IDENTIFICATION OF THE CELLULAR PROTEINS WHICH INTERACT WITH THE ESSENTIAL HSV-1 PROTEIN IE63

BY

SARAH WADD

A thesis presented for the degree of
Doctor of Philosophy

in

The Faculty of Science
at the University of Glasgow

Institute of Virology
Church Street
Glasgow G11 5JR

January 2000

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ACKNOWLEDGEMENTS

I would like to thank my joint supervisors Professor Barklie Clements and Dr Anne Phelan for their support and advice throughout the course of this project. I am particularly grateful to Dr Roger Everett, my second supervisor who gave encouragement, sound advice and hours of his time. Thanks must also go to Dr Margy Koffa and Jim Scott for their assistance and to Dr Jane Parkinson for reading this manuscript.

Finally, I would like to thank Andrew for keeping me sane and for whisking me off to the hills whenever things became difficult. I would like to dedicate this thesis to my parents and sister Victoria for their unfailing support and encouragement.

The author was in receipt of a Medical Research Council Industrial Collaborative Studentship (with Cruachem Ltd). The author carried out all the work presented in this thesis.

ABSTRACT

Herpesviruses are involved in a range of prominent medical or veterinary diseases making this one of the most significant virus families. Their ubiquitous occurrence, genetic complexity and differing biological properties have stimulated considerable research efforts worldwide particularly in the area of molecular genetics. Herpesviruses differ widely in their pathogenic potential but share the ability, after primary infection, to establish latent infections throughout the lifetime of the host.

To date, eight different herpesviruses whose natural host is man have been identified: herpes simplex virus (HSV) types 1 and 2, varicella-zoster virus, Epstein-Barr virus, human cytomegalovirus and human herpesviruses (HHV) types 6-8. A large proportion of people worldwide have been exposed to, and may be latently infected with, one or more herpesvirus. Generally, these viruses are controlled by an intact immune system but they cause life-threatening disease in immunocompromised individuals and some are implicated in different types of cancer. With increased use of immunosuppressive drugs and the advent of AIDS, herpesvirus infections pose a growing problem . Although a number of anti-herpesvirus drugs have been marketed, drug resistance can develop and researchers continue to search for novel targets for drug development at key points in the virus lytic cycle.

HSV-1 proteins are divided into three regulatory classes, immediate-early (IE), early and late depending upon their kinetics of expression and requirement for ongoing DNA synthesis. The HSV-1 protein IE63 (ICP27) is essential for viral replication, and is the only HSV-1 IE protein which has homologues in every herpesvirus sequenced so far which infect mammals and birds, indicative of its key regulatory role and making it an excellent antiviral target .

IE63 is a 512 amino acid phosphoprotein which has previously been demonstrated to be required for virus late gene expression and DNA replication, and appears to exert its effects at both transcriptional and post-transcriptional levels. The mechanism of IE63 action at the transcriptional level is not known but IE63 appears to interact with

the essential HSV-1 transcriptional regulator IE175 and may alter the promoter binding of this protein.

At the post-transcriptional level, IE63 increases RNA 3' processing at late vital poly(A) sites, a process which may allow IE63 to regulate the switch between early and late viral gene expression. IE63 has also been shown to inhibit pre-mRNA splicing in HSV-1 infected cells. During HSV-1 infection, splicing factors redistribute within the nucleus and IE63 is necessary and sufficient for this effect, leading researchers to postulate that IE63 may inhibit splicing by facilitating the removal of splicing factors to dysfunctional aggregates. The inhibition of splicing is highly advantageous because the majority of HSV-1 transcripts do not contain introns and therefore do not require splicing, in contrast to the majority of cellular transcripts. Although the redistribution of splicing factors has a role in the inhibition of splicing, studies have shown that another, as yet unidentified mechanism, acts in concert with the redistribution of splicing factors to disrupt the splicing process in HSV-1 infected cells.

Finally, IE63 facilitates the nuclear retention of intron-containing transcripts which also favours the translation of viral and not cellular RNAs. IE63 binds RNA, shuttles between the nucleus and the cytoplasm, and it is thought that IE63 acts to selectively export viral RNAs from the nucleus. Other viral proteins notably the human immunodeficiency virus-1 (HIV-1) Rev protein shuttle between the nucleus and the cytoplasm transporting spliced or partially spliced HIV-1 RNAs. Interestingly, a number of cellular and viral proteins have an altered phosphorylation state in the presence of IE63 although IE63 itself does not exhibit kinase activity. This indicates that IE63 may exert effects by recruiting a kinase to phosphorylate proteins required for its various functions.

IE63 clearly exhibits several activities and it seemed likely that it exerted its effects by interacting with various cellular proteins involved in these processes and modifying their activities. The aim of this study was therefore to screen a library of cellular proteins to identify those which interacted with HSV-1 IE63. This information was likely to shed light on the mode of action of this protein.

The yeast two-hybrid system was used to identify interacting proteins. This system exploits the ability of a 'bait' protein fused to the binding domain of a yeast transcription factor to interact with a target protein fused to the activation domain of the same yeast transcription factor, bringing the two domains into close enough proximity to reconstitute activity of the transcription factor and subsequent transcription of reporter genes in yeast cells. Proteins that do not interact do not permit activation of reporter gene expression.

In this study, a truncated IE63 protein was used to screen proteins expressed from an uninfected HeLa cell cDNA library. The protein was truncated because it was found that that the N-terminal ten amino acids of IE63 caused transcription of reporter genes in the absence of an interacting protein. The screen identified eight cellular proteins which interacted with IE63, four of which were of initial interest. These were heterogeneous nuclear ribonucleoprotein K (hnRNP K), casein kinase II (CKII) β subunit, and the splicing factors Spliceosome Associated Protein 145 (SAP145) and p32:- these interactions were all subsequently confirmed in virus infected cells by others in our laboratory. Truncation mutants of IE63 were constructed and using these in the two-hybrid assay, the regions of IE63 required for these various interactions were mapped as well as the IE63 region required for dimersation. Similarly, using hnRNP K truncations, regions required for interaction with IE63 were identified

A region containing the zinc finger domain of IE63, located towards the C-terminus, was required for interaction with the CKIIβ subunit, with SAP145 and for IE63 dimerisation. Dimerisation would permit IE63 to form interactions with multiple partners in a complex of proteins with RNA, allowing cross-talk between these molecules. An IE63 region involved in the interaction with hnRNP K contained a five amino acid sequence (SADET) which is well conserved throughout the alphaherpesviruses, and which is present in a region of hnRNP K required for interaction with IE63 and is necessary for the nuclear export of hnRNP K. This region of IE63 could have a role in nuclear export, perhaps recognising a nucleoporin also recognised by hnRNP K. A further IE63 region with runs of prolines was identified as

involved in interactions with hnRNP K and CKIIβ, and was the only region necessary for the interaction with p32.

It is highly probable that IE63 inhibits splicing by interacting with SAP145 and/or p32, and it appears that CKII is recruited by IE63 to phosphorylate itself, hnRNP K and probably other proteins. The function of hnRNP K in uninfected cells is not completely understood but it is thought that the protein has a role in transcription, pre-mRNA processing, RNA transport and translation. It is proposed that IE63 may alter transcription, mRNA processing and transport in the HSV-1 infected cell via its interaction with hnRNP K.

The study also demonstrated that IE63 can form interactions with a transcriptional activator ALY, a glutamine-rich protein, the protein human JTV-1 of unknown function and the nucleoporin RIP1. Furthermore, the nucleoporin CRM1 which interacts with HIV-1 Rev and facilitates its export to the cytoplasm did not interact with IE63 in the two-hybrid assay, suggesting that IE63 and Rev may use different pathways to cross the nuclear pore. Finally, the HHV-8 (Kaposi's sarcoma herpesvirus) ORF57 homologue of IE63, which also acts post-transcriptionally, did not interact with the hnRNP K, CKIIβ, SAP145 and p32 clones identified in the library screen suggesting that these distantly related homologues, which belong to different herpesvirus subgroups, are likely to possess certain different activities.

CONTENTS

ABBREVIATIONS

3-AT	3-amino-1,2,4-triazole
aa	amino acid
AD	activation domain
ADH1	alcohol dehydrogenase
Amp	ampicillin
ATP	adenosine triphosphate
BD	binding domain
β-gal	β-galactosidase
BHK	baby hamster kidney cells
bp	base pairs
C	cytosine
cAMP	adenosine 3',5'-cyclic monosphosphate
CAT	chloramphenicol acetyl transferase
cDNA	complementary DNA
cfu	colony forming units
CKII	casein kinase II
CPSF	cleavage and polyadenylation specificity factor
CstF	cleavage stimulation factor
cyh^R	cycloheximide resistance
dATP	2'-deoxyadenosine-5'-triphosphate
dCTP	2'-deoxycytosine-5'-triphsophate
dGTP	2'-deoxyguanosine-5'-triphosphate
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
dNTP	2'-deoxyribonucleoside-5'-triphosphate
DO medium	dropout medium
DRB	5'6-dichloro-1-β-ribofuranosylbenzimidazole
dTTP	2'-deoxythymidine-5'-triphosphate
E	early

E. coli	*Escherichia coli*
EDTA	sodium ethylenediamine tetra-acetic acid
G	guanine
GAR	glycine-arginine-rich domain
GARP	glutamine/glutamic acid rich protein
gB	glycoprotein B
GTP	guanosine-5'-triphosphate
h	hour(s)
His	histidine
HIV-1	human immunodeficiency virus
hnRNP	heterogeneous nuclear ribonucleoprotein particle
HSV-1	herpes simplex virus type 1
HSV-2	herpes simplex virus type 2
ICP	infected cell protein
IE	immediate-early
IGC	interchromatin granule cluster
k	kilo (ie 10^3)
KS	Kaposi's sarcoma
kb	kilobase pair(s)
kDa	kilodalton
KNS	K nuclear shuttling domain
KPK	K protein kinase
L Particles	light particles
Lac Z	β-galactosidase
LAT	latency associated transcript
Leu	leucine
LOX	erythroid 15-lipoxygenase
LPF	late processing factor
MDa	megadalton
mg	milligram(s)
MHC	major histocompatability complex
min	minute(s)
ml	millilitre(s)

mM	millimolar
mm	millimetre(s)
MOI	multiplicity of infection
mRNA	messenger RNA
NES	nuclear export signal
ng	nanogram
NLS	nuclear localisation signal
nm	nanometre(s)
NPC	nuclear pore complex
NuLS	nucleolar localisation signal
OD	optical density
ORF	open reading frame
PABII	poly(A) binding protein II
PAP	poly(A) polymerase
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PFs	perichromatin fibrils
PKI	protein kinase inhibitor
pmol	picomole(s)
pol	polymerase
poly (A)	polyadenylation
PREP	pre-DNA replication particles
RNA	ribonucleic acid
rpm	revolutions per minute
RRE	rev response element
S. cerevisiae	*Saccharomyces cerevisiae*
SAP145	Spliceosome Associated Protein 145
SD	synthetic dropout
SDS	sodium dodecyl sulphate
sec	second(s)
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein particle
SV40	Simian virus 40
TBP	TFIID TATA-binding protein

Trp	tryptophan
UAS	upstream activitation signal
μg	microgram(s)
UTR	untranslated region
v:v	volume : volume
Vhs	virion host shut off
Vmw	virion-specific polypeptide of apparent molecular weight
Vol	volume(s)
w:v	weight : volume
wt	wild type
X-GAL	5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside

INTRODUCTION

1.1 THE UBIQUITOUS HERPES SIMPLEX VIRUS TYPE-1

Herpes simplex virus (HSV) causes an acute, infectious disease which is characterised by the development of small fluid-filled vesicles in the skin and mucous membranes. The virus exists in two forms, herpes simplex type 1 (HSV-1) and herpes simplex type 2 (HSV-2), both of which are members of the family Herpesviridae. More than 80 herpesviruses have been isolated from a diverse range of species including mammals, fish, reptiles, birds, amphibians and marsupials. Eight of the herpesviruses have been shown to infect humans. Table 1.1 lists the eight human herpesviruses identified together with the illnesses that they cause and the subgroups of mammalian herpesviruses to which they belong.

All herpesviruses (reviewed in Roizman & Sears, 1996) have three distinct characteristics. First, all have a large genome consisting of a single molecule of double stranded DNA ranging in size between 120 and 250 kilobase pairs (kb). Second, all have the ability to enter a latent phase in the host cell which ensures survival of the viral genome throughout the lifetime of an infected individual. Third, all show the same basic morphology and this morphology is the primary basis of identification.

1.1.1 Virion morphology

The HSV-1 virion is approximately 150-200 nm in diameter and consists of four primary structural features which are illustrated diagramatically in

Table 1.1. HUMAN HERPESVIRUSES

Virus	Abbreviation	Lineage	Pathology
Herpes simplex virus type 1	HSV-1	α1	Recurrent epithelial lesions; rarer, serious neural disease
Herpes simplex virus type 2	HSV-2	α1	As for HSV-1
Varicella-zoster virus	VZV	α2	Primary infection chickenpox; recurrence as shingles
Human cytomegalovirus	HCMV	β1	Cause of congenital abnormalities; spontaneous abortion; severe infections in immunocompromised
Human herpesvirus 6	HHV6	β2	Primary infectious exanthem subitum
Human herpesvirus 7	HHV7	β2	No definitive disease association
Epstein-Barr virus	EBV	γ1	Primary infection mononucleosis; associated with several Burkitt's lymphoma and other neoplasias
Human herpesvirus 8	HHV8	γ2	Associated with Kaposi's sarcoma and other neoplasias

Figure 1.1. The capsid has an icosahedral structure (100-110 nm in diameter) which consists of 162 capsomers. Enclosed inside the capsid is an electron opaque structure, the core, which contains the viral DNA. Both capsid and core are enveloped by a lipid containing membrane, the envelope, which contains a number of virus encoded glycoprotein spikes (HSV-1 encodes at least 11) and is acquired by the budding of capsids through the lipid bilayer membrane of the host cell. The tegument is an electron dense material which is located between the capsid and the envelope and is comprised of viral and possibly some cellular proteins, the properties and functions of which are largely unknown, but they include the proteins VP16 and UL41 which contribute to efficient infection.

1.1.2 Clinical features and epidemiology of HSV-1 infection

The severity of the illness caused by the HSV-1 virus can vary greatly, ranging from the asymptomatic to a severe systemic illness which is occasionally fatal. The majority of childhood HSV-1 infections are asymptomatic but they may be followed by subsequent periods of symptomatic disease. In children 1-3 years of age, the major symptom of HSV-1 infection is gingivostomatitis, an infection of the gums, tongue, mouth, lip, facial area and pharynx. Reactivated HSV-1 is associated with mucosal ulcerations or lesions around the lips that last 4-7 days and are often referred to as cold sores or fever blisters. Such lesions can be both irritating and distressing to the sufferer. Acute HSV-1 rhinitis is a primary infection of the nose and is recognised by the appearance of tiny vesicles in the nostrils

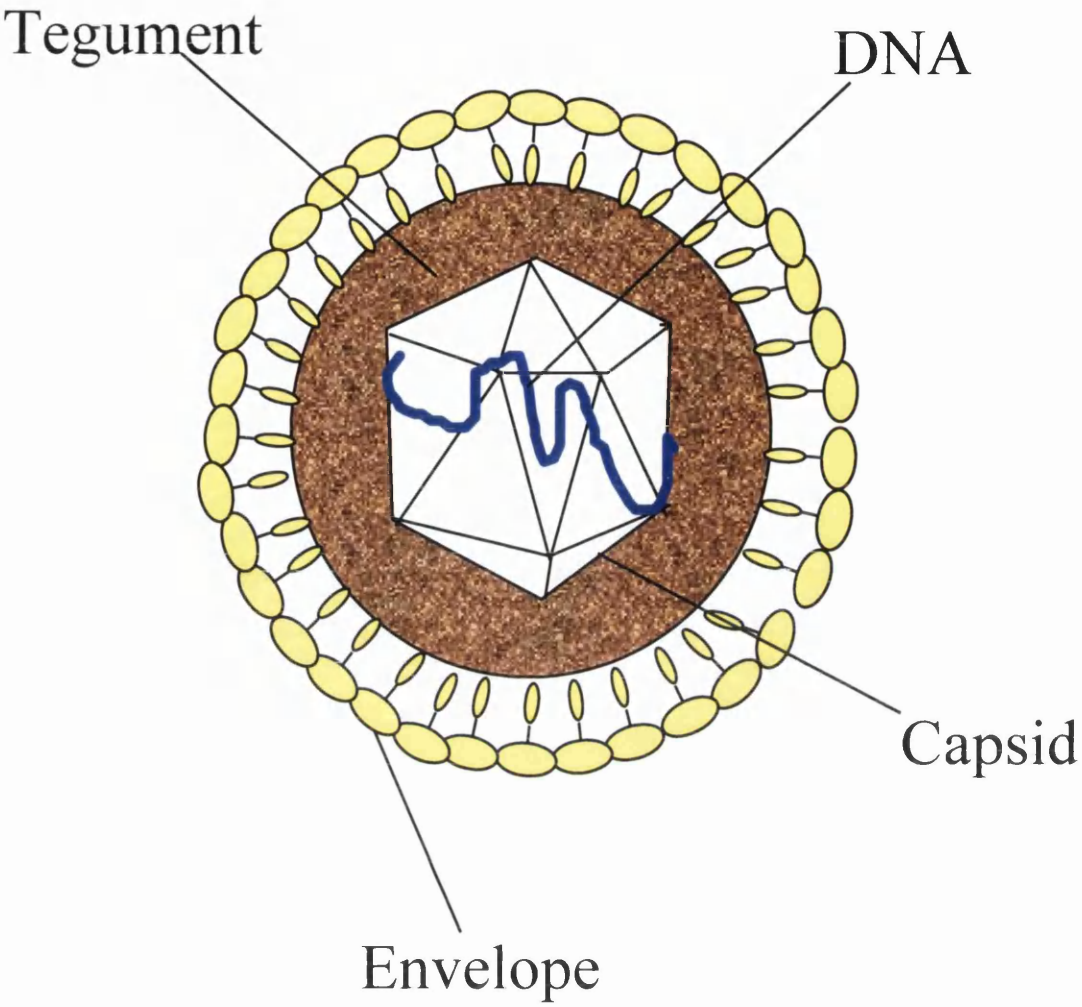


Figure 1.1
Representation of virion morphology of Herpesvidae, showing the outer envelope layer containing glycoprotein molecules, the tegument and icosahedral capsid surrounding the central core of double stranded DNA.

usually associated with fever and enlarged lymph nodes in the neck. Other HSV-1 induced diseases include primary herpes dermatitis, eczema herpeticum and traumatic herpes. Herpetic whitlow results from infection of broken skin, often on fingers after coming into contact with virus from another individual. HSV-1 infections of the eye can lead to keratoconjunctivitis which can, in rare instances, cause blindness.

HSV-1 infections occur throughout the world with approximately one third of the population in developing countries and the lower socio-economic classes of developed countries seroconverting to HSV-1 by 5 years of age, and the frequency increases to 70-90% during adolescence. For HSV-1 infection to occur, the virus must come into contact with mucosal surfaces or grazed skin. Following replication at the site of infection, the virus is transported via the neurons to the dorsal root ganglia where latency is established. Replication in the ganglia occasionally results in a life threatening central nervous system infection, however, latency generally predominates.

1.1.3 The HSV-1 genome and life cycle

(a) Genome arrangement

The HSV-1 genome is comprised of linear, double-stranded DNA and has an unusually high G+C content (67%). The HSV-1 genome arrangement together with the position of latency associated transcripts (LATs) is

illustrated in Figure 1.2. The genome consists of two unique components designated unique long (U_L) and unique short (U_S) which comprise 82 and 18% of the DNA respectively. Each is bounded and linked by a set of internal (IR) and terminal (TR) repeats; TRL/IRL and IRS/TRS which are also designated ab/b'a' and a'c'/ca. A sequence which varies in length between 250-500 bp (the *a* sequence) is present as a direct repeat at both termini and in an inverted orientation at the L-S junction (Sheldrick & Berthelot, 1974, Davison & Wilkie, 1983, Delius & Clements, 1976, Hayward *et al.*, 1975). During infection the L and S components invert relative to each other so that the progeny viral DNA consists of equimolar amounts of four isomers differing from each other solely in the relative orientation of the two components. These isomers are termed: P (prototype), IL (inversion of L), IS (inversion of S) and ISL (inversion of both L and S) (Delius & Clements, 1976, Hayward *et al.*, 1975, Roizman, 1979).

(b) The life cycle of the virion, L particles and PREPS

Figure 1.3 illustrates the HSV-1 life cycle which takes approximately 18 h to 20 h in fully permissive tissue culture cells (reviewed in Roizman & Sears, 1996). The first stage of the life cycle is entry of the virus into the host cell and this occurs in two phases, viral attachment with rapid transition to penetration of virus into the cell. Entry of the virus is thought to require the sequential or simultaneous interaction of the virion envelope glycoproteins with cell surface receptors (Fuller & Lee, 1992). The de-enveloped capsid is transported to the nuclear pores via the cellular cytoskeleton where its DNA

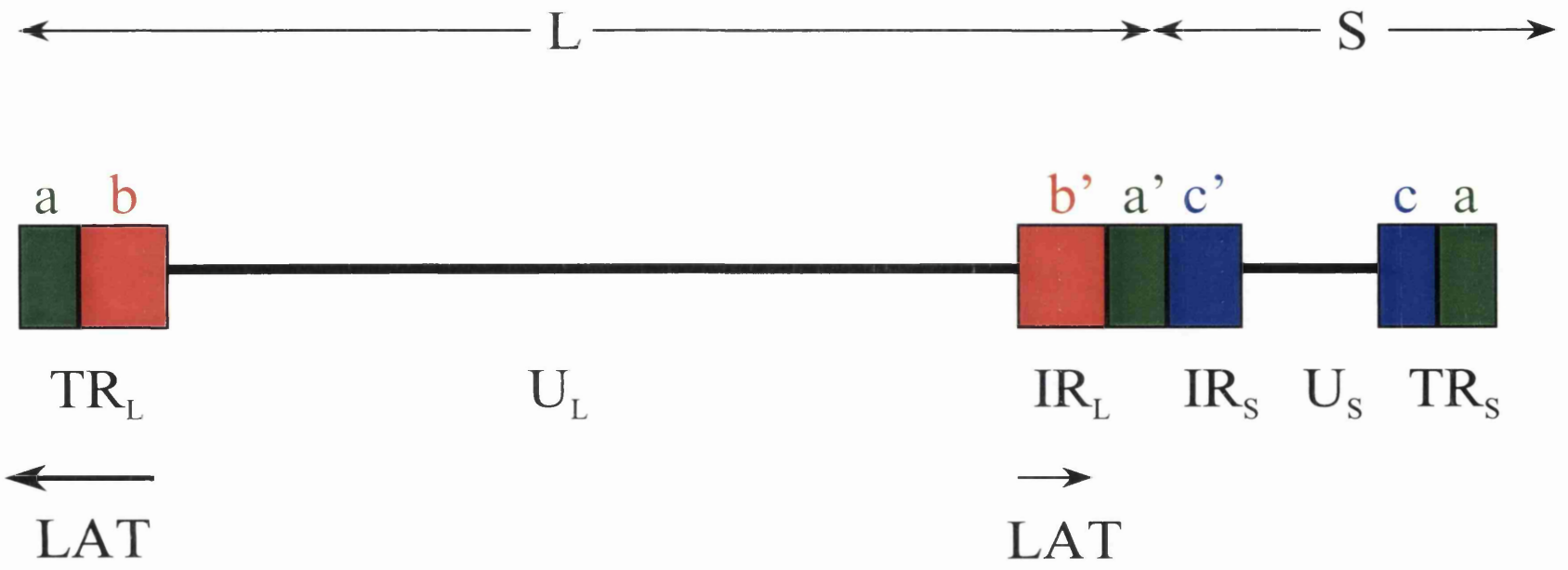


Figure 1.2
The structure of the HSV-1 genome including regions encoding latency associated transcripts (not to scale). The viral genome which consists of two components L (long) and S (short) is shown. Each component contains unique sequences U_L and U_S and internal (IR_L and IR_S) and terminal (TR_L and TR_S) repeat elements. A direct repeat, termed the 'a' sequence, is present at each end of the genome and also in an inverted orientation at the L-S junction.

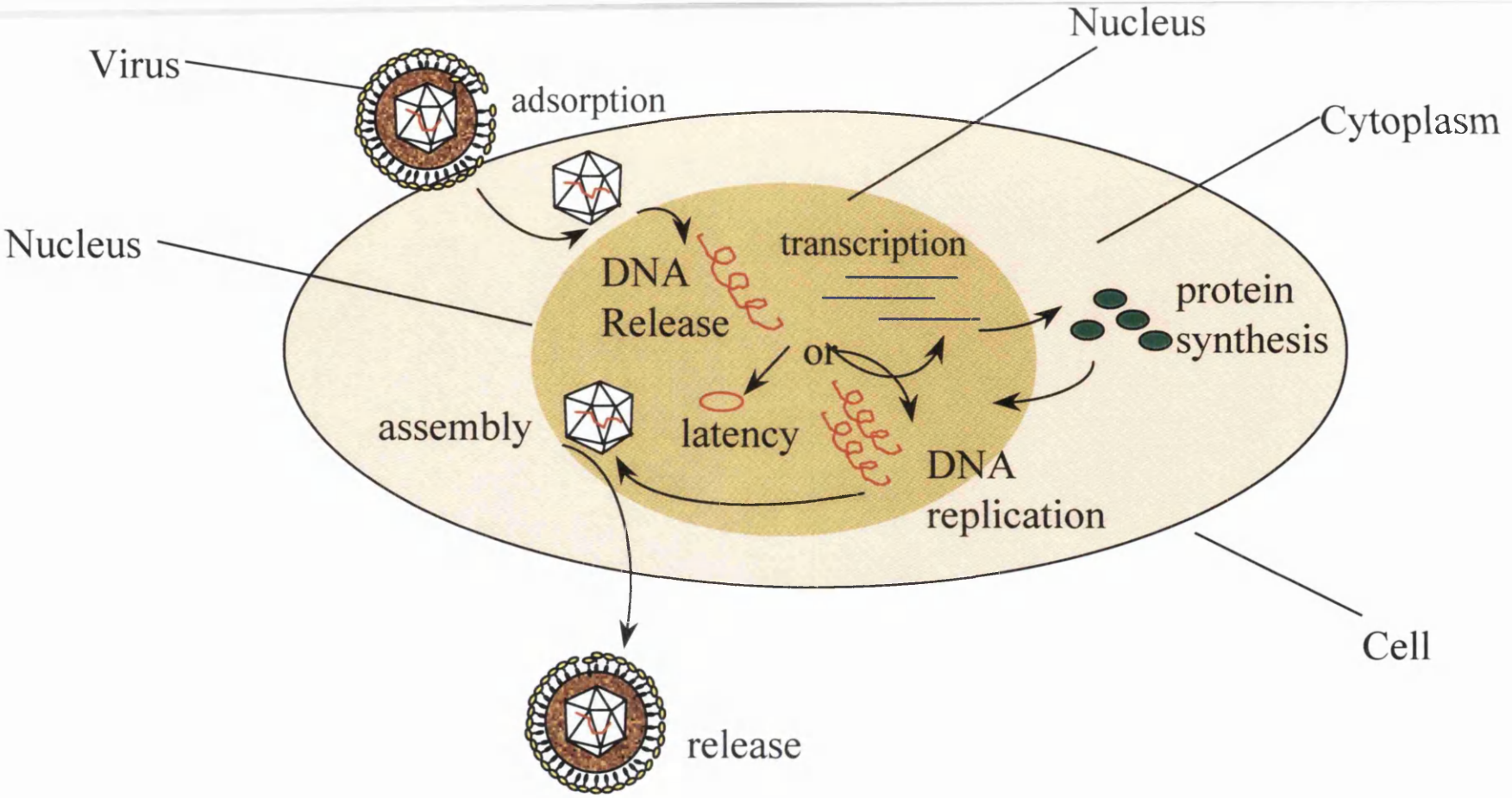


Figure 1.3
Life cycle of HSV-1. The virus adsorbs to the host cell via cell surface receptors and the virion is released into the cytoplasm. Next, DNA is released into the nucleus where it is either transcribed into RNA or becomes latent. Mature transcripts are transported to the cytoplasm where protein synthesis is carried out. Following immediate-early and early gene expression, DNA replication occurs and the newly synthesized DNA and proteins are packaged into capsids. Finally the virion acquires its envelope and glycoproteins by budding through the plasma membrane of the cell and the new virus particle is released.

is released into the nucleus. Host macromolecular synthesis is subsequently shut off. Transcription of the viral DNA by host cell RNA polymerase II takes place in the nucleus of the cell and the mature transcripts are transported to the cytoplasm where protein synthesis is carried out (Honess & Roizman, 1974). Viral DNA synthesis occurs in the host cell nucleus and can be detected from 3h post infection (pi) and continues for at least another 9 h to 12 h pi (see Section 1.1.4). In most instances, infected cells release two types of HSV-1 particles, but only one, the virion, contains HSV DNA and is thus capable of infecting another cell. L particles (Szilagyi & Cunningham, 1991) lack the nucleocapsid which contains the viral DNA and are thus non-infectious. They can be separated from the heavier HSV virions by banding in a Ficoll gradient, hence their name L (light) particles. Although the function of L particles in productive infection is not known at present, it is possible that they act as a decoy, enhancing the virion's chances of successful infection by adsorbing to a proportion of the circulating antibodies.

'Pre-DNA replication' or PREP particles (Dargan *et al.*, 1995) are synthesised by HSV-1 infected cells only when viral DNA synthesis is blocked either genetically in viruses with a mutation in a DNA replication gene or biochemically by growth of virus in the presence of inhibitors such as acyclovir or phosphonoacetic acid. PREP particles are not, therefore, normally produced during infection. PREP particles are similar to L particles but they either lack true late proteins or possess them in minute quantities.

The relative proportions of many other tegument and envelope proteins are also altered.

It is thought that the virion is enveloped by patches on the inner nuclear membrane. The virions then accumulate in the endoplasmic reticulum and transverse the Golgi apparatus and the trans Golgi network to the extracellular space.

(c) Latency

HSV-1 is able to exist in a non-replicative state which is often followed by virus reactivation and symptomic disease (reviewed in Rock, 1993). Following replication in the epithelia, virions or capsids are transported along neurons to the nerve cell bodies in the tri-geminal ganglia (Kristensson *et al.*, 1986). Here HSV DNA is thought to exist in an extrachromosomal state, probably as a circular episome (Mellerick & Fraser, 1987). During latent infection the only detectable viral gene expression is that of a major class of transcripts called the latency associated transcripts (LATs) (reviewed in Block & Hill, 1997). Although LATs are not essential for the establishment and maintenance of latency it appears that both the LAT region of the genome (Block *et al.*, 1993, Deshmane *et al.*, 1993, Dobson *et al.*, 1989, Hill *et al.*, 1990, Javier *et al.*, 1988, Leib *et al.*, 1989) and a functional IE110 protein (Chen *et al.*, 1991, Harris *et al.*, 1989, Lieb *et al.*, 1989, Zhu *et al.*, 1990) affect the efficiency of reactivation of latent virus. Stimuli that cause

virus reactivation are thought to include fever, exposure to sunlight, stress and possibly hormonal irregularities.

1.1.4 DNA replication

During lytic infection with HSV-1 in cell culture, as much as 50% by mass of the host cell DNA content is viral following the peak rates of viral DNA replication; this typically occurs between 10h and 20h post infection. Viral DNA replication is initiated at one of the 3 origins of replication by interaction with the ORI binding protein UL9. The viral helicase/primase complex consisting of UL5, 8 and 52 subsequently associates with the origin and the hydrolysis of ATP creates an initiation 'bubble' which permits association with the polymerase UL30/UL42 DNA-binding protein complex. DNA synthesis is continuous along one strand and discontinuous on the lagging strand with the major DNA-binding protein UL29 maintaining the growing forks. Figure 1.4 shows a schematic summary of the biochemical activities of the HSV encoded proteins involved in DNA replication as proposed by Challberg (Challberg, 1991).

