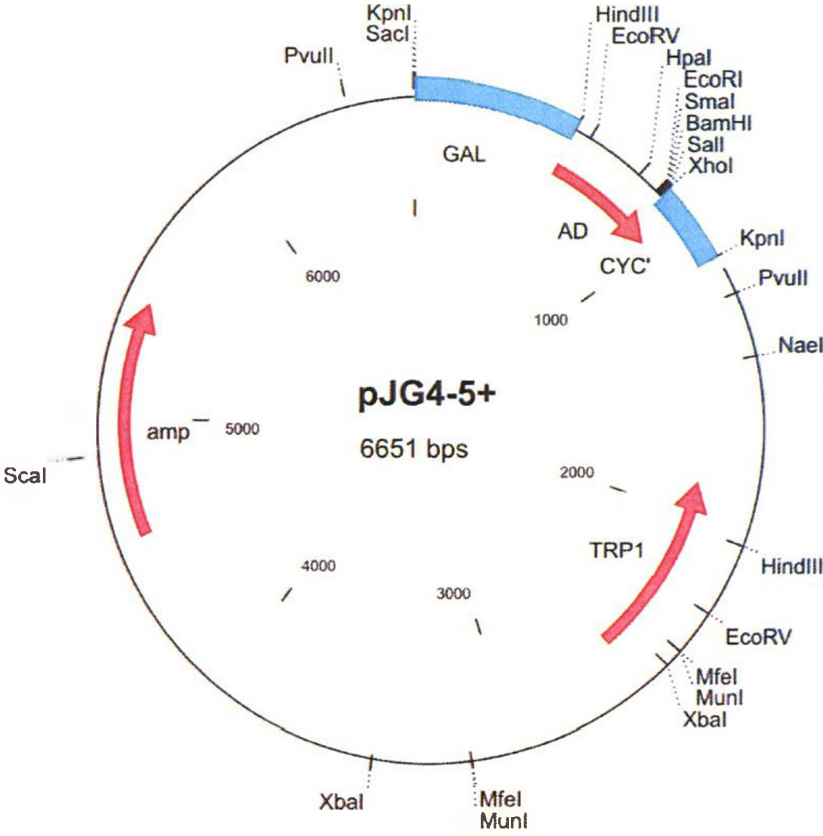
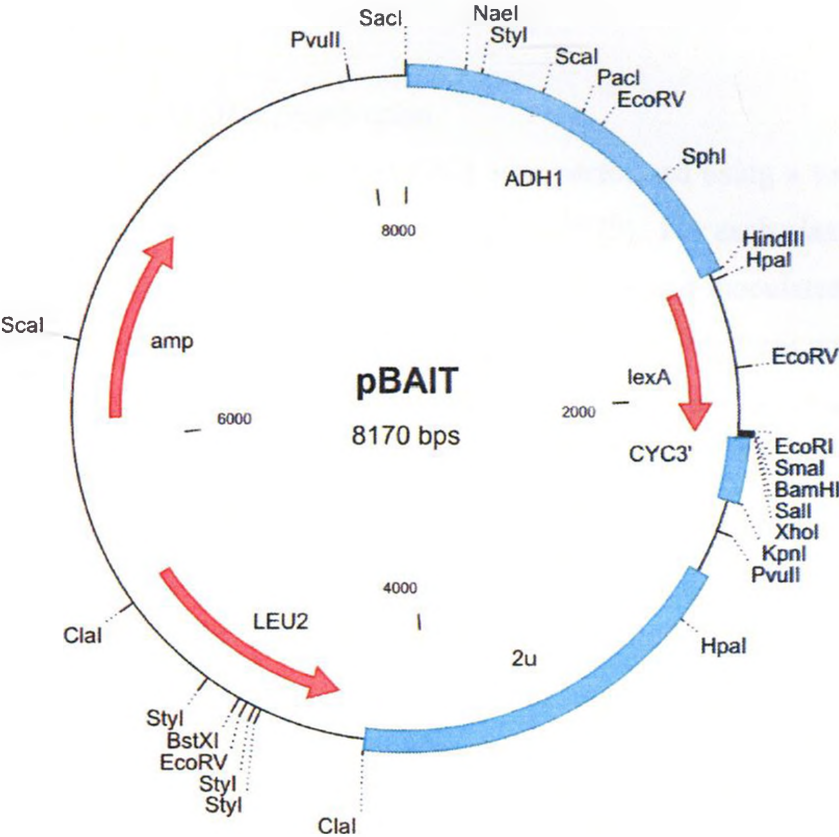
HPV16 V55T	GGACAGAGCCCATTACAATATTACAACCTTTTGTGCAAGT GTGAC (F, JMO 720) GTCACACTTGCAACAAAAGGTTGTAATATTGTAATGGGCTC TGTCC (R, JMO 700)
HPV16 F57A	GAGTCACACTTGCAACAAGCGGTTACAATATTGTAATGGGC (JMO 699)
HPV16 C59S	CCCATTACAATATTGTAACCTTTTGTAGCAAGTGTGACTCTA CGCTT (F, JMO 721) GAAGCGTAGAGTCACACTTGCTACAAAAGGTTACAATATTG TAATGGG (R, JMO 698)
HPV16 K60E	CCGAAGCGTAGAGTCACACTCACAACAAAAGGTTACAATAT TG (R, JMO 697)
HPV16 D62K	GTAACCTTTTGTGCAAGTGTAATCTACGCTTCGGTTGTGC (F, JMO 718) GCACAACCGAAGCGTAGATTTACACTTGCAACAAAAGGTTA C (R, JMO 696)
HPV16 S63D	GCACAACCGAAGCGTATCGTCACACTTGCAACAAAAGG (R, JMO 687)
HPV16 T64D	GCTTTGTACGCACAACCGAAGATCAGAGTCACACTTGCAAC (JMO 695)
HPV16 R66E	GCAAGTGTGACTCTACGCTTGAGTTGTGTGTGCAAAGC (F, JMO 723) GCTTTGCACACACAACTCAAGCGTAGAGTCACACTTGC (R, JMO 686)
HPV16 T72D	CCAAAGTACGAATGTCTACGTGGTCGCTTTGTACGCACAAC C (R, JMO 694)
HPV16 H73E	CCAAAGTACGAATGTCTACTTCTGTGCTTTGTACGC (R, JMO 685)

HPV16 V74T	GGTCTTCCAAAGTACGAATGTCTGTGTGTGTGCTTTGTACGC (R, JMO 693)
HPV16 R77E	CCCATTAACAGGTCTTCCAAAGTCTCAATGTCTACGTGTGTG C (R, JMO 692)
HPV16 E80K/D81K	CGTACTTTGAAAAAGCTGTTAATGGGC (F, JMO 724) GCCCATTAACAGCTTTTTCAAAGTACG (R, JMO 714)
HPV16 M84S	CGTGTCGACTTATGGTTTTTGAGAACAGATAGGACACACAA TGCCTAGCGTGCCACTTAACAGGTCTTCC (R, JMO 690)
HPV16 G85A	GCTGTCGACTTATGGTTTTTGAGAACAGATAGGACACACAA TGCCTAGCGTAGCCATTAACAGG (R, JMO 689)
HPV16 T86D	GCTGTCGACTTATGGTTTTTGAGAACAGATAGGACACACAA TGCCTAGATCGCCCATTAACAGG (R, JMO 688)
HPV16 P92A	CGTGTCGACTTATGGTTTTTGAGAACAGATGGCGCACAC (R, JMO 684)
HPV16 I93T	CGTGTCGACTTATGGTTTTTGAGAACAGGTGGGGCACAC (R, JMO 683)
HPV16 QKP96-98EEA	CGTGTCGACTTATGCCTCTTCAGAACAGATGGGG (R, JMO 682)

* JMO = Joe Mymryk Oligonucleotide

Mutants in bold were created by the megaprimer method. Unbolded mutants were generated by overlap extension.

Figure 2.3. Maps for yeast expression vectors pBAIT and pJG4-5+. In both vectors, the gene encoding full-length HPV16 E7 was inserted between the *EcoRI* and *SalI* restriction sites within the multiple cloning site (MCS) to generate either LexA-DBD or B42-AD fusions. Transcription from the pBAIT and pJG4-5+ parent vector is regulated by the ADH1 and the GAL1 promoter respectively. Both vectors also contain a 2-micron (2u) origin of replication to allow selection and propagation in yeast; and the pUC origin (not shown) and the ampicillin resistance gene (amp) to allow propagation and selection in *E.coli*. The pBAIT and pJG4-5+ vectors contain a *LEU2* and *TRP1* selectable marker, respectively. The pJG4-5+ vector also contains an HA epitope (not shown).



2.3 DNA preparation

2.3.1 *Small scale plasmid DNA preparation*

Small scale preparation of plasmid DNA was performed using a variation of the original technique described by Birnboim and Doly (1979). For each plasmid isolation, a single bacterial colony of transformed *E.coli* was isolated and inoculated into 5mL of LB with 0.5mg/mL ampicillin (Bioshop, Burlington, ON, CA). Cultures were grown 16-18 hours at 37°C with agitation. A 1.5mL aliquot of each culture was transferred to a microcentrifuge tube, and the bacteria were pelleted by centrifugation for 30 seconds at 14,000g at room temperature. The supernatant was aspirated, and the pellet was resuspended with 100µL of Solution I (see section 2.7 for composition of all buffers) by vortexing for 10 seconds. Two hundred microliters of Solution II was added, the sample was mixed by inversion and then incubated on ice for 5 minutes. One hundred and fifty microliters of Solution III was added, and vortexed for 10 seconds. Samples were then incubated for 5 minutes on ice and centrifuged at room temperature at 14,000g for 5 minutes to pellet debris. The supernatant was transferred to another microcentrifuge tube containing 300µL of a 1:1 mix of phenol (Sigma, St. Louis, MO, USA) and chloroform (Bioshop, Burlington, ON, CA). Samples were vortexed for 30 seconds and spun at 14,000g for 5 minutes. The upper phase of each sample was transferred to a fresh microcentrifuge. DNA was then precipitated with the addition of 700µL of 100% ethanol, and pelleted at 14,000g for 10 min at 4°C. The DNA pellet was then rinsed with ice-cold 70% ethanol to remove remaining salts. The dried pellet was dissolved in 30µL of sterile distilled water. The samples were frozen at -20°C until required.

2.3.2 *Screening potential recombinant plasmids with restriction digestion analysis*

Five microliter aliquots of the small scale plasmid DNA preparations was digested using the appropriate restriction endonucleases and buffers using the conditions specified by the suppliers (New England Biolabs, Mississauga, ON, CA; and Invitrogen, Carlsbad, CA, USA). Thirty microliter reactions were prepared containing 0.3µg/ µL RNase A (Bioshop, Burlington, ON, CA) and analyzed by horizontal 1-D gel electrophoresis. Agarose gels contained 0.5µg/mL ethidium bromide in 1x TBE buffer

and ranged from 1-2% agarose depending on the desired separation characteristics. After electrophoresis, the gels were viewed under ultraviolet light and photographed using a Foto/Eclipse Fotodyne “gel doc system” (Fotodyne Inc.: Havard, Wisconsin, USA).

2.3.3 *Large scale plasmid DNA preparation*

Large scale preparations of plasmid DNA were performed using a variation of the method of Birnboim and Doly (1979). Two hundred microliters of bacterial culture containing the plasmid of interest was used to inoculate 250mL of LB containing 0.5mg/mL ampicillin (Bioshop, Burlington, ON, CA) in 500mL flasks. Cultures were grown 16-18 hours at 37°C with agitation. Bacteria were pelleted by centrifugation at 4,000g for 15 minutes at 4°C and the supernatant was discarded. The pellets were resuspended in 5mL of Solution I and then transferred to 50mL oakridge tubes. 10mL of freshly made Solution II was added and samples were gently mixed by inversion six times and incubated at room temperature for 10 minutes. Following incubation, 7.5mL of cold Solution III was added to each preparation and mixed vigorously by shaking. Samples were incubated on ice for 10 minutes, then centrifuged at 14,500g at 4°C for 15 minutes. Supernatants were filtered out using cheesecloth into new oakridge tubes containing isopropanol (Bioshop, Burlington, ON, CA) to precipitate DNA. Samples were centrifuged again at 10,500g at 4°C for 15 minutes to pellet DNA. DNA pellets were resuspended with 600µL STE and transferred to 1.5mL microcentrifuge tubes. 10µg/µL of RNase A (Sigma, St. Louis, MO, USA) was added, and the samples were vortexed and set at 65°C for 10 minutes. Samples were extracted twice with 500µL of 1:1 mixture of phenol (Sigma, St. Louis, MO, USA) and chloroform (Bioshop, Burlington, ON, CA) and then once with 500µL chloroform. For each extraction, samples were vigorously mixed for 30 seconds after the addition of phenol:chloroform, centrifuged at 14,000g for 5 minutes at room temperature, with the final upper phase transferred to fresh centrifuge tubes. After the last extraction, 1mL of 100% ice-cold ethanol was added to precipitate DNA. The samples were then centrifuged at 14,000g at 4°C for 10 minutes. The DNA pellet was rinsed with ice-cold 70% ethanol and air dried.

The DNA pellet was re-suspended in 200 μ L of sterile distilled water and the samples were frozen at -20°C until needed.

2.4 Yeast transformation

Yeast transformations were performed using a modified lithium acetate procedure (Gietz *et al.*, 1995). Yeast cultures were inoculated into 5mL of YEP supplemented with 5% glucose and selective media, and grown overnight at 30°C with agitation. 1mL of grown culture was inoculated into a 20mL of fresh media and growth for another 2-4 hours to obtain cells at mid-log phase. A 1.5mL aliquot of culture was transferred to microcentrifuge tubes and pelleted at 14,000g for 10 seconds. After supernatant was discarded, cells were washed with 1mL, washed with 1mL of sterile distilled water, pelleted, re-suspended in 1mL of 100mM lithium acetate (LiAc) and incubated at 30°C for 5 minutes. The cells were pelleted at 14,000g for 8 seconds, and the following components were added into each sample in order: 240 μ L of 50% polyethylene glycol (PEG), 36 μ L of 1M LiAc, 25 μ L of 2mg/mL of salmon sperm DNA, 2-4 μ L of each desired plasmid and 40 μ L of sterile distilled water. The pellet was re-suspended by vigorously mixing for 1 minute, and the cell suspension was incubated at 42°C for 30 minutes. Cells were then pelleted at 14,000g for 8 seconds, and the supernatant was discarded. The pellet was re-suspended in 200 μ L of sterile distilled water, plated on selective SC omission plates and incubated at 30°C for 2-4 days.

2.5 Liquid β -galactosidase assay in yeast

This “quick” method of assaying for β -galactosidase activity from yeast is described by Adam *et al.*, 1998. Transformed yeast colonies were inoculated in SC selection media overnight at 30°C with agitation. A 1.5mL aliquot of grown culture was centrifuged at 14,000g to pellet the cells. The cell pellet was re-suspended in 1mL of Z-buffer containing 2.7 μ L/mL of β -mercaptoethanol (BME). The absorbance of each sample was measured at 600nm to obtain the cell density. 300 μ L of cells with an average

OD₆₀₀ reading of 0.5-1 were utilized for each assay. 300µL triplicates of each sample was added to 700µL of Z-buffer (+ BME). In addition, 40µL of 0.1%SDS and 20µL of chloroform were added to the sample mixture. After vigorously mixing for 30 seconds, samples were placed in the 30°C water bath for 15 minutes. After incubation, 200 µL of *ortho*-Nitrophenyl-β-galactoside (ONPG; Bioshop, Burlington, ON, CA) stock 4mg/mL solution was added to each sample and vigorously mixed. The samples were placed in the 30°C water bath until they turned pale yellow, at which time the reaction was terminated with 0.4mL of 1M Na₂CO₃ and the reaction time recorded. The samples were centrifuged for 10 minutes at 14,000g at room temperature to pellet cell debris, and 200µL of each sample was used to measure absorbance at 420nm. The β-galactosidase activity was calculated in Miller units utilizing the following formula:

$$= \frac{\text{OD}_{420}}{\text{OD}_{600} \times 0.3\text{mL (culture volume)} \times \text{reaction time (minutes)}}$$

The experiments performed via this quick method were done separately twice from independent yeast transformation and performed in triplicate. The error was calculated as standard deviation and the student t-test was performed, using Sigma Plot, with a 95% confidence interval ($p < 0.05$).

2.6 Yeast cell protein extraction and Western blot analysis

Yeast colonies were isolated from selection plates and inoculated in 5mL of selective SC liquid medium. Cultures were grown at 30°C with agitation overnight. Each 5mL culture was collected in a 15mL conical tube, and cells were pelleted by centrifugation for 5 minutes at 4,500g. The cell pellet was resuspended in 250µL of 0.5% NP-40 Lysis Buffer with complete protease inhibitor cocktail (Sigma, St. Louis, Missouri, USA). An equal volume of acid-washed glass beads (Sigma, St. Louis, MO, USA) was added to each sample. Samples were vigorously mixed in cycles of 3 minutes, and then placed on ice for 3 minutes, three times. Protein concentrations of the lysates were measured using the Bradford DC Protein assay kit (BioRad, Hercules, CA, USA). 200µL of the supernatant was transferred to clean, chilled 1.5 mL microcentrifuge tubes.

Twenty-five micrograms of total protein from each sample was prepared in 4X loading dye with 0.1M DTT. The samples were heated at 100°C for 10 minutes and resolved on Novex pre-cast 5-20% gradient Tris-Glycine Polyacrylamide gel (Invitrogen, Carlsbad, CA, USA). Transfer was performed onto a polyvinylidene fluoride (PVDF; Amersham, Piscataway, HJ, USA) membrane according to suppliers' recommendations. The membrane was then incubated with Tris Buffered Saline containing 0.1% Tween-20 (TBS-T) and 5% (w/v) Carnation non-fat dry milk for 1 hour at room temperature. The blot was then incubated with primary antibody resuspended in TBS-T +5% milk, for 1 hour at room temperature or overnight while shaking. The membrane was washed with TBS-T 3 times for 10 minutes each, to remove unbound primary antibody. The samples containing LexA DBD fusions were blotted with rabbit anti-LexA (Upstate, Charlottesville, VA, USA) at 1:500 dilution. Samples containing the B42 AD fusions were blotted with rat anti-HA (Hoffmann-La Roche, Mississauga, ON, CA) After washing, the membrane was treated with secondary antibody conjugated to horseradish peroxidase re-suspended in TBS-T + 5% milk for 30 minutes at room temperature while shaking. For the anti-LexA primary antibody, the goat anti-rabbit IgG secondary conjugated to horseradish peroxidase (HRP), was utilized (Amersham, Piscataway, HJ, USA). For the anti-HA primary antibody, the goat anti-rat IgG-HRP secondary was utilized (Amersham, Piscataway, HJ, USA). The bands were detected using the Amersham ECL kit (Piscataway, HJ, USA). Images were developed using a Kodak M35A X-OMAT automatic film processor and viewed on Kodak film (Amersham, Piscataway, HJ, USA). To ensure equal loading, membranes were stained with Ponceau S (Sigma, St. Louis, MO, USA) for 15 minutes, destained and dried.

2.7 Solutions

4X Loading Dye

250mM Tris-HCl pH 6.8
40% (v/v) glycerol
8% (w/v) SDS
20% (v/v) β -mercaptoethanol
0.2% (w/v) bromophenol blue

0.5% NP-40 Lysis Buffer

0.5% NP-40
150mM NaCl
50mM Tris-Cl pH 7.8

Luria Bertani Broth (LB)

0.5% (w/v) yeast extract
85mM NaCl
1% (w/v) tryptone
 \pm 0.5mg/ml ampicillin
 \pm 2% (w/v) agar
Sterilized by autoclaving

ortho-Nitrophenyl- β -galactoside (ONPG)

4mg/mL stock solution

Synthetic Complete (SC) Medium

5% (w/v) ammonium sulphate
1.7% (w/v) yeast nitrogen base (without amino acids)
0.02% (w/v) adenine sulphate
0.02% (w/v) arginine
0.03% (w/v) methioine
0.03% (w/v) tyrosine
0.03% (w/v) isoleucine
0.05% (w/v) phenylalanine
0.10% (w/v) glutamic acid
0.10% (w/v) aspartic acid
0.15% (w/v) valine
0.20% (w/v) threonine
0.37% (w/v) serine
 \pm 0.01% (w/v) leucine
 \pm 0.02% (w/v) tryptophan
 \pm 0.01% (w/v) lysine
 \pm 0.01% (w/v) histidine
 \pm 0.01% (w/v) uracil
 \pm 5% (w/v) glucose
 \pm 2% (w/v) agar

Adjusted to pH 5.5 and sterilized by autoclaving.