1.1.5 Transcription of the genome

(a) Kinetic classes of HSV-1 transcripts

Herpesvirus genes are transcribed by the host cell RNA polymerase II (Constanzo *et al.*, 1977). The genes can be divided into at least 3 regulatory

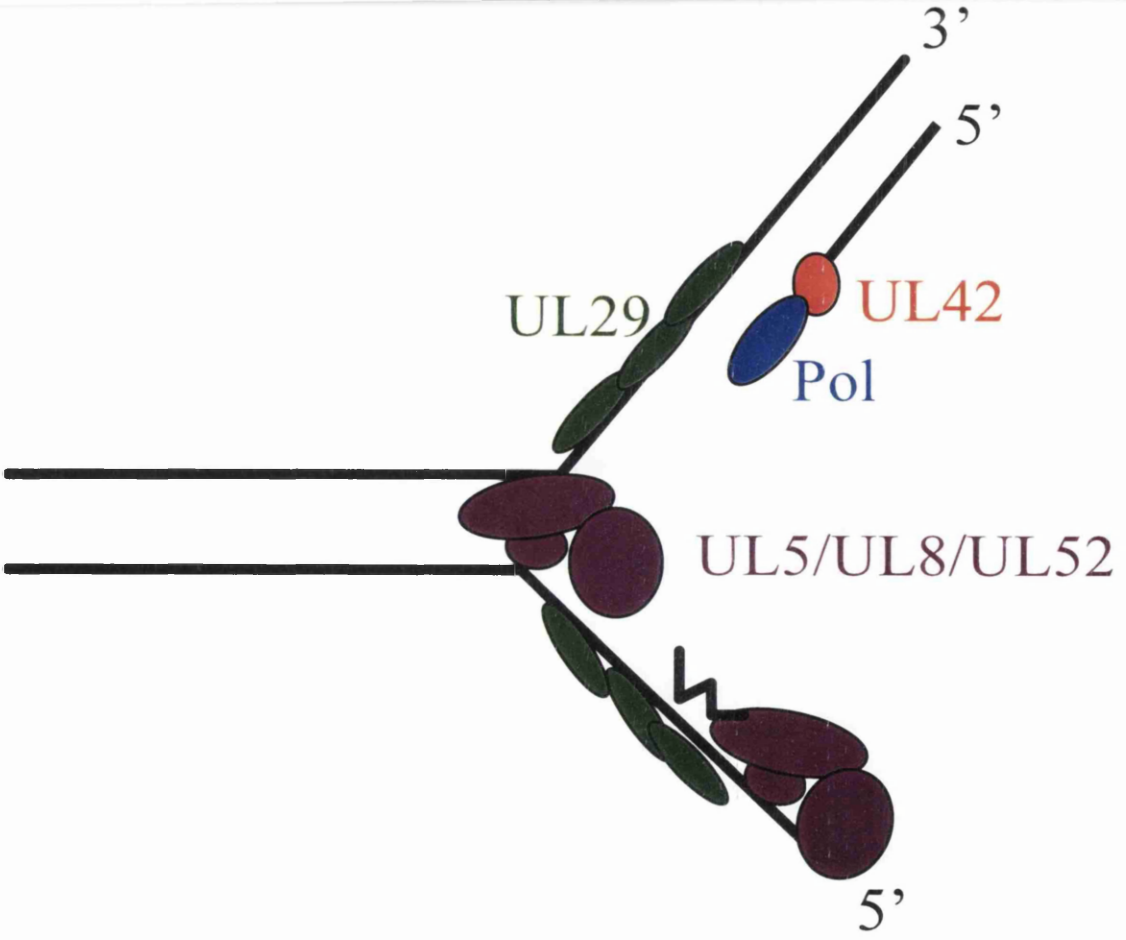


Figure 1.4
DNA replication as proposed by Challberg (1991). The DNA polymerase exists as a heterodimer of the UL30 and UL42 polypeptides. UL29 is a DNA binding protein which unwinds the parental duplex DNA thus facilitating the use of these strands as templates for the DNA polymerase. The UL5/UL8/UL52 complex functions both as a primase and a helicase and it appears to prime lagging strand synthesis as it unwinds DNA at the replication fork.

classes, immediate early (also known as α or IE), early (also known as β; delayed-early or E) and late (also known as γ or L) (Clements *et al.*, 1977, Honess & Roizman, 1974). This classification system is based on the kinetics of their expression and requirements for ongoing DNA synthesis. Five IE proteins (reviewed by Roizman & Sears, 1996) are produced from transcripts that can be expressed in the absence of viral protein synthesis and four of these regulate expression of early and late viral genes. Early proteins are involved in DNA metabolism and DNA replication, their expression signalling the onset of viral DNA synthesis. Expression of late viral genes can be divided into two phases based on the stringency of their requirement for DNA synthesis; the leaky late class of genes (γ_1) expressed prior to DNA replication, and the strict late genes (γ_2) expressed once viral DNA replication has commenced. These proteins comprise the majority of virus structural components. HSV-1 transcribes some 80 genes but only four of the virus pre-mRNAs contain introns.

The HSV-1 genome is estimated to encode at least 80 unique proteins. Table 1.2 lists the functions or properties of the proteins encoded by the HSV-1 genes that have been identified to date.

(b) Systems of HSV-1 protein nomenclature

There are a number of systems that have been designed to describe the proteins of HSV-1. These include:-

(i) the order of the open reading frames (ORFs) occurring in the various regions of the genome e.g. UL19.

(ii) a system based on protein function e.g. ribonucleotide reductase.

(iii) a system based on the apparent molecular weight of the protein as determined by migration of viral proteins in denaturing polyacrylamide gels e.g. viral protein of molecular weight 63 kDa (Vmw63) or immediate-early 63 (IE63). This system is known as the Glasgow system.

(iv) The Chicago or Infected Cell Protein (ICP) system. This is distinct from the Glasgow system and represents a number corresponding to the location on a polyacrylamide gel rather than the actual molecular weight of the protein e.g. ICP27 corresponds to IE63.

In this thesis, the IE proteins have been referred to using the Glasgow system. Table 1.3 shows the corresponding Chicago system nomenclature for the immediate-early genes. The majority of other proteins have been referred to on the basis of their ORF location or function.

Table 1.2. FUNCTIONS AND PROPERTIES OF HSV-1 PROTEINS

Gene	Function of Protein
RL1	Neurovirulence factor (ICP34.5)
RL2	IE protein; modulator of cell state and gene expression (ICP0, Vmw110)
LAT	Latency associated transcripts
UL1	Glycoprotein L; complexes with glycoprotein H (UL22)
UL2	Uracil-DNA glycosylase
UL3	Function unknown
UL4	Function unknown
UL5	Component of DNA helicase-primase complex; possesses helicase motifs
UL6	Minor capsid protein
UL7	Function unknown
UL8	Component of DNA helicase-primase complex
UL8.5	Encodes a 486 amino acid protein; potential role in DNA replication
UL9	Ori-binding protein essential for DNA replication
UL9.5	Function unknown
UL10	Virion surface glycoprotein M
UL11	Myristylated tegument protein; role in virion envelopment
UL12	Deoxyribonuclease; role in maturation/packaging of DNA
UL13	Tegument protein, protein kinase
UL14	Function unknown
UL15	Role in DNA packaging, putative terminase component
UL16	Function unknown
UL17	Function unknown
UL18	Capsid protein (VP23); component of intercapsomeric triplex
UL19	Major capsid protein (VP5); forms hexons and pentons
UL20	Integral membrane protein; role in egress of nascent virions; host range phenotype; syn locus
UL21	Tegument protein
UL22	Virion surface glycoprotein H; complex with glycoprotein L (UL1)l; role in cell entry
UL23	Thymidine kinase
UL24	Function unknown, syn locus
UL25	Capsid-associated tegument protein
UL26	Protease, acts in virion maturation; N-terminal portion is capsid protein (VP24)
UL26.5	Internal protein of immature capsids (VP22a); processed by UL26 protease
UL27	Virion surface glycoprotein; role in cell entry; syn locus
UL27.5	ORF is antisense to UL27; function unknown
UL28	Role in DNA packaging
UL28.5	Function unknown
UL29	Single-stranded DNA-binding protein (ICP8)
UL30	Catalytic subunit of replicative DNA polymerase; complexes with UL42 protein

Gene	Function of Protein
UL31	Function unkown
UL32	Function unknown
UL33	Role in DNA packaging
UL34	Membrane-associated phosphoprotein; substrate for US3 protein kinase
UL35	Capsid protein (VP26); located on tips of hexons
UL36	Very large tegument protein
UL37	Tegument protein
UL38	Capsid protein (VP19C); component of intercapsomeric triplex
UL39	Ribonucleotide reductase large subunit (ICP6, Vmw136, R1)
UL40	Ribonucleotide reductase small subunit (Vmw38, R2)
UL41	Tegument protein, host-off factor
UL42	Subunit of replicative DNA polymerase; increases processivity; complexes with UL30 protein
UL43	Function unknown, probable integral membrane protein
UL43.5	ORF is antisense to UL43. Localizes with capsid proteins in the nucleus; may play a role in capsid assembly
UL44	Virion surface glycoprotein C; role in cell entry
UL45	Tegument/envelope protein
UL46	Tegument protein; modulates IE gene transactivation by UL48 protein
UL47	Tegument protein; modulates IE gene transactivation by UL48 protein
UL48	Tegument protein; transactivates IE genes (VP16, Vmw65, α-TIF)
UL49	Tegument protein
UL49A	Envelope protein disulphide-linked to tegument
UL50	Deoxyuridine triphsophatase
UL51	Function unknown
UL52	Component of DNA helicase-primase complex
UL53	Glycoprotein K
UL54	IE protein; post-translational regulator of gene expression (ICP27, Vmw63)
UL55	Function unknown
UL56	Function unknown
RS1	Nucleotidylated, polyADP-ribosylated immediate early phosphoprotein (Vmw175, ICP4) regulates positively most early and late genes and negatively itself and RL2. It binds to DNA in sequence specific manner
US1	Nucleotidylated immediate early protein (Vmw68, ICP22); function unknown; host range phenotype
US1.5	Gene 3' co-terminal with US1
US2	Function unknown
US3	Protein kinase; phosphorylates UL34
US4	Virion surface glycoprotein G; function unknown
US5	Virion surface glycoprotein J; function unknown
US6	Virion surface glycoprotein D; role in cell entry and cell-cell spread

Gene	Function of Protein
US7	Virion surface glycoprotein I; complexed with glycoprotein E (US8) to form Fc receptor; role in cell-cell spread
US8	Virion surface glycoprotein E; complexed with glycoprotein I (US7) to form Fc receptor; role in cell-cell spread
US9	Virion protein; function unknown
US10	Virion protein; function unknown
US11	Abundant tegument protein; localizes to the nucleolus and ribosomes in infected cells; RNA binding protein
US12	Immediate early protein (Vmw12, ICP47); role in preventing antigen presentation
LAT	Latency-associated transcript; probably not protein-coding
ORF P	Encodes a 248 amino acid protein of unknown function. ORF is antisense to RL1. The gene is expressed under conditions in which Vmw175 is not functional
ORF O	ORF is partially antisense to RL1. The protein is expressed under conditions in which Vmw175 is not functional.

Table 1.3. NOMENCLATURE OF HSV-1 IE GENES AND PROTEINS

Gene		Protein	
Glasgow	**Chicago**	**Glasgow**	**Chicago**
IE1	α0	IE110 (Vmw110)	ICP0
IE2	α27	IE63 (Vmw63)	ICP27
IE3	α4	IE175 (Vmw 175)	ICP4
IE4	α22	IE68 (Vmw68)	ICP22
IE5	α47	IE12 (Vmw12)	ICP47

(c) Promoter sequences

Transcription of the immediate-early genes proceeds from enhanced promoters by interaction of the virion protein α-TIF (also known as Vmw65 or VP16) (Campbell *et al.*, 1984) with cellular transcription factors (Preston *et al.*, 1988) and does not require *de novo* protein synthesis in the infected cell (reviewed by O'Hare, 1993). The complex containing α-TIF activates gene transcription by recognition of 'TAATGARAT' sequences upstream of immediate early gene promoters (Gaffney *et al.*, 1985). The basic promoter for all early viral transcripts is a RNA polymerase II promoter similar to those that exist in a number of eukaryotic genes and other DNA-containing viruses. The promoters controlling the leaky late transcripts also include an 80-120 bp pol II promoter element, but some have transcriptional regulatory elements near and downstream of the cap site (Wagner *et al.*, 1995). The functional portion of 'strict' late promoters extends 5' of the cap site 25-30 bases to a 'TATA' box, but elements at and 3' of the cap site are also present.

(d) Proteins that function in the regulation of viral transcription

The α-TIF protein (VP16, or Vmw65) has a relative molecular weight of 65,000 Daltons and is comprised of two separable functional domains (reviewed by Flint & Shenk, 1997). The N-terminal 80% interacts with the cellular transcription factor Oct-1 and at least one additional cellular protein to form a DNA-protein complex which binds to the consensus

'TAATGARAT' upstream of immediate-early gene promoters (Gaffney *et al.*, 1985). The C-terminal 20% of the protein is extremely acidic and acts as a generalised 'acidic blob' type of transcriptional activator. Although the transactivation function of the α-TIF protein is dispensable at least in tissue culture, the protein probably has a role in the extremely rapid lytic replication cycle of HSV.

Five HSV genes IE175, IE110, IE68, IE12 and IE63 are expressed at immediate-early times post infection and function in the earliest stages of the lytic cycle. With the exception of IE12, these proteins appear to have a role in the regulation of viral gene expression.

The protein encoded by the IE175 transcript is a large multifunctional protein that has a predicted unmodified molecular weight of 140,000 and possesses an autoregulatory function (Muller, 1987). IE175 is strictly required for virus replication and is continuously required for transcription of all but the immediate-early class of viral mRNA (Dixon & Schaffer, 1980, Watson & Clements, 1980). IE175 acts as a transcription factor that either represses or activates transcription through contacts with the general transcriptional machinery. It can be located in the nucleus of infected cells by immunofluorescence and multiple phosphorylated forms of the protein have been identified (Pereira *et al.*, 1977, Wilcox *et al.*, 1980). The protein has a sequence-specific DNA-binding ability, with preference for the consensus ATCGTCN$_4$YGCRC.

The nonessential nuclear IE110 protein (primary predicted molecular weight 79,000 Da) enhances the transcriptional activation of IE175 in transient assays and can act as a powerful transcriptional activator on some viral promoters by itself (Everett, 1984, O'Hare & Hayward, 1985, Quinlan & Knipe, 1985). It also facilitates the reaction of virus from latency in mice (Cai *et al.*, 1993, Lieb *et al.*, 1989). HSV-1 mutants deleted for the IE110 gene grow poorly in cell culture, particularly at low multiplicities of infection (MOIs) but remain viable (Sacks & Schaffer, 1987, Stow & Stow, 1986).

Much less is known about the IE68 protein. It is dispensable for virus replication in cell culture, but is required for HSV replication in certain types of cultured cells (Post & Roizman, 1981, Purves *et al.*, 1993) and may have a role in maintaining the virus's ability to replicate in a wide range of cells in the host. This protein also appears to be required for modification of RNA polymerase II and establishment of viral transcription (Rice *et al.*, 1995). The IE12 protein has a role in inhibiting the presentation of peptide fragments of viral proteins at the cell surface by the MHC system, and thus acts to repress recognition of virus-infected cells (York *et al.*, 1994). Finally, the structure and function of the IE63 protein which is the subject of this thesis is described in detail in Section 1.2.

HSV regulatory proteins from other kinetic classes also exist. These include three leaky late proteins: the UL46 and 47 proteins which modulate the activity of α-TIF by unknown mechanisms, and the virion-associated host shut-off protein (vhs, UL41) which functions by generally destabilising all

mRNA molecules in the infected cell. The major DNA-binding protein (UL29) which has an essential role in viral DNA replication, also appears to function in gene regulation by mediating the shut-off of transcription of at least some early genes and activating expression of late genes (Roizman & Sears, 1996).

1.1.6 Virion host shutoff function

Infection of cells with HSV-1 results in the shutoff of host protein synthesis. The inhibition of host protein synthesis is a multiphase process. In brief, a primary phase of the shutoff of host protein synthesis is mediated by the virion component known as the virion-associated host shutoff protein (vhs) (UL41) and at least one other virion protein (reviewed in Roizman & Sears, 1996). This results in dissociation of mRNA from the polysomes and mRNA degradation. A late (secondary) shutoff function reduces the remaining levels of host protein synthesis and requires the presence of viral IE proteins such as IE68 and IE63. The mode of action of the secondary shutoff function is not, as yet, understood.

1.1.7 Posttranslational modification of HSV-1 proteins

The majority of HSV proteins examined to date appear to have extensive post-translational modifications. Such modifications include cleavage (e.g. UL26), phosphorylation (a number of HSV-1 proteins are phosphorylated including the IE proteins IE175, IE110, IE68 and IE63), sulphation

(glycoproteins), glycosylation (glycoproteins), myristylation (UL11) and poly(ADP-ribosyl)ation (IE175, UL18) (Roizman & Sears, 1996).

Unusually, only a relatively small proportion of HSV mRNAs (IE110, IE68, IE12, UL15) have been shown to be derived by splicing.

1.2 THE ESSENTIAL HSV-1 PROTEIN IE63

IE63 is required for early and late viral gene expression (McCarthy *et al.*, 1989, Rice & Knipe, 1990, Sacks *et al.*, 1985). How IE63 activates the expression of early genes is not known but it appears that IE63 carries out two distinct functions which contribute to late gene activation. First, IE63 increases the level of viral DNA replication up to 10-fold (Rice & Knipe, 1990), possibly by transactivating a subset of E genes involved in DNA replication (McGregor *et al.*, 1996, Uprichard & Knipe, 1996) or by altering the redox state and subsequent folding and/or localisation of the essential DNA replication protein ICP8 (Curtin & Knipe, 1993). Second, IE63 enhances L gene expression by a mechanism that does not involve the stimulation of viral DNA replication (Hibbard & Sandri-Goldin, 1995, Rice & Knipe, 1990, Rice & Lam, 1994).

Although the exact mechanism(s) by which IE63 regulates viral and also cellular genes during HSV-1 infection is unknown, evidence suggests that the protein exerts its effects both at the level of transcription and via a number of posttranscriptional control mechanisms.

1.2.1 IE63 can affect the transcription of viral and cellular genes during infection

Studies investigating the effect of IE63 mutants on the expression of a number of HSV-1 immediate-early, early and late genes have shown that the expression of the early glycoprotein B (gB) gene is increased significantly by IE63(Rice & Knipe, 1988). Several other early and late genes such as ICP5 show lesser dependence on IE63 expression. This effect appears to act at the level of transcription since when the gB gene is fused to the reporter gene encoding chloramphenicol acetyltransferase (CAT), induction of CAT activity correlates with an increase in the amount of CAT transcripts initiated from the transcriptional start site of the gB gene. However, increased stability of mRNA in the presence of IE63 may occur.

Examination of transcription patterns of genes in cells infected with mutant and wt HSV-1 suggests that IE63 effectively down-regulates transcription of the cellular GAPDH and histone H2b genes (Pangiotidis *et al.*, 1997). However, c-myc, c-fos, and γ-actin genes are repressed transcriptionally by both wild-type virus and viruses with mutations in IE63. The repression of histone H2b gene transcription requires IE63 yet this gene does not contain introns and its mRNA is not polyadenylated. This suggests that IE63 is not exerting control in this instance via its previously characterised posttranscriptional mechanisms (see Section 1.2.2) but that it may indeed have a role in control of gene expression at the transcriptional level.

IE63 may also exert an effect on the transcription of gene expression by interacting with and modulating the principal viral gene regulator IE175. Transient expression systems have shown that IE63 can increase or decrease the ability of IE175 to activate HSV-1 genes (Su & Knipe, 1989) and GST fusion assays have suggested that IE175 and IE63 interact directly (Pangiotidis *et al.*, 1997). Although the level of expression of IE175 does not appear be altered by IE63, an alteration in the electrophoretic mobility of IE175 is observed in the presence of IE63. Since differently phosphorylated forms of IE175 have been shown to have different DNA binding activities it is plausible that IE63 either directly or indirectly alters the phosphorylation of IE175 leading to the different promoter-binding properties of the IE175 protein (Michael *et al.*, 1988).

Finally, IE63 appears to have a direct effect on RNA polymerase III transcription *in vitro* and this effect is mediated by the transcription factor TFIIIC (Jang & Latchman, 1992). However, the significance of this observation is as yet unclear since no HSV-1 genes are known to be transcribed by RNA polymerase III.

1.2.2 IE63 acts at the posttranscriptional level to regulate gene expression

(a) IE63 modulates virus poly(A) site usage

(i) *Polyadenylation is an essential step in the generation of mature eukaryotic mRNA*

An essential step in the generation of mature eukaryotic mRNAs is the formation of a specific 3' terminus. In the majority of mRNAs this requires a sequence-specific endonucleolytic cleavage followed by the addition of ≤250 adenylate residues to the newly generated 3' end. The purposes of this polyadenylation appears to be: (1) to aid the export of mature mRNA from the nucleus ; (2) to increase the stability of at least some mRNAs in the cytoplasm; and (3) for efficient translation of the mRNA (Atwater *et al.*, 1990, Bachvarova, 1992, Jackson & Standart, 1990, Wickens, 1990). Many genes encode multiple polyadenylation sites and in some mammalian and viral systems, alternative or regulated usage of poly(A) sites is used as a means of controlling gene expression .

Two sequence elements that specify pre-mRNA cleavage and poly(A) addition have been identified. The sequence AAUAAA is located 10-30 bases upstream of the cleavage/polyadenylation site and the downstream sequence a G+U/U-rich motif is located 20-40 bases downstream of the cleavage site (McLauchlan *et al.*, 1989). These two sequences and their relative spacing with one another determine the site and strength of poly(A)

signal (Chen *et al.*, 1995, MacDonald *et al.*, 1994), with pre-mRNA cleavage occurring after an 'A' located between them.

Figure 1.5 shows the proteins required for cleavage and polyadenylation of mRNA 3' ends. The proteins required for recognition of a poly(A) signal and cleavage of the RNA include the Cleavage Poly(A) Specificity Factor (CPSF) and Cleavage Stimulatory Factor (CstF) together with two Cleavage Factor proteins (CFI and CFII). Poly(A)Polymerase (PAP) adds poly(A) to the newly formed 3' end and this activity requires RNA-bound CPSF and another polypeptide, Poly(A) Binding Protein II (PABII), to assist in making the full poly(A) tail. The 160 kDa protein of CPSF binds to the AAUAAA sequence and the 64 kDa protein of CstF binds to the G+U/U-rich sequence.

(ii) *IE63 increases RNA 3' processing at inherently weak poly(A) sites*

HSV-1 transcripts which have common 5' termini but different poly(A) sites have been identified (McGeoch *et al.*, 1988). This observation raises the possibility that HSV-1 gene expression could be influenced by the efficiency of poly(A) site usage. Studies using precursor RNAs containing 3' processing sequences from different HSV genes have shown that an activity which was named late processing factor (LPF) and which is present in HSV-1 infected nuclear extracts can stimulate RNA 3' processing specifically at a late HSV poly(A) site of the UL38 gene (McLauchlan *et al.*, 1989). Increased usage of the late UL38 poly(A) site was shown to be dependent upon IE63 expression alone.

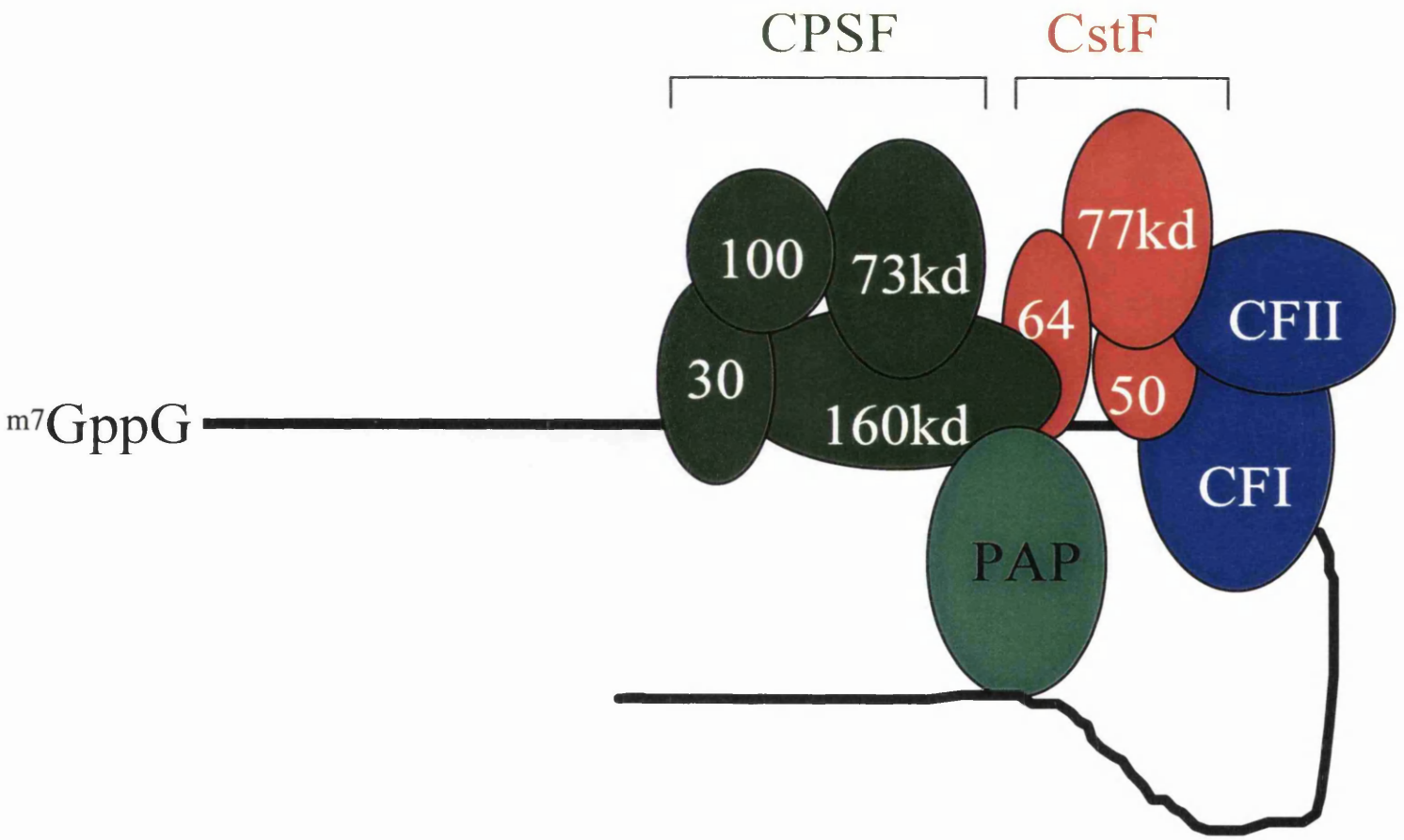


Figure 1.5
Assembly on pre-mRNA of the 3' processing complex, showing the factors required for cleavage and polyadenylation. CPSF is cleavage and polyadenylation specificity factor, CstF is cleavage stimulation factor, PAP is poly(A) polymerase and the two poorly characterised cleavage factors CF1 and CFII are shown in blue. ^{m7}GpppG is the trimethyl cap.

This work has been extended by McGregor *et al* (McGregor *et al.*, 1996) who showed that the 3' processing of the two late genes UL38 and UL44 was increased by an IE63-induced activity whereas there was no detectable change in the processing of a number of immediate-early and early transcripts in the presence of IE63 expression. *In vitro* polyadenylation assays determined that the 3' processing sites of the late genes that were tested were predominantly weak and that the 3' processing sites of the immediate-early and early genes were relatively strong.

Previous UV cross-linking studies have shown CPSF and CstF bind directly to pre-mRNA templates and the binding of CstF is though to determine the efficiency of polyadenylation (Takagaki *et al.*, 1989). Using HSV-1 infected cell extracts, UV cross-linking of the 64 kDa CstF component to poly(A) sites of all temporal classes is increased and this binding appears to be dependent upon the expression of IE63 (McGregor *et al.*, 1996).

These findings suggest that IE63 increases 3' processing at inherently weak poly(A) sites, perhaps by increasing the binding of factors required for processing. Immediate-early and early poly(A) sites which are efficiently used are unaffected by this increased binding of processing factors. Conceivably, IE63 could regulate the switch between early and late gene expression by increasing CstF binding at late times during infection when the activity of such proteins is likely to be significantly reduced.

(b) IE63 inhibits pre-mRNA splicing

(i) *Mechanisms of mammalian pre-mRNA splicing*

Splicing is the process by which non-coding sequences, or introns, are removed from pre-mRNA molecules and exon sequences are 'spliced' together. The splicing reaction is catalysed by the spliceosome, a complex made up of five small nuclear ribonucleoprotein particles (snRNPs) and more than 50 non-snRNP proteins (Will & Luhrmann, 1997). Spliceosomes assemble in a highly ordered and stepwise manner (see Figure 1.6). Assembly begins with the association of the U1 snRNP with pre-mRNA. Next the U2 snRNP and then the U4/U6.U5 triple snRNP bind, U1 and U4 become destabilised and the spliceosome is activated for catalysis. The splicing reaction takes place within the spliceosome in two distinguishable steps via a branched lariat intermediate. In the initial step, the 2' hydroxyl of a conserved, intronic adenosine attacks the phosphate at the 5' splice site, producing a free 5' exon and branched species, termed the lariat intermediate. In the second reaction, the 3' hydroxyl of the 5' exon attacks the phosphate at the 3' splice site, yielding ligated mRNA and a lariat intron.

In the nucleus of mammalian cells, splicing factors are preferentially accumulated in 20-40 nuclear domains or speckles when visualised by fluorescence microscopy (Figure 1.7A). In some cell types, the snRNP proteins localize at small nuclear structures termed coiled bodies. The majority of speckles correspond to interchromatin granule clusters (IGCs).

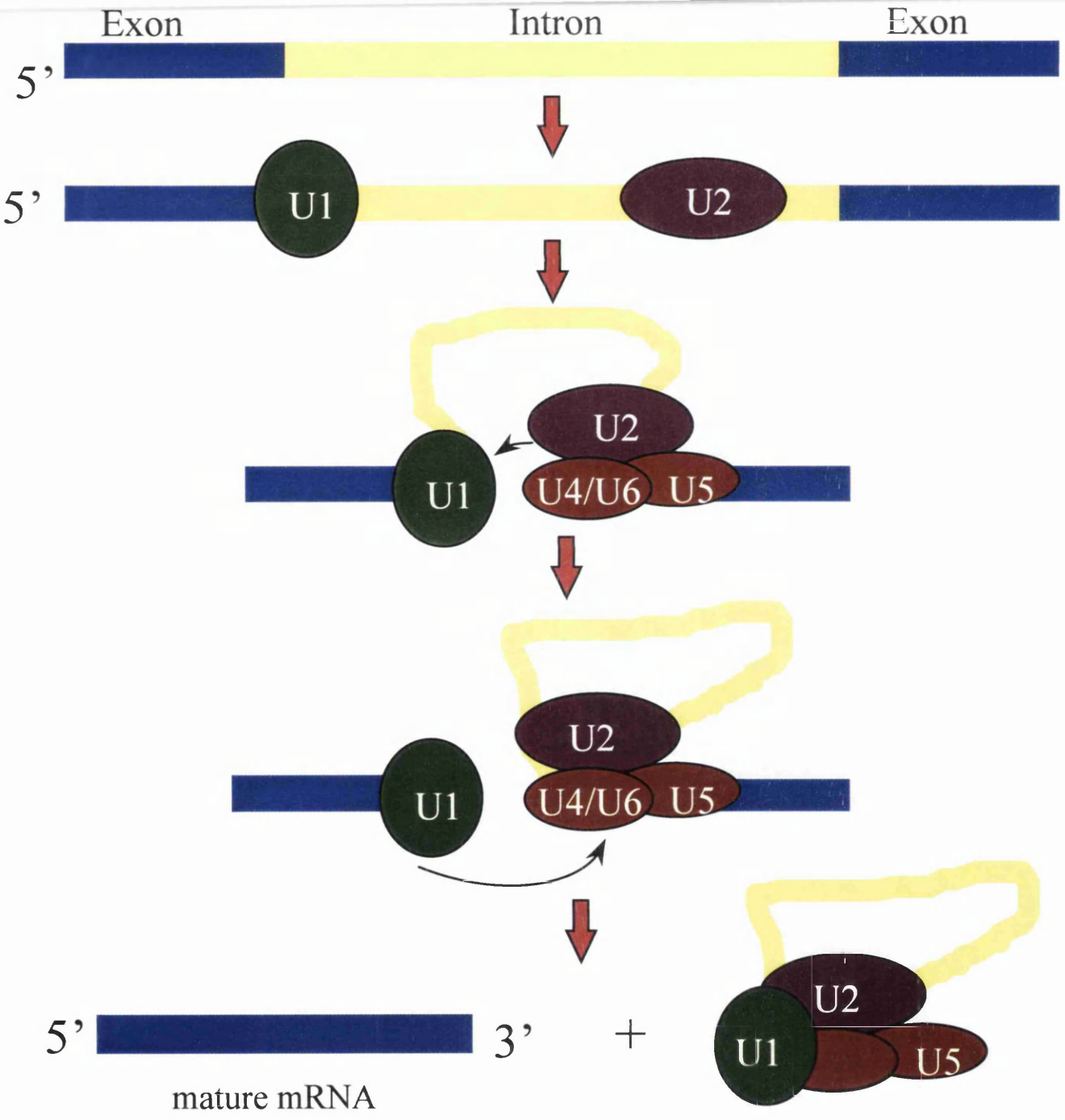


Figure 1.6
The RNA splicing mechanism

U1 and U2 snRNPs bind to 5' splice site and branch site respectively

Branchpoint A nucleotide in the intron sequence attacks the 5' splice site and cleaves it

The cut 5' end of the intron sequence becomes covalently linked to this A nucleotide

The ends of the exons are joined cleaving the RNA molecule at the 3' splice site. The intron sequence is released as a lariat

Fibrillar structures, called perichromatin fibrils (PFs) extend from the periphery of IGCs. It is thought that IGCs may function as storage/assembly sites for splicing factors and that factors are recruited from IGCs to PFs when necessary.

Splicing is thought to be regulated by the SR proteins (Fu, 1995). These proteins appear to be required for splice site selection and spliceosome assembly and are so named because they contain a carboxy-terminal serine/arginine rich domain which mediates interactions with other SR proteins and facilitates the recruitment and stabilisation of snRNPs with pre-mRNA during splicing. SR proteins interact with RNA enabling them to bind to specific splicing enhancer or repressor sequences (reviewed in Manley & Tacke, 1996).

ASF/SF2 is a typical SR protein which has constitutive splicing activity *in vitro* (Krainer *et al.*, 1990) and influences splice site selection. ASF/SF2 interacts with at least two other SR proteins, the U1 specific protein SC35 and the essential splicing factor U2AF. ASF/SF2 interacts with RNA via an RNP domain and is thought to be typical of SR protein control of splicing.

(ii) *IE63 expression causes the redistribution of splicing factors and inhibits host cell splicing*

Upon HSV-1 infection the snRNPs and the non-snRNP protein SC35 redistribute within the nucleus to form clusters which colocalize with the IGCs (Martin *et al.*, 1987, Sandri-Goldin *et al.*, 1995). These clusters

condense and migrate to the nuclear periphery as infection proceeds (depicted in Figure 1.7B). Subsequent studies have shown that the IE63 protein is both necessary and sufficient for this effect (Phelan *et al.*, 1993, Sandri-Goldin *et al.*, 1995) and that IE63 colocalises with the redistributed snRNPs at later times post-infection (Phelan *et al.*, 1993, Sandri-Goldin *et al.*, 1995).

Sandri-Goldin *et al* have further shown that IE63 coimmunoprecipitates with anti-Sm antisera during wild-type HSV-1 infection (Sandri-Goldin & Hibbard, 1996). Since Sm is an epitope which is common to a number of snRNPs it was not possible to identify any specific RNP which was coimmunoprecipitating with IE63. The snRNPs and SC35 redistribution may reflect a movement from the active splicing sites to inactive storage centres.

IE63 viral mutants are defective in host shutoff (Hardwicke & Sandri-Golding, 1994). While the majority of mammalian genes contain introns, HSV-1 encodes relatively few spliced products and these gene products, (IE110, IE68, IE12 and UL15), are predominantly expressed at immediate-early times post-infection. Therefore, it is plausible that the IE63-induced redistribution of splicing factors which occurs quite early in infection could contribute to host shutoff whilst having little or no detrimental effect on viral gene expression. Hardwicke and Sandri-Goldin (Hardwicke & Sandri-Goldin, 1994) have analysed the levels of spliced cellular mRNAs during HSV-1 infection and found that during infection with IE63 viral mutants, the levels of at least 3 spliced host mRNAs were higher than those seen with wild type infected cells. These effects could not be attributed to effects on transcription or RNA stability. Further studies have investigated the

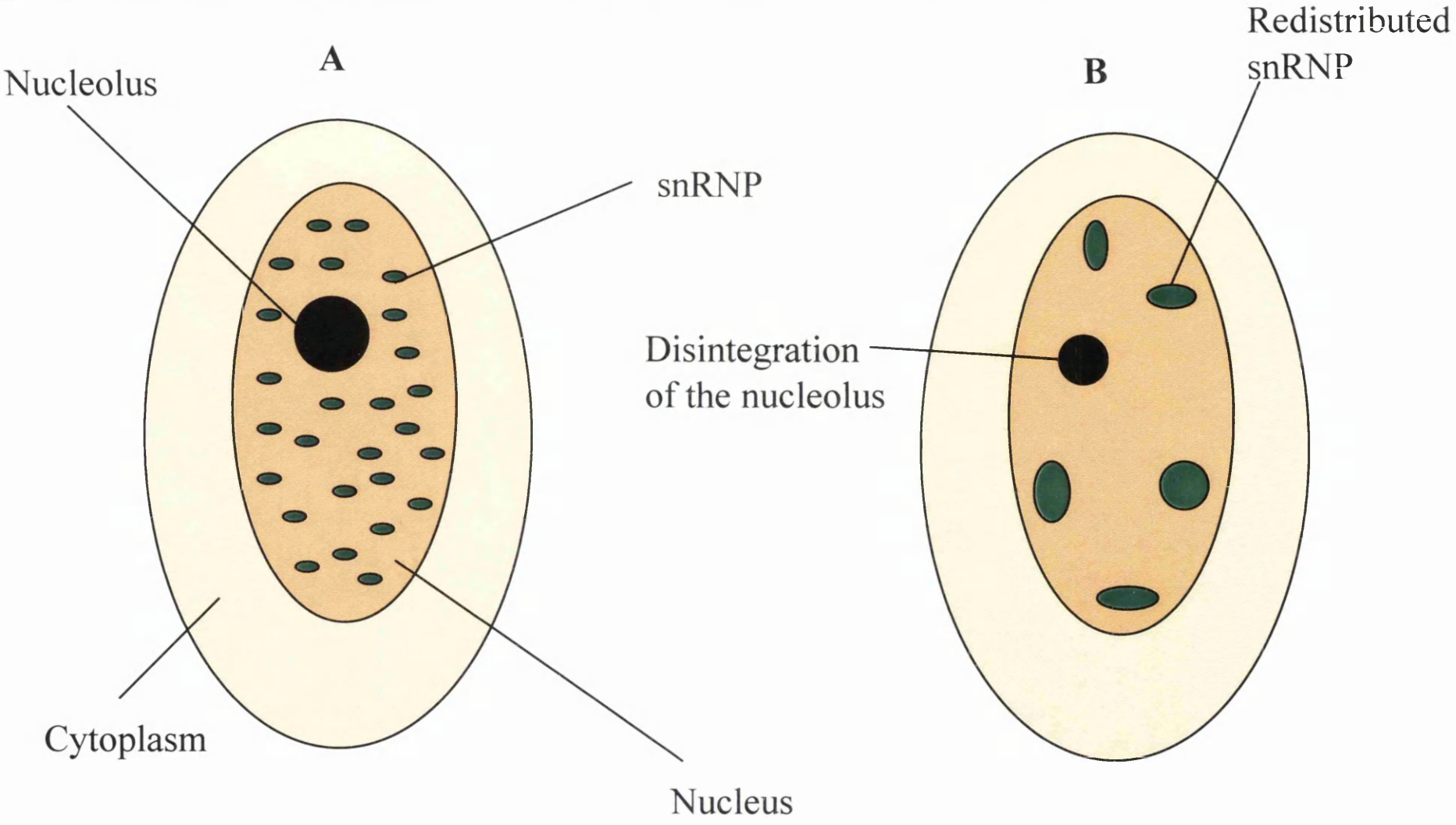


Figure 1.7
Redistribution of snRNPs upon HSV-1 infection. (A) uninfected cell, (B) redistribution of snRNPs and disintegration of nucleolus following HSV-1 infection, labelled with an anti-snRNP antibody.

consequences of the proposed impaired splicing efficiency on the nuclear and cytoplasmic accumulation of two viral transcripts which contain introns; an immediate early gene product (IE110) and a late gene product (UL15) (Hardy & Sandri-Goldin, 1994). These studies showed that pre-mRNA accumulated in the nucleus of the cell suggesting that splicing and not transport of the gene products was being affected. Further, these studies showed that splicing was inhibited in extracts from wt infected cells but not in extracts from cells infected with an IE63 null mutant.

An IE63 repressor mutant with a base substitution at aa 480 has been characterised and this mutant showed a redistributed snRNP staining pattern but splicing was not inhibited (Sandri-Goldin *et al.*, 1995). This suggests that although the redistribution of snRNPs and SC35 correlates with the impairment of splicing, these alterations do not fully inhibit splicing and sufficient snRNPs may be left distributed throughout the nucleoplasm to perform the splicing of UL15 and other cellular intron-containing transcripts.

Studies have been carried out to determine the effect IE63 could have on the splicing proteins which might cause their redistribution (Sandri-Goldin & Hibbard, 1996). These studies showed that there was an increase in the phosphorylation state of two cellular proteins during wt HSV infection, one of which is likely the U1 snRNP 70 kDa protein. This effect was lost when an IE63 null mutant was used. The movement of splicing factors between active and storage sites has been shown to be dependent upon their phosphorylation state (Mermoud *et al.*, 1994). Therefore it is plausible that IE63 causes snRNP redistribution by directly or indirectly causing a change

in the phosphorylation state of the snRNP molecules. It appears that this redistribution is not sufficient to cause a total inhibition of splicing but it may complement a more direct mechanism such as disruption of spliceosome complex formation.

It is of interest to note that infections with adenovirus and influenza virus also cause a redistribution of snRNPs (Bridge *et al.*, 1993, Fortes *et al.*, 1995). Adenovirus RNAs are extensively spliced and the splicing factors are recruited into virus transcription and processing sites; in contrast, influenza virus has a negligible requirement for splicing, indicating that redistributed snRNPs in influenza-virus infected cells may represent dysfunctional aggregates.

Finally, it is of interest to note that the Wilms' tumour gene product, which has at least sixteen isoforms and is thought to be a transcriptional regulator acting at both the transcriptional and post-transcriptional levels (reviewed in Ladomery, 1997), also colocalises with the snRNPs suggesting a role in splicing. Furthermore, the different isoforms of the Wilm's tumour gene product appear to have varying properties raising the possibility that different phosphoisoforms of IE63 may also act in different ways or exist at different cellular locations.

(c) IE63 and nucleocytoplasmic transport of mRNAs

(i) *Nuclear import and export of proteins and mRNA*

In eukaryotic cells, the nucleus is separated from the cytoplasm by a double membrane system known as the nuclear envelope. Therefore the transport of various RNA and protein molecules between their sites of synthesis and function is required, and this allows the regulation of certain cellular processes by control of the nucleocytoplasmic distribution of these molecules.

Transport between the cytoplasm and the nucleus occurs through protein-lined aqueous channels called nuclear pore complexes (NPC). A diagrammatic representation of a nuclear pore is shown in Figure 1.8. The proteins that make up the NPC are called nucleoporins and they have a combined molecular mass of approximately 125 MDa (Rout & Wente, 1994). Current models for nucleocytoplasmic transport suggest that: (1) distinct types of cargo contain specific molecular signals; (2) translocation of cargo through the NPC is mediated by shuttling carrier proteins and; (3) most if not all transport processes require the small GTPase Ran.

The specific sequence motifs for import are called nuclear localization sequences (NLSs), and for export, nuclear export signals (NESs). The first NLS to be identified was a short stretch of basic amino acids (PKKKRKV) in simian virus 40 (SV40) large T-antigen (Kalderon *et al.*, 1984). Similar

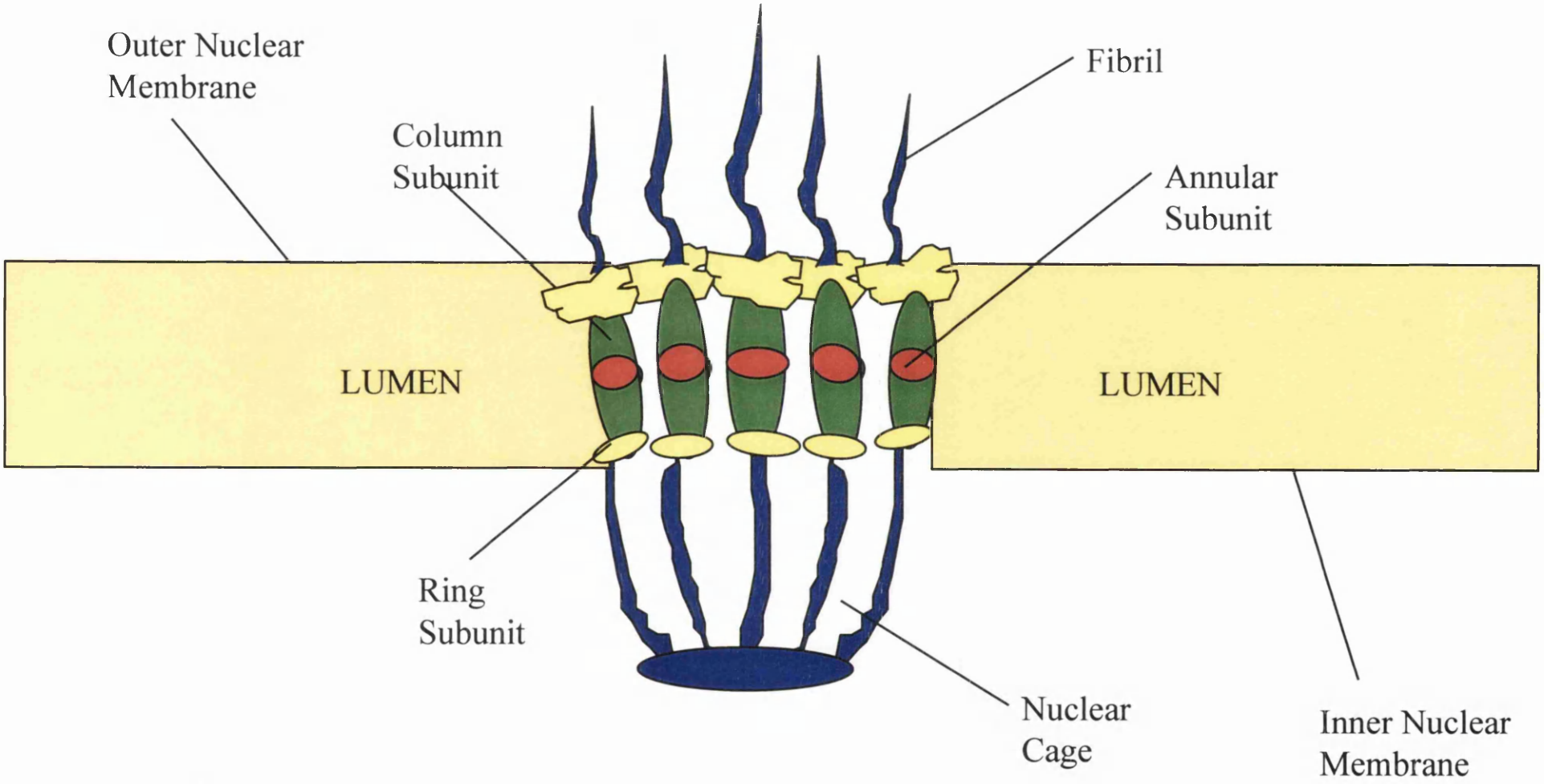


Figure 1.8
A cut-away model of a nuclear pore complex. The nuclear pore complex is composed of three parts, a column component which forms the bulk of the pore wall, an annular component which extends 'spokes' towards the centre of the pore and a lumina component which is formed by a large transmembrane glycoprotein which helps anchor the complex to the nuclear membrane.

sequences were later found in a large number of viral and cellular proteins. These import signals are commonly referred to as 'classical' NLSs as they were the first to be discovered. Recently, two NLSs have been identified during studies of heterogenous nuclear RNPs (hnRNPs). A subset of these proteins, including hnRNPs A1, A2, D, E, I and K, shuttle continuously between the nucleus and the cytoplasm. For hnRNP A1 the shuttling signal is present in a glycine-rich segment of 40 amino acids (termed M9) (Siomi & Dreyfuss, 1995) and for hnRNP K the signal is an entirely different 40 amino-acid signal (termed KNS) (Michael *et al.*, 1997a). The snRNPs also have a unique signal that is composed of the Sm core protein and the trimethylguanosine cap (Fischer *et al.*, 1991, Hamm & Mattaj, 1990). Finally, Truant *et al.* have recently shown that the human Tap protein which mediates the sequence-specific nuclear export of RNAs contains an NLS which is transportin-specific but that lacks nuclear export signal function (Truant *et al.*, 1999).

Two types of NESs have been identified. First, the nuclear import sequences in hnRNP A1 and K also function as NESs (Michael *et al.*, 1997a). Second, sequences rich in leucine residues act as NESs in the HIV-1 Rev molecule (LPPLERLTL) (Bogerd *et al.*, 1995, Fischer *et al.*, 1995) and in the cAMP-dependent protein kinase inhibitor, PKI (LALKLAGLDI) (Wen *et al.*, 1995). These leucine-based export signals have been identified in a variety of exported proteins.

Four independent protein import pathways have been identified which rely on interaction with carrier proteins called importins although some proteins without an NLS can be carried 'piggy back' through the nuclear pore complex via strong interaction with NLS-containing proteins. The importins that have been identified to date are listed in Table 1.4. The importin α/β complex recognises classical NLSs and the transportins mediate import of hnRNPs that contain a non-classical NLS. Independent pathways that are used to transport mRNA-binding proteins also exist but the NLSs responsible are so far unknown.

The export carrier proteins that have been identified to date are also shown in Table 1.4. The first export carrier to be identified was a nuclear protein termed RIP1 (also known as Rab) (Bogerd *et al.*, 1995, Fritz *et al.*, 1995). A family of importin related molecules has recently been identified and shown to be required for the export of proteins containing leucine-based NESs (Neville *et al.*, 1997). These exportins may mediate the protein export of a wide variety of proteins. For the hnRNP-based export pathway it is likely that transportin, the receptor for nuclear import also mediates export.

(ii) *IE63 affects the nuclear export of virus intron-containing transcripts and shuttles between the nucleus and the cytoplasm*

The regulation of the nucleocytoplasmic transport of mRNAs has been identified in a number of viral systems. For example, the HIV-1 protein Rev, in combination with the cellular factor, RIP1 (Rab), induces the accumulation

Table 1.4. IMPORTINS AND EXPORTINS IDENTIFIED TO DATE.

Family name	Specific proteins	Homologues	Molecule(s) transported, e.g.	Signal
Import				
Importin α	Rch 1	hSRP1	SV40 T-antigen	'Classical' NLS
	NPI-1	p54/56	nucleoplasmin	basic basic
	Qip 1	Karyopherin α		bipartite
	hSRP1γ	OHO31		
	ySRP1/Kap60	At IMPα		
		NBP70		
		PTAC58		
Importin β		p97		
		Karyopherin β/Kap95		
		RSL1		
		PTAC97		
Transportin	Transportin 1	Karyopherin β2	hnRNP A1, F	'Nonclassical' NLS/NES (glycine rich)
	Transportin 2	Kap 104		
		MIP		
			hnRNP K	'Nonclassical' 40-aa domain
	Mtr10		mRNA-binding proteins	
	Karyopherin β3	Yrb 4	Ribosomal proteins	Novel basic NLS
	Sxm 1		Ribosomal proteins	
	Importin β		U snRNAs	m_3G cap Sm core protein
Export				
Exportin 1	CRM1	Xpo1	HIV-1 Rev PKI	Leucine-based NES
	Pse1			
Exportin 2	CAS		Importin α	
	Kap123			
	Sxm1			
	hRIP	Rab	HIV-1 Rev	Leucine-based NES
Transportin	Transportin1		mRNA	'Nonclassical' NES/NLS
CBC	CBP20		U snRNA mRNA	m^7G cap
	Exportin t		tRNA	
			rRNA	NES on protein?

of unspliced and partially spliced virus transcripts in the cytoplasm. In contrast, the influenza virus NS1 protein inhibits the transport of spliced polyadenylated RNAs and the adenovirus E1B protein has been shown to facilitate the export of viral mRNAs to the cytoplasm.

Hibbard and Sandri-Goldin (Hibbard & Sandri-Goldin, 1995) , using Northern hybridisation assays, found that cellular mRNA levels were reduced and unspliced pre-mRNAs accumulated in the nucleus in the presence of IE63. This is to be expected since IE63 inhibits splicing. Unexpectedly, these workers also found that the inhibition of splicing resulted in the nuclear export of some unspliced RNAs suggesting that IE63 may also have a role in the regulation of RNA transport.

Further experiments have shown that intron-containing immediate-early and late transcripts (IE110 and UL15) increasingly accumulate in clumps in the nucleus as infection proceeded whereas intronless IE and late transcripts (IE63, UL38 and UL44) are detected predominantly in the cytoplasm (Phelan *et al.*, 1996). IE63 was shown to be required for the nuclear retention of intron-containing transcripts since in the absence of IE63, all transcripts were predominantly cytoplasmic. Unspliced RNAs that are retained in the nucleus were shown to colocalise with both IE63 and the redistributed snRNPs. These findings strongly indicate that IE63 inhibits the nuclear export of intron-containing transcripts in HSV-1 infected cells giving a distinct advantage to a virus that encodes few transcripts which contain introns.

IE63 can bind RNA directly through an RGG box (Mears & Rice, 1996) and this RGG box is post-translationally modified by methylation. Methylation is a modification that is common to a number of hnRNP RNA-binding proteins suggesting that IE63 may have a similar role to hnRNP proteins. The hnRNP proteins A1, K and E and the HIV-1 protein Rev shuttle between the nucleus and the cytoplasm and it seemed likely that IE63 may also be a shuttling protein. Extensive studies have been carried out using HSV-1 infected cells treated with actinomycin D, a drug which prevents the nuclear import of shuttling proteins whilst allowing export to continue (Phelan & Clements, 1997, Soliman *et al.*, 1997). Such studies make it possible to visualize the cytoplasmic accumulation of shuttling proteins. These studies have shown that the IE63 protein does shuttle between the nucleus and the cytoplasm at late times postinfection. The shuttling ability of IE63 is not dependent upon viral infection or other viral proteins since the protein shuttles when expressed in uninfected cells (Mears & Rice, 1998).

Consistent with IE63's shuttling function, the protein has recently been shown to bind intronless mRNA in both the nucleus and the cytoplasm (Sandri-Goldin, 1998). Interestingly, the binding of hnRNP A1 to cellular mRNA is reduced significantly in wt infection. This suggests that IE63 favours the export of viral, intronless mRNA whilst blocking the export of cellular mRNA via the hnRNP A1 pathway.

1.2.3 A number of herpesviruses encode IE63 homologues

The essential HSV-1 protein IE63 is the only HSV IE protein with homologues throughout the herpesviridae family. The varicella-zoster virus encodes ORF4, a multifunctional protein that transactivates gene expression by a number of mechanisms including a mechanism that appears to act at the posttranscriptional level (Defechereux *et al.*, 1997). The herpesvirus saimiri encodes ORF57, a protein which activates and represses expression of viral genes and redistributes splicing products in the infected cell nucleus (Whitehouse *et al.*, 1998). The human cytomegalovirus protein UL69 can transactivate a number of viral and cellular genes (Winkler *et al.*, 1994) as can the Epstein-Barr virus protein BMLF1 which regulates gene expression at the post-transcriptional level via its effects on splicing and RNA export and which shuttles between the nucleus and the cytoplasm (Semmes *et al.*, 1998). Interestingly, BMLF1 interacts with CRM1 and this interaction is essential for its ability to shuttle (Boyle *et al.*, 1999). Finally, bovine herpesvirus type 1 expresses BICP27, a protein which appears to increase the efficiency of mRNA processing at weak poly (A) sites (Singh *et al.*, 1996).

1.2.4 Functional domains of IE63

IE63 and a number of other regulatory proteins encoded by nuclear-replicating DNA viruses have been shown to contain several functional domains. Such proteins are essential to these viruses since they have a limited coding capacity but must interact with multiple components of the

cell . The functional domains that have been identified in the IE63 protein to date are illustrated in Figure 1.9.

(a) Potential zinc finger and activation and repressor regions

The regions for the activator and repressor activity of IE63 reside in the carboxy terminus of the IE63 protein. Metal chelate affinity chromatography has shown that IE63 binds zinc *in vitro* and that the region of IE63 which binds zinc also resides in the carboxy terminal spanning a putative metal binding motif (amino acids 407-512) (Vaughan *et al.*, 1992). Zinc finger domains are associated with protein-nucleic acid and protein-protein interactions. Interestingly, the region of IE63 which is required for coimmunoprecipitation with anti-Sm antisera begins after amino acid 450 and extends at least to residue 504 (Sandri-Goldin & Hibbard, 1996). This data indicates that the potential zinc finger domain of IE63 is required for interaction with a component of the cellular snRNPs.

(b) Essential acidic domain

IE63 contains a high proportion of acidic amino acid residues in its amino terminal region (amino acids 4 to 64). A deletion mutant lacking this domain was unable to complement an IE63 null mutant suggesting that the acidic domain is essential for IE63 function (Rice *et al.*, 1993).

Similar highly acidic regions have been found in a number of transcription activators such as the HSV-1 virion protein VP16 (Struhl, 1991). It has been proposed that these acidic domains stimulate transcription by interacting with one or more components of the RNA polymerase II transcription complex.

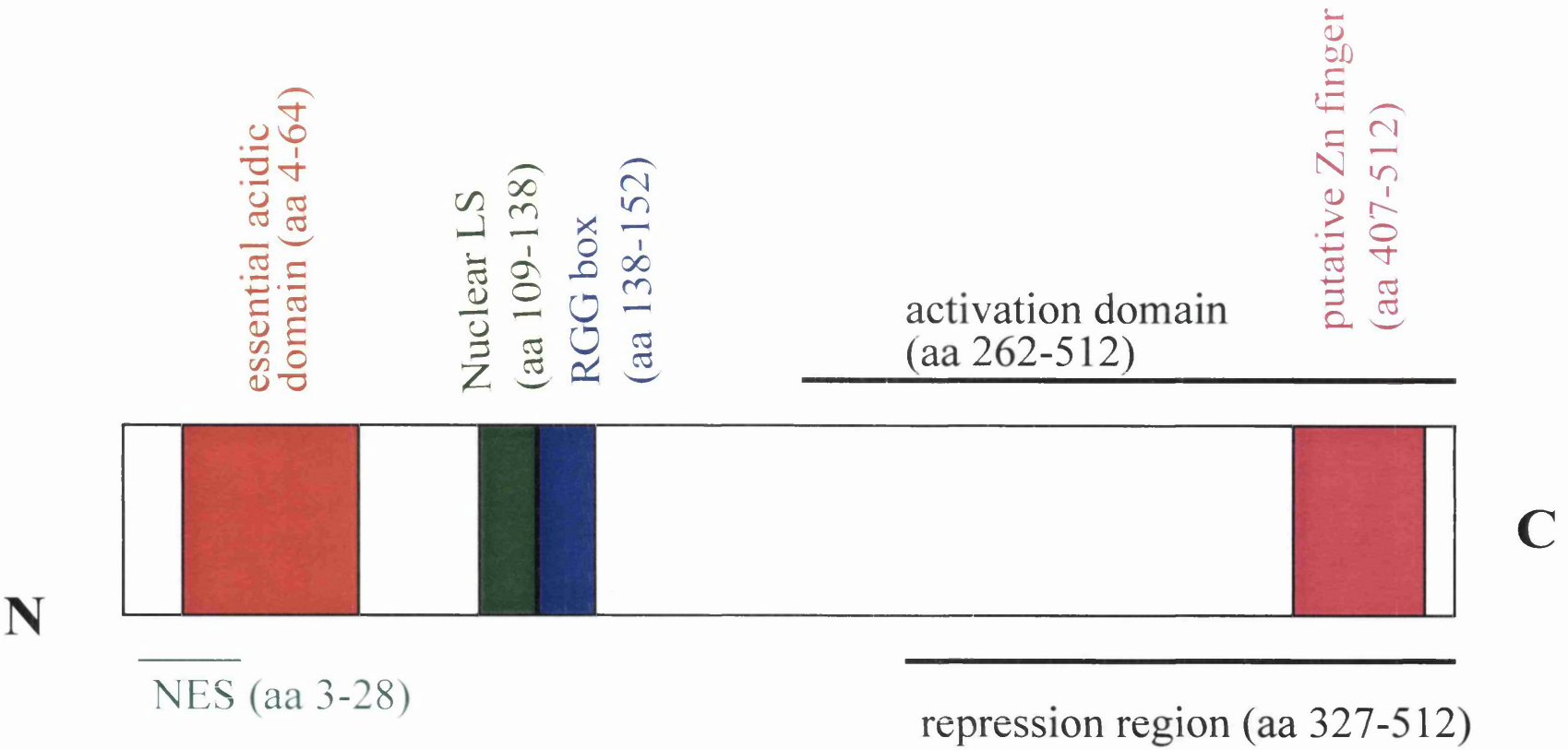


Figure 1.9

A diagramatic illustration of IE63 showing its principal functional domains, for details see Section 1.2.4.

The cluster of highly acidic residues in the amino terminal region of IE63 is a feature conserved within all of the IE63 alphaherpesvirus homologues suggesting that the acidic region of IE63 is involved in a function common to diverse herpesviruses.

(c) Nuclear localisation signals

Cell fraction and immunofluorescence studies have shown that IE63 is predominantly localised to the cell nucleus both during HSV-1 infection and in transfected cells. Evidence suggests that IE63 also interacts with the cell nucleoli and shuttles between the nucleus and the cytoplasm. Mears, Lam and Rice (Mears *et al.*, 1995) have used fusion protein constructs to identify the specific regions of IE63 which can direct a normally cytoplasmic protein, pyruvate kinase, to the nuclei of transfected cells. These workers have found that IE63 contains multiple signals, both strong (aa 110-137) and weak (mapping to the C-terminal portion of the protein), which mediate its nuclear loclalisation. In addition they identified a region of IE63 which possesses a relatively short sequence that functions as a nucleolar targeting signal (aa 110-152).

(d) Nuclear export sequence

IE63 contains a leucine-rich region (amino acids 7 to 27) which exhibits good homology with export signals of other proteins known to shuttle between the nucleus and cytoplasm such as the HIV-1 Rev protein (Kalland *et al.*, 1994, Meyer & Halim, 1994, Richard *et al.*, 1994), the heat-stable inhibitor of cAMP-dependent protein kinase (PKI) (Wen *et al.*, 1995), the

Rex protein of HTLV-1 (Bogerd *et al.*, 1995, Kim *et al.*, 1996, Palmer & Malim, 1996), the Rev proteins of visna virus and equine infectious anemia virus (Meyer *et al.*, 1996), adenovirus E3-34-kD (Dobbelstein *et al.*, 1997), transcription factor IIIA (Fridell *et al.*, 1996), and the yeast mRNA transport protein Gle 1 (Murphy & Wente, 1996).

Sandri-Goldin (Sandri-Goldin, 1998) has constructed a mutant that lacks a region in the IE63 protein which resembles the putative leucine-rich NES. Cells were transfected with this mutant or wt IE63 and were either actinomycin treated or untreated. The wt IE63 moved to the cytoplasm following addition of actinomycin D. However, deletion of the NES resulted in a protein that remained nuclear in the presence of actinomycin D. Further experiments showed that the NES of IE63 was sufficient for export of a heterologous protein (Sandri-Goldin, 1998). These experiments suggest that the leucine-rich region of IE63 is both necessary and sufficient for the export of the IE63 protein.

Another group has used the interspecies heterokaryon assay to examine IE63's shuttling function (Mears & Rice, 1998). This assay requires the fusion of HSV-1 infected Vero monkey cells that have been cycloheximide treated with uninfected mouse cells which have also been treated with cycloheximide. One hour later the localization of viral nuclear proteins can be assessed by immunofluorescence and, since protein synthesis is inhibited, the appearance of an HSV-1 nuclear protein in the mouse nucleus of a fused cell suggests that the viral protein has shuttled out of the infected monkey nucleus and has been re-imported into the uninfected mouse nucleus. The

results of these studies show that deletion of the putative NES of IE63 (aa 7-20) has only a modest detrimental effect on shuttling. However, a mutation which alters amino acids 465 and 466 of IE63 from aa PG to LE completely abrogates shuttling.

The conflicting findings of these studies suggest that IE63 may contain at least two regions that enable the protein to shuttle between the nucleus and the cytoplasm. The leucine-rich NES does appear to promote shuttling but another domain which maps to the carboxy terminal 46 amino acids appears to be absolutely essential for shuttling. This carboxy terminal region may contain an additional NES or it may be required for the ability of IE63 to interact with a cellular shuttling protein which allows IE63 to access an alternative shuttling pathway utilised by the host cell.

(e) RGG box and KH domains

IE63's Nucleolar Localisation Signal (NuLS) is similar to the short basic NuLSs of the HIV proteins Rex, Rev and Tat in that it is composed of an NLS (residues 110 to 137) and an arginine-rich basic region (residues 138-152) (Mears *et al.*, 1995). The arginine-rich region in the IE63 NuLS is also very rich in glycine residues (Mears & Rice, 1996). The motif, termed the RGG box or GAR (glycine-arginine-rich) domain is particularly predominant in hnRNPs and proteins which function in pre-rRNA processing such as nucleolin and fibrillarin. One of the best matches with IE63's RGG box is a sequence of the SmD protein which is a constituent of snRNPs (Mears & Rice, 1996). This observation suggests that IE63's RGG box may mediate an

interaction with a pre-mRNA or small nuclear RNA involved in mRNA processing.

IE63 also contains an RNA binding motif in the C-terminal half of the protein and this motif has homology to a series of hnRNP K homology (KH) domains (T B Soliman and S J Silverstein, Columbia University, New York, personal communication). It appears that IE63 contains three KH domains and that these are required for the expression of HSV-1 L genes.

Using IE63 RNA binding assays, Ingram *et al.* (Ingram *et al.*, 1996) have shown that IE63 binds to RNA in a weak and non-specific manner. This is in contrast to the findings of Brown *et al.* (Brown *et al.*, 1995) who found that IE63 binds to certain RNAs specifically. The RGG box of IE63 appears to mediate this RNA binding and the RGG box is post-translationally modified by methylation (Mears & Rice, 1996).

(f) Regions of IE63 required for interaction with IE175

In vitro pull down assays with GST fusion proteins have shown that IE63 can interact with the viral IE175 protein directly and that this interaction is mediated by aa 27-103 and aa 179-406 of IE63 (Pangiotidis *et al.*, 1997).

However, it should be noted that there was no interaction between IE175 and IE63 in extensive co-immunoprecipitation experiments carried out by our group.