Solution I

50mM glucose
25mM Tris-HCl
10mM EDTA (pH 8.0)

Solution II

0.2M NaOH
1% (w/v) SDS

Solution III

3M Potassium Acetate
0.2% (v/v) acetic acid

Salmon sperm DNA (SS-DNA)

2mg/ml of deoxyribonucleic acid (DNA) sodium salt type III from salmon testes
10mM Tris-HCl pH 8.0
1mM EDTA

Salt Tris, EDTA (STE) Buffer

100mM NaCl
10mM Tris-HCl pH 8.0
1mM EDTA
Sterilized by autoclaving

Tris, Borate, EDTA (TBE) 10x

5.5% (w/v) boric acid
0.58% (w/v) EDTA
10.8% (w/v) Tris Base
Sterilized by autoclaving

Yeast peptone (YEP) medium

20% (w/v) peptone
10% (w/v) yeast extract
±2% (w/v) D-glucose
± 2% (w/v) agar
Sterilized by autoclaving

Z-Buffer

60mM Na_2PO_4
40mM NaH_2PO_4
10mM KCl
1mM MgSO_4
Adjusted to pH 7.0

CHAPTER 3: RESULTS

3.1 Development of a model for HPV16 E7 CR3

Using information obtained from the most recent structural studies of HPV E7 CR3, our laboratory has generated a structural model of CR3 for the HPV16 E7 protein, with the assistance of Dr. Gary Shaw (Department of Biochemistry, University of Western Ontario). Upon initial analysis of the dimeric CR3 structures published for HPV45 (PDB ID: 2F8B) and HPV1 (PDB ID: 2B9D), it was found that the two structures superimpose to a backbone root mean square (RMS) deviation of 1.61 Å, which is comparable to the RMS value of 1.4 Å reported by Ochleschlager, *et al.* Subsequent studies focused on the HPV45 dimeric structure of CR3 which served as a template in the prediction of a three-dimensional structure for this region in HPV16. The HPV45 structure was chosen as a starting point instead of the HPV1 structure for several reasons. Firstly, both the high-risk HPV45 and HPV16 belong to the supergroup A family of HPVs and are commonly associated with human cancers. Secondly, CR3 of the E7 protein of HPV16 has approximately 40% sequence identity and 71% sequence similarity to the E7-CR3 of HPV45, suggesting that significant structural similarity may exist between the two proteins. In contrast, the CR3 region of the non-oncogenic HPV1 E7 protein has the 40% sequence identity and 60% sequence similarity to HPV16 E7-CR3. Furthermore, there are no gaps in the sequence alignment of CR3 from HPV45 and HPV16 E7. Based on these criteria, a homology modeling method was chosen, as it can produce high-quality structural models when the target (HPV16) and the template (HPV45) are closely related (Baker & Sali, 2001).

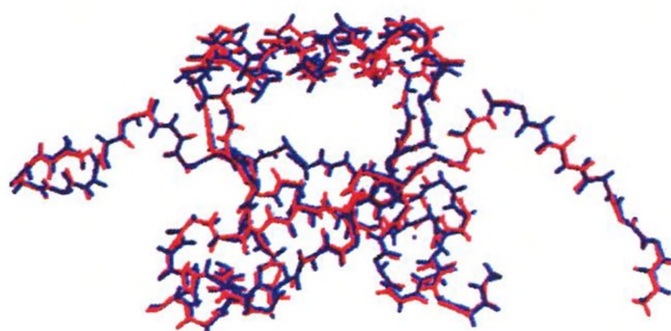
The CR3 dimer of HPV45 was then analyzed by the web server program VADAR (Volume, Area, Dihedral Angle Reporter; <http://redpoll.pharmacy.ualberta.ca/vadar/>). VADAR is a compilation of more than 15 different algorithms and programs for analyzing and assessing peptide and protein structures from the atomic coordinate data, also known as protein data bank (PDB) files (Willard *et al.*, 2003). These programs calculate and evaluate a large number (>30) of key structural parameters of the proteins as well as provide specific descriptions for individual atoms, side chains, backbone and residues, making it possible to quantitatively describe protein structures. VADAR is

specifically designed to assess protein structures determined by X-ray crystallography, NMR spectroscopy, 3D-threading or homology modelling. The NMR structure of the C-terminal domain of the HPV45 E7 dimer was analyzed to determine the accessible surface area of the protein and the chi angles (rotation of the R group substituent around the α -carbon and β -carbon). Appropriate residues in HPV45 were then swapped to match those in HPV16. Oftentimes, the quality of the homology model is complicated by the presence of alignment gaps, indicating that a structural region may be present in the target but not in the template, and vice-versa. The lack of gaps in CR3 in the alignment between HPV45 and HPV16 E7 proteins made the direct substitution of residues possible. The chi angles in the HPV16 model were then manually adjusted to superimpose with those in HPV45. This was followed by a series of energy minimization steps that were applied to the model to refine its structure. Atomic overlaps and unnatural strains in the structure were removed, and strong hydrogen bonds were reinforced while weak ones were broken. This was all done while keeping the critical zinc ions locked in place. Because VADAR accepts either a PDB accession number (for previously determined structures) or a PDB formatted file (for newly determined structures) as input, the structure of HPV16 E7 CR3 could be validated with the template. In particular, the accessible surface areas of the side-chains of residues were re-checked to ensure that residues that were not appreciably exposed to solvent in the HPV45 dimer remained buried. A schematic of the final structural model of CR3 for the E7 protein of HPV16 is shown in Figure 3.1. As expected from the high level of sequence similarity it closely resembles that of HPV45 E7, and the two structures superimpose to a backbone RMS deviation of 0.34 Å.

3.2 Application of the CR3 model to identify targets for site-directed mutagenesis

E7 does not bind DNA nor possesses any intrinsic enzymatic activity. Instead, it performs its numerous functions by interacting with other proteins. Given that the CR3 region of E7 has a specific folded tertiary structure, various surfaces of the folded conformation of E7 are likely to mediate specific interactions with these target proteins.

Figure 3.1. HPV45 and HPV16 E7 CR3 dimers. A) Solution structure of HPV45 determined by NMR spectroscopy (top). Predicted structure of the HPV16 E7 CR3 dimer based on homology modeling to HPV45 (bottom, residues 43-98). Each monomer is coloured in blue and yellow. B) Superimposed carbon backbone of HPV45 (red) and HPV16 (blue) CR3 dimer, illustrating the predicted conserved fold. Images generated by Swiss PDB Viewer.

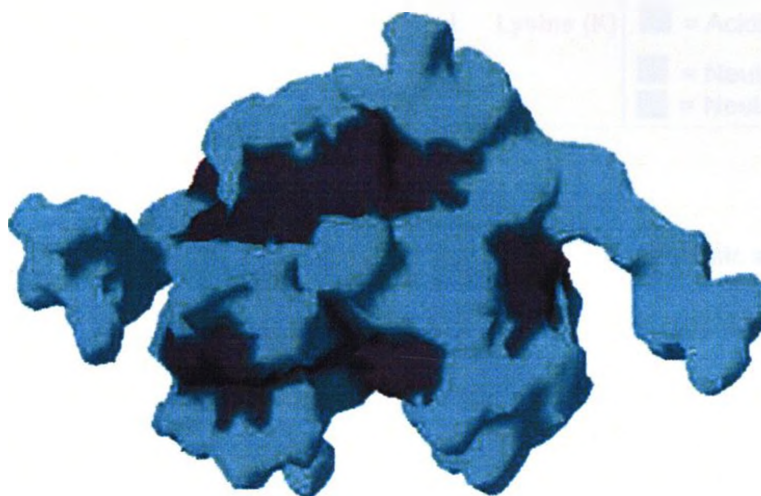
A**HPV 45****HPV 16****B.**

Identification and characterization of these surfaces will contribute to our understanding of E7 function, and these interactions could potentially be targeted for therapy of HPV induced diseases, including cancer. Thus, it was decided that all amino acid residues which are at least 25% surface exposed should be systematically mutated to assess their contribution to E7 function (Figure 3.2). As described above, these residues are likely to be the most accessible for interaction with cellular proteins. In addition, they are not likely to be involved in mediating dimer formation. The construction of such a panel of CR3 mutants will provide an invaluable collection of reagents for HPV research, and allow a detailed structure/function analysis of CR3 to be performed for the first time.

3.2.1 *Construction of CR3 mutants*

The overall goal of the substitution scheme was to mutate residues in such a manner that would retain the folded state of the CR3 monomer and dimer and preserve solvent interactions. The conventional technique for mapping functional epitopes in proteins is alanine-scanning mutagenesis. This site-directed approach systematically removes the side chain atoms past the β -carbon of each targeted amino acid without altering the main-chain conformation and without imposing extreme electrostatic or steric effects (Cunningham & Wells, 1989). In this way, substitution with alanine effectively eliminates a side chain contact and allows for functional epitopes to be mapped in proteins. However, given the structure of CR3, mutating polar or charged amino acids to alanine will quite likely disrupt the interaction of these residues with solvent, and may lead to precipitation or aggregation. Consequently, a mutational strategy was devised in collaboration with Dr. James Omichinski (University of Montréal, Montréal, Québec) for the surface-exposed residues that were to be targeted for mutagenesis. The exact amino acid substitutions that were made in this study are summarized in Table 3.1. In general, basic or acidic amino acids were substituted with residues of opposite charge; and the polarity and size of the side chain group were maintained. In the cases of proline, glycine, phenylalanine and tyrosine, there was no clear substitution choice; and these amino acids were substituted with alanine. PCR mediated site-directed mutagenesis was used to generate 21 CR3 mutations in the full-length sequence of HPV16 E7, and the

A.

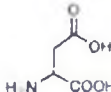
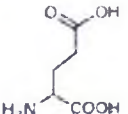
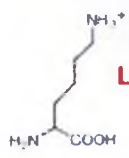
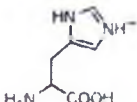
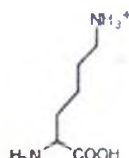
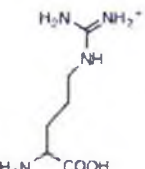
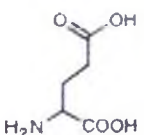
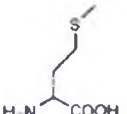
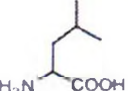
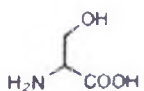
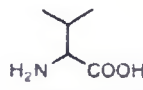
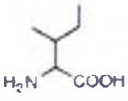
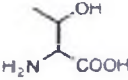
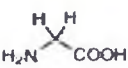
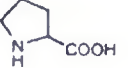
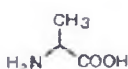
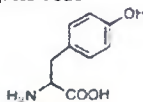
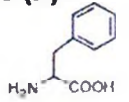
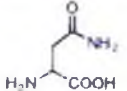
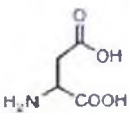
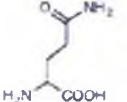
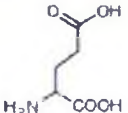
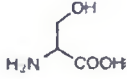
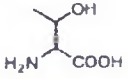
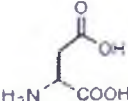
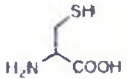
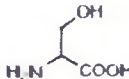


B.



Figure 3.2. Molecular surface of the HPV16 E7 CR3 dimer. A) Highlighted in green are residues that are at least 25% solvent exposed. Residues coloured in black are less than 25% solvent exposed. B) Residues used in construction of the CR3 model are indicated, with the solvent exposed residues mutated by site-directed mutagenesis highlighted in red. These same amino acid residues are highlighted in the sequence alignment of E7 proteins in Figure 1.6.

Table 3.1. Substitution Scheme for CR3 Mutagenesis

HPV16 CR3 Amino Acid		Substitution	
 Aspartic acid (D)	 Glutamic acid (E)	 Lysine (K)	<div> <div></div> = Basic <div></div> = Acidic <div></div> = Neutral, polar <div></div> = Neutral, non-polar </div>
 Histidine (H)	 Lysine (K)	 Arginine (R)	 Glutamic acid (E)
 Methionine (M)	 Leucine (L)	 Serine (S)	
 Valine (V)	 Isoleucine (I)	 Threonine (T)	
 Glycine (G)	 Proline (P)	 Alanine (A)	
 Tyrosine (Y)	 Phenylalanine (F)		
 Asparagine (N)	 Aspartic acid (D)		
 Glutamine (Q)	 Glutamic acid (E)		
 Serine (S)	 Threonine (T)	 Aspartic acid (D)	
 Cysteine (C)	 Serine (S)		

sequence of each mutant was confirmed by the Sequencing Core Facility at York University (Toronto, Ontario). Using the Swiss PDB Viewer, the residues of CR3 targeted for mutagenesis are highlighted on the predicted molecular surface of the dimeric structure of CR3 (Figure 3.3). Additional mutants were already available in the Mymryk laboratory, many of which have been extensively characterized in the literature and were included in my studies. These constructs include mutations in CR1 and CR2, and include the pRb binding mutant (Δ DLYC21-24), the CKII phosphorylation mutant (S31G/S32G), mutations that affect the transformation functions of E7 (H2P, Δ PTLHE6-10); as well as mutations in the highly-conserved hydrophobic residues within CR3, including mutations that disrupt the Cys-X-X-Cys motifs (C58G/C91G and C91G). These are highlighted onto the monomer structure of CR3 (Figure 3.4). A complete list of all the mutations used in my experiments is summarized in Table 3.2.

3.3 Characterizing the collection of E7 mutations

It is possible, that despite our mutational strategy, some of the CR3 mutations may affect the ability of E7 to properly fold and/or dimerize. A number of the other pre-existing mutants in the Mymryk laboratory contain mutations that target the highly conserved small hydrophobic residues that comprise the monomer core or dimer interface. Many of these mutations are expected to perturb folding, which would likely disrupt dimerization, or may directly alter the dimer interface to interfere with dimerization. I initially characterized the complete set of mutants for their ability to dimerize. Mutants unable to dimerize could then be excluded from future protein interaction analysis, as their effects go beyond simple changes on the protein surface. Thus, I could focus my attention on mutations that only influence potential binding surfaces of HPV16 E7 CR3.

Figure 3.3. Predicted molecular surface of the HPV16 E7 CR3 dimer. Each monomer is coloured in blue and yellow. Solvent-exposed residues of CR3 targeted by mutagenesis are coloured in black, grey and white on the yellow monomer, with the visibility of each residue shown in three different views.

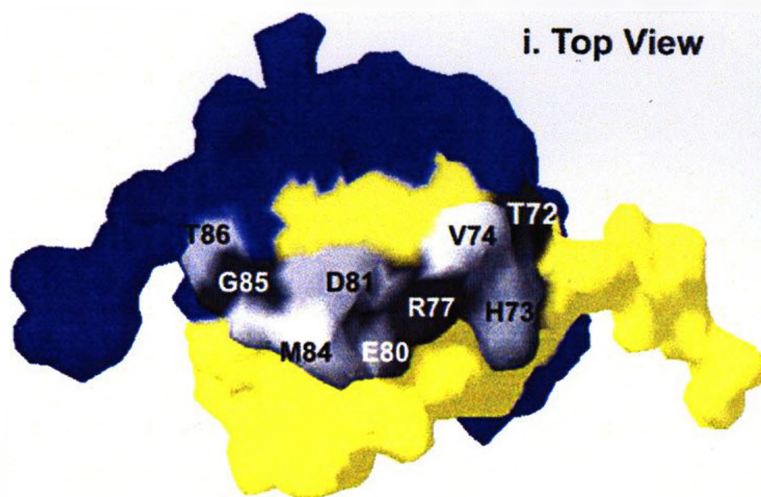
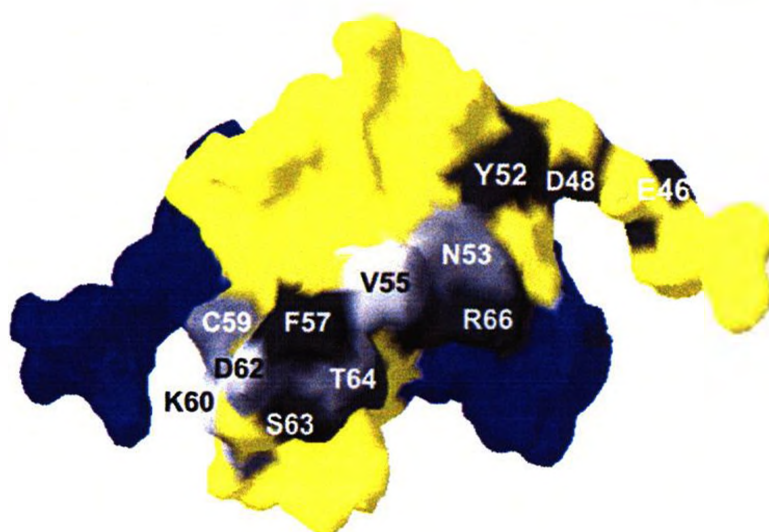
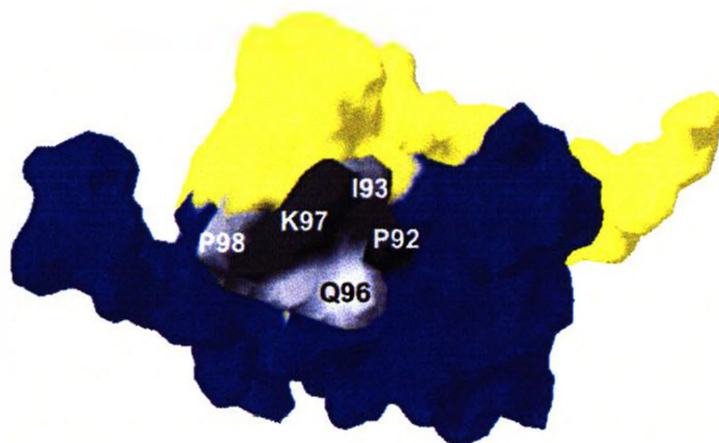
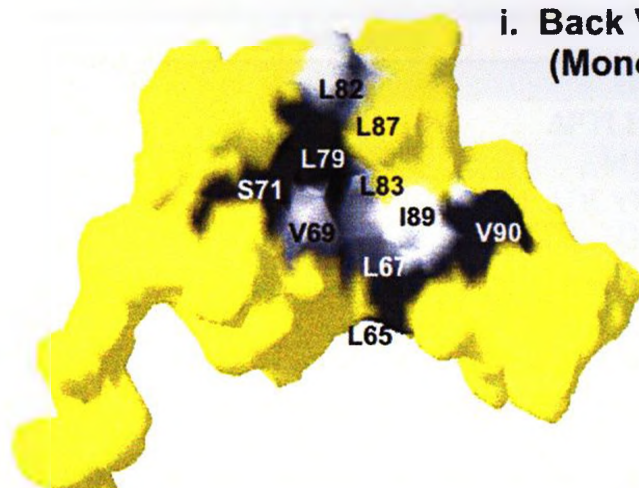
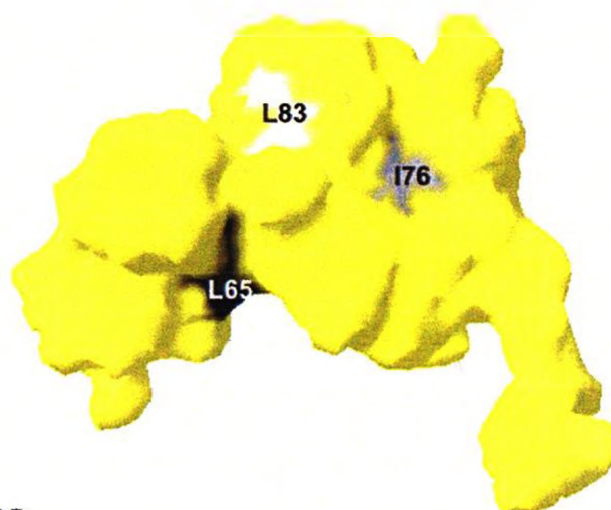
i. Top View**ii. Front View****iii. Bottom View**

Figure 3.4. Predicted molecular surface of HPV16 E7 CR3 monomer. Residues of CR3 targeted by mutagenesis that are predicted to be less than 25% solvent exposed in the dimer, are coloured in black, grey and white. The degree of visibility for each residue is shown in three different views.

**i. Back View
(Monomer Core)**



ii. Front View



iii. Co-ordinating Cysteines

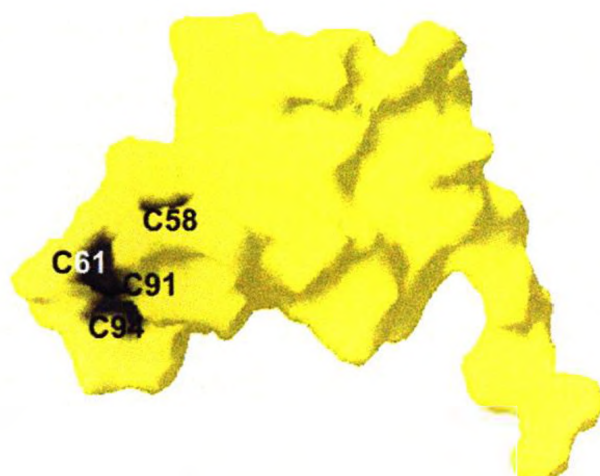


Table 3.2. Collection of E7 Mutants

E7 Mutants	
CR1	H2P
	Δ PTLHE 6-10
	M12V
CR2	Δ DLYC 21-24
	S31G/S32G
	E46Q/D48N
	Y52A
	N53D
CR3	V55T
	F57A
	C58G/C91G
	C59S
	K60E
	D62K
	S63D
	T64D
	L65A
	R66E
	L67R
	V69A
	S71I
	T72D
	H73E
	V74T
	I76A
	R77E
	L79A
	E80K/D81K
	L82A/L83A/M84A/G85A
	M84S
	G85A
	T86D
	L87A
	I89A/V90A
	C91G
	P92A
	I93T
	Q96E/K97E/P98A

*Mutants in bold represent mutations specifically made for this study (i.e surface exposed targets). All other mutants were previously constructed and many have been used extensively in the literature.