1.3 OVERVIEW OF THE CELLULAR PROTEINS IDENTIFIED IN THIS STUDY WHICH INTERACT WITH IE63

1.3.1 The hnRNP K Protein

(a) hnRNP K is a nuclear protein which exists in four isoforms

The K protein has an apparent molecular weight of 65 kDa and was originally discovered as a component of the heterogenous nuclear ribonucleoprotein (hnRNP) particle, hence its name (Matunis *et al.*, 1992). The K protein can be recovered from both cytoplasmic and nuclear extracts but it is predominantly nuclear and it shuttles between the nucleus and the cytoplasm (Buchenau *et al.*, 1997, Michael *et al.*, 1997b). The gene which encodes the protein maps in humans to chromosome 9 and four different isoforms of the protein have been identified. The differences in the isoforms are due to the presence or absence of a 15 bp and a 60 bp exon (Dejgaard *et al.*, 1994). The presence of two different polyadenylation sites result in additional variants and the variation in the 3' non-translated region together with the differences in the coding region may result in up to 12 different putative mRNAs. The four coding region variants separate into at least 3 spots following 2-D gel electrophoresis and this is due to differences in phosphorylation (Dejgaard *et al.*, 1994).

(b) hnRNP K is a phosphoprotein

All isoforms of hnRNP K contain potential phosphorylation sites for casein kinase II (CKII), protein kinase C and tyrosine kinase. The hnRNP K binds

to and is phosphorylated by a serine/threonine inducible interluekin 1-responsive kinase termed K protein kinase (KPK) (Ostrowski *et al.*, 1991, Seuningen *et al.*, 1996) which recent evidence suggests may be protein kinase Cδ (Schullery *et al.*, 1999). The kinase itself is thought to be regulated by phosphorylation (Seuningen, 1995). Although hnRNP K can be phosphorylated by CK II *in vitro* (Ostrowski *et al.*, 1994), the kinase which copurifies with the hnRNP K protein is not CKII [Seuningen, 1995 #145]. Phosphopeptide maps from *in vitro* and *in vivo* experiments overlap but are not identical, suggesting that hnRNP K may be phosphorylated by more than one kinase (Ostrowski *et al.*, 1991).

(c) Functional Domains of hnRNP K

The functional domains of hnRNP K are illustrated in Figure 1.10. These include three KH domains which are involved in RNA binding (Siomi *et al.*, 1994), a bipartite-basic nuclear localisation signal, three proline-rich SH3-binding regions which consitute an SH3-binding domain (SBD) and interact *in vitro* with a subset of SH3-containing proteins such as Src and Vav (Bustello *et al.*, 1995, Taylor & Shalloway, 1994, Weng *et al.*, 1994). It is of interest to note that all three K proline-rich regions have an adjacent RGG motif, a site that is methylated by a specific methyltransferase and it has been suggested that the SH3 interactions might be regulated by the methylation of these residues (Liu & Dreyfuss, 1995). The K protein also contains a novel shuttling domain (termed KNS) which confers bi-directional transport across the nuclear envelope (Michael *et al.*, 1997b). The N-terminus of the K

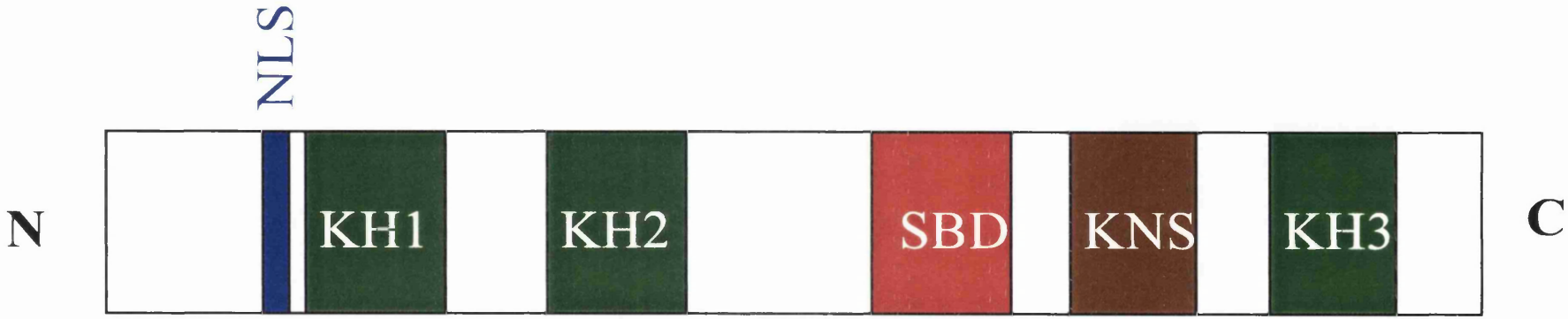


Figure 1.10
Functional domains of hnRNP K protein; nuclear localisation signal (NLS), the KH domains which are involved in RNA binding, a proline-rich SH3-binding domain (SBD) which interacts with a number of SH3-containing proteins and a bi-directional nuclear shuttling domain (KNS).

protein is highly acidic and has been reported to contain transcriptional activity (Tomonaga & Levens, 1995).

(d) hnRNP K binds to DNA and RNA and may have a role in pre-mRNA processing and transport

The K protein has been shown to bind to specific sequences in RNA (Matunis *et al.*, 1992), and to single and double-stranded DNA which, in at least some instances, occurs through distinct DNA motifs such as the κB motif (Ostrowski *et al.*, 1994) and the homopurine/homopyrimidine tract of the CT motif, present in the c-*myc* promoter P1 (Takimoto *et al.*, 1993). Recent evidence suggests that the CT motif of the c-*myc* promoter adopts an under-wound and flexible conformation in the presence of hnRNP K allowing the DNA to bend bringing together *cis* elements which facilitate transcription (Tomonaga *et al.*, 1998).

Binding to RNA diminishes when the K protein is phosphorylated indicating that binding may be modulated by phosphorylation (Dejgaard *et al.*, 1994). Since the K protein binds to RNA and shuttles between the nucleus and the cytoplasm it is thought that the K protein may have a function in pre-mRNA processing or transport.

(e) hnRNP K interacts with a number of cellular proteins involved in transcription, translation and signal transduction

K protein has been shown to interact with the SH3 domains of a number of proteins involved in signal transduction such as the tyrosine kinases Src, Fyn

and Lyn and the proto-oncoprotein Vav which is exclusively expressed in hematopoietic cells and is involved in expansion of T and B cells in antigen-mediated proliferative responses and in the induction of intrathymic T cell maturation (Bustello *et al.*, 1995). Although the K protein can engage the tyrosine kinases Src, Fyn and Lyn, it is not itself tyrosine-phosphorylated (Ostrowski *et al.*, 1991, Weng *et al.*, 1994). However, in the presence of K protein, KPK can be reactivated by c-Src suggesting that the K protein may facilitate cross-talk between tyrosine and serine/threonine kinases (Seuningen, 1995).

The K protein also interacts with a number of proteins involved in gene expression. K protein co-immunoprecipitates with epitope-tagged TFIID TATA-binding protein (TBP) and [^{35}S]TBP translated *in vitro* binds to recombinant K protein (Michelotti *et al.*, 1996). The K protein acts synergistically with TBP and the association is DNase- and RNase-resistant (Michelotti *et al.*, 1996). K protein also interacts with the zinc-finger transcriptional repressor Zik 1 (Denisenko *et al.*, 1996). Interestingly, poly(C) RNA which binds K protein, abrogates the Zik1-K protein interaction, while poly(A) RNA which does not bind K protein has no effect on this interaction. The Zik1-K protein interaction can also be blocked by the κB motif which like poly(C) RNA also binds to the K protein (Denisenko *et al.*, 1996). This suggests that the interaction of Zik1 with K protein is regulated by cognate nucleic acid motifs.

Recently, K protein has been shown to interact with and repress transcription factor C/EBP$^{\beta}$, a factor which is know to regulate a number of genes including those involved in the acute-phase response (Miau *et al.*, 1998). It has been suggested that K protein does this by disrupting the synergistic interaction with C/EBP$^{\beta}$ and its coactivator Nopp140 and that phosphorylation of hnRNP K may affect the stability and activity of the hnRNPK-C/EBP$^{\beta}$ complex.

Functional studies also indicate that the K protein has a role to play in the regulation of transcription. The K protein increases transcription both *in vivo* (Michelotti *et al.*, 1996) and *in vitro* (Gaillard *et al.*, 1988). Since K protein both activates and represses transcription of a reporter gene linked to a heterologous promoter which does not contain any obvious K protein-binding sequence-specific DNA motifs, it is possible that the effect of the K protein on transcription may in some instances occur indirectly through interactions with cellular transcription factors (Gaillard *et al.*, 1988).

Finally, K protein appears to have a role in translation since it interacts with the elongation factor-1α (O N Densienko and K Bomsztyk, Maria Sklodowska-Curie Memorial Centre and Institute of Oncology, Warsaw, unpublished observations) and, together with hnRNP E1, interacts with and confers translational silencing on erythroid 15-lipoxygenase (LOX) mRNA (Ostareck *et al.*, 1997). LOX is a key enzyme during erythroid cell differentiation and it must be silenced in early erythroid cells in the bone marrow and the peripheral blood. Evidence suggests that hnRNP K and

hnRNP E1 regulate the translation of LOX by interacting with a C+U-rich, repetitive sequence motif in the 3' UTR of LOX mRNA.

(f) hnRNP K interacts with the Hepatitis C Virus Core Protein

The K protein has been shown to weakly suppress the human thymidine kinase promoter (Hsieh *et al.*, 1998). The hepatitis C virus (HCV) core protein has recently been shown to interact with the K protein and to relieve the suppression effect of K on the activities of the human thymidine kinase gene (Hsieh *et al.*, 1998). The promoter of the human thymidine kinase gene contains a C+T motif approximately 100 bp upstream of its TATA element. The HCV core protein is a component of the viral nucleocapsid and it is associated with the endoplasmic reticulum and the nucleus. It is phosphorylated and it binds to DNA, RNA and ribosomes. The HCV core protein can modulate the expression of several viral and cellular genes including *c-myc* and p53 indicating that the HCV core protein may, like IE63, alter normal cellular functions. The finding that HCV core protein binds K protein, which, itself stimulates the *c-myc* promoter but suppresses the TK promoter indicates that the core protein exerts its effects on gene expression via interactions with cellular factors.

(g) hnRNP K interacts with human papillomavirus L2

Recent studies have shown that hnRNP K can interact with human papillomavirus 16 (HPV-16) late viral protein L2 in a sequence-specific

manner effectively inhibiting translation of L2 mRNA (Collier *et al.*, 1998). The HPV life-cycle is intricately linked to the differentiation of the stratified epithelium and the L2 protein is restricted to terminally differentiated epithelial cells in the superficial layers of the squamous epithelium. It seems plausible that HPV-16 may control gene expression at the level of translation, at least in part, via its interaction with hnRNPK.

(h) hnRNP K may act as a docking platform to facilitate molecular interactions

That the K protein can interact with multiple proteins and nucleic acids has led some to suggest that it may act as a linker molecule between DNA and RNA and proteins that cannot bind nucleic acids directly, thus allowing cross-talk between these factors. If this were true it would explain the ability of K protein to act on such diverse cellular functions as transcription, translation and signal transduction.

1.3.2 Casein Kinase II

(a) Casein kinase II phosphorylates proteins required for cell growth, cell division, signal transduction and transcription

Casein kinase II (CKII) phosphorylates serine or threonine residues that are N-terminal to acidic amino acids. CK II uses either ATP or GTP as phosphoryl donors and is located in the cytosol, nuclei and nucleolus. The holoenzyme

(130 kDa) consists of two catalytic α subunits and two regulatory β subunits in the tetrameric configuration $\alpha_2\beta_2$, $\alpha\alpha'\beta_2$ or $\alpha'_2\beta_2$ (See Table 1.5) (Allende & Allende, 1995). The α and α' subunits are highly homologous but are the products of different genes. CKII can interact with its substrates either in its holoenzyme form or as individual subunits.

CKII is a regulator of cell growth, cell division and signal transduction pathways and it plays an important role in positive and negative transcriptional regulation by specific phosphorylation of the transcription factors Sp1 (Pugh & Tijian, 1990), Ap1 (Lin *et al.*, 1992), serum response factor (Marais *et al.*, 1992), upstream binding factor (Voit *et al.*, 1992) and steroid hormone receptor (Picard *et al.*, 1990). Over 30 other cellular proteins have been identified which interact with CKII and these include nuclear proteins such as myc, p53 and Nopp140, cytosolic proteins such as ornithine decarboxylase and tubulin and membrane bound proteins such as the LDL receptor and the androgen receptor (Allende & Allende, 1995).

(b) Casein Kinase II phosphorylates and modulates the function of a number of viral proteins

Interestingly CKII also phosphorylates a number of viral proteins including influenza virus PA polymerase (Sanz-Ezquerro *et al.*, 1998), adenovirus E1A protein (Massimi *et al.*, 1996) and human papilloma virus E7 (Phillips & Vousden, 1997). Furthermore, the HIV-1 protein Rev has recently been shown to be a potent activator of CKII and the activated CKII phosphorylated several

Table 1.5. SUBUNIT STRUCTURE AND ISOFORMS OF CKII

Subunit	Amino Acids	Mol. Wt.	SDS-PAGE
α	391	45,160	42 kDa
α'	350	41,400	40 kDa
β	215	24,925	25 kDa

cellular and viral proteins in HIV-1-infected cells (Ohtsuki *et al.*, 1998). The binding of Rev to the Rev Response Element (RRE) is significantly reduced when it is phosphorylated suggesting that the HIV-1 Rev-RRE complex may be regulated by CKII (Ohtsuki *et al.*, 1998). It also seems likely that interactions between Rev and its cellular mediators, such as RIP1 may also be regulated by CKII in HIV-1 infected cells. Clearly CKII is utilised by a number of viruses to phosphorylate viral regulatory proteins enabling their activities to be tightly regulated.

1.3.3 The p32 Protein

(a) p32 interacts with the essential splicing factor ASF/SF2 and regulates splicing

p32 was originally isolated as a protein tightly associated with the essential splicing factor ASF/SF2 (Krainer *et al.*, 1991). ASF/SF2 is a member of the SR family of splicing factors and it has a role in the enhancement of splicing and the regulation of alternative splicing (reviewed in Fu, 1995). ASF/SF2 has also been shown to shuttle between the nucleus and the cytoplasm and the phosphorylation state of the protein appears to regulate this activity (Caceres *et al.*, 1998). The phosphorylation status of other SR proteins has been shown to affect their function in pre-mRNA recognition (Cao *et al.*, 1997), spliceosome assembly (Xiao & Manley, 1997) and splicing catalysis (Kanopka *et al.*, 1998) as well as the subcellular localisation (Caceres *et al.*, 1998) and sub-organelle distribution (Misteli & Spector, 1997).

Recent evidence suggests that p32 is not a general splicing factor but that it exerts its effect on splicing by modulating ASF/SF2 function (Petersen-Mahrt *et al.*, 1999). In fact it appears that p32 inactivates ASF/SF2 as a splicing repressor or activator protein by inhibiting the ability of ASF/SF2 to bind RNA and hence its function in initiation of pre-spliceosome formation and also by blocking ASF/SF2 phosphorylation. Since this post-translational modification regulates the subcellular localisation and sub-organelle distribution of ASF/SF2 it has been suggested that p32 may exert its effect on this protein by sequestering it into an inhibitory complex. However, the exact mechanism by which p32 regulates splicing has yet to be determined.

(b) The cellular localisation of the p32 protein

There is some controversy as to the exact localisation of the p32 protein. Various groups have reported that it can be found in the mitochondria (Muta *et al.*, 1997, Matthews & Russell, 1998), cytoplasm (Luo *et al.*, 1994, Simos & Georgatos, 1994, Tange *et al.*, 1996, Wang *et al.*, 1997), cell surface (Eggleton *et al.*, 1995, Ghebrehiwet *et al.*, 1997, Guo, 1997, Peterson *et al.*, 1997) and nucleus (Luo *et al.*, 1994, Simos & Georgatos, 1994, Tange *et al.*, 1996, Wang *et al.*, 1997, Matthews & Russell, 1998). Recent work carried out by our group suggests that p32 exists predominantly in the cytoplasm of uninfected cells but that it redistributes to the nucleus in the presence of IE63 alone and during HSV-1 infection (H E Bryant, D A Matthews, S Wadd, J E Scott, S

Graham, W C Russell, J B Clements, Institute of Virology, University of Glasgow, submitted for publication).

(c) p32 interacts with a number of diverse viral proteins

p32 interacts with a number of viral proteins including the HSV-1 ORF-P protein (Bruni & Roizman, 1996), the adenovirus polypeptide V (Matthews & Russell, 1998), Epstein-Barr virus EBNA I protein (Wang *et al.*, 1997, Chen *et al.*, 1991) and the HIV Rev (Luo *et al.*, 1994, Tange *et al.*, 1996) and Tat proteins (Yu *et al.*, 1995). It has been postulated that the interaction of p32 with Rev which is bound to ASF/SF2 at the 5'-splice site stabilises the interaction of U1 snRNP with the 5'-splice site and inhibits assembly of a functional spliceosome (Tange *et al.*, 1996). The dysfunctional spliceosome could then function as a substrate for Rev-mediated nuclear transport. However, p32 also appears to bridge the HIV-1 protein Tat to the general transcription machinery via its interaction with TFIIB and appears to act synergistically with Tat to stimulate transcription (Yu *et al.*, 1995), suggesting that p32 plays multiple roles in HIV-1 replication. Deletion of the regions of EBNA-1 required for interaction with p32 abrogated the transcriptional activity of EBNA-1, implicating p32 as a possible mediator of EBNA-1-dependent transactivation. In addition, by examining infected cells using confocal microscopy and immunofluorescence, Matthews and Russell have concluded that p32 may play a role in importing cellular proteins to the nucleus and that adenovirus protein V hijacks this protein to deliver the adenovirus genome to the nucleus (Matthews & Russell, 1998).

1.3.4 Spliceosomal Associated Protein 145

(a) Spliceosomal Associated Protein 145 is a subunit of the essential splicing factor SF3b

Subsequent to E-complex formation, prespliceosomes are formed by interaction of the U2 snRNP with the branch site. This interaction involves base pairing between the U2 snRNA and the branch-point sequence and is stabilised by a number of proteins (Figure 1.6). These proteins include the heteromeric splicing factors SF3a and SF3b, which are essential for prespliceosome assembly (Brosi *et al.*, 1993b). SF3a is comprised of subunits of 60, 66 and 120 kDa (designated spliceosomal-associated proteins SAP61, SAP62, and SAP114, respectively) whereas SF3b contains subunits of 53, 120, 150 and 160 kDa which are thought to correspond to SAPs 49, 130, 145 and 155 respectively (Brosi *et al.*, 1993b, Brosi *et al.*, 1993a, Hodges & Beggs, 1994).

(b) Spliceosmal Associated Protein 145 is required for the binding of the U2 snRNP to pre-mRNA

All SF3a and SF3b subunits, with the exception of $SF3b^{120}$, can be cross-linked to a 20-nucleotide region just upstream of the branch site (Champion-Arnaud & Reed, 1996, Gozani *et al.*, 1996). Although the binding of these proteins occurs in a sequence-independent manner, the interaction of the SF3a

and SF3b subunits with this site is absolutely essential for tethering the U2 snRNP to the branch site. Therefore it appears that the primary role of SAP145 is to mediate the binding of the U2 snRNP to the pre-mRNA.

1.3.5 ALY

(a) ALY stimulates transcription of the components of the T-cell receptor alpha enhancer complex

ALY is a ubiquitously expressed nuclear protein which associates with the proteins LEF-1 and AML-1 (Bruhn *et al.*, 1997). These proteins are components of the T-cell receptor alpha enhancer complex and it is thought that ALY stimulates transcription by facilitating the collaboration of these proteins in this complex.

(b) ALY is a spliceosome-associated protein

A recent study using mass spectrometry and expressed-sequence tag databases (EST) has allowed the characterization of the mammalian spliceosome complex (Neubauer *et al.*, 1998). This study has identified 19 proteins which were not previously thought to be associated with the spliceosome, one of which was ALY (mouse). This evidence strongly suggests that ALY has a role in mammalian splicing.

1.4 AIM AND SCOPE OF PROJECT

1.4.1 Evidence suggests that IE63 interacts with a number of cellular and viral proteins

IE63 plays an essential role during HSV-1 infection. The preceding text has presented evidence that IE63 has an effect on:-

(i) the transcription of viral and some cellular genes

(ii) polyadenylation

(iii) splicing

(iv) transport of viral and cellular mRNAs across the NPC

(v) viral DNA replication

Clearly, IE63 is not a simple protein with one function and a single domain required to carry out this function. It appears that IE63 has multiple effects on the infected cell nucleus and the protein requires many functional domains to perform its essential role in the virus life cycle. It is difficult to envisage a mechanism by which IE63 could carry out such diverse functions in the host cell. The most convincing explanation for this is that IE63 interacts with and modifies essential viral and cellular regulatory proteins involved in these processes.

The finding that IE175 and the U1 SnRNP 70 kDa protein have an altered phosphorylation state and ICP8 exhibits a conformational modification in the presence of IE63 supports this hypothesis. Further, IE63 must interact with a number of cellular proteins to allow it to shuttle back and forth across the NPC. IE63 has been shown to interact with the IE175 protein and to coimmunoprecipitate with anti-Sm antisera. It is likely that identification of the other proteins with which IE63 interacts will provide the key to determining the mechanism of action of this essential HSV-1 protein.

Considering this evidence, it seemed appropriate to investigate IE63-protein interactions. It is the aim of this study to identify the cellular proteins with

which IE63 interacts. The yeast two-hybrid system is the method that has been chosen to carry out this project. This method and the reason for choosing this approach has been described in detail in Section 2.1 of this thesis.

1.4.2 The significance of interactions of IE63 with host-cell proteins

Since IE63 is essential for viral replication, determining the mode of action of this protein is clearly critical for an explicit understanding of HSV-1 biology. In addition, any indispensable viral protein is a possible target for antiviral therapy. For a viral infection that is so widespread and that has, as yet, no definitive cure the value of such studies becomes obvious.

That IE63 has homologues in viruses such as the Kaposi's sarcoma virus, varicella-zoster virus and Epstein-Barr virus suggests that determining the interactions of IE63 with cellular proteins may improve understanding of these viruses. Finally, as has been the case with many studies of viruses, it is possible that determining IE63 interactions with cellular proteins may shed light on some poorly defined or as yet unidentified cellular pathways.

MATERIALS AND METHODS

2.1. THE YEAST TWO-HYBRID SYSTEM

2.1.1 Basic Principles

The two-hybrid system exploits the ability of a pair of interacting proteins to bring the activation domain (AD) of a transcription factor in close proximity with the DNA-binding domain (BD) of that same transcription factor bound to a *cis*-acting regulatory element, thus stimulating the expression of adjacent reporter genes (see Figure 2.1) (reviewed in Mendelsohn & Brent, 1999).

A schematic diagram of the two-hybrid system screening protocol is shown in Figure 2.2. In this study, the yeast transcription factor is GAL4. Two different cloning vectors are used to generate fusions of: (a) AD + gene encoding the target protein and, (b) BD + gene encoding the protein used to screen the library. The hybrid proteins are co-expressed in a strain of the yeast *Saccharomyces cerevisiae* that contains the reporter genes HIS3 and LacZ. Since the cloning vector that encodes the activation domain/target protein also carries the TRP1 gene and the vector that encodes the DNA-binding domain/bait protein encodes the LEU2 gene, yeast colonies containing both plasmids which encode interacting proteins can be selected for by culture on a medium lacking tryptophan, leucine and histidine. An assay for β-galactosidase activity (lacZ expression) confirms that colonies are candidates for interaction of the two-hybrid proteins.

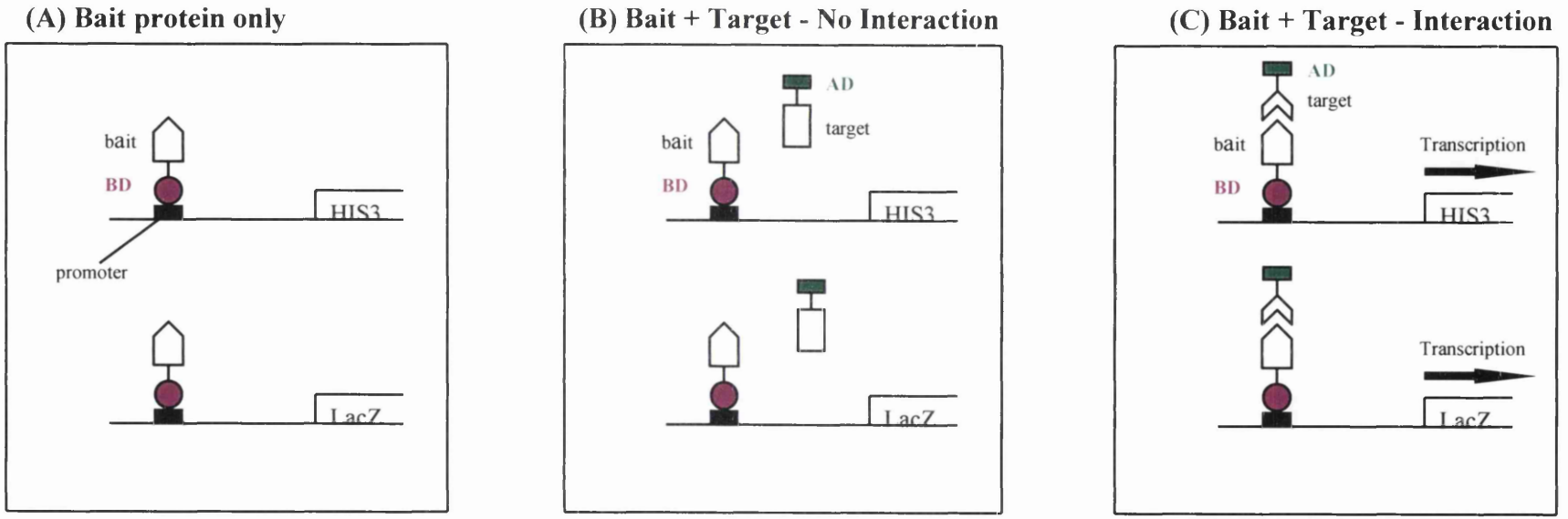


Figure 2.1
Illustration of the principle of the yeast two-hybrid system where (A) is the bait protein transformed into yeast cells with no target protein present, (B) is a bait and target protein that do not interact (hence no transcription of reporter genes) and (C) is interacting proteins and concomitant transcription of the reporter genes HIS3 and LacZ.

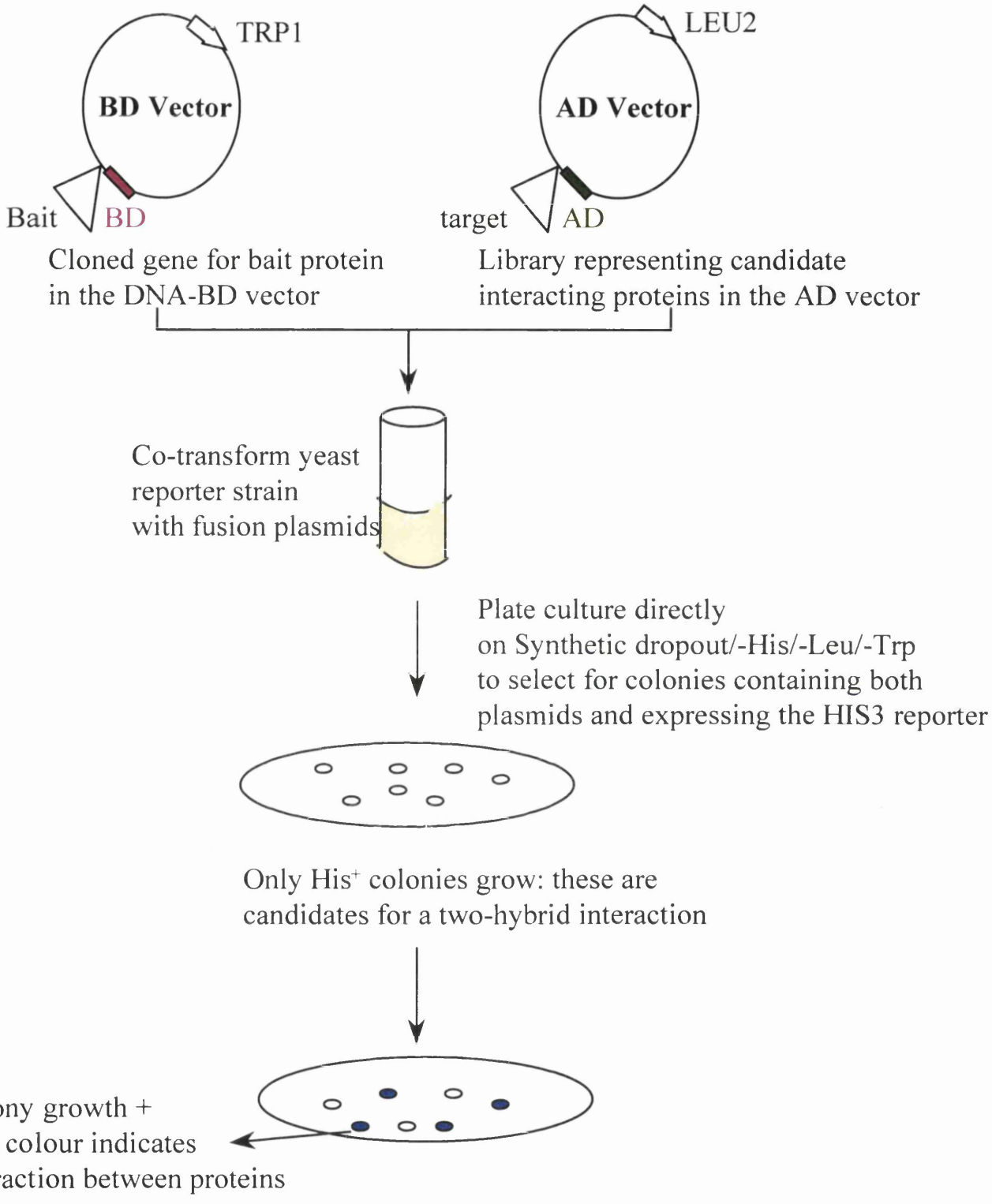


Figure 2.2
Schematic diagram of two-hybrid system. Plasmids encoding the bait are cotransformed into yeast cells with plasmids encloding a cellular library of proteins. Yeast colonies that grow on medium lacking histidine, leucine and tryptophan and that turn blue in a β-gal assay contain putative interacting proteins.

The two-hybrid system has been extended to screen activation domain libraries for proteins that interact with a bait protein. cDNA sequences are fused to the DNA encoding an activation domain. This library and a plasmid encoding a the BD-bait protein fusion are co-expressed in yeast and the resulting transformants are screened for the two reporter genes. Plasmids responsible for reporter gene expression are then isolated from the yeast cells and DNA sequencing is used to identify the proteins encoded by the library plasmids. Figure 2.3 illustrates the procedures used to screen an activation domain/library.

2.1.2 Development of the Two-Hybrid System

(a) Historical perspective

Findings from three types of experiments led to the development of the two-hybrid system. First, the two functional domains of a number of transcription factors were identified. Certain transcription factors, such as the yeast GAL4 and GCN4 proteins, were shown to consist of two distinct and separable domains: a DNA-binding domain, and an acid region which is required for transcriptional activation (Fields & Song, 1989, Hope & Struhl, 1986).

Second, the ability to generate hybrid transcription factors was demonstrated by generating a hybrid LexA-GAL4 fusion protein which was found to activate transcription in yeast cells containing a reporter gene under the

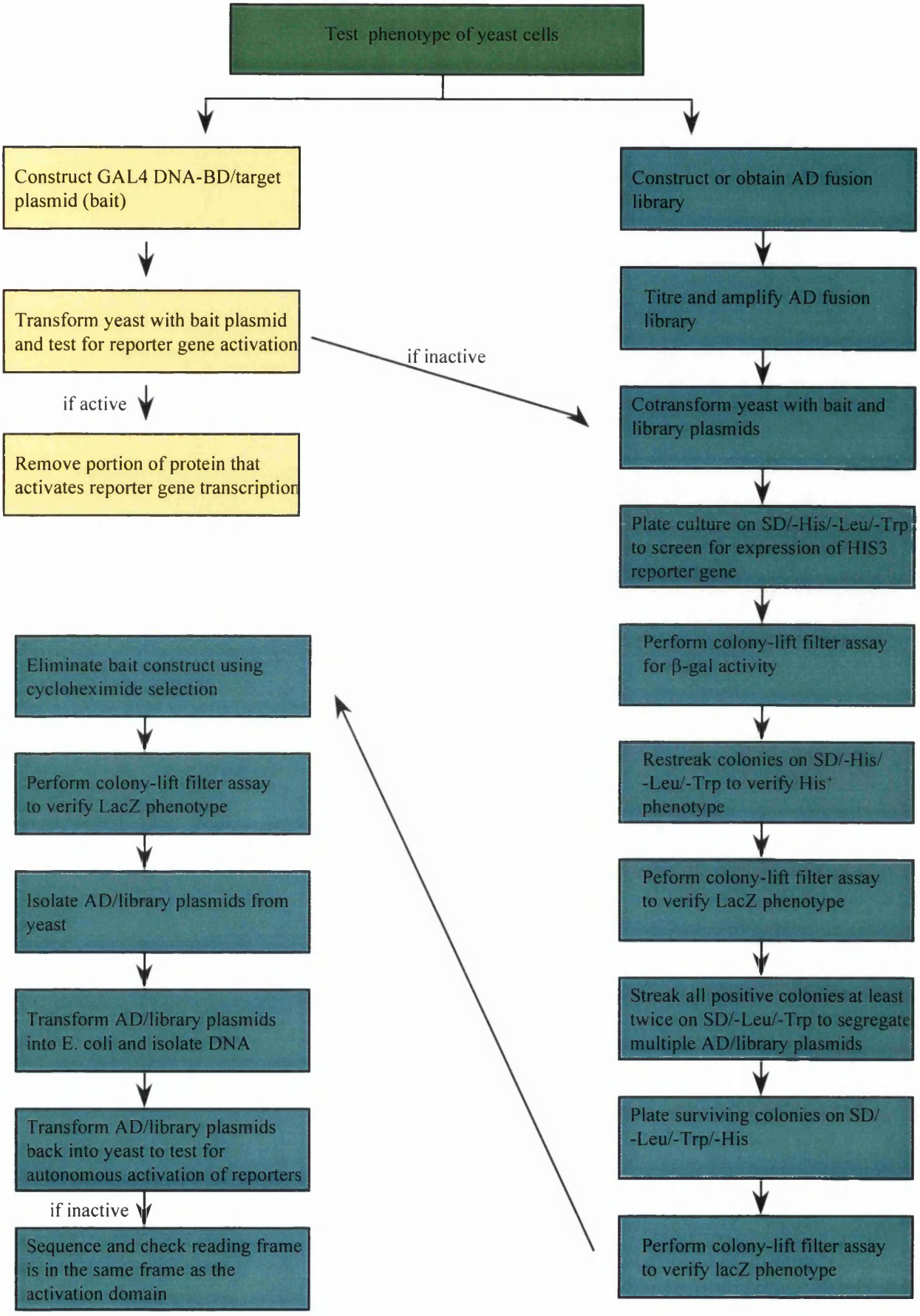


Figure 2.3
Procedure used to screen an activation domain library using the yeast two-hybrid system.

control of LexA operator sequences (Brent & Ptashne, 1985). In these experiments, the DNA-binding activity was provided by the LexA component and the activation function was provided by the GAL4 component.