3.3.1 Dimerization properties of all HPV16 E7 mutants

The yeast two-hybrid test has been previously utilized to demonstrate that E7 can form dimers *in vivo* (Clemens *et al.*, 1995; Zwerschke *et al.*, 1996). This relatively simple assay was used to assess the oligomerization properties of the entire set of HPV16 E7 mutants. In the yeast-two hybrid system, wild-type or mutant E7 protein were fused to the DNA binding domain of the bacterial LexA repressor as “bait” or to the synthetic B42 transcriptional activation domain as “prey”. When expressed in the yeast *Saccharomyces cerevisiae* strain L-40, interaction between the bait and prey brings the activation domain to the promoter of a reporter gene, where the DNA binding domain is anchored via LexA binding sites (Figure 3.5). The activation domain is then able to activate transcription of the β -galactosidase gene downstream of the promoter. In this way, dimerization can be quantitatively measured from yeast extracts using the colorimetric liquid β -galactosidase assay. The lack of pRb orthologues in yeast, and the use of a LexA-based hybrid transcription factor allow the study of E7-driven transcriptional activation, independent of E2F/pRb complexes.

All PCR generated HPV16 E7 CR3 mutants were initially constructed in frame fused to the LexA DNA binding domain. Oligonucleotide primers generating *Eco*R1 and *Sal*I restriction endonuclease sites at the 5' and 3' termini of the E7 ORF allowed all 21 new E7 CR3 mutants to be subsequently subcloned into the prey vector to create B42-E7 hybrid proteins. All other mutants previously constructed in the laboratory were also subcloned into these vectors to create corresponding bait and prey fusion proteins. The experimental design for each mutant is outlined in Table 3.3. All dimerization experiments were conducted in L-40 yeast, which contain an integrated LexA responsive β -galactosidase reporter (Figure 3.5). The first two combinations (wild-type/vector, vector/wild-type) represent necessary controls for this assay, as HPV16 E7 has been reported to possess an intrinsic ability to activate transcription (Zwerschke *et al.*, 1996) and interacts with a variety of transcriptional regulators such as TBP (Phillips & Vousden, 1997). I first determined the ability of each mutant to transactivate the reporter gene. L-40 strain yeast were transformed with vectors expressing a LexA-E7 mutant fusions alongside an empty prey vector. In a similar fashion, L-40 strain yeast were transformed with a vector expressing the B42-E7 mutant fusion with an empty bait vector

Table 3.3 Experimental Design for Testing Dimerization in Yeast 2-Hybrid Assay

Bait	Prey
Vector	Vector
Wild-type E7	Vector
Vector	Wild-type E7
Mutant	Vector
Mutant	Wild-type E7
Vector	Mutant
Wild-type E7	Mutant
Mutant	Mutant

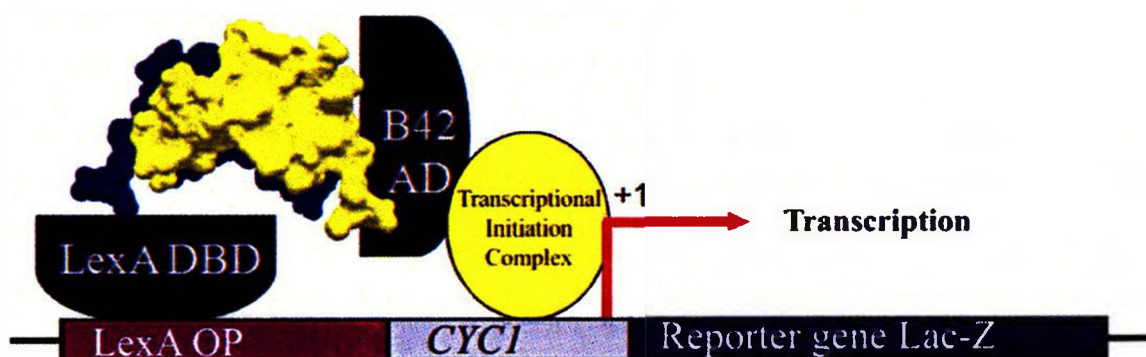


Figure 3.5. Schematic of the yeast two-hybrid system for dimerization assays. In this system, either wild-type or mutant E7 will be fused to the LexA DNA binding domain as “bait” or the B42 activation domain as “prey”. Specific interaction of the bait and prey chimeras reconstitutes a functional transcriptional activator that can bind to LexA operator sites upstream of the promoter. A transcriptional initial complex (including RNA polymerase II) is recruited to the yeast minimal promoter (*CYC1*), leading to transcription of the reporter gene. The experimental design for each mutant is outlined in Table 3.3.

to determine if they influenced basal expression of the yeast reporter gene. Lastly, each mutant was tested for dimerization in three different combinations: LexA-E7 mutant fusions were expressed in combination with the B42-WT E7 fusion, the LexA-WT E7 fusion together with the B42-E7 mutant fusions and finally, the same mutant expressed as both LexA and B42 fusions.

To ensure that the yeast two-hybrid assay could reproducibly and reliably determine dimer formation, a number of control experiments were conducted using empty vectors and wild-type E7 fusions (Figure 3.6). Wild-type E7, when fused to LexA, activates transcription very poorly in the absence of an interacting prey. From these results, it is clear that E7 is a weak transactivator on its own under these conditions. This intrinsic transcriptional activation activity does not significantly contribute to the ability of the E7 homodimer to initiate transcription, since co-expression of wild-type E7 as both bait and prey proteins results in a dramatically higher reporter gene activity. Expression of wild-type E7 as prey had no effect on reporter gene activity.

Following these experiments, the entire collection of 38 E7 mutants were tested for their ability to form heterodimers with wild-type E7 or homodimers with itself. The assay was performed twice from two independent yeast transformations and the results from one trial are summarized in Figure 3.7. The intrinsic transcriptional activation property of each mutant is depicted by the first bar graph for each mutant (striped bars). Although the intrinsic ability of the various mutants to activate reporter gene expression varied over a substantial range, autoactivation by each mutant is insignificant in comparison to the transcriptional activity induced by the E7 dimer (black bar). Interestingly, when a mutant was expressed in combination with wild-type E7, as fusions to either the LexA or B42 moiety (gray vs. striped bars), the two sets of yeast two-hybrid data did not always correlate with one another. This suggests that each type of fusion could have a different level of protein expression and/or stability. The difficulties associated with analyzing heterodimer data led me to focus on data from yeast that were transformed with CR3 mutants expressed both as LexA and B42 fusion proteins. These results were further analyzed by performing the Student's *t*-test to determine whether the differences in mean activity between the wild-type E7 homodimer

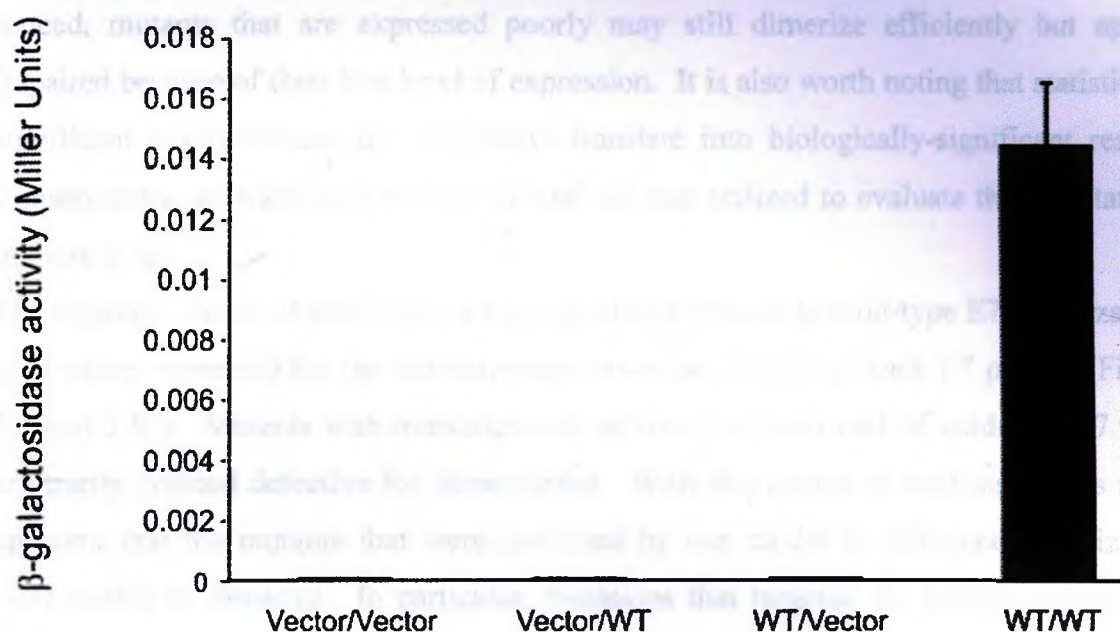
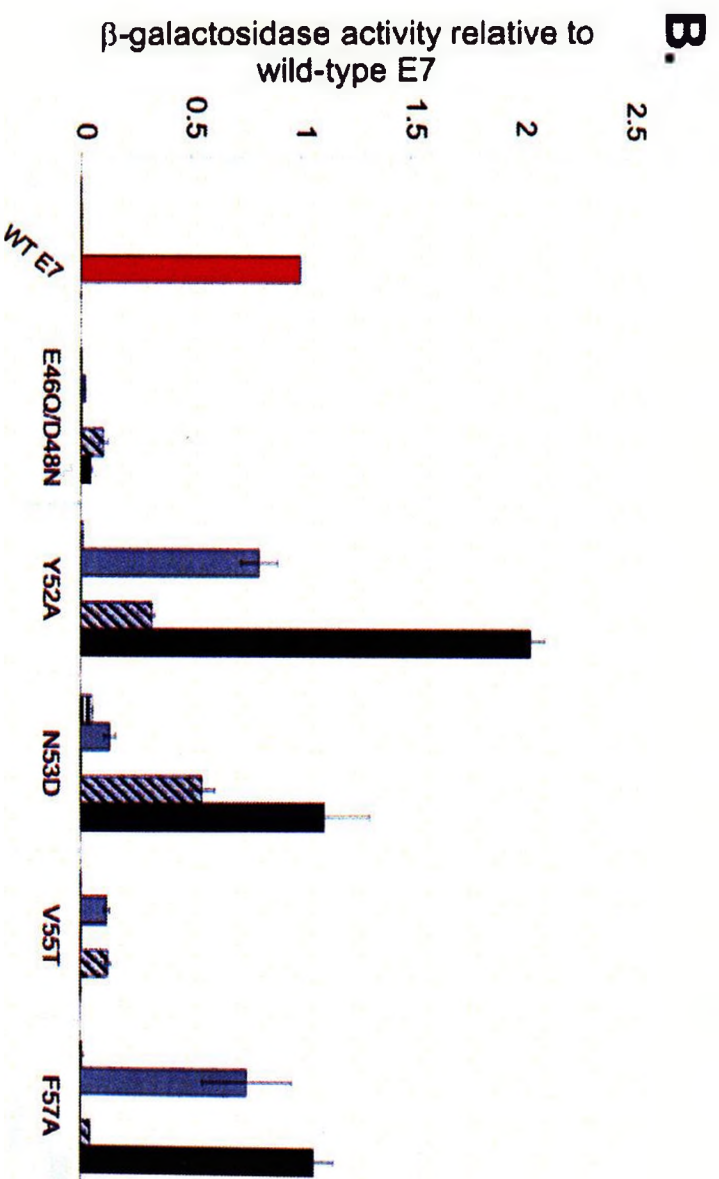
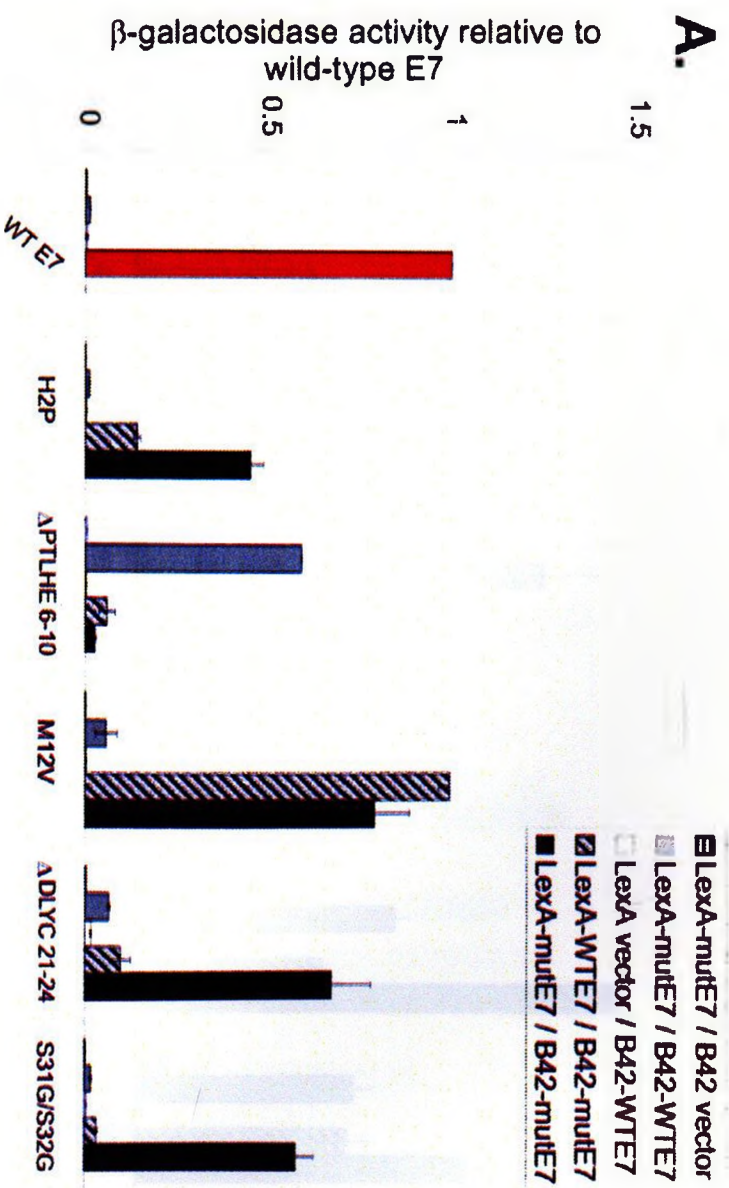


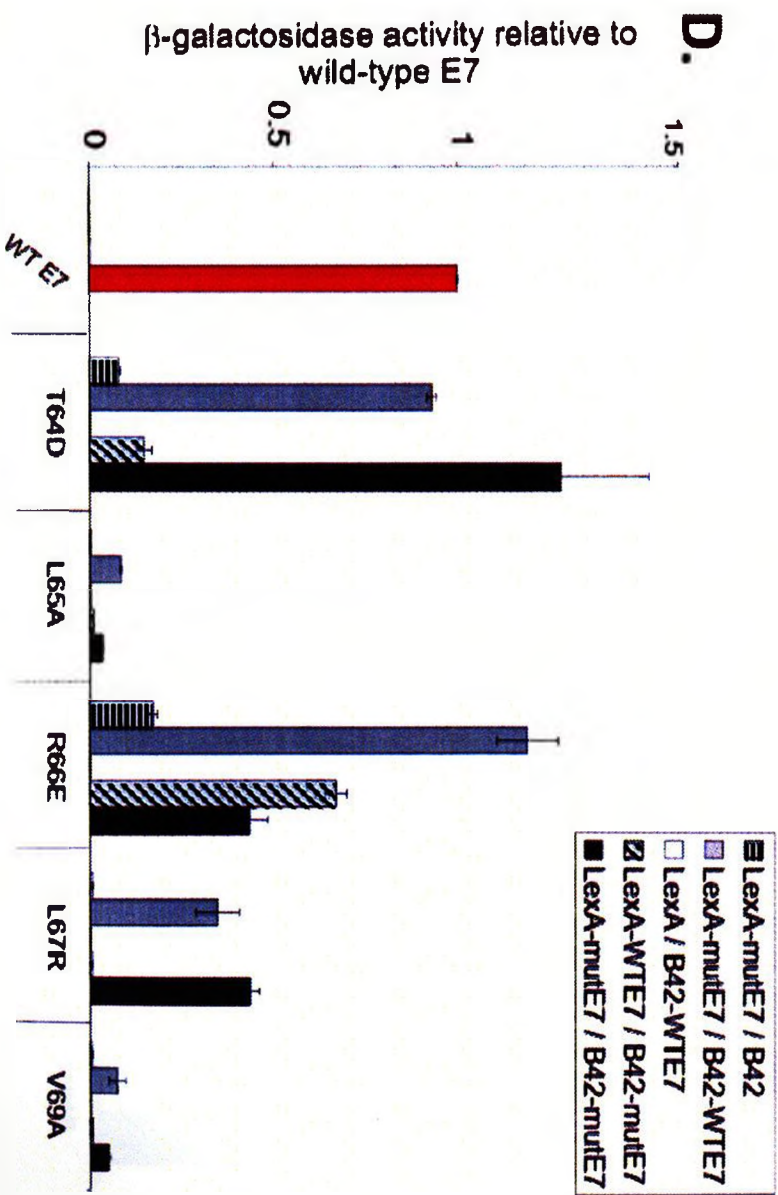
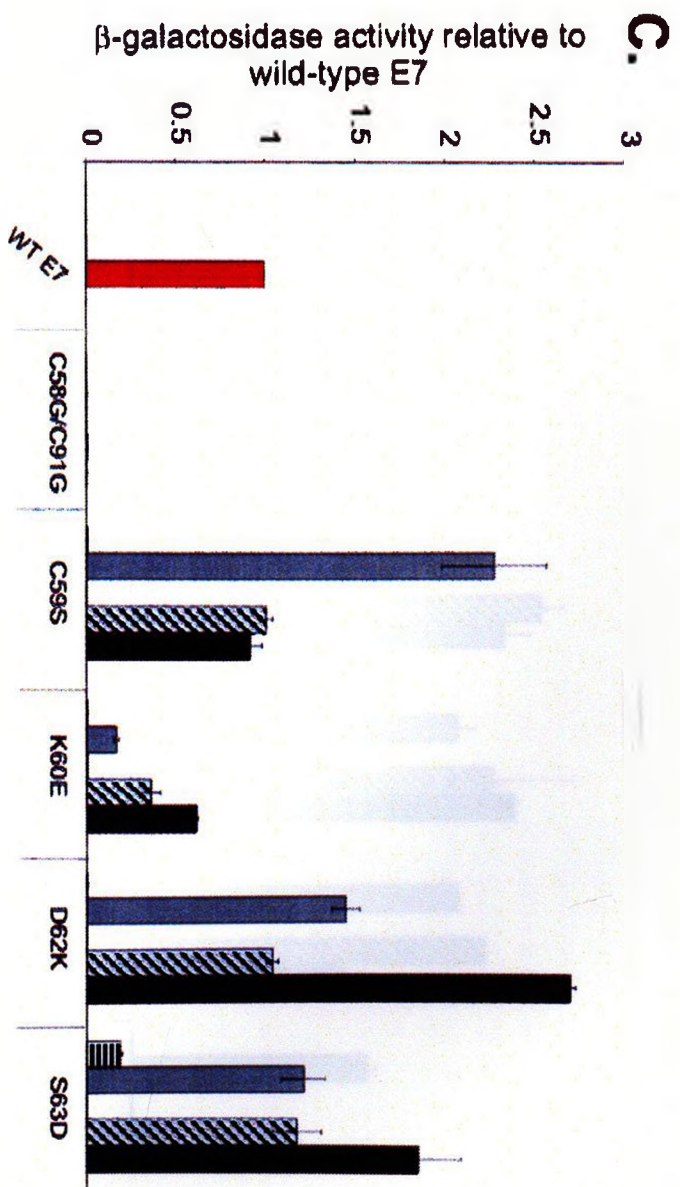
Figure 3.6. Controls for the yeast two-hybrid assay. L-40 yeast cells were cotransformed with a LexA DNA binding domain (bait) and a B42 activation domain (prey). The combination of vectors used are indicated as bait/prey on the x-axis, with 'vector' representing an empty yeast vector and 'WT' indicating a vector expressing wild-type HPV16 E7. Vectors expressing wild-type (WT) E7 fused to the LexA DNA binding domain or B42 activation domain were cotransformed alongside an empty bait or prey vector into L-40 yeast cells containing an integrated LexA responsive β -galactosidase reporter. Cell extracts were prepared and assayed for β -galactosidase activity by standard methods. Values represent the average of triplicate results for one trial.