Third, the concept of the DNA-binding and activation domains functioning when non-covalently joined via protein-protein interactions was recognized. The yeast protein GAL80 was shown to bind to the transcription factor GAL4 in a process which obscured the activation domain of GAL4 (Ma & Ptashne, 1988). Ma and Ptashne deleted the activation domain from GAL4 and fused an activation domain to GAL80. They showed that when the two proteins interacted, the modified GAL80 protein caused the inactive GAL4 to stimulate transcription.

As a result of these observations, Fields and his coworkers suggested that protein interactions could be detected if two potentially interacting proteins were expressed as fusion proteins (Fields & Song, 1989). They proposed that the first protein contain a DNA-binding domain which binds to DNA upstream of a reporter gene, and that the second protein contain an activation domain. Field's group went on to demonstrate the feasibility of this concept using the two yeast proteins SNF1 and SNF4 which had previously been shown to interact *in vitro*. They fused SNF1 to a DNA-binding domain and SNF4 to an activation domain and together these proteins were shown activate transcription of a reporter gene.

(b) Yeast two-hybrid systems

All two-hybrid systems developed to date share a number of common features. First, all use a plasmid which encodes a 'bait' protein that is brought to a specific DNA sequence by fusion to a DNA-binding domain. Second, they employ one or more reporter genes with upstream binding sites for the DNA-binding domain/bait protein. Third, they all use a plasmid that encodes 'target' proteins fused to an activation domain.

However, these systems also differ in a number of features. Bait proteins differ in their DNA-binding domains: some systems use baits that contain the *E. coli* LexA repressor protein which binds to appropriate operators, whilst other systems use bait proteins that contain a portion of the yeast GAL4 protein which binds to the GAL upstream activation sequence (UAS). Reporter genes used in two-hybrid systems also vary. The products of some reporter genes (e.g. HIS3 and LEU2) allow yeast cells expressing them to be selected by growth on appropriate medium, whereas the products of others (e.g. LacZ) allow cells expressing them to be visually screened. The number and affinity of upstream binding sites for the bait proteins can also differ and these differences affect the strength of protein interactions that can be detected.

Target proteins differ in their activation domains and whether they contain other useful regions such as NLS's and epitope tags. Although some

activation domains are stronger than others, their expression may damage yeast cells when strong activation domains come into contact with the DNA.

A number of developments in recent years have greatly improved the efficacy and viability of yeast two-hybrid systems. The first generation of DNA-binding domain and activation domain plasmids were driven by a truncated 410-bp ADH-1 promoter which, at one point, was believed to be sufficient for high-level expression (Beier & Young, 1982). However, in most vector constructs this truncated promoter led to low or very low levels of fusion protein expression. The high-level expression reported by Beier & Young was apparently due to a segment of DNA derived from pBR322 which, by chance, enhances transcriptional activity in yeast (Tornow & Santangelo, 1990). Therefore in the second generation of plasmids, strong constitutive fusion protein expression is driven by the 410-bp truncated ADH1 promoter adjacent to this enhancing pBR322 segment or by the full-length ADH1 promoter. Fusion proteins expressed by the second generation of plasmids can be detected by Western blots and higher levels of expression allow the detection of weak or transient protein-protein interactions.

The development of a cycloheximide resistant yeast strain precludes the requirement for the lengthy recovery of the plasmids through yeast. The plasmid encoding the DNA-binding domain fusion protein has the CYH2 gene conferring cycloheximide sensitivity and if the yeast are grown in cycloheximide, the plasmid expressing the DNA-binding domain fusion

protein will be eliminated. The resulting strain can then be mated to a strain that has been transformed with all the appropriate control plasmids.

Finally, the commercial availability of Synthetic Dropout Media which contain all but one or more amino acids essential for yeast growth has abrogated the laborious and time-consuming task of preparing the different amino acid combinations for yeast culture.

(c) Derivatives of the yeast two-hybrid system

Although yeast cells are still used in the majority of two-hybrid sytems, a number of two-hybrid mammalian systems have recently been developed. In one system, the DNA binding domain of GAL4 and the activation domain of the HSV-1 VP16 protein provide functional transcriptional activation (Fearon *et al.*, 1992). This system utilises five GAL4 binding sites positioned upstream of the firefly luciferase gene to provide a sensitive reporter system in mammalian cells. In another system, the interaction of a VP16-tagged derivative with a GAL4 DNA-binding domain/bait protein drives the synthesis of simian virus 40 T antigen. This antigen in turn promotes the replication of the target plasmid which carries a SV40 origin (Vasavada *et al.*, 1991).

The mammalian two-hybrid system can be used to confirm putative protein-protein interactions that were initially identified either by the yeast two-hybrid system or by *in vitro* biochemical methods. A two-hybrid system for

mammalian cells allows characterization of mammalian protein-protein interactions within a cellular environment that more closely mimics the native protein environment. In addition, differences exist between yeast and mammalian cells in post-translational modifications such as glycosylation, phosphorylation and acylation, as well as in the intracellular localization of proteins. These types of protein modifications may influence the ability of protein domains to interact. Therefore by using a mammalian two-hybrid system a number of false positives obtained with the yeast two-hybrid system can be eliminated and true positive results can be confirmed in a more natural environment. In its present state, however, the efficiencies of the mammalian two-hybrid system do not allow for screening of complete mammalian libraries.

The yeast-based system retains a number of advantages, including the ease of transformation of yeast cells and the availability of nutritional and well-characterized reporter genes for direct selection. Furthermore, yeast proteins are less likely to bind a mammalian target protein, preventing its interaction with a protein encoded by a library. Finally, mammalian cell-based systems require significant experimental resources whereas the yeast-based system is comparatively rapid and cost-effective.

The yeast one-hybrid system is an adaptation of the two-hybrid system and is used for studying DNA-protein interactions. In this system cDNA candidates that may encode the protein of interest are expressed as fusions with the activation domain. If the activation domain/target protein interacts with a

DNA-binding sequence the HIS3 and LacZ reporter genes are expressed. DNA binding can then be confirmed independently by sequencing, database searches, and in vitro DNA-binding assays.

The yeast three-hybrid system (reviewed in Zhang *et al.*, 1999) is used for studying RNA/protein interactions. Here, a known RNA-binding protein/Lex A fusion protein is used to tether an RNA containing the appropriate binding site. This hybrid protein brings the bound RNA to a reporter gene which is regulated by LexA binding sites. The RNA is bifunctional in that it also contains a binding site for a second RNA-binding protein. This protein is fused to a GAL4 transcriptional activation domain. Thus, the interaction of the RNA with both RNA-binding proteins should result in transcriptional activation of the reporter. This method is capable of defining domains that are necessary for RNA-binding proteins that have been previously characterized. It can also be used to identify the genes for RNA-binding proteins that recognize biologically important RNA sequences. Proteins that interact with such sequences may be identified using the large variety of existing libraries of genomic and cDNA sequences cloned into activation domain vectors.

(d) Future developments

Future developments of the two-hybrid system are likely to involve the design of DNA-binding domain and activation domain plasmids with compatible and less restrictive cloning sites. In addition, a significant

development of the two-hybrid system would be to couple the two-hybrid system with a biochemical assay to test for hybrid protein interactions. One possibility would be to extend the fusion of the DNA-binding domain and target gene to include sequences encoding a region widely used in binding studies such as glutathione-S-transferase (GST) or maltose-binding protein The inclusion of bacterial promoter sequences in the constructs would allow the production of bacterial fusion proteins that could be assayed for biochemical interaction both *in vivo* and *in vitro.*

2.1.3 *In Vitro* or *In Vivo* - Why Choose the Two-Hybrid System?

A variety of biochemical methods have been used to study protein-protein interactions. Some of these methods include protein affinity chromatography, affinity blotting, co-immunoprecipitation and chemical cross-linking.

However, the two-hybrid system is a very sensitive method for detecting protein-protein interactions and whilst it cannot be guaranteed to identify all proteins which interact with the bait protein, it can detect binding that is often beyond the limits of *in vitro* detection methods such as immunoprecipitation. For example, interaction of the mammalian Ras protein with the protein kinase Raf was detected in the two-hybrid system, but had not been detected by co-immunoprecipitation (Aelst, 1993). The sensitivity of the method is partly due to the fact that the hybrid proteins are generally produced at a high level by strong promoters on high-copy vectors

and this favours complex formation. In addition, interaction of the activation domain with the basal transcription machinery may help stabilize a weak protein-protein interaction. In the two-hybrid system interacting complexes do not need to survive several washes which diminish the signal but are necessary to prove specificity in many *in vitro* methods. Furthermore, the detection method effectively amplifies the signal because it produces mRNA which generates multiple stable enzymes per transcript. The resultant phenotype, blue colour in the β-galactosidase assay and the synthesis of the HIS3 protein, is extremely sensitive.

As the two-hybrid system is performed *in vivo*, conditions are similar to those in which protein interactions normally occur. In addition, a number of protein-protein interactions are mediated by mRNA and this molecule will be present in the two-hybrid system but not in many biochemical systems. Perhaps the greatest advantage of the system is that when positive plasmids are identified, a portion of the gene encoding the interacting protein is immediately available for sequencing and the identity of interactors can be determined very quickly. Furthermore, purified target protein or antibody against the protein is not required for the two-hybrid assay and these may not be available to those studying proteins that are not well characterized.

The high sensitivity of the yeast two-hybrid system combined with the other advantages described above made it an attractive system for use in this study. However it is important to note that the system has a number of potential

drawbacks and these were considered carefully before a decision was made to use the two-hybrid system here.

For example, some hybrid proteins may not be stably expressed in yeast cells or localized to the nucleus and a number of hybrid genes such as the cell-cycle regulators cyclin A and cyclin B are either harmful or lethal when expressed in yeast (reviewed in Estojak *et al.*, 1995). In addition, since proteins are targeted to the nucleus, interactions which require post-translational modifications that take place within the endoplasmic retiticulum such as glycolysation and disulfide bond formation may not occur in this system. Proteins which require other modifications for association such as phosphorylation or aceytlation by non-yeast proteins also may not be suitable. Further, GAL4 domains may interfere with the ability of some tests proteins to interact. Although these factors can be problematic in the yeast two-hybrid system, in practice they rarely cause two-hybrid screens to be abandoned. A much more significant and common problem to be encountered is that some DNA-binding-protein hybrids activate transcription autonomously. Such proteins can of course be truncated to remove the regions of the protein which are causing transcription without disrupting the domains required for protein-protein interactions.

Finally, the high sensitivity of the two-hybrid system can create a large number of false positives. Screening assays can eliminate many of these but some proteins appear to have a low affinity for a large number of proteins and such proteins are frequently described as being 'sticky'. Thus perhaps

the greatest drawback of the two-hybrid system is the difficulty in determining whether positives are representative of *in vivo* cellular interactions.

Following careful consideration the two-hybrid system was chosen for the detection of protein-protein interactions in this project with the proviso that any putative positives arising from the this work would be screened independently using an *in vitro* method and that the two-hybrid system would be used only as a first step to determine biologically significant IE63-protein interactions.

2.1.4 Applications of the Two-Hybrid System

The three principal uses of the two-hybrid system at present are the confirmation of protein-protein interactions between two known partners, detection of protein-protein interactions between a bait protein and a library of random proteins and the mapping of sequences required for a given protein-protein interaction.

In the future, two-hybrid systems may be used routinely to identify new targets for pharmaceuticals. For example a two-hybrid system has been used to identify new nuclear hormone receptors by expressing a LexA-retinoic acid receptor bait and coexpressing this with an activation-tagged liver cDNA library (cited in Mendelsohn & Brent, 1994). Four new retinoic acid interacting proteins were identified, two of which were novel nuclear

hormone receptors. The ligands for these receptors are likely to be biologically active and may have a pharmaceutical significance.

An additional industrial use of the two-hybrid system is to find peptides or compounds that interact with a target protein to disrupt its normal interactions. This strategy was used to identify peptides that interact with cyclin-dependent kinase 2 to inhibit its kinase activity (Colas *et al.*, 1996). In this study, an activation domain library was constructed that encoded a number of random peptides displayed in the active site of *E. coli* thioredoxin. The peptides that interacted with cyclin-dependent kinase 2 were identified, and biochemical analysis demonstrated the ability of these peptides to inhibit the kinase activity of this protein.

2.2 MATERIALS

2.2.1 Yeast Cells

Table 2.1 describes the yeast cells used in this project. All yeast strains were purchased from Clontech.

In the absence of a positive two-hybrid interaction; CG-1945 and HF7C exhibit varying levels of constitutive leaky histidine expression. Leaky histidine expression can be controlled by including 5mM 3-amino-1,2,4-triazole (3-AT) in the medium which suppresses background growth. 3-AT is a competitive inhibitor of the yeast HIS3 protein.

2.2.2 Yeast Culture Medium

YPD Medium	2% (w:v) peptone, 1% (w:v) yeast extract, 1.8% (w:v) agar (plates only), 2% (w:v) glucose. pH 5.8
Synthetic Dropout (SD) Medium	0.67% (w:v) yeast nitrogen base without amino acids, 2% (w:v) agar (plates only), 2% (w:v) glucose, 0.077% (w:v) -His dropout (DO)

	supplement or 0.067% (w:v) -Leu DO supplement or 0.074% (w:v) -Trp DO supplement or 0.064% (w:v) -Leu/-Trp DO supplement or 0.062% (w:v) -Leu/-Trp/-His DO supplement or 0.077% (w:v) -His DO supplement or 0.077% (w:v) -Ura DO supplement. pH 5.8. Dropout supplement used depends on intended use of the medium.
Dropout Supplements	Purchased from Clontech. Includes all but one or more of the amino acids adenine, arginine, aspartic acid, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, tyrosine, valine and uracil depending on intended use of medium.

Table 2.1. GENOTYPES AND USAGE OF THE YEAST HOST STRAINS

Strain	Required for	Genotype	Reporter Gene(s)	Transformation Markers
Saccharomyces cerevisiae Y187	Testing for an interaction between two proteins Matchmaker Two-Hybrid System 2	MATα, ura3-52, his3-200, ade 2-101, trp 1-901, leu 2-3, 112, gal4Δ, met$^-$, gal80Δ, URA3::GAL1$_{UAS}$-GAL1$_{TATA}$-lacZ	lacZ	trp1, leu2
Saccharomyces cerevisiae CG-1945	Screening cDNA libraries for proteins that interact with bait protein Matchmaker Two-Hybrid System 2	MATa, ura3-52, his3-200, lys2-801, ade2-101, trp1-901, leu2-3, 112, gal4-542, gal80-538, cyh^{r}2, LYS2::GAL1$_{UAS}$-GAL1$_{TATA}$-HIS3, URA3::GAL4$_{17\text{-mers(x3)}}$-CyC1$_{TATA}$-lacZ	HIS3, lacZ	trp1, leu2, cyh^{r}2
Saccharomyces cerevisiae SFY526	Testing for an interaction between two proteins Matchmaker Two-Hybrid System 1	MATα, ura3-52, his3-200, ade 2-101, lys 2-801, trp 1-901, leu 2-3, 112, can^r, gal4-542, gal80-538, URA3::GAL1-lacZ	lacZ	trp1, leu2
Saccharomyces cerevisiae HF7C	Screening cDNA libraries for proteins that interact with bait protein Matchmaker Two-Hybrid System 1	MATa, ura3-52, his3-200, lys2-801, ade2-101, trp1-901, leu2-3, 112, gal4-542, gal80-538, LYS2::GAL1-HIS3, URA3::GAL4$_{17\text{-mers(x3)}}$-CyC1-lacZ	HIS3, lacZ	trp1, leu2

2.2.3 Bacterial Cells

Escherichia coli K12 strain DH5α (*hsd* R17, *lac* U169, *gyr* A69, *thi*-1, *rel* A1, *rec* A1, *end* A1, *sup*E44, φ80, *lacZ* M15) was used for the propagation of plasmids.

2.2.4 Bacterial Culture Media

Luria-Bertani (LB) Medium	1% (w:v) bacto-tryptone, 0.5%(w:v) bacto-yeast extract, 177 mM $NaCl_2$, 1.5% (w:v) agarose (plates only). pH 7.0.
Terrific Broth	1.2% (w:v) bacto-tryptone, 2.4% (w:v) bacto-yeast extract, 4% (v/v) glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4.

2.2.5 Two-Hybrid System

The two-hybrid systems used were the Matchmaker Two-Hybrid System I and the Matchmaker Two-Hybrid System II both of which were purchased from Clontech.

2.2.6 Plasmids

pGBT9 (Clontech) is a shuttle vector that can be propagated in *E. coli* and yeast. The plasmid was supplied with the Matchmaker Two-Hybrid System I and was used to generate fusions of IE63 with the GAL4 DNA-binding domain. pGBT9 carries the tryptophan gene for selection in yeast cells that lack the ability to produce this amino acid. This plasmid is under the control of the truncated ADH1 promoter.

pGAD424 (Clontech) is a shuttle vector that can be propagated in *E. coli* and yeast. The plasmid was supplied with the Matchmaker Two-Hybrid System I and was used to generate fusions of known proteins with the GAL4 AD. pGAD424 carries the leucine gene for selection in yeast cells that lack the ability to produce this amino acid. This plasmid is under the control of the truncated ADH1 promoter.

pAS2-1 (Clontech) is a shuttle vector that can be propagated in *E. coli* and yeast. The plasmid was supplied with the Matchmaker Two-Hybrid System II and was used to generate fusions of IE63 with the GAL4 DNA-binding domain. pAS2-1 carries the tryptophan gene for selection in yeast cells that lack the ability to produce this amino acid. This plasmid is under the control of the full-length ADH1 promoter and expresses protein at a higher level than pGBT9.

pACT2 (Clontech) is a shuttle vector that can be propagated in *E. coli* and yeast. The plasmid was supplied with the Matchmaker Two-Hybrid System II and was used to generate fusions of known proteins with the GAL4 AD. pACT2 carries the leucine gene for selection in yeast cells that lack the ability to produce this amino acid. This plasmid is under the control of the truncated ADH1 promoter that is adjacent to a section of pBR322 which acts as a transcriptional enhancer in yeast. This plasmid expresses protein at a higher level than pGAD424.

pGAD GH (Clontech) is a shuttle vector that can be propagated in *E. coli* and yeast. The plasmid was used to generate fusions of a human HeLa cDNA library with the GAL4 AD. pGAD GH carries the leucine gene for selection in yeast cells that lack the ability to produce this amino acid. This plasmid is under the control of the full-length ADH1 promoter.

pCL1 (Clontech) was used as a positive control plasmid in the two-hybrid system vector pACT2 and it encodes and expresses the full-length, wild-type GAL4 protein.

PTD1-1 (Clontech) encodes the SV40 large T-antigen protein fused to the GAL4 activation domain in pACT2. The protein was known to interact with the murine p53 protein which is expressed fused to the GAL4 binding domain of the plasmid pVA3-1. The two plasmids, cotransformed into yeast cells, were used as a positive control in the two-hybrid system.

pLAM5'-1 (Clontech) encodes the human Lamin C protein fused to the GAL4 binding domain in pAS2-1. The plasmid was known not to interact with the SV40 Large T-antigen which is encoded by the pTD1-1 plasmid. The two plasmids, cotransformed into yeast cells, were used as a negative control in the two-hybrid system.

pSG130 is a derivative of the pUC18 vector with the full-length IE63 gene cloned under the control of its own promoter. The plasmid was supplied by Rozanne Sandri-Goldin and was used as a template from which portions of the IE63 gene were generated by PCR.

The regions of hnRNP K required for interaction with IE63 were mapped using the following truncations (Tsieh *et al.*, 1998) which were a kind gift from Tsai-Yuan Hsieh (University of Southern California School of Medicine, Los Angeles) pGAD424-hnRNP K (aa 327-463), pGAD424-hnRNP K (aa 1-327), pGAD424-hnRNP K (aa 250-463), pGAD424-hnRNP K (276-463), pGAD424-hnRNP K (aa 1-276).

The RIP1 (Rab) and CRM1 clones were a kind gift from Dr Bryan Cullen (Howard Hughes Memorial Institute, Durham, North Carolina) and were cloned into pGAD424 and pGBT9 respectively.

2.2.7 cDNA Libraries

A human HeLa cDNA library cloned into the pGAD GH vector (Clontech) was used to generate target proteins in a two-hybrid screen where the bait protein was the IE63 truncation 502C expressed by pAS2-1. The library was a gift from Paul Yeo (Institute of Virology, Glasgow) and it had been amplified by Dr Yeo. The library contained an estimated 93% of colonies with inserts and the total number of independent clones was 6.0×10^6. The average cDNA size was 1.5 kb with a cDNA size range of 0.4-4.0 kb. The titer of the library was $\geq 10^8$ cfu (of DH10B)/ml in LB broth and 25% sterile glycerol. The cloning sites used were *Eco*RI and *Xho* I.

A BHK cDNA library cloned into pAD-GAL4 (Stratagene) was used to generate target proteins in a two-hybrid screen where the bait protein was the IE63 truncation 442C BD expressed by pGBT9. The library contained an estimated 90% of colonies with inserts and the titre was 2×10^6 pfu/ml. The cloning sites used were *Eco*RI and *Sma* I.

2.2.8 Enzymes

Restriction enzymes, *Taq* polymerase and Klenow were obtained from Boehringer Mannheim. Vent Polymerase was purchased from New England Biolabs and T4 Ligase was purchased from Gibco BRL.

2.2.9 Antibodies

Mouse monoclonal anti-IE63 H1113 (Ackerman *et al.*, 1984) was supplied by the Goodwin Institute for Cancer Research.

GAL4 DNA-BD and AD monoclonal antibodies were purified from serum-free media of mouse hybridoma cultures and were purchased from Clontech.

2.2.10 Oligonucleotides

Oligonucleotides for PCR synthesis were obtained from Cruachem.

2.2.11 Sequencing Primers

The primer that was used for sequencing towards the junction of the GAL4 DNA-binding domain and target proteins cloned into the plasmids pGBT9 and pAS2-1 was the 17-mer 5'-TCATCGGAAGAGAGTAG-3' which binds 59 bp from the multiple cloning start site. This primer was purchased from Clontech.

The primer that was used for sequencing towards the junction of the GAL4 activation domain and candidate proteins cloned into pGAD424, pACT2 and pGAD GH plasmids was the 17-mer 5'-TACCACTACAATGGATG-3' which binds 107 bp from the multiple cloning start site. This primer was purchased from Clontech.

2.2.12 Molecular Weight Markers

A 1 kb ladder DNA marker and a 220 kDa rainbow protein marker were obtained from Amersham.

2.2.13 Chemicals and Miscellaneous Reagents

Chemical/Reagent	Manufacturer
3-amino-1,2,4-triazole (3-AT)	Sigma
Cycloheximide	Sigma
Herring Testes Carrier DNA	Clontech
Ampicillin	Smith Kline Beecham Research
5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-GAL)	Melford Laboratories
Zymolase	ICN Biochemicals

2.2.14 Commonly used Buffers and Solutions

DNA loading buffer	50% (w:v) glycerol, 445 mM Tris, 10 mM EDTA, 445 mM boric acid, 0.1% bromophenol blue, 0.1% xylene cyanol
2x Laemmli sample buffer	10% (v:v) 2 mercaptoethanol, 6%

	(w:v) SDS, 20% (v:v) glycerol, 0.2 mg/mL bromophenol blue, 0.025 x Laemmli stacking buffer
2.5x Laemmli stacking buffer	0.3 M Tris-HCl (pH 6.8), 0.25% (w:v) SDS
STET buffer	1% Triton X-100, 10 mM Tris-HCl (pH 8.0), 234 mM sucrose, 50 mM EDTA
Z buffer	113 mM Na_2HPO_4, 40 mM $NaH_2PO_4.H_2O$, 10 mM KCl, 1 mM $MgSO_4.7H_2O$
PBS	170 mM NaCl, 3.4 mM KCl, 10 mM $NaHPO_4$, 2 mM KH_2PO_4 (pH 7.2)
TAE	40 mM Tris, 0.1% acetic acid, 50 mM EDTA
Yeast protoplasting buffer	0.01% (v:v) 2 mercaptoethanol, 100 mM Tris HCl (pH 7.5), 10 mM EDTA (pH 7.5), 0.2 mg/ml

	Zymolase (20,000 units/mg)
Yeast lysis solution	0.2 M NaOH, 1% (w:v) SDS
X-gal stock solution	6% (w:v) X-gal in N,N-dimethylformamide (DMF)
Z buffer/X-gal solution	0.27% (v:v) 2-mercaptoethanol, 1.67% (v:v)X-gal stock solution, 98% (v:v) Z buffer
PEG/LiAc solution	40% (w:v) PEG 4,000, 1x TE buffer, 1x LiAc
Tris-glycine electrophoresis buffer	25 mM Tris, 250 mM glycine (pH 8.3), 0.1% (w:v) SDS
Protein transfer buffer	39 mM glycine, 48 mM Tris base, 0.037% (w:v) SDS, 20% (v:v) methanol

2.2.15 Miscellaneous

Gene Clean Kit	BIO-101
90 mm filter paper #5	Whatman

Miniprep Kit	QIAGEN
Maxiprep Kit	QIAGEN

2.3 METHODS

2.3.1 Yeast Culture and Manipulation

(a) Growth and maintenance of yeast host strains

Yeast host strains were stored in YPD medium with 25% glycerol at -70°C. To recover a strain, cells were scraped from the frozen stock with an inoculation loop and streaked onto YPD plates. The plates were inverted and incubated for 3-5 days at 30°C until the colonies reached approximately 2 mm in diameter. The working stock plates were then stored at 4°C for a maximum of two weeks. After two weeks a single colony was taken from the plate and streaked onto a fresh YPD plate.

(b) Phenotype verification

The phenotypes of all of the yeast strains were verified before transforming them with plasmids. A single colony was taken from the stock plate and streaked onto one of four SD plates lacking either tryptophan, leucine,

histidine or uracil. Plates were incubated at 30°C for 3-5 days and scored for growth.

(c) Control transformations

To ensure that the two-hybrid system was working efficiently and to verify the transformation markers on the plasmids, the control transformations described in Table 2.2 were carried out using the small scale transformation method described in Section 2.2.1 (d). Plates were scored for growth and a filter lift was carried out to determine the LacZ phenotype of the transformants.

(d) Small scale transformation of yeast cells

The small scale transformation procedure was used to verify that the DNA-BD/target protein does not autonomously activate reporter genes, to perform control experiments and to do single plasmid transformations for subsequent mating assays.

To make yeast competent cells, several colonies of the appropriate yeast strain were inoculated into 1 ml of YPD medium and vortexed vigorously to disperse any clumps of cells. The yeast cells were then transferred into a flask containing 100 ml of YPD medium and incubated at 30°C for 16-18 h with shaking at 250 rpm to stationary phase (OD_{600}>1.5). The overnight culture was added to 300 ml of YPD media to a final OD_{600} of 0.2-0.3 and

Table 2.2. CONTROL TRANSFORMATIONS CARRIED OUT TO ENSURE THAT THE TWO-HYBRID SYSTEM WAS WORKING EFFECTIVELY

DNA-BD plasmid	**AD plasmid**	**Selection Medium**	**Expected LacZ phenotype**
-	pCL1	-Leu	blue
pAS2-1	-	-Trp	white
-	pACT2	-Leu	white
pAS2-1	pACT2	-Leu, -Trp	white
pVA3-1	-	-Trp	white
pVA3-1	pACT2	-Leu, -Trp	white
-	pTD1-1	-Leu	white
pAS2-1	pTD1-1	-Leu, -Trp	white
pVA3-1	pTD1-1	-Leu, -Trp	blue
pLAM5'-1	-	-Trp	white
pLAM5'-1	pTD1-1	-Leu-Trp	white

incubated for 3 h at 30°C with shaking at 230 rpm. The cells were then centrifuged at 3,000 rpm for 5 min, the supernatant was discarded and the cells were resuspended in 25-50 ml of water by vortexing. The cells were centrifuged once more at 3,000 rpm for 5 min and the supernatant discarded. The cell pellet was resuspend in 1.5 ml of sterile 1x TE/LiAc solution.

To transform the competent cells, 1 μg of DNA-BD vector construct, 1 μg of AD vector construct and 0.1 mg of herring testes carrier DNA were added to each reaction tube together with 0.1 ml of the yeast competent cells. 0.6 ml of sterile PEG/LiAc solution was added to each tube and vortexed to mix. The reactions were incubated at 30°C for 90 min with shaking at 200 rpm. DMSO was added to 10% and the cells were heat shocked for 15 min in a 42°C water bath. The cells were then chilled on ice and pelletted by centrifugation at 14,000 rpm and 4°C for 5 sec. The supernatant was discarded and the cells were resuspended in 0.5 ml of 1x TE buffer. 100 μl of the transformation mixture was plated onto each 100-mm SD agar plate. Plates were incubated for 5 days at 30°C.

(e) Library scale transformation of yeast cells

To make the yeast competent cells, several colonies of the appropriate yeast strain were inoculated into 1 ml of YPD medium and vortexed vigorously to disperse any clumps of cells. The yeast cells were then transferred into a flask containing 300 ml of YPD medium and incubated at 30°C for 16-18 hs

with shaking at 250 rpm to stationary phase (OD_{600}>1.5). The overnight culture was added to 1 L of YPD media to a final OD_{600} of 0.2-0.3 and incubated for 3 hs at 30°C with shaking at 230 rpm. The cells were then centrifuged at 3,000 rpm for 5 min, the supernatant was discarded and the cells were resuspended in 500 ml of water by vortexing. The cells were centrifuged once more at 3,000 rpm for 5 min and the superatant discarded. The cell pellet was resuspend in 8 ml of sterile 1x TE/LiAc solution.

To transform the competent cells, 1 mg of DNA-BD vector construct, 0.5 mg of AD vector construct and 20 mg of herring testes carrier DNA were added to 8 ml of the yeast competent cells. 60 ml of sterile PEG/LiAc solution was added to each tube and vortexed to mix. The reaction was incubated at 30°C for 90 min with shaking at 200 rpm. DMSO was added to 10% and the cells were heat shocked for 15 min in a 42°C water bath, with occasional swirling to mix the cells. The cells were then chilled on ice and pelletted by centrifugation at 14,000 rpm for 5 min. The supernatant was discarded and the cells were resuspended in 10 ml of 1x TE buffer. 0.3 ml of the transformation mixture was plated onto each of 50 140-mm plates SD/-Trp/-Leu-His plates. In addition, 100 μL of a 1:1000, 1:100 and 1:10 dilution were spread onto SD/-Trp/-Leu plates for co-transformation efficiency controls. 1 μL (diluted in 100 μL H_2O) was spread onto 100 mm SD/-Trp and SD/-Leu plates to check the transformation efficiency of each plasmid individually. Plates were incubated for 10 days at 30°C.

(f) Calculation of co-transformation efficiency and number of clones screened

To calculate the co-transformation efficiency of a library screen, the number of colonies (cfu) growing on the SD/-Leu/-Trp dilution plate that had 30-300 cfu was counted and the number inserted into the following equation.

$$\frac{\text{cfu x total suspension vol. (µl)}}{\text{Vol. Plated (µl) x dilution factor x amt. DNA used (µg)*}} = \text{cfu/µg DNA}$$

* In a cotransformation, this is the amount of the limiting plasmid

To estimate the number of clones screened:

cfu/µg x amt. DNA used (µg) = # of clones screened

(g) False positive testing

In a yeast two-hybrid library screen, true positives exhibit reporter gene expression only when they contain both the DNA-BD/target plasmid and the AD/library plasmid which encode fusion proteins that interact with each other. False positives are His^+ or $LacZ^+$ cotransformant colonies that carry plasmids that do not encode hybrid proteins that interact. False positive were eliminated as described below.

Colonies growing on SD/-Leu/-Trp/-His/+3AT 10 days after a two-hybrid library screen were re-streaked onto a fresh SD/-Leu/-Trp/-His/+3AT plate and incubated at 30°C for 4 days. A colony-lift filter assay was carried out on these plates and any colonies with a $LacZ^+$ phenotype were plated onto SD/-Leu/-Trp and incubated at 30 °C for 4 days to segregate multiple AD/library plasmids within a single colony. Colonies growing on the SD/-Leu/-Trp plates were streaked onto fresh SD/-Leu/-Trp once again and incubated at 30 °C for 4 days. Surviving colonies were then transferred back onto SD/-Leu/-Trp/-His/+3AT and incubated at 30 °C for 4 days. A colony-filter lift assay was carried out to confirm the positive LacZ phenotype.

Following elimination of the DNA-BD/bait construct with cycloheximide (see Section 2.3.1, i) and isolation of the plasmid from the yeast plasmids were transformed back into yeast cells using the small-scale transformation method and yeast mating in the combinations described in Table 2.3.

(h) Colony-lift filter assay

Colony-lift filter assays were used to determine the LacZ phenotype of yeast cells.

Fresh colonies that were 1-3 mm in diameter were used for colony-lift assays. For each of the transformants to be assayed, a sterile Whatman #5 filter was

Table 2.3 COMBINATIONS OF PLASMIDS COTRANSFORMED INTO YEAST TO VERIFY POSITIVES

Plasmid 1 (in Y187) (MATa)	**Plasmid 2** (in CG-1945) (MATa)	**Expected LacZ phenotype for a true positive**	**Expected His3 phenotype for a true positive**
DNA-BD/no insert	AD/no insert	White	no growth
DNA-BD/bait	AD/no insert	White	no growth
DNA-BD/no insert	AD/library	White	no growth
DNA-BD/bait	AD/library	Blue	positive growth
DNA-BD/lamin C	AD/library	White	no growth

pre-soaked in 2 ml of Z Buffer/X-gal solution in a sterile 100-mm plate. A clean , dry filter was placed over the surface of the plate of the colonies to be assayed and holes were cut in the filter in three asymmetric locations to orient the filter to the agar. The filters were then lifted from the agar plate with forceps and transferred, colony side up, into a pool of liquid nitrogen to permeabilize the cells. The filters were completely submerged in the liquid nitrogen for at least 15 sec and then removed and allowed to thaw at room temperature. The filter was then placed, colony side up, on the filter pre-soaked in Z Buffer/X-gal solution and incubated at 30°C for up to 8 hs.

Blue colonies indicate a positive LacZ phenotype and white colonies indicate a negative LacZ phenotype.

(i) Cycloheximide selection for colonies that had lost the DNA-BD/target plasmid

The yeast host strain CG-1945 is cycloheximide resistant (Cyh^R) because it carries the cyh^r2 mutant allele. The wild-type CYH^s2 gene is dominant to the cyh^r2 mutant allele. Therefore, when transformed with the DNA-BD target plasmid which contains the wild-type CYH^s2 gene, CG-1945 becomes sensitive to cycloheximide. CG-1945 cotransformed with the DNA-BD target plasmid and the AD/library plasmid will also be sensitive to cyclohexmide. Thus, it is possible to select for yeast cells that have spontaneously lost the DNA-BD target plasmid while retaining the AD/library plasmid by plating the cotransformants onto medium containing

cycloheximide as described below. This allows quick isolation of yeast containing AD/library plasmids only which can then be isolated and sequenced.

Colonies that were 1-3 mm in diameter were restreaked onto SD/-Leu/+ 0.3 mM cycloheximide plates and incubated for 5 days at 30°C until individual Cyh^R colonies appeared. The loss of the DNA-BD target plasmid was verified by plating the putative Cyh^R colonies onto SD/-Trp, where colonies lacking the DNA-BD target plasmid should not grow, and SD/-Leu where growth indicates that the DNA-BD target plasmid has been successfully eliminated. A colony-filter lift assay was then carried out on colonies growing on SD/-Leu to ensure that the AD/library constructs could not activate gene expression autonomously.

(j) Yeast mating

Yeast mating allows the introduction of two different plasmids into the same host cells. It can therefore be used to verify that candidate AD/library proteins identified in the two-hybrid library screen can activate the reporter genes only in the presence of the DNA-BD/target protein. It can also be used to test for an interaction between two previously cloned proteins and to map regions of the proteins required for interactions.

The yeast host strain Y187 was transformed with plasmids that express DNA-BD truncation proteins and the host strain CG-1945 was transformed

with plasmids expressing DNA-AD truncation proteins. One colony of each type was picked and inoculated together into 0.5 ml of YPD medium in a 1.5-ml microcentrifuge tube. Each tube was then vortexed and incubated at 30°C with shaking at 250 rpm overnight. 50 µl of each mating culture was plated onto SD/-Leu/-Trp/-His/+3-AT to select for interacting proteins and also onto SD/-Leu/-Trp to ensure that the mating had been successful. Plates were incubated at 30°C for 5 days and scored for growth. A colony-filter lift assay was performed on the diploid cells to assay for β-galactosidase activity.

(k) Plasmid isolation from yeast cells

A yeast culture was grown overnight in 2 ml of YPD at 30°C to early stationary growth phase (~2 x 10^8 cells/ml). 1.3 ml of the culture was transferred to a 1.5 ml microfuge tube and the cells were harvested by centrifugation at 6,000 rpm for 3 min. The supernatant was discarded and the cells resuspended in 1 ml of sterile water by vortexing. To lyse the yeast, cells were harvested once again by centrifugation at 6,000 rpm for 3 min and the supernatant was discarded. The cells were resuspended in 0.2 ml of protoplasting buffer by vortexing and incubated at 37°C for 1-2 h with occasional inversion of the tube to prevent sedimentation of the cells. 0.2 ml of lysis solution was added and the cells were mixed gently by inversion. The sample was then incubated at 65°C for 20 min and cooled rapidly on ice. 0.2 ml of 5 M potassium acetate (pH 5.4) was added to the solution and incubated on ice for 15 min. Insoluble potassium dodecylsulphate and

denatured proteins precipitate from the solution at this stage and these were removed by centrifuging for 3 min in a microcentrifuge at room temperature.

The superantant was transferred to a fresh tube without disturbing the pellet and 0.6 vol. of isopropanol was added to the tube and mixed gently by inversion. The solution was allowed to stand at room temperature for 5 min to allow complete precipitation of DNA. DNA was isolated by centrifugation for 30 sec and the supernatant was discarded. 1 ml of 70% ethanol was added to the pellet and mixed by inversion of the microfuge tube. The tube was allowed to stand at room temperature for at least 10 min for salt ions entrapped in the DNA aggregate to diffuse out. The solution was once more centrifuged for 30 sec in a microfuge tube and the supernatant was discarded. The pellet was dried briefly and the DNA was dissolved in 20-50 µl of TE buffer.

(l) Preparation of yeast extracts for SDS-PAGE

Hybrid-proteins were isolated from yeast cells and run on SDS-PAGE gels to show that they were being expressed in yeast and that no degradation was occurring.

A fresh yeast colony for each bait being tested together with a positive control for protein expression, was inoculated into 5 ml of the appropriate selective medium. The colonies were incubated with shaking at 30°C overnight. The cultures were added to fresh selective medium to an OD_{600} =

0.15 and then incubated at 30°C until OD_{600} = 0.7 (~ 7 hs). The culture was then spun at 3,000 rpm for 5 min and the supernatant was removed. 100 µl of 2x Laemmli sample buffer was immediately added to the pellet. The mixture was vortexed and stored at -70°C until required. Prior to analysing by SDS/PAGE/Western blotting, the frozen samples were boiled for 5 min. Samples were then spun at 13K for 5 sec to pellet large debris immediately before loading onto the gel.

(m) Mapping regions of proteins required for interactions

Having determined that the bait protein IE63 interacts with a cellular protein it is of great interest to identify the regions of IE63 which mediate this interaction. This permits previously identified functional domains IE63 to be associated with interactions of specific proteins thus allowing a putative function to be attributed to the interaction. Disruption of the interaction by using truncated IE63 proteins which do not contain the domain required for interaction acts as a convincing negative control.

The small scale transformation method was used to transform yeast with plasmids encoding truncated IE63 proteins and the full-length protein identified to interact with IE63. The same method was also used to transform yeast with the positive control plasmids pVA3-1 and pTD-1 and the negative controls pAS2-1 and pLam5'-1. The LacZ phenotype of the yeast cells was determined using a colony-lift filter assay, the strength of the blue colour being indicative of the strength of protein-protein interaction. True

quantitative comparison of the intensity of blue colour is not justifiable as it depends on colony size and degree of expression of the fusion protein. However, here a semi-quantitative scale was used with the strength of the blue colour being graded by comparison to the positive and negative controls with '++' being equivalent to the positive control, '+' being less than the positive control and '-' being equivalent to the negative control.

To determine the regions of proteins required for interactions, truncations were analysed in seven separate experiments.

2.3.2 Nucleic Acid Manipulation and Cloning Procedures

(a) Preparing bacterial calcium chloride competent cells

A colony of DH5α bacterial cells was inoculated into 2ml of LB and incubated overnight at 37°C with shaking. The entire overnight culture was then added to 100 ml L-broth and grown to an OD_{600} of 0.2 (aproximately 2 hs). The culture was centrifuged at 3,000 rpm for 5 min and the supernatant discarded. The pellet was resuspended in a half volume of 50 mM calcium chloride and was centrifuged once more at 3,000 rpm for 5 min. The supernatant was discarded once more and resuspended in a half volume of 75 mM calcium chloride. The cells were incubated on ice for 30 min. Cells were then centrifuged at 3,000 rpm for 5 min and resupended in 3 ml of 75 mM calcium chloride or 50 mM calcium chloride + 15% glycerol if the cells were to be stored at -70°C. Cells were aliquoted into 0.2 ml volumes and

either used immediately or frozen immediately on dry ice for storage at -70°C.

(b) Bacterial cell transformations

5-10 μg of plasmid DNA was added to a 0.2 ml aliquot of bacterial competent cells and incubated on ice for 30 min. The bacterial cells were heat shocked at 42°C for 2 min and transferred to a sterile universal containing 2 ml of L-broth. The cells were incubated at 37°C for 2 h with shaking. 1.5 ml of the culture was centrifuged at 13,000 rpm for 5 min and resuspended in 100 μl of TE buffer for ligations or low copy number plasmid or 500 μl of TE buffer for high copy number plasmids. 100 μl of cells were plated onto LB-agar plates containing the appropriate antibiotic and incubated overnight at 37°C.

(c) Glycerol stocks of bacterial cultures

For cultures growing in liquid media, 0.85 ml of bacterial culture was added to a 1.5-ml microfuge tube containing 0.15 ml of sterile glycerol. The culture was vortexed to ensure that the glycerol was evenly dispersed and immediately frozen on ethanol-dry ice for storage at -70°C. To recover the bacteria, the frozen surface of the culture was scraped with an inoculation loop and immediately streaked onto LB-agar plates containing the appropriate antibiotic.

For cultures growing on agar plates, a single colony of bacteria was scraped into 2 ml of L-broth in a sterile 1.5-ml microfuge tube. An equal volume of L-broth containing 30% sterile glycerol was added and the mixture vortexed to ensure that the glycerol was completely dispersed. The cells were then dispensed into aliquots and frozen and revived as described above.

(d) DNA gel electrophoresis (minigels)

100 ml of 1% (w:v) agarose and TAE was prepared and heated to boiling. The molten agarose was cooled to ~50°C and 5 µl of ethidum bromide solution (10 mg/ml) was added. The agarose was swirled to mix, poured into a gel cast and allowed to cool. 5 µl of loading buffer was added to each sample and gels were run at 100 V for 30-60 min in 1 x TAE. Gels were visualized under either short-wave UV light for visualization only or long wave UV light if the DNA was to be removed from the gel and used in subsequent experiments.

(e) Small scale preparations of DNA (minipreps)

The method used for small-scale preparation of DNA was dependent upon the intended use of the DNA. The STET preparation method described below was used when the DNA was to be used for diagnostic digests and DNA gel electrophoresis only. The alkaline-lysis/PEG precipitation method was used to yield high-quality, super-coiled plasmid DNA which was

relatively free of contaminating chromosomal DNA and RNA which was used for sequencing. QIAGEN miniprep kits were used for quick DNA preparation for transformation into yeast or bacterial cells. The QIAGEN minipreps were carried out in accordance with the manufacturer's recommended protocol.

To isolate DNA using the STET preparation method, a single colony of bacteria was inoculated into 2 ml of LB containing the appropriate antibiotic and incubated at 37°C for 5-6 hs. 1.5 ml of this culture was centrifuged at 13,000 rpm for 1 min and the supernatant was removed by aspiration. The bacterial pellet was resuspended in 200 µl of STET buffer and 20 µl of lysozyme (10 mg/ml) was added. The mixture was boiled for 40 sec and centrifuged for 10 min at 15,000 rpm. The supernatant was then removed to a clean 1.5-ml microfuge tube and the DNA was precipitated with $1/10^{th}$ volume 3 M sodium acetate and an equal volume of 100% isopropanol on dry ice for 5 min. The DNA was pelleted by centrifugation at 15,000 rpm for 10 min and the DNA was allowed to dry. DNA was then resuspended in 20 µl of water. For restriction enzyme digestions, 5 µl of the sample was digested and RNase A was added to a final concentration of 20 µg/µl during the last h of the digest.

To isolate the DNA using the alkaline-lysis/PEG precipitation method, a single colony of bacteria was inoculated into 2 ml of Terrific Broth containing the appropriate antibiotic overnight at 37°C with shaking. 1.5 ml

of the overnight culture were centrifuged at 13,000 rpm for 1 min. The supernatant was removed by aspiration and the pellet was resuspended in 200 µl of GTE buffer. 300 µl of freshly prepared 0.2 M NaOH/1% SDS solution was added and the contents of the tube were mixed by inversion until the solution cleared. The mixture was then incubated on ice for 5 min and the pH of the solution was neutralised by adding 300 µl of 3.0 M potassium acetate, pH 4.8. The mixture was incubated on ice for 5 min and the cellular debris was removed by centrifugation at 15,000 rpm for 10 min and the supernatant removed to a clean tube. DNase-free RNase was added to a final concentration of 20 µg/µl and the tube was incubated at 37°C for 20 min. After the RNase treatment, the supernatant was extracted four times with 400 µl of chloroform, mixing the layers by hand for 30 sec after each extraction. The tubes were centrifuged for 1 min to separate the phases and the aqueous phase was removed to a clean tube. DNA was precipitated by adding an equal volume of 100% isopropanol and immediately centrifuging the tube at 15,000 rpm for 10 min at room temperature. The DNA pellet was then washed with 500 µl of 70% ethanol and dried under vacuum for 3 min. The pellet was dissolved in 32 µl of deionized H_2O and the DNA was precipitated by adding 8.0 µl of 4 M NaCl and 40 µl of autoclaved 13% PEG and incubating on ice for 20 min. The DNA was pelleted by centrifugation for 15 min at 15,000 rpm at 4°C in a fixed-angled rotor. The supernatant was removed by careful aspiration and the pellet was rinsed with 500 µl of 70% ethanol. The DNA was dried under vacuum for 3 min and resuspended in 20 µl of H_20.

(f) Large scale preparations of plasmid DNA (maxipreps)

Maxipreps were carried out using QIAGEN columns. QIAGEN Maxipreps were carried out in accordance with the manufacturer's recommended protocol. In brief, bacterial cells were lysed under alkaline conditions and the crude lysates were cleared using centrifugation. The cleared lysate was then loaded onto the QIAGEN column where plasmid DNA bound to the QIAGEN resin. RNA, proteins, metabolites, and other low molecular weight impurities were removed by a medium salt wash. Plasmid DNA was eluted in high salt buffer and subsequently concentrated and desalted by isopropanol precipitation.

(g) Nucleic acid purification

Proteins were removed from aqueous solutions of nucleic acids by extracting with phenol:chloroform and chloroform.

An equal volume of phenol:chloroform was added to the nucleic acid sample and the contents of the tube were mixed until an emulsion formed. The mixture was then centrifuged at 15,000 rpm for 15 sec and the aqueous phase was removed to a fresh tube. An equal volume of chloroform was added to the solution and the cells centrifuged once more at 15,000 rpm for 15 sec. The aqueous phase was removed to a fresh tube and the nucleic acid was precipitated by precipitation with ethanol as described below.

(h) DNA precipitation

Nucleic acids were concentrated by precipitation with ethanol. $1/10^{th}$ volume of 3 M sodium acetate (pH 5.2) and 2.5 volumes of ice-cold ethanol were added to the DNA solution. The solution was incubated on dry ice for 10 min and the DNA was recovered by centrifugation at 15,000 rpm for 10 min. The supernatant was removed by aspiration and 500 µl of 70% ethanol was added to the pellet. The sample was re-centrifuged at 15,000 rpm for 2 min and the supernatant was once again removed by aspiration. The DNA was air dried and dissolved in the desired volume of buffer.

(i) DNA quantification

The amount of DNA in a solution was measured by spectrophotometric measurement of the amount of UV irradiation absorbed by the bases or more approximately by running the sample on a minigel and comparing the intensity of ethidium bromide fluorescence of the unknown DNA to the fluorescence of a series of standard DNA solutions.

For spectrophotometric determination readings were taken at wavelengths of 260 nm and 280 nm. The reading at 260 nm allowed calculation of the concentration of the nucleic acid in the sample since an OD of 1 corresponds to approximately 50 µg/ml of double-stranded DNA. The ratio between the

readings at 260 nm and 280 nm provided an estimate of the purity of nucleic acid since pure preparations of DNA have an OD_{260}/OD_{280} reading of 1.8.

Minigel analysis was carried out as described in Section 2.3.2 (d).

(j) Isolation of DNA from agarose gels

DNA to be isolated was run on a minigel as described previously and visualized using long-wavelength UV irradiation . The band of interest was cut from the gel using a scalpel and the DNA was isolated from the gel fragment using the BIO-101 gene clean kit according to the manufacturer's protocol.

(k) Restriction endonuclease digestions

Restriction endonuclease digests were performed using the buffers and conditions recommended by the supplier, the volume of enzyme never exceeding 10% of overall reaction volume.

(l) Ligation reactions - cohesive termini

The plasmid and insert DNA were digested with appropriate restriction enzymes. The DNAs were then purified by extraction with phenol:chloroform and ethanol precipitation as described in sections 2.3.2 (g) and 2.3.2 (h). The DNAs were redissolved in 1x TE and approximately 0.5

μg of plasmid DNA and an equimolar amount of insert DNA were added to a 1.5-ml microfuge tube. 2 μl of 10x bacteriophage T4 DNA ligase buffer and 2 μl of T4 DNA ligase were added to reaction and the final reaction volume was made up to 20 μl with H_2O. The reactions were incubated overnight at room temperature and 5 μl of the reaction mix was used to transform competent *E. coli.*

(m) Sequencing

Sequencing reactions were carried out on an automated ABI -Prism sequencer (Perkin-Elmer) using 4 pm of 1 pm/μl primer and 0.5μg of DNA.

2.3.3 Polymerase Chain Reaction (PCR)

PCR was used to generate IE63 deletion mutants which, when cloned into the two-hybrid system expression vectors, would be expressed as truncated IE63 proteins. Seven mutants were constructed (N400, N250, N397, 270C, 346C, 440C, 502C) and these are illustrated diagramatically in Figure 2.4. All mutants were sequenced to ensure that they did not contain mutations and that they lay in the same reading frame as the two-hybrid system expression vectors.

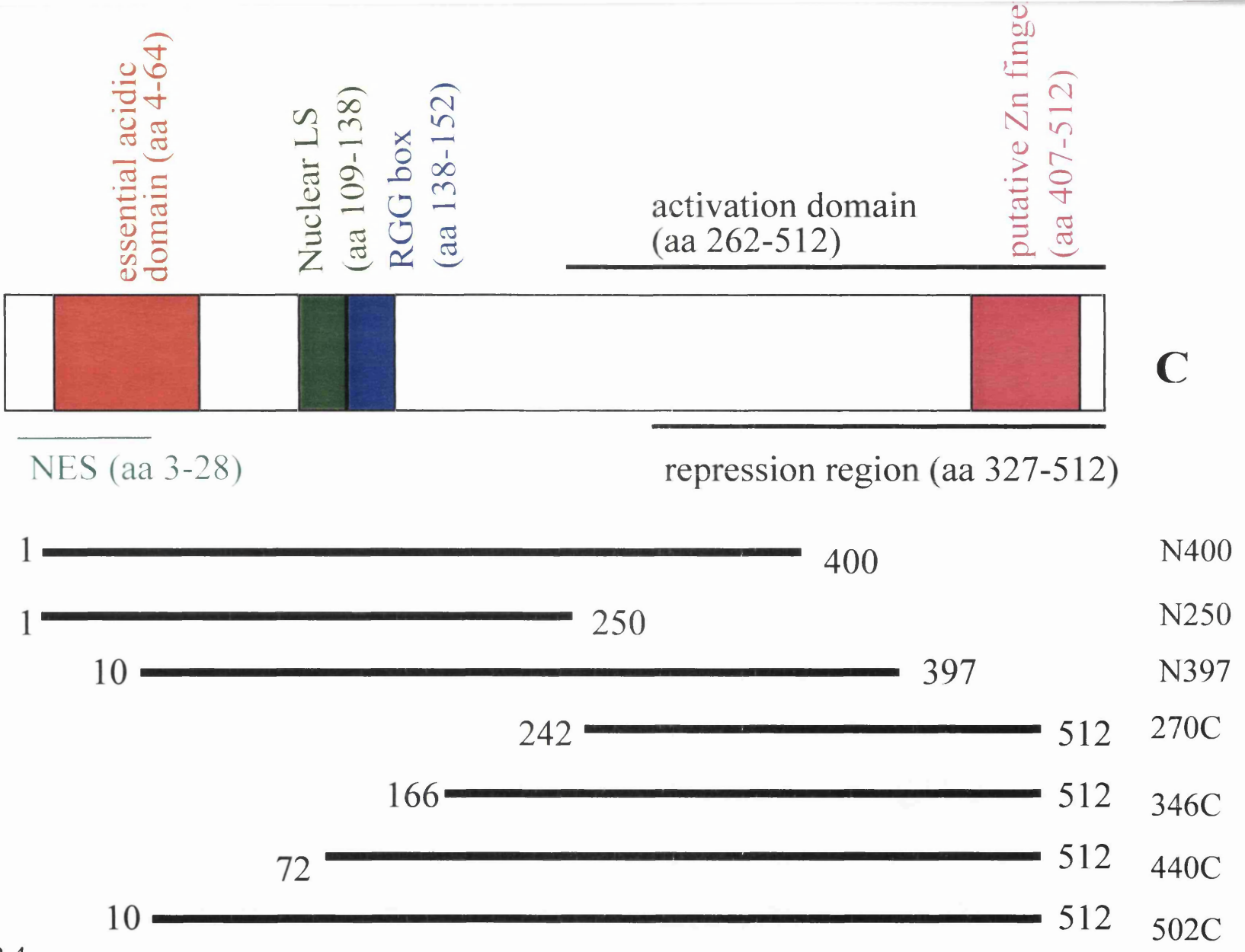


Figure 2.4
Schematic representation of IE63 truncation mutants for details see Section 2.3.3.

(a) Primers

Oligonucleotide primers were designed so that they would amplify specific regions of IE63 and restriction enzyme sites were incorporated into the primers to allow them to be cloned into the two-hybrid system expression vectors. Restriction enzyme sites were chosen which would allow the PCR product to be inserted into the same reading frame as the GAL4 DNA-BD or AD thereby creating the appropriate fusion protein once translated. The primers used are described in Table 2.4 and were obtained from Cruachem.

(b) Reaction conditions

The IE63 mutants 270C and 346C were amplified by PCR using Taq DNA polymerase and the IE63 expression plasmid pSG130 was used as a template. The reaction mixture consisted of PCR buffer (no $MgCl_2$), 100 μM of each dNTP (dATP, dCTP, dGTP, dTTP), 100 ng DNA template, 40 pmol of each oligonucleotide primer, 2.5 mM $MgCl_2$ and 2 units of Taq DNA polymerase in a total volume of 100 μL sterile H_2O.

The PCR programme used was as follows:-

Table 2.4. PRIMERS USED TO SYNTHESISE IE63 TRUNCATIONS. BOLD NUCLEOTIDES INDICATE RESTRICTION SITES

Mutant	5' primer	3' primer	Restriction Sites
270C	CG**GAATTC**GACACCATCGACGCCACCAC	CGC**CTGCAG**GAGGGGCGATTTGTTTGAAAT	*Eco* RI and *Pst* I
346C	CGCCA**GAATTC**CCCAGAACCAATCG	CCC**GGATCC**GAGGGGCGATTTGTTTGAAAT	*Eco* RI and *Bam* HI
440C	CGCCA**GAATTC**TCTCGTCCAGAAGA	CCC**GGATCC**GAGGGGCGATTTGTTTGAAAT	*Eco* RI and *Bam* HI
502C	CGCGA**GAATTC**GGCCTGGACCTCTC	CCC**GGATCC**GAGGGGCGATTTGTTTGA	*Eco* RI and *Bam* HI
N250	CGCGA**GAATTC**CCACCAGAGGCCAT	CCG**GGATCC**GTGGTGGCGTCGATGGTGTC	*Eco* RI and *Bam* HI
N397	CGCGA**GAATTC**GGCCTGGACCTCTC	CCG**GGATCC**CAGGCCTCGCCCTTCAGGTA	*Eco* RI and *Bam* HI
N400	CGCGA**GAATTC**CCACCAGAGGCCAT	CCG**GGATCC**CGTGTCTAGGATTTCGATCAG	*Eco* RI and *Bam* HI

	Temp (°C)	Time
Stage 1 (1 cycle)		
Step 1	95	5 mins
Step 2	55	1 min 30 secs
Step 3	72	2 mins
Stage 2 (30 cycles)		
Step 1	95	1 min 30 secs
Step 2	55	1 min 30 secs
Step 3	72	2 mins
Stage 3 (1 cycle)		
Step 1	72	10 mins

The IE63 mutants 440C and 502C were amplified by PCR using Vent DNA polymerase and the IE63 expression plasmid pSG130 was used as a template. The reaction mixture consisted of PCR buffer, 200 μM of each dNTP (dATP, dCTP, dGTP, dTTP), 50 ng DNA template, 80 pmol of each oligonucleotide primer, 5 % DMSO and 2 units of Vent DNA polymerase in a total volume of 100 μL sterile H_2O.

The PCR programme used was as follows:-

	Temp (°C)	Time
Stage 1 (1 cycle)		
Step 1	94	1 min
Stage 2 (15 cycles)		
Step 1	96	35 sec
Step 2	67	30 sec
Step 3	68	4 min
Stage 3 (10 cycles)		
Step 1	96	35 sec
Step 2	67	30 sec
Step 3	68	4 min
Stage 4 (15 cycles)		
Step 1	96	35 sec
Step 2	67	30 sec
Step 3	68	4 min
Stage 5 (1 cycle)		
Step 1	72	30 min

All PCR reactions were covered with a layer of mineral oil to prevent evaporation and amplification was carried out on an automated PCR machine (Hybaid, Omnigene).

2.3.4 SDS-PAGE and Western Blot Analysis of Proteins

Proteins were separated by electrophoresis on 10% SDS-polyacrylamide mini gels. Transfer of the proteins to a nitrocellulose membrane was carried out for 2 hs at 200 mA in transfer buffer, using a Bio-Rad blotting apparatus according to the manufacturer's protocol. To prevent non-specific binding of the antibody, the membrane was blocked by incubation in PBS containing 5% dried milk and 0.1% tween-20 at 4°C overnight. The membrane was then washed 6x in PBS with 0.1% tween-20, prior to incubation for 2-4 hs with the appropriate dilutions of specific primary antibodies in fresh blocking buffer with gentle agitation. Six 5 min 0.1% PBS-tween washes removed excess primary antibody and the membrane was incubated for 1 h at room temperature with a 1:1000 dilution of the appropriate HRP conjugated secondary antibody in PBS-tween containing 2% dried milk with gentle agitation. A further six 5 min TBS-tween washes removed excess secondary antibody and detection of antibody labelled proteins was performed using the ECL detection kit of Amersham in accordance with the recommended protocol.

RESULTS AND DISCUSSION

3.1 PRE-LIBRARY SCREEN ASSAYS ON YEAST STRAINS, EXPRESSION PLASMIDS AND cDNA LIBRARIES

3.1.1 The Yeast Strains CG1945 and Y187 have a phenotype suitable for use in the yeast two-hybrid system

A critical step in two-hybrid system library screens requires the transformation of the yeast strain CG1945 with plasmids expressing the GAL4 AD fused to a cDNA library together with plasmids expressing the GAL4 BD fused to the gene encoding the bait protein. Yeast colonies that have been successfully transformed with both plasmids and which contain interacting proteins are selected for on medium lacking histidine, leucine and tryptophan. Clearly it is essential that the phenotype of the yeast be -Leu - Trp, -His. Therefore untransformed yeast cells were plated onto combinations of selective media lacking these amino acids to ensure that they were unable to replicate.

Table 3.1 shows the results of these experiments. These clearly show that CG1945 had the phenotype -Leu, -Trp - His demonstrating that CG1945 was suitable for screening a library with the bait protein.

In addition, testing for interaction between two individual proteins necessitates that the yeast strain Y187 be transformed with plasmids expressing the GAL4 AD fused to the target gene together with a plasmid expressing the GAL4 BD fused to the gene encoding the bait protein. Yeast colonies that have been transformed with both plasmids are selected for on medium lacking leucine and tryptophan. The phenotype of this yeast strain must, therefore, be -Leu, - Trp, -His. The yeast were plated onto combinations of selective media lacking these amino acids to verify that they were of the correct phenotype.

Table 3.1 also shows the results of these experiments. These clearly demonstrate that Y187 had the phenotype -Leu, -Trp , -His and was suitable for these experiments.

3.1.2 Full-length IE63 causes transcriptional activation of the two-hybrid system's reporter genes and cannot be used to screen for interacting proteins

A variety of proteins which have a role *in vivo* in the regulation of gene expression have been shown to activate the expression of the reporter genes LacZ and HIS3 which are used as markers of interacting proteins in the two-hybrid system (Bartel *et al.*, 1993). These proteins cannot be used in the two-hybrid system because detection of interacting proteins in the system is based entirely on assays of reporter gene activity. Clearly, a protein which autonomously activates reporter genes would produce multiple false positives and it would be impossible to differentiate between false positives and true interacting proteins.

Since IE63 is known to up-regulate HSV-1 gene expression and function at the level of transcription, a plasmid encoding full-length IE63 fused to the GAL4 BD was transformed into the yeast cells in the absence of a target plasmid alongside a series of positive and negative controls to test if the IE63 protein could autonomously activate reporter gene expression. The yeast cells were plated onto medium lacking tryptophan to select for yeast clones which contained the IE63 plasmid and these cells were allowed to grow for 3 days before testing for activity of the LacZ reporter gene using a filter-lift assay.

β-galactosidase activity was detected. This experiment demonstrated that full-length IE63 can induce transcription of the two-hybrid system's reporter genes. This result was consistently reproducible and demonstrated that the two-hybrid system could not be used to detect proteins which interact with the full-length IE63 protein.

Table 3.1. Results of experiments to determine the phenotype of yeast strains CG1945 and Y187; '+' indicates growth of yeast, '-' indicates no growth

Strain	-TRP	-LEU	-HIS	YPD
CG1945	-	-	-	+
Y187	-	-	-	+

3.1.3 The N-terminal ten amino acids of the IE63 protein are responsible for activation of the two-hybrid system's reporter genes

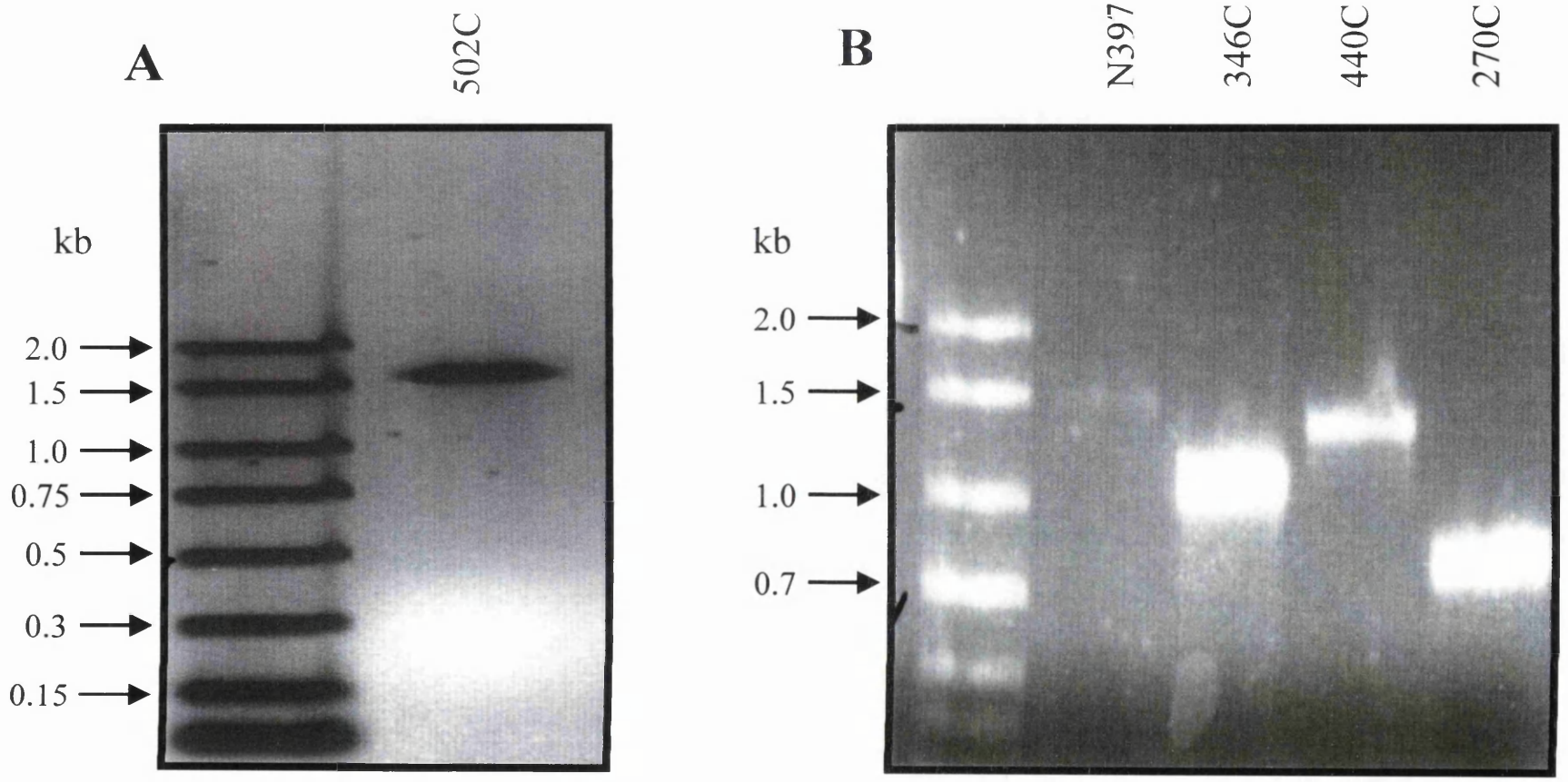
The two-hybrid system is a powerful and attractive method for identifying interacting proteins. Therefore it seemed appropriate to identify the region of the IE63 protein which was transactivating the LacZ reporter gene and delete it, thereby enabling the largest possible truncated form of IE63 to be used to screen for interacting proteins.