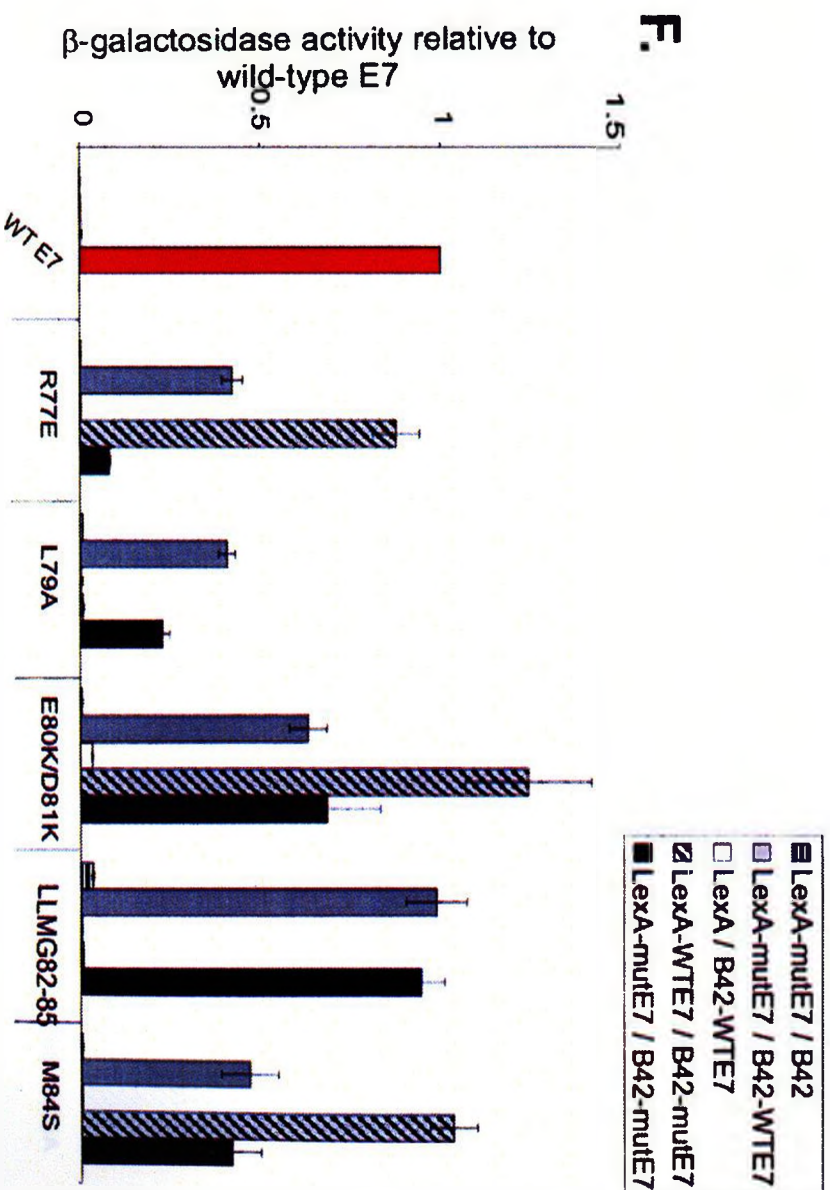
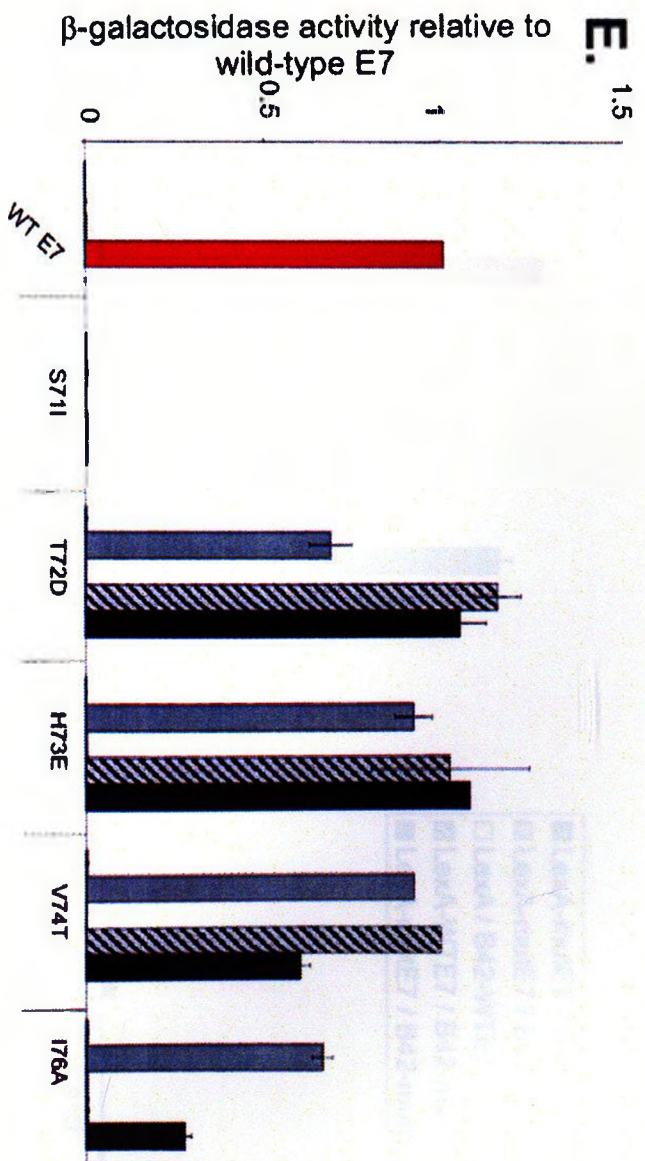
and the mutant E7 homodimer were statistically significant or arose due to random sampling (Figure 3.8). Based on these results, it would appear that a large proportion of these mutants are less able to form a dimer than wild-type E7. However, the yeast-two hybrid assays comprise only the initial phase in the characterization of these mutants. Indeed, mutants that are expressed poorly may still dimerize efficiently but appear impaired because of their low level of expression. It is also worth noting that statistically significant interpretations do not always translate into biologically-significant results. Consequently, an additional method of analysis was utilized to evaluate these mutations in more detail.

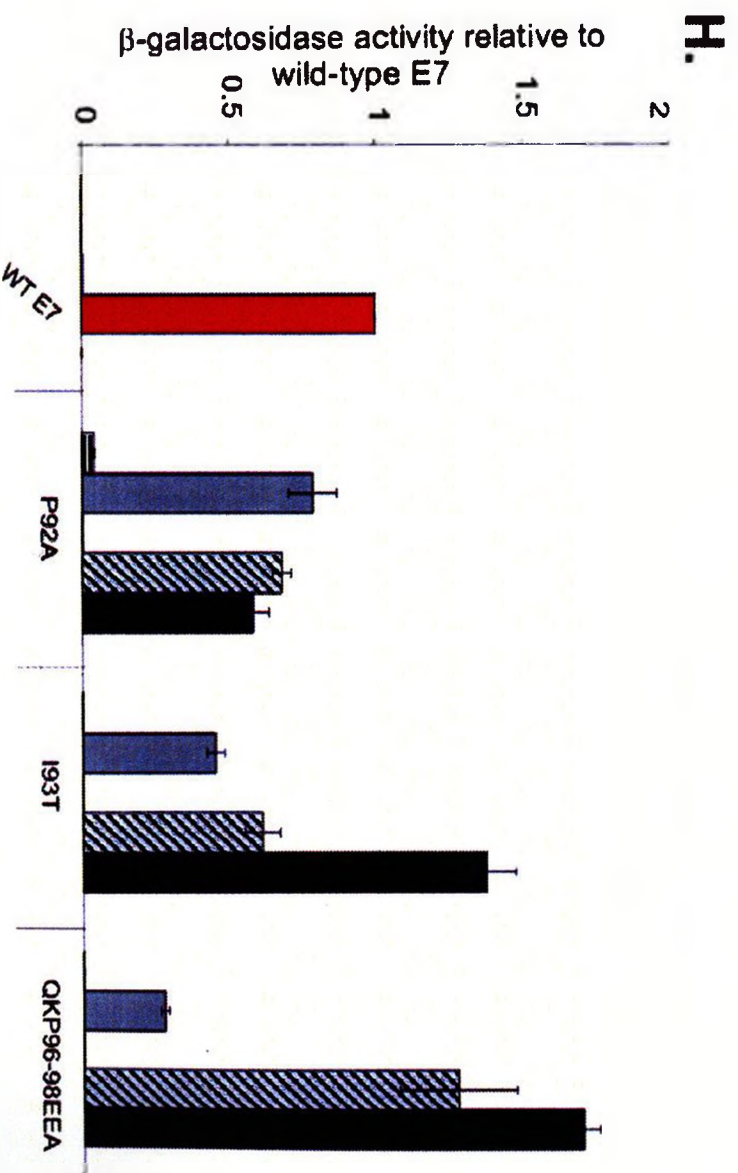
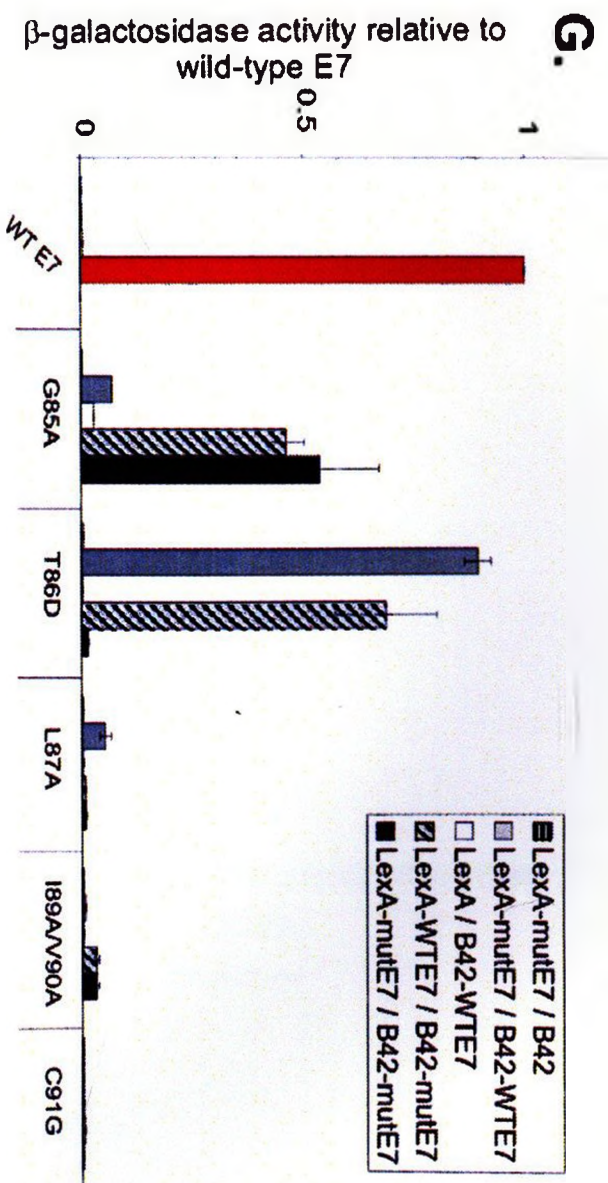
The oligomerization of each mutant was calculated relative to wild-type E7 dimerization, with values corrected for the intrinsic transactivation activity of each E7 protein (Figure 3.9 and 3.10). Mutants with transcriptional activity less than half of wild-type E7 were arbitrarily deemed defective for dimerization. With this course of analysis, it was more apparent that the mutants that were predicted by our model to influence dimerization were unable to dimerize. In particular, mutations that targeted the buried hydrophobic residues and zinc co-ordinating cysteines had more than a 50% reduction in transcriptional activity relative to wild-type E7 (Figure 3.10B). Despite these promising results, a number of dimerization-negative mutants were found in the N-terminus and the residues predicted by our model to be solvent exposed (Figure 3.10A and C). This was an unexpected result considering that the N-terminus has no apparent structure and that the surface-exposed residues should not interfere with folding of the monomer or contribute to stabilizing the dimer.