Seven truncated IE63 proteins were synthesised using PCR, each one deleting or disrupting one or more of the identified functional domains of IE63 (see Figure 2.4). The PCR products of the clones are shown in Figure 3.1. Next, the truncations were cloned into a plasmid expressing the GAL4 BD and were sequenced to ensure that they were in frame.

Each truncation was subsequently transformed into yeast cells and plated onto medium lacking tryptophan to select for yeast clones which contained the IE63-truncation plasmids. The plates were incubated for 3 days before testing for activity of the LacZ reporter gene using a filter-lift assay.

The results of these experiments are illustrated in Figure 3.2. The interpretation of these results is that the N-terminal ten amino acids of IE63, which includes the first seven amino acids of IE63's essential acidic domain, was responsible for the reporter gene activation and that deleting this region in construct 502C entirely abrogated β-galactosidase activity. The IE63 protein with the N-terminal 10 amino acids deleted was, therefore, suitable for screening. This truncation was designated 502CBD since it encoded the C terminal 502 amino acids of IE63 fused to the BD of GAL4.

Figure 3.1
PCR of IE63 Truncation Mutants **(A)** 502C and **(B)** N397, 346C, 440C and 270C.



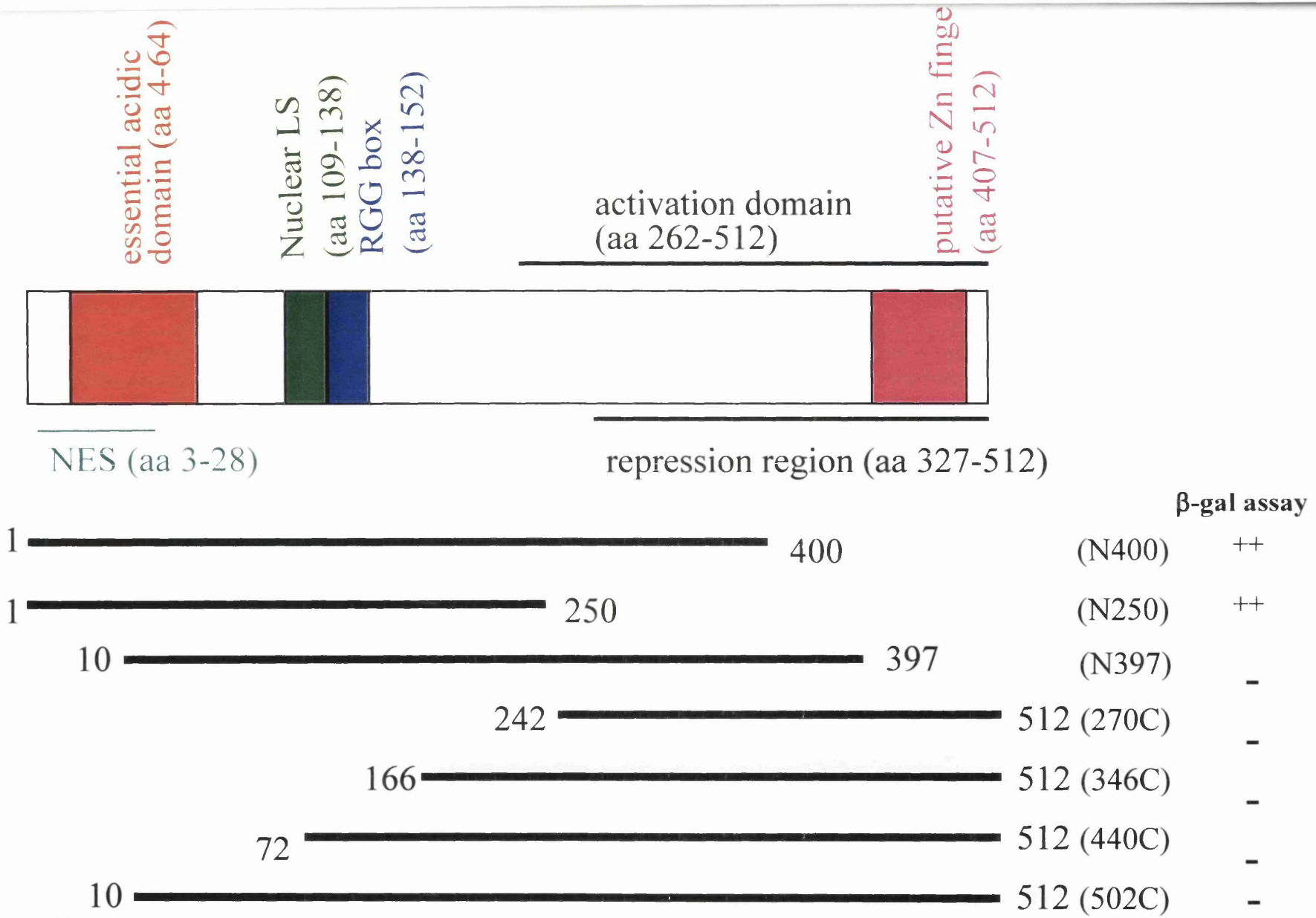


Figure 3.2
Results of experiment to determine which region of IE63 causes transactivation of the two-hybrid system reporter gene β-gal relative to the positive control pVA3-1 and pTD-1 (++) and the negative control pAS2-1 and pLam5'1 (-).

3.1.4 The Matchmaker Two-Hybrid System I expression plasmids did not express protein at a level high enough to be detected by Western Blotting

The expression plasmids pGBT9 and pGAD424 represent the BD and AD vectors respectively of the Matchmaker Two-Hybrid System I. It was necessary to check via western blotting that the bait protein was expressed *in vivo* and that it was not degraded by proteolytic enzymes within the yeast cells.

IE63 truncations 346C, 440C and 502C ligated into pGBT9 were transformed into HF7C cells and the samples were analysed in a western blot alongside lysed untransformed HF7C cells as a negative control and infected cell extract as a positive control (data not shown). This experiment clearly showed that little or no protein was expressed or detected by these expression plasmids. Further western blots also failed to detect IE63 protein.

Subsequently the Matchmaker Two-Hybrid System II was developed by Clontech and it became evident that pGBT9 and pGAD424 express protein at a very low level and that the small quantity of protein *cannot* be detected by western blot despite previous information to the contrary. In view of this and the fact that plasmids supplied with the Matchmaker Two-Hybrid II kit were reportedly high-expression vectors, a decision was made to excise the IE63 truncations from pGBT9 and ligate them into pAS2-1 with the Matchmaker Two-Hybrid System II kit being used for all subsequent experiments.

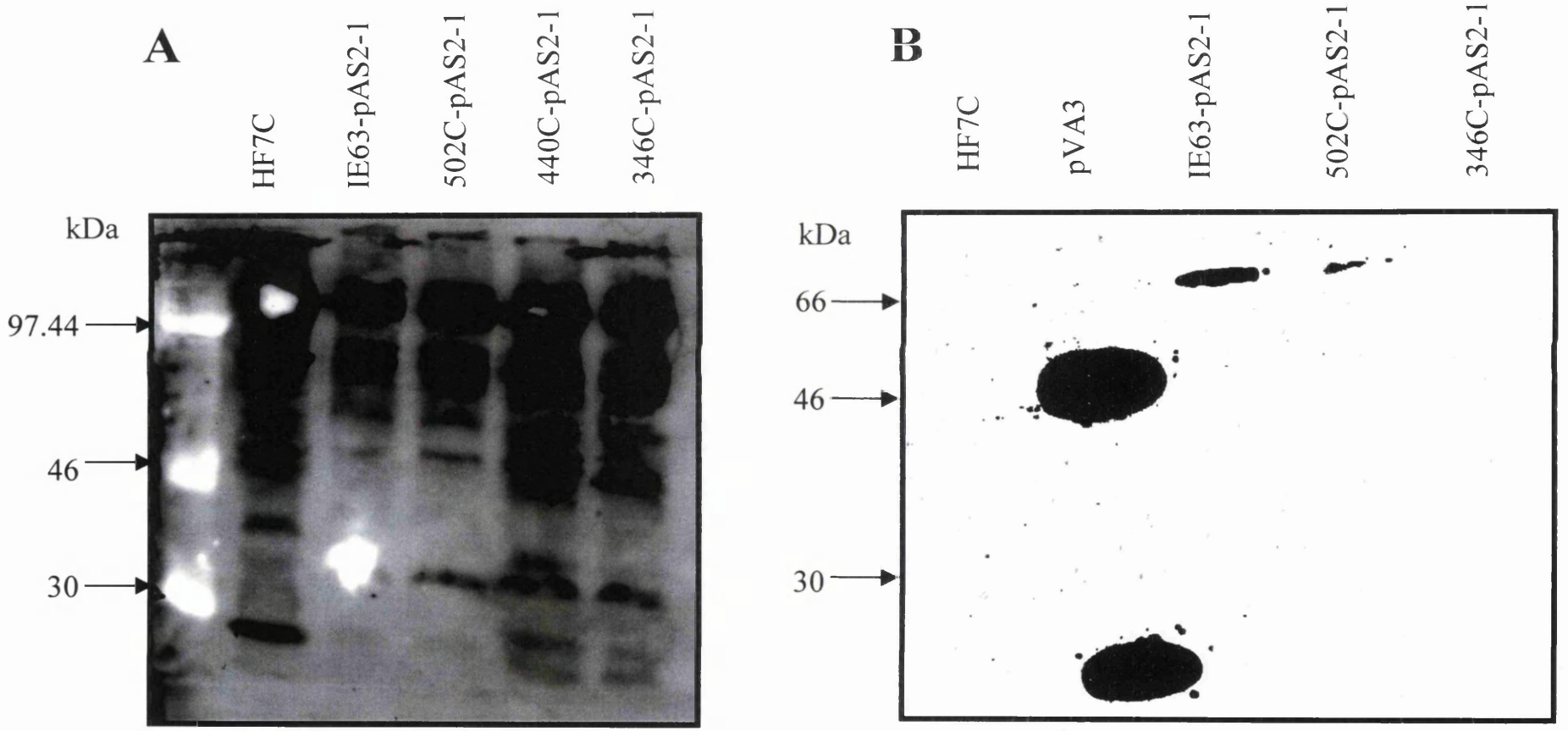
3.1.5 An IE63 truncation cloned into the Matchmaker Two-Hybrid System II expression plasmids was expressed in yeast cells

Figure 3.3 (A) shows a western blot to examine expression of IE63 subclones using anti-IE63 H113 antibody. There is a significant amount of cross-reactivity using this antibody, therefore figure 3.3 (B) shows a western blot using CG1945 yeast cells transformed with either 502C-pAS2-1 or 346C-pAS2-1 and blotted with GAL4 DNA-BD mAb. pAS2-1 is the high expression BD plasmid which is supplied with the Clontech Matchmaker

Figure 3.3

(A) Western blot using mouse monoclonal anti-IE63 H113 to examine expression of IE63 subclones 502C-pAS2-1, 440C-pAS2-1 and 346C-pAS2-1 along with positive control; IE63 in pAS2-1 and negative control; untransformed yeast cells (HF7C).

(B) Western blot using GAL4 DNA-BD mAb (Clontech) to examine expression of IE63 subclones 502C-pAS2-1 and 346-pAS2-1 along with positive controls; IE63 in pAS2-1 and the two hybrid system positive control pVA3, and negative control; untransformed yeast cells (HF7C).



Two-Hybrid System II. The lane loaded with the 502C-pAS2-1 sample contains a protein band of approximately 71 kDa which is the anticipated size of the 502C-pAS2-1 construct, indicating that the fusion protein is being expressed in yeast cells and that it is not degraded. However, there was no band in the 346C lane of approximately 55 kDa, nor could the truncations 440C and 270C be visualised using the same procedure (data not shown). Considering that the sequence of these constructs had been confirmed and the 502C construct was expressed from pAS2-1 a decision was made to screen a library with 502CBD. It seemed likely that 440C, 346C and 270C were genuinely being expressed from pAS2-1 but that the concentration of protein was too low to be visualised on the Western blot.

3.1.6 The BHK cDNA Library did not represent a full complement of BHK cell proteins and could not be used to screen for IE63 interactions

A HSV-1 infected BHK cDNA library was the original material for a library screen to determine IE63 protein-protein interactions. On construction, the library had been titred by a co-worker and contained an estimated 90% of colonies with a titre of 2×10^6. However, two library screens were performed using this library and 502CBD as a bait but in both instances, no positive clones were identified. Further experiments by a co-worker revealed that on storage the titre had declined to 2×10^4 prior to use in the library screen. For this reason an uninfected HeLa cDNA library which had been synthesised by Clontech and which had an estimated 93% of colonies with inserts and a total number of independent clones of 6.0×10^6 was used for further library screens.

3.2 SCREENING A cDNA LIBRARY WITH TRUNCATED IE63 PROTEIN IDENTIFIED FIVE PROTEINS WHICH INTERACT WITH IE63 AND WHICH ARE LIKELY TO BE BIOLOGICALLY SIGNIFICANT

To identify the proteins which interact with IE63, the yeast strain CG1945 was simultaneously transformed with the IE63 bait plasmid 502CBD and the target plasmids encoding the HeLa cDNA library fused to the AD. A total of 2.3×10^6 independent clones were screened. Colonies which had a $Trp^+Leu^+His^+$

phenotype were selected and assayed for β-galactosidase activity. β-galactosidase activity combined with the ability to grow on a medium lacking histidine indicated interacting proteins. Colonies with a $His^+ LacZ^+$ phenotype were further screened to identify false positives as described below.

3.2.1 150 yeast colonies grew on the medium lacking histidine indicating that they were expressing interacting proteins

9 days post-transformation there were 150 colonies on the -Trp,-Leu, -His plates. As the yeast strain used in these experiments has a -His phenotype only histidine-expressing yeast cells could be cultured on the histidine deficient medium. Histidine is expressed if the bait protein IE63 interacts with a protein from the HeLa cDNA library. Clearly, any colonies growing on this medium were candidates for expressing proteins that interact with IE63. These colonies were picked and re-plated onto -Trp,-Leu, -His medium. Following 5 days of incubation, 6 (4%) of the 150 colonies no longer grew on this medium indicating that they were false positives.

A filter lift assay was then carried out to assay for the reporter gene LacZ. 123 (85%) of the remaining 144 colonies had a $LacZ^+$ phenotype. These colonies were picked and further screened for false positives as described in Section 2.3.1 (g).

3.2.2 99 yeast colonies passed a screen for false positives

False positives are an inherent problem in the yeast two-hybrid system and it is essential to carry out tests which will eradicate the majority of these. They may be due to library proteins which can autonomously activate the reporter genes or they may be due to two different library proteins being expressed in one yeast colony, one of them being a true interacting protein and another which does not interact with the bait.

The yeast colonies with a $His^+ LacZ^+$ phenotype were replated onto -Leu-Trp medium and allowed to grow for three days. These colonies were transferred once more onto fresh -Leu-Trp plates and allowed to grow for a further three days. This replating onto -Leu-Trp medium was performed to aid the segregation of any yeast colony which may contain two plasmids encoding different library proteins. The yeast colonies were then replated onto -Leu-Trp, -His medium and incubated for three days to check that the $His^+ LacZ^+$ phenotype was retained. Colonies that grew on this medium were tested for β-galctosidase activity. 109 (89%) of the remaining 123 colonies retained their $His^+ LacZ^+$ phenotype.

Next, the IE63 bait plasmid 502CBD was deleted from the yeast cells by cycloheximide selection and the yeast were plated onto both -Leu medium and -Trp medium to ensure that the 502CBD plasmids had been deleted and that the library plasmids had been retained in the yeast cells. As expected, no colonies grew on the -Trp medium showing that they had not retained the 502CBD plasmid. 4 (4%) of the remaining 109 colonies would not grow on the -Leu medium indicating that these colonies no longer contained a library plasmid. These colonies were discarded.

A definitive test for β-galactosidase activity was carried out on the remaining 105 colonies to ensure that the library plasmids alone could not activate the LacZ and HIS3 reporter genes. None of these colonies exhibited β-galactosidase activity demonstrating that the target plasmids could not autonomously activate the two-hybrid system's reporter genes.

The library plasmids were successfully isolated from all but six of the yeast colonies. Therefore 99 (66%) plasmid DNAs isolated from the original 150 colonies were sequenced.

Interestingly, an entire library screen which was carried out with a 440C truncation which later proved to be out of frame identified only 15 positive colonies and subsequent testing revealed all of these were false positives. This

clone was discarded and a 440C clone which was in-frame was used for all future experiments.

3.2.3 Sequence analysis identified five positive clones with a putative biological sigificance

The 99 remaining plasmids were sequenced using a 19 oligonucleotide sequencing primer. 16 either could not be sequenced or gave inadequate sequence for further analysis. This was probably due to a contaminating factor in the DNA. The sequences from the remaining 83 plasmids were entered into the EMBL/Genbank database using GCG Fast A.

The results of the EMBL/Genbank database screen showed that six of the 83 plasmids had no homology to any proteins in the database. A further plasmid was an empty expression vector.

One clone had a 99.5% homology with a human quinone oxidoreductase over 388 bp. Another had 87.8% homology with L Apoferritin over 303 bp and a further two had homology to the BiP protein (a member of the heat shock protein family) with the highest homology being 97.5% in 406 bp. A number of human proteins are known to form non-specific protein-protein interactions in the two-hybrid system although the mechanism for this is not yet fully understood. Mitochondrial proteins such as human quinone oxidoreductase, ferritins and heat shock proteins are proteins which frequently give this type of false positive (http://www.fccc.edu/research/labs/golemis/ InteractionTrapInWork.html, 1999). For this reason these clones were discarded.

A further two clones were identified as Gallus Gallus Glutamine Rich Protein (80.6% identify in 453 bp) and another as human JTV-1 (85.8% identity in 387 bp). These clones did not have an immediately obvious biological significance according to current understanding of the IE63 protein (see Section 4.1.1 and 4.1.2. For this reason these clones were set aside for follow-up studies.

Sequence analysis identified five other proteins which interact with IE63 and the previously determined function of these proteins suggested that their interaction with IE63 may have a critical biological function in HSV-1 infected cells. Comparisons of the sequences of these clones to the genes with which they have homology can be found in Appendix 1. These clones and the possible significance of these interactions are discussed below. The results of the library screen are summarised in Table 3.2.

3.2.4 The IE63 protein interacts with the human hnRNP K and CKII β Subunit

Analysis of the sequences of the positive clones using the GenEMBL database revealed that one of the clones, containing an 1.3kb insert, had a sequence which precisely corresponded to aa 215-464 and the 3' UTR of human hnRNP K and was present in the vector in the correct reading frame. A further fifty-one of the clones could be identified as the casein kinase II beta subunit (CKIIβ). Full sequence analysis of 5 clones showed 4 independent clones, each of which contained the entire CKIIβ gene but beginning at different positions in the 5' UTR.

(a) The C terminal half of the IE63 protein is sufficient for the interaction with hnRNP K

In order to map the region of IE63 which mediated the interaction with hnRNP K a series of IE63 truncations was expressed as hybrids with the GAL4 BD. Yeast cells transformed with plasmids encoding the truncations were mated with yeast cells transformed with the clone identified as hnRNP K from the library screen. Figure 3.4 illustrates the region of IE63 encoded by each construct and the degree of activation of the two reporter genes HIS3 and LacZ in the presence of the hnRNP K protein relative to the positive control pVA3-1 and pTD-1 and the negative control pAS2-1 and pLam5'1.

Table 3.2 Summary of results of two-hybrid uninfected HeLa cell cDNA library screen with IE63 truncation 502CBD.

	No. of clones identified	Clone with highest homology	Function
Proteins found to interact with IE63			
hnRNP K	1	96.3%	Pre-mRNA processing & transport (?), transcription, translation, signal transduction, linker for molecular interactions (?)
Casein Kinase II	51	99.2%	Ubiquitous serine/threonine protein kinase
SAP145	1	89.4%	Binding of U2 snRNP to pre-mRNA
ALY	4	91.9%	Transcription, splicing (?)
p32	12	99.3%	Splicing, transcription, mRNA transport (?)
Proteins found to interact with IE63 but no obvious biological significance			
Gallus gallus glutamine rich protein	2	80.6%	Putative transcription factor in chickens
Human JTV-1	1	85.8%	Function unknown
False positives			
Human quinone oxidoreductase	1	99.5%	Oxidative phosphorylation
L Apoferritin	1	87.8%	Regulation of iron binding
BiP	2	97.5%	Heat shock protein

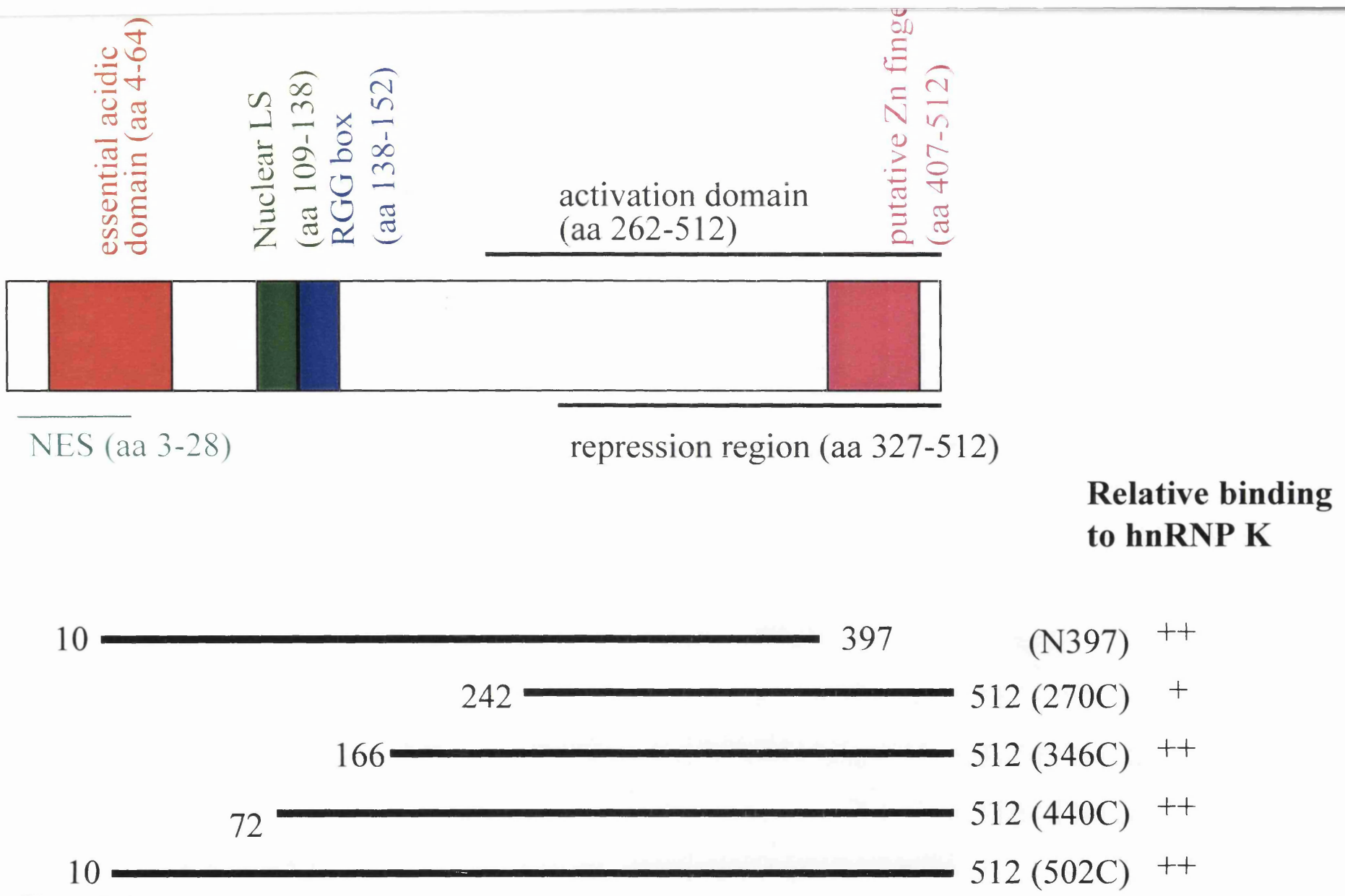


Figure 3.4

Schematic diagram (not to scale) of the IE63 truncations which interacted with hnRNP K in a yeast two-hybrid assay and their binding to hnRNP K relative to the positive control pVA3-1 and pTD-1 (++) and the negative control pAS2-1 and pLam5'1 (-).

The results show that sequences within aa 242-397 of IE63 were both necessary for the interaction with hnRNP K. However, the binding of aa 242-512 was weaker than that of aa 166-512 suggesting that aa 166-242 also contribute to the interaction. Although this latter region does not include the RGG domain necessary for RNA binding it does include adjacent proline-rich sequences. Interestingly, Rice and Lam have shown that point mutations within the region absolutely required for interaction with hnRNPK abolish the ability of wt virus to grow suggesting that the interaction may be essential for HSV-1 function (Rice & Lam, 1994).

(b) The putative zinc finger of IE63 together with amino acids in the N-terminus of the IE63 protein are essential for the interaction with Casein Kinase II

To map the region of IE63 required for interaction with CKIIβ, the IE63 truncations were used in mating assays with the clone identified as CKIIβ from the library screen. Figure 3.5 shows the region of IE63 present in each construct and the degree of activation of the two reporter genes HIS3 and LacZ in the presence of the CKIIβ protein relative to the positive control pVA3-1 and pTD-1 and the negative controls pAS2-1 and pLam5'1.

By comparison of truncations 270C and 346C, the results demonstrate that a portion of the argine-rich region (aa 166-242) and aa 397-512 which includes the zinc finger of IE63 and the region required for inhibition of splicing were sufficient for the interaction. However, the binding of aa 166-512 was weaker than that of aa 10-512 and this indicates that aa 10-166 also contribute to the interaction. Interestingly, the IE63 truncation 440CBD (aa72-512) did not interact with CKIIβ in contrast to the smaller IE63 truncation 346BD (aa 166-512) which interacted with CKIIβ weakly. The reason for disruption of the binding of the larger protein could be due to conformational changes in the protein masking an IE63 binding site which is exposed in the smaller truncation.

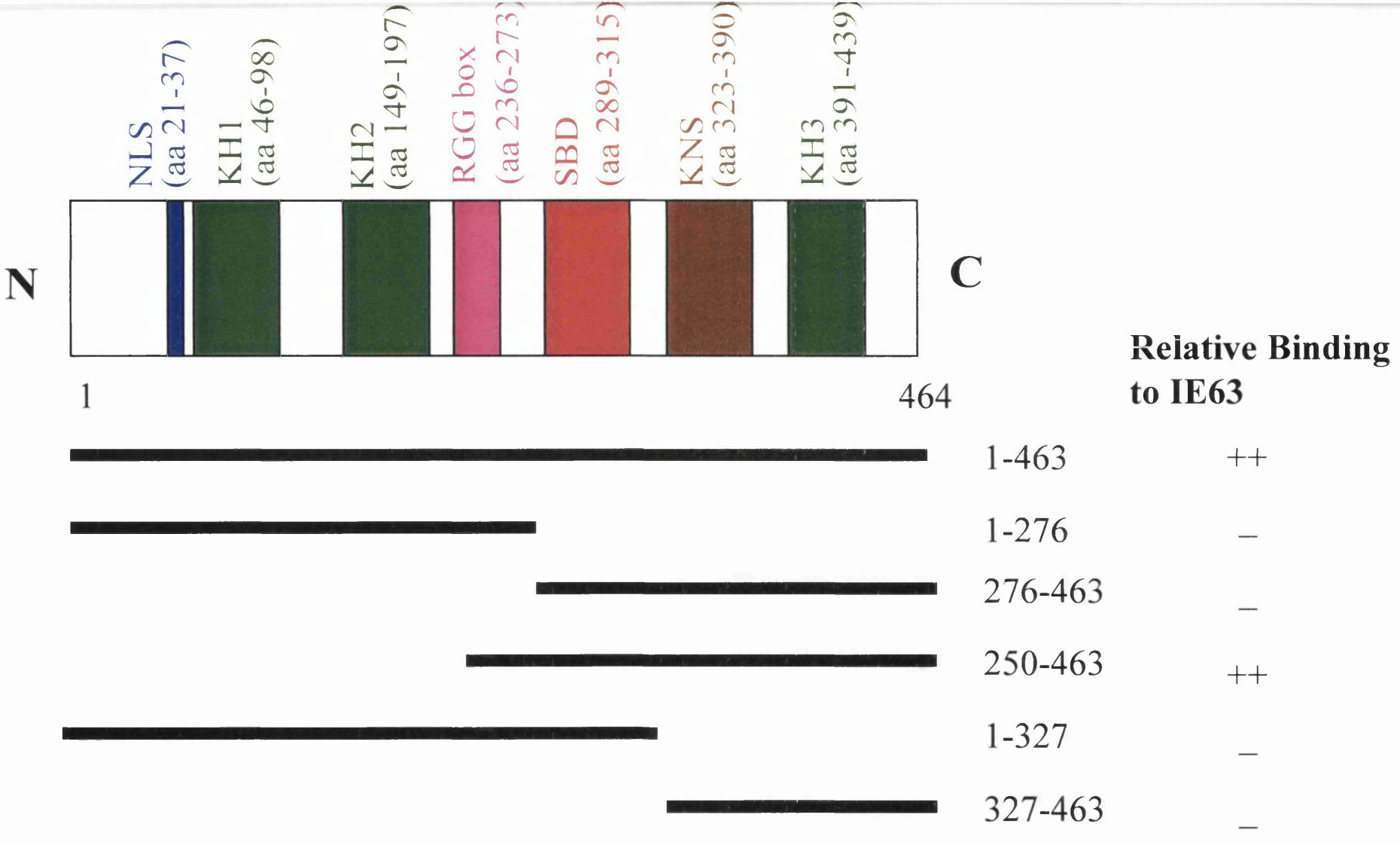


Figure 3.6
Schematic diagram (not to scale) of the hnRNP K truncations which interacted with IE63 in a yeast two-hybrid assay and their binding to IE63 relative to the positive control pVA3-1 and pTD-1 (++) and the negative control pAS2-1 and pLam 5'1 (-). SBD is the SH3-binding domain, the KNS is the bidirectional nuclear shuttling domain and the KH domains are involved in RNA binding (see Section 1.3.1 (c)).

(c) Mapping the regions of hnRNP K required for interaction with IE63

In order to map the region of hnRNP K required for interaction with IE63 a series of hnRNP K truncations was expressed as hybrids with the GAL4 AD. This exact series of hnRNP K truncations was also used by Tsieh *et al.* (Tsieh *et al.*, 1998) to map the regions of hnRNP K which interact with the hepatitis C virus core protein. Yeast cells transformed with these truncations were mated with cells transformed with the IE63 bait 502CBD fused to the GAL4 DNA-binding domain. Figure 3.6 illustrates the region of hnRNP K present in each construct and its relative strength of activation of the reporter gene LacZ in the presence of the IE63 bait protein relative to the positive control pVA3-1 and pTD-1 and the negative controls pAS2-1 and pLAM5'1. Interpretation of these results is that regions within aa 250-276 and aa 327-463 are required for the interaction. Interestingly, the region of hnRNP K required for interaction with IE63 contained the RGG box and the C-terminal portion known to bind the hepatitis C virus core protein whose expression has been shown to relieve hnRNP K supression of the cellular thymidine kinase promoter (Tsieh *et al.*, 1998) and also is required for interacting with the transcriptional repressor Zik 1 and the hnRNP K protein kinase. Although the *Xenopus* hnRNP K RGG box is not required for poly(rC) binding (Ladomery, 1997) a contribution to RNA binding is possible. Therefore, it is possible that G+C rich viral RNA (HSV-1 DNA is 68% G+C overall) could contribute to the interaction of IE63 and hnRNP K and this could occur in yeast two-hybrid assays and co-immunoprecipitations where RNA:protein binding may protect against RNase treatment.