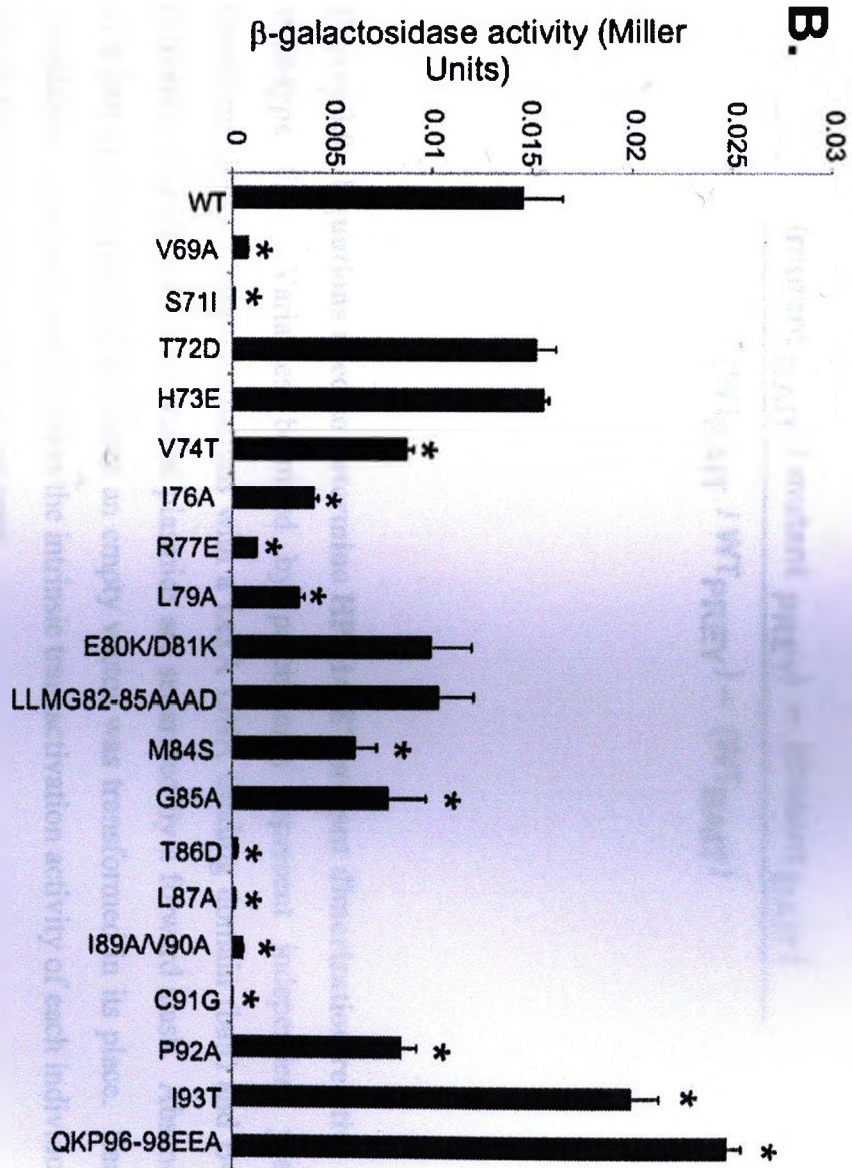
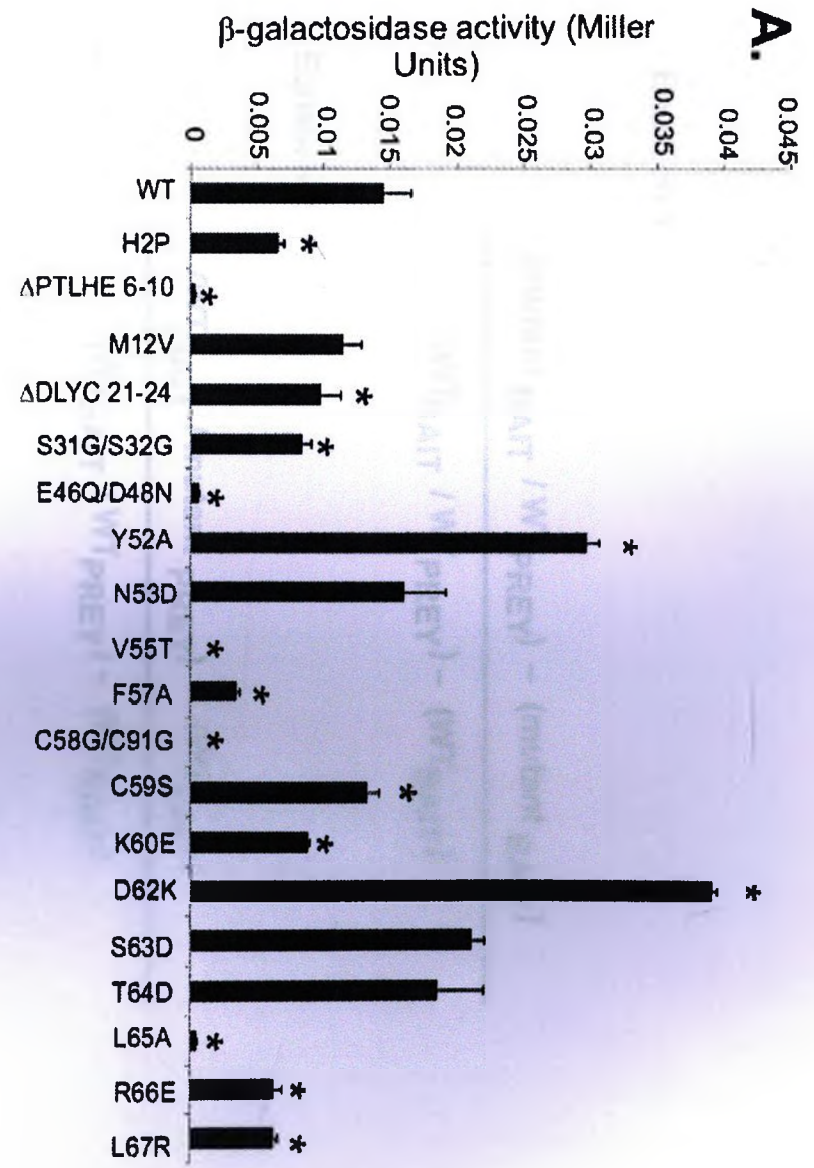
Figure 3.7 A-H. Dimer interactions by wild-type and mutant HPV16 E7 proteins. Wild-type (WT) HPV16 E7 or the indicated E7 mutant were expressed as fusions to the LexA DNA binding domain or B42 activation domain in L-40 yeast cells containing an integrated LexA responsive β -galactosidase reporter. Appropriate controls (refer to Figure 3.6) are indicated on the far left of each graph. Cell extracts were prepared and assayed for β -galactosidase activity by standard methods. Values represent the average of triplicate results for one trial.











Equation 1

$$\frac{(\text{mutant}_{\text{BAIT}} / \text{WT}_{\text{PREY}}) - (\text{mutant}_{\text{BAIT}})}{(\text{WT}_{\text{BAIT}} / \text{WT}_{\text{PREY}}) - (\text{WT}_{\text{BAIT}})}$$

Equation 2

$$\frac{(\text{WT}_{\text{BAIT}} / \text{mutant}_{\text{PREY}}) - (\text{WT}_{\text{BAIT}})}{(\text{WT}_{\text{BAIT}} / \text{WT}_{\text{PREY}}) - (\text{WT}_{\text{BAIT}})}$$

Equation 3

$$\frac{(\text{mutant}_{\text{BAIT}} / \text{mutant}_{\text{PREY}}) - (\text{mutant}_{\text{BAIT}})}{(\text{WT}_{\text{BAIT}} / \text{WT}_{\text{PREY}}) - (\text{WT}_{\text{BAIT}})}$$

Figure 3.9. Equations used to determine HPV16 E7 mutant dimerization relative to wild-type E7. Variables bounded by parentheses represent independent yeast transformations. Cotransformations with a LexA DNA binding domain (bait) and B42 activation domain (prey) E7 fusion plasmids are separated by a forward slash. Absence of a bait or prey plasmid indicates an empty vector was transformed in its place. Each calculation takes into consideration the intrinsic transactivation activity of each individual bait in the absence of an interacting prey.

Figure 3.10. A -C. Dimerization ability of selected HPV16 E7 mutants relative wild-type E7. Data from Figure 3.8 was subjected to the equations presented in Figure 3.9 and categorized according to the spatial location of the mutated residues in the model of HPV16 E7 CR3: A) N-terminus, B) hydrophobic core and C) predicted solvent exposed residues. Data represents the average of calculated values from two independent trials.