(d) Further evidence that IE63 interacts with hnRNP K and Casein Kinase II β subunit

The interaction of IE63 with both hnRNP K and CKII has since been independently confirmed by Helen Bryant using GST-pull down and immunoprecipitation experiments (Wadd *et al.*, 1999). Helen has also used

GST and MBP fusions with different truncations of CKIIβ to map the region of CKII which interacts with IE63. These experiments demonstrated that aa 150-215 of CKIIβ are required for the interaction with IE63 and this is the region of the protein involved in α:β subunit heterodimerisation.

(e) Casein Kinase can phosphorylate IE63 and the form of hnRNP K which co-immunoprecipates with IE63

Interestingly, Helen Bryant has also shown that the predominant form of hnRNP K which co-immunoprecipitates with IE63 using mouse monoclonal anti-IE63 serum migrates more rapidly in gels than hnRNP K recognised by rabbit anti-K serum raised against a synthetic peptide representing the C-terminal aa 452-464, conserved in the murine and human hnRNP K (Wadd *et al.*, 1999). Similarly, GST-IE63 pulled down forms of hnRNP K present in cells infected with HSV-1 wt which migrated faster in gels than the predominant band of hnRNP K found in mock-infected extracts. This indicates that the form of hnRNP K which interacts with IE63 is a less processed or possibly a smaller form. Phosphorylation of the slower-migrating form of hnRNP K precipitated by hnRNP K antiserum was not inhibited *in vitro* by the CKII specific inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB). In contrast, phosphorylation of the hnRNP K fraction immunoprecipitated by anti-IE63 serum from wt-infected cells was capable of being partially inhibited by DRB. These *in vitro* phosphorylation data further demonstrate a difference between the fraction of hnRNP K immunoprecipitated with anti-hnRNP K serum compared with anti-IE63 serum.

Finally, the phosphorylation of IE63 was decreased significantly in the presence of DRB indicating that IE63 can itself be phosphorylated by CKII *in vitro*.

3.2.5 The proposed role of IE63-hnRNPK and Casein Kinase II interactions in the HSV-1 infected cell

There are striking similarities between the structures and functions of hnRNP K and IE63. IE63 and hnRNP K are of similar size, IE63 consisting of 512 amino acids and hnRNP K 464 amino acids. Both proteins form oligomers, are phosphorylated, contain RGG motifs which in both instances are methylated and both have a highly acidic N terminus. hnRNP K contains three KH regions that are required for RNA binding activity and three KH domains have recently been identified towards the C terminus of IE63 (T M Soliman and S J Silverstein, Columbia University, New York, unpublished data). Furthermore, the two proteins bind RNA, shuttle between the nucleus and the cytoplasm, have a classical NLS in addition to a shuttling domain(s) and both interact with multiple molecular partners. Finally, IE63 and hnRNP K have a role in the regulation of transcription and both proteins are thought to be required for pre-mRNA processing and transport.

However, IE63 and hnRNP K also have dissimilar characteristics. For example, hnRNP K binds to DNA whilst there is no evidence of IE63 binding DNA. Additionally, hnRNP K is thought to exert some of its effects at the level of translation and IE63, apparently unlike hnRNP K, has an effect on polyadenylation and splicing.

The similarities between the structures and functions of IE63 and hnRNP K suggest that they may utilise common pathways in the infected cell. This raises the possibility that IE63 interacts with hnRNP K preventing its access to these pathways so that IE63 can implement its effect on the cell without having to compete with hnRNP K. Alternatively, it is plausible that many of the functions previously attributed to IE63 may in fact be mediated by its interaction with hnRNP K. The merits of both the proposed scenarios are considered in the text below.

(a) IE63 may repress hnRNP K activity in the HSV-1 infected cell

(i) *IE63 may disrupt the transport of cellular mRNA by hnRNP K*

Work carried out in our group demonstrates that hnRNP K can be phosphorylated by CKII in the presence of IE63 but not in its absence and Dejgaard *et al* have shown that phosphorylation of hnRNP K diminishes its ability to bind RNA (Dejgaard *et al.*, 1994). A number of workers have proposed that hnRNP K may transport cellular mRNAs from the nucleus into the cytoplasm for translation to occur. Therefore in principle, the IE63-mediated phosphorylation of hnRNP K by CKII may disrupt hnRNP K-mRNA binding, effectively abrogating the transport of cellular mRNAs into the cytoplasm via this pathway. In support of this is the observation that intron-containing cellular mRNAs accumulate during HSV-1 infection in the nucleus only in the presence of IE63. The transport of cellular mRNAs out of the nucleus and into the cytoplasm is essential for cell protein synthesis and therefore critical for host-cell function. Interestingly, the binding of hnRNP A1 to mRNA is also reduced significantly in the presence of IE63 raising the possibility that IE63 may also disrupt the transport of cellular mRNA by hnRNP A1 (Sandri-Goldin, 1998).

(ii) *IE63 may alter the ability of hnRNP K to act as a transcriptional regulator*

hnRNP K also has a role in the regulation of transcription of a number of cellular genes and interaction with IE63 may either relieve the repressive effect of hnRNP K on certain genes or inhibit the activation of other genes by hnRNP K. Indeed, the expression of some cellular proteins has been shown to be up-regulated during HSV-1 infection whilst others are down-regulated. This latter observation, known as the host-shut-off effect can be divided into a primary and secondary phase, the secondary phase requiring the presence of IE63 (see Section 1.1.6).

Presumably IE63 could alter the transcriptional activity of hnRNP K by disrupting its interactions either with transcription factors or promoter sequences.

(iii) *The sub-cellular localisation of hnRNP K may be altered by IE63*

Finally, phosphorylation of hnRNP K by CKII in the presence of IE63 may alter the sub-cellular localisation of hnRNP K, perhaps sequestering it to inactive sites within the nucleus. Immmunofluoresence studies which look at the localisation of hnRNP K in the presence and absence of IE63 will be carried out in our laboratory to investigate this possibility. Facilitating the removal of hnRNP K into inactive sites would enable IE63 to fully utilise a shuttling pathway which it may share with hnRNP K.

(b) hnRNP K may be critical for IE63 function

(i) *IE63 may 'hitch a ride' from hnRNP K to cross the nuclear pore*

Both IE63 and hnRNP K shuttle between the nucleus and the cytoplasm. It may be that the ability of IE63 to shuttle relies on its ability to bind to hnRNP K and subsequently to ride 'piggy-back' on the hnRNP K protein to cross the nuclear pore. Against this is the finding that the NES of IE63 has been shown to be sufficient for export of a heterologous protein (Sandri-Goldin, 1998) and this study shows that the NES of IE63 does not bind hnRNP K. However, the possibility remains that binding to hnRNP K is required for optimal shuttling of IE63 or shuttling via a specific pathway.

(ii) *hnRNP K may transport IE63 to specific nuclear domains or create a 'bridge' between IE63 and other molecules*

The interaction between hnRNP K and IE63 may be required to target IE63 to specific nuclear domains or to allow IE63 access to one of the hnRNP K's molecular partners such as DNA or another cellular protein. Interestingly,

DSEF-1, a member of the hnRNP H family of RNA-binding proteins, has recently been shown to increase the level of cross-linking of the 64 kDa protein of cleavage stimulation factor (CstF) to polyadenylation substrate RNAs (Bagga *et al.*, 1998). During HSV-1 infection, binding of this 64 kDa component of CstF to poly(A) sites of all temporal classes of viral mRNA is increased and this binding is dependent upon the expression of IE63 (McGregor *et al.*, 1996). Whether hnRNP K has a similar role in polyadenylation to the hnRNP H protein DSEF-1 or whether IE63 may also interact with DSEF-1 is a topic for further investigation.

(iii) *IE63 may recruit hnRNP K for viral gene expression*

Perhaps the most probable biological role for the interaction between hnRNP K and IE63 is that IE63 recruits hnRNP K to facilitate viral gene expression. This study shows that the region of IE63 which interacts with hnRNP K is the region which confers its ability to activate and repress viral genes. It is also interesting that hnRNP K interacts with and represses C/EBP$^{\beta}$, a transcription factor which is active during acute-phase or immune response (Miau *et al.*, 1998). The co-activator of this transcription factor, Nopp140, is phosphorylated by and interacts with CKII. Therefore hnRNP K and CKII colocalize at sites of transcription. That IE63 interacts with both proteins supports the hypothesis that hnRNP K-CKII-IE63 complexes may have a role in the regulation of transcription of HSV-1 genes.

In addition, hnRNP K interacts with TBP (Michelotti *et al.*, 1996) and the hnRNP K:TBP:IE63 complex may be crucial for viral transcription. In support of this theory, tandem copies of a CT-rich DNA sequence similar to known hnRNP K DNA binding sites are present in a HSV-1 domain that has been proposed to act as a transcriptional regulator of virus immediate early genes (Quinn *et al.*, 1998).

(c) The role of CKII in the function of IE63

Phosphorylation regulates the function and localisation of many viral as well as cellular proteins. The text above has hinted at the possible significance of the interaction between CKII and IE63 with relation to the CKII-mediated phosphorylation of the hnRNP K protein. However, a number of viral and cellular proteins such as IE175 and the U1 snRNP 70 kDa protein also have an altered phosphorylation state in the presence of IE63 although IE63 itself does not exhibit kinase activity. Therefore, CKII may be 'hijacked' by IE63 to phosphorylate these and other viral and cellular proteins. Phosphorylation of IE175 may be one of the mechanisms by which IE63 regulates viral transcription. In addition, the phosphorylation state of the U1 snRNP has been proposed to alter the localization of this protein which in turn may be associated with splicing repression in HSV-1 infected cells. Since phosphorylation of U1 70 kDa protein by CKII may inhibit splicing, it is of interest to note that the region of IE63 that binds CKII is also the region of IE63 responsible for splicing repression.

What is the significance of phosphorylation of the IE63 protein itself by CKII? IE63 contains several consensus sites for phosphorylation by CKII and other cellular kinases. Two serine residues located at aa 16 and 18 have been shown to serve as targets for CKII phosphorylation *in vivo* and the serine at aa 114, which lies within the nuclear localisation signal, was phosphorylated by protein kinase A (Zhi & Sandri-Goldin, 1999). Phosphorylation of serine 114 seemed to modulate the efficiency of IE63 nuclear import, however plasmids containing these kinase consensus site mutants were still capable of complementing the growth of an IE63 null mutant virus. Since IE63 is capable of being phosphorylated *in vitro* by co-immunoprecipitated CKII activity (Wadd *et al.*, 1999) CKII could modify IE63 activity or modify partner proteins in the complex. Interestingly CKII activity has been reported in preparations of HSV-1 virions (Steveley *et al.*, 1985) and several other HSV-1 proteins are phosphorylated by CKII (Guerra & Issinger, 1999).

A number of other viral proteins are phosphorylated by CKII . For example, phosphorylation of Simian Virus 40 large T antigen by CKII is required for import into the nucleus (Vancurova *et al.*, 1995) whereas phosphorylation of the Sendai virus P protein by CKII is required to maintain its structural integrity (Byrappa *et al.*, 1996). Furthermore, the HIV-1 protein Rev has been shown to be down-regulated by CKII phosphorylation at late times in infection (Meggio *et al.*, 1996). Finally, work carried out by Helen Bryant suggests that phosphorylation of IE63 may be important for the interaction between IE63 and hnRNP K (Wadd *et al.*, 1999).

It is interesting that in HSV-1 infected cells, the partner kinase of hnRNP K is altered such that a fraction of K is associated with CKII. In uninfected cells hnRNP K forms a complex with an interleukin-1 responsive kinase shown to be the protein kinase C delta isoform (Schullery *et al.*, 1999) and not with CKII although K protein is capable of being phosphorylated by CKII *in vitro* (Seuningen *et al.*, 1995). Thus phosphorylation by CKII could influence the molecular partners interacting with hnRNP K as well as those interacting with IE63. Further studies must be carried out to determine the role of phosphorylation *in vivo* in modulating IE63 activity.

Finally, CKII has been shown to nucleotidylate the HSV-1 protein IE68 (Mitchell *et al.*, 1997) and IE63 contains nucleotidylation consensus sites (Blaho *et al.*, 1993) raising the possibility that this modification is carried out by CKII. Although, Helen Bryant has shown that CKII does not nucleotidylate IE63 *in vitro* (unpublished data) it remains possible that CKII may be involved in the processing of HSV-1 proteins and that this processing is not limited to phosphorylation.

3.2.6 The IE63 protein interacts with the Spliceosomal Asociated Protein 145

Analysing the sequences of the positive clones using the GenEMBL database revealed that one of the clones could be identified as the Spliceosome Associated Protein 145 (SAP145). The clone had a 89.4% homology in 499

bp of this 2,839 bp gene. The sequence of this clone lay in the same reading frame as the SAP145 gene. Helen Bryant has gone on to confirm the interaction between IE63 and SAP145 by both co-immunoprecipitation and GST pull down assays. hnRNP K and CKII were also found in the complex, but were not associated with SAP145 in the absence of IE63. Significantly, the interactions of SAP145 and p32 with IE63 appear to be mutually exclusive of each other. Helen has also shown that SAP145 and IE63 co-localise and that SAP145 co-localises with the splicing factor SC35 which itself co-localises with IE63. These findings confirm that IE63 interacts with SAP145 and the spliceosome.

(a) The proposed zinc finger of IE63 is essential for the interaction with Spliceosomal Associated Protein 145

To map the region of IE63 required for interaction with SAP145, a series of IE63 truncations was expressed as hybrids with the GAL4 DNA-binding domain. These were mated into host cells transformed with the clone identified as SAP145 from the library screen. Figure 3.7 shows the region of IE63 present in each construct and the degree of activation of the two reporter genes HIS3 and LacZ in the presence of the SAP145 protein relative to the positive control pVA3-1 and pTD-1 and the negative controls pAS2-1 and pLam5'1.

The results show that sequences within aa 397-512 of IE63 which encode the zinc finger and the region required for splicing inhibition were essential for the interaction with SAP145. However, the binding of the smaller IE63 truncations to SAP145 was noticeably weaker than the binding of the larger truncation 502CBD. This may be because other sites of IE63 are involved in the interaction or there may be a change in conformation of the smaller truncations which affect the ability of these subclones of IE63 to bind to SAP145.

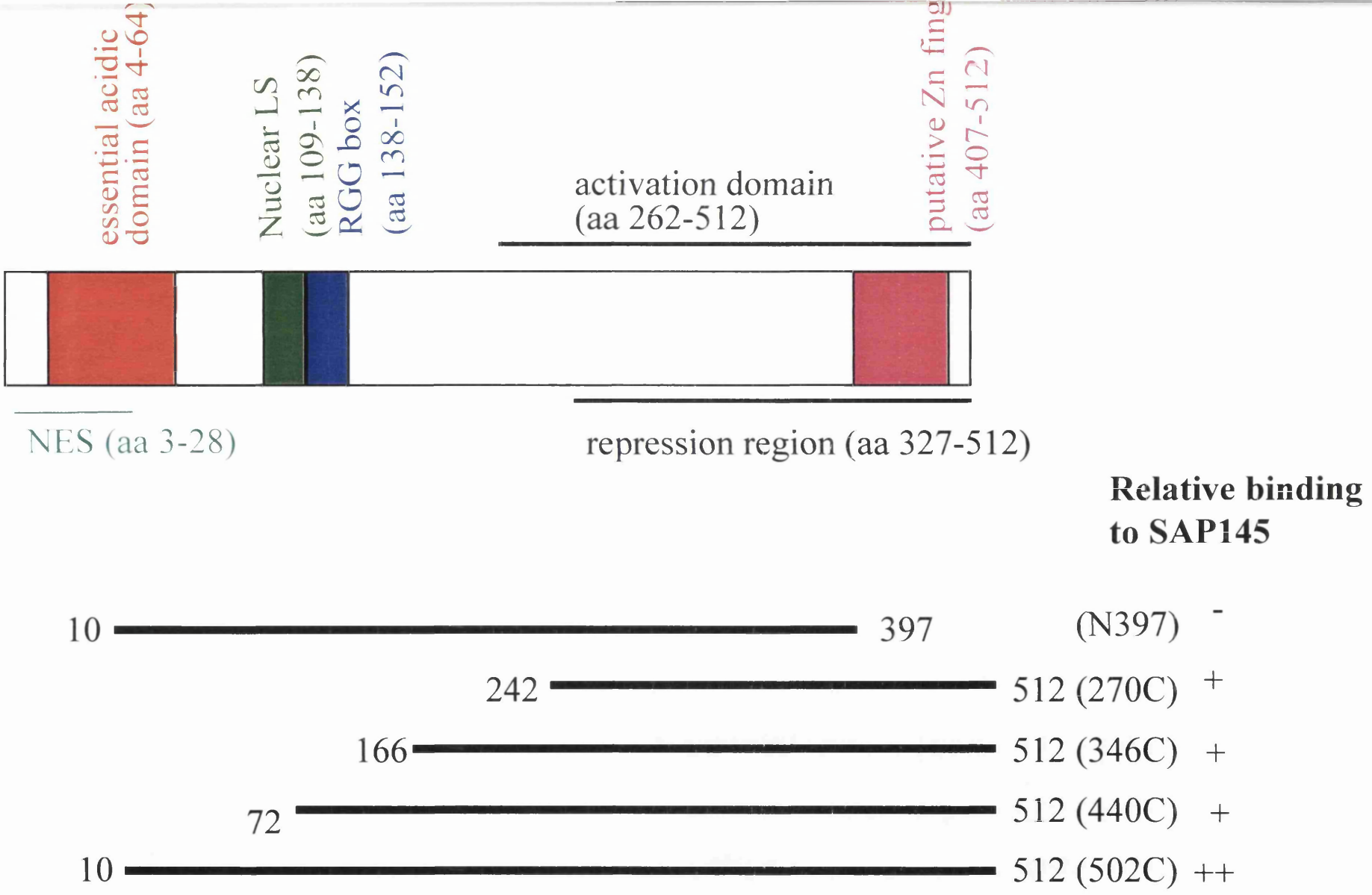


Figure 3.7
Schematic diagram (not to scale) of the IE63 truncations which interacted with SAP145 in a yeast two-hybrid assay and their binding to SAP145 relative to the positive control pVA3-1 and pTD-1 (++) and the negative control pAS2-1 and pLam5'1 (-).

(b) The proposed role of IE63-SAP145 interactions in the HSV-1 infected cell

IE63 promotes the redistribution of splicing factors and inhibits splicing in the HSV-1-infected cell. However, Sandri-Goldin *et al.* have shown that an IE63 temperature sensitive mutant which contains a base substitution in the repressor region (aa 480) causes a redistribution of the snRNPs but splicing was not significantly inhibited (Sandri-Goldin *et al.*, 1995). Thus whilst the redistribution of snRNPs may inhibit splicing, additional as yet unidentified mechanisms contribute to the disruption of splicing.

This study shows that IE63 interacts with SAP145 - a subunit of the splicing factor SF3b and that the region of IE63 which interacts with SAP145 maps between amino acids 397-512. This region contains the putative zinc finger of IE63 and also encompasses amino acid 405-512 of IE63 which are essential for splicing inhibition and for the redistribution and interaction with the snRNPs (Sandri-Goldin *et al.*, 1995). That IE63 mutated at amino acids 480 redistributes the snRNPs but does not inhibit splicing, suggests that an additional interaction - likely being mediated between amino acids 405 and 480 of the IE63 protein, is also critical for the inhibition of splicing. We propose that this interaction is the interaction with SAP145. The principal role of SAP145 is to tether the U2 snRNP to the pre-mRNA which is a critical step in pre-spliceosome formation. By disrupting this interaction IE63 could effectively inhibit splicing as illustrated in Figure 3.8.

Why then does IE63 interact with so many splicing regulatory proteins? This study has shown that IE63 interacts with the splicing proteins SAP145, ALY and p32. Other workers (K S Sciabica, R M Sandri-Goldin, University of California, Irvine, unpublished observations) have shown that IE63 interacts with the splicing regulatory protein Srp20 (reviewed by Kramer, 1996) and have confirmed the interaction with p32. Presumably by targeting a number of points in the splicing process, IE63 could impose maximum disruption of this cellular process. Efficient inhibition of splicing in HSV-1 infected cells may be necessary as splicing signals are highly degenerate and cryptic splice donor

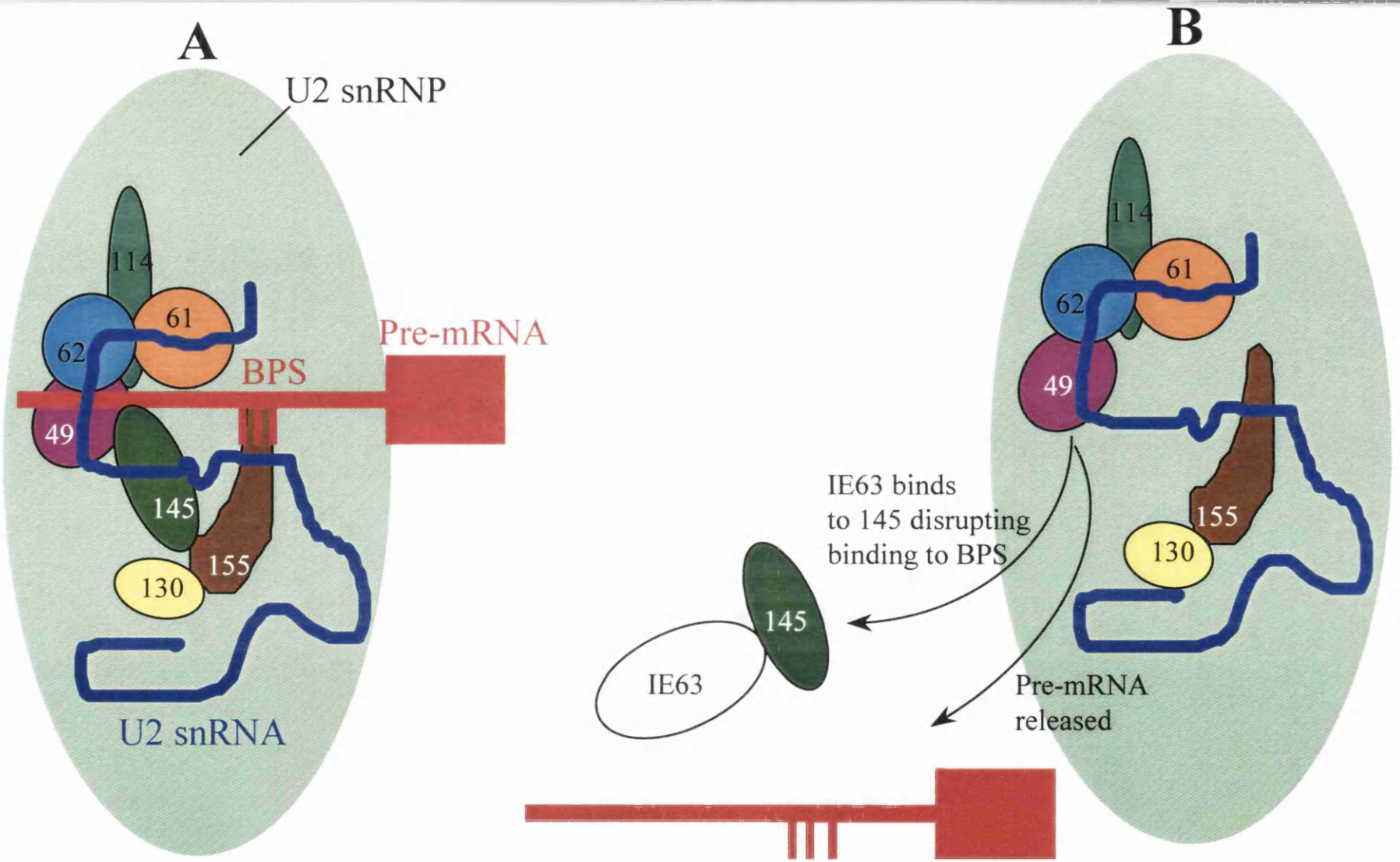


Figure 3.8
Diagram illustrating (A) binding of U2 snRNP with associated protein factors to pre-mRNA in the uninfected cell and (B) proposed role of IE63-SAP145 binding in HSV-1 infected cell. In infected cells IE63 binds to SAP145 and this may disrupt binding of U2 snRNA to the branch point sequence (BPS) of pre-mRNA.

and acceptor signals are therefore likely to appear by chance in HSV transcripts that lack efficient or bona-fide introns. Although insufficient for splicing, such signals are predicted to promote non-productive interactions with the splicing machinery, leading to nuclear retention (Chang & Sharp, 1990) or degradation through the discard pathway. IE63 could counteract this effect acting to 'rescue' virus transcripts. Intriguingly, a small fraction of HSV-1 glycoprotein C mRNA is unspliced, indicating the presence of weak splicing signals in this transcript (Frink *et al.*, 1983). As gC protein is translated from the unspliced form of the mRNA, it is perhaps significant that accumulating of (unspliced) gC mRNA stringently requires IE63 (Rice & Knipe, 1990).

In this regard, Buisson *et al.* have shown that the Epstein Barr virus IE63 homologue BMLF1 inhibited the cytoplasmic accumulation of mRNA generated from precursors containing weak (cryptic) 5' splice sites but had no effect on the cytoplasmic accumulation of mRNA generated from precursors containing constitutive splice sites (Buisson *et al.*, 1999). Additional interactions of IE63 with the splicing machinery could be mediated by a recently identified aa sequence motif (T M Soliman and S J Silverstein, Columbia University, New York, unpublished observations) located just upstream of the zinc finger region which is homologous to the Sm motif conserved in the eight human proteins common to snRNPS U2, U2, U4/U6 and U5 (reviewed in (Luhrman *et al.*, 1990)). The Sm motif, which is also present in non-spliceosomal proteins, potentiates Sm protein-protein interactions (Hermann *et al.*, 1995).

3.2.7 The IE63 protein interacts with the spliceosome associated protein ALY

Checking the sequences of the positive clones against the GenEMBL database revealed that four clones of similar size could be identified as the splicing factor ALY, the largest of which contained an insert of 990 bp. The clone with the highest homology has 91.9% homology in 456 bp of this 1,100 bp gene.

Previous ALY studies have been on the murine gene, however recently Wichmann *et al.* cloned the 1.1. kb human ALY cDNA which encodes a 28 kDa protein. This protein has 94% amino acid homology with murine ALY, and autoantibodies to ALY were found in a patient with systemic Lupus Erythematosus (Wichmann *et al.*, 1999).

(a) Significance of the interaction between ALY and IE63

Since IE63 inhibits splicing and ALY is a spliceosome associated protein (Neubauer *et al.*, 1998) it is possible that IE63 may mediate this effect via interaction with ALY. However, at present little is know about the role of ALY in splicing and further studies of ALY will be required before any definite biological role for this interaction can be determined.

Alternatively, evidence suggests that ALY is a cellular transcription factor (Bruhn *et al.*, 1997) and it is plausible that IE63 either recruits ALY for viral transcription or disrupts the transcription of cellular genes by removing ALY from sites of cellular transcription.

3.2.8 The IE63 protein interacts with the splicing factor p32

Checking the sequences of the positive clones against the GenEMBL database revealed that twelve independent clones could be identified as the splicing factor p32, the largest insert of which was 1,100 bp. The clone with the highest homology had 99.3% homology in 446 bp sequenced of this 1,163 bp gene. Helen Bryant has confirmed that IE63 interacts with p32 using co-immunoprecipitation and has also shown that p32 is phosphorylated *in vitro* by coimmunoprecipitated CKII activity and is redistributed in the presence of IE63 (H E Bryant, D A Matthews, S Wadd, J E Scott, S Graham, W C Russell, J B Clements, Institute of Virology, University of Glasgow, submitted for publication).

(a) A region in the N-terminus of IE63 is responsible for the interaction with p32

In order to map the region of IE63 required for interaction with p32 a series of IE63 truncations were expressed as hybrids with the GAL4 DNA-binding domain. These were mated into host cells transformed with the clone identified as p32 from the library screen. Figure 3.9 shows the region of IE63 present in each construct and the degree of activation of the two reporter genes HIS3 and LacZ in the presence of the p32 protein relative to the positive control pVA3-1 and pTD-1 and the negative controls pAS2-1 and pLam5'1.

The results show that sequences within aa 166-242 of IE63 are essential for the interaction with p32. However, the binding of the smaller IE63 truncations to p32 was weaker than the binding of the larger truncation 502CBD indicating that other regions of IE63 may contribute to the interaction with p32. Interestingly, the IE63 region aa 166-242 was also necessary for the interaction with hnRNP K and CKIIβ.

(b) Significance of the interaction between p32 and IE63

p32 has putative roles in splicing, regulation of transcription and the transport of proteins across the nuclear pore. These are all functions which it apparently shares with IE63. However, the region of IE63 which appears to be critical for the interaction with p32 (aa 166-242) is not thought to be required for splicing inhibition, activation and repression of viral genes or nuclear transport. Using immunofluorescence with a standard fixation method, workers in our laboratory saw p32 predominantly in the cytoplasm of uninfected cells but on HSV-1 infection and with expression of IE63 alone, the distribution altered with the bulk of p32 no longer being present in the nucleus (H E Bryant, D A Matthews, S Wadd, J E Scott, S Graham, W C Russell, J B Clements, Institute of Virology, University of Glasgow, submitted for publication). A similar redistribution was observed in adenovirus infected cells where p32 has been suggested to play a role in adenovirus splicing regulation (Matthews &

Russell, 1998). Furthermore, the HIV-1 protein Rev which upregulates the cytoplasmic appearance of singly and unspliced viral RNAs during infection interacts with p32 and it has been proposed that p32 functions as a link between p32, Rev and the cellular splicing apparatus (Tange *et al.*, 1996). This raises the possibility that during HSV-1 infection the IE63-dependent redistribution of p32 may contribute to the disruption of splicing and the region of IE63 required for interaction with p32 (aa 166-242) may modulate this effect representing a further mechanism of IE63-induced splicing inhibition.

3.3 THE REGIONS OF IE63 THAT INTERACT WITH hnRNP K, SAP145 AND CKII ARE CONSERVED THROUGHOUT THE ALPHAHERPESVIRINAE

Figure 3.10 (A) shows the amino acid conservation of IE63 and its seven alphaherpesvirus homologus as aligned using the Pileup program. This Figure shows that IE63 aa 242-397, essential for the interaction with hnRNP K, is the most conserved region throughout the alphaherpesviruses with twenty two amino acids in this region being absolutely conserved throughout the group. Of particular interest in this region is the sequence SADET (IE63 aa 355-359) which is well conserved throughout the alphaherpesviruses. This sequence is homologous to part of the KNS domain (GFSADETWDSAIDTWSPSEWQMAY) located between hnRNP K aa 338-361 and shown to be sufficient for its NES activity (Michael *et al.*, 1997a). The data presented here show that this region was not required for interaction of IE63 with CKIIβ, p32 and SAP145.

The IE63 region aa 397-512 containing the putative zinc finger was required for the interaction with CKIIβ and SAP145. This region contains fourteen amino acids which are completely conserved through the alphaherpesviruses, six of which form the zinc finger and, strikingly, their relative spacing is completely conserved (C-X_{10}-H-X_3-C-X_4-C-X_{14}-H-X_6-C) located between IE63 aa 468-508. Zinc fingers are well known to be involved in DNA binding

Figure 3.10 (A) Amino acid conservation of IE63 and homologues in the alphaherpesviruses. Conserved amino acid in bold. Amino acids coloured blue represent aa 166-242 which contributes to the interaction of IE63 with hnRNP K, CKIIβ and p32. Amino acids coloured red represent aa 242-397 which is the region required for interaction with IE63 and hnRNP K. Amino acids marked in green represent aa 397-512 and this region is required for interaction of IE63 with CKIIβ and SAP145, and for IE63 self-dimerization.

(A) Alphaherpesviruses

```
HSV-1     MATDIDMLIDLGLDLSDSDLDEDPPE.PAESRRDDLESDSSGECSSSDEDMEDPHGEDGPEPILDAARPAVRPSRPEDPG 79
HSV-2     MATDIDMLIDLGLDLSDSELEEDALERDEEGRRDDPESDSSGECSSSDEDMEDPCGDGGAEAI.DAAIPKGPPARPEDAG
EHV-1                                                      MALSSVSSCEPMEDEMSIMGSDTEDNFTGGDTCA
EHV-4                                                      MALSSVSSCEPMDDEMSIMGSDTDDTL.GGSCVE
BHV-1                                                          MADPEIATLSTASESDDLSLFGSDRE
PRV
VZV                                                                                MASASIPTD
MDV                                                  MSVDAFSRESDDMMSLLDYDFIEGSSSDENAEVTEMETSAK

HSV-1     VPSTQTPRPTERQGPNDPQPAPHSVWSRLGARRPSCSPEQHGGKVARLQPPPTKAQPARGGRRGRRRGRGRGGPGAADGL 159
HSV-2     TPEASTPRPAARRGADDPPPATTGVWSRLGTRRSASPREPHGGKVARIQPPSTKAPHPRGGRRGRRRGRGRYGPGGADST
EHV-1     EATRGLVNKSAFVPTQTVGTVSALRNVVGDPPKSVVVSFSASPQRAQPSNPKSERPAFGHGRRNRRRPFRRNNWKQQQRG
EHV-4     AAQSAVVNKRAFEMSESTGTMSTIRNVVSEVPKSLVVSFAASPKNPKPQNTTSERSAFPHGRKNRRRPFRRNNWK..QRA
BHV-1     EDDEAPSLAPALRSVVGQVRKRKLEGAEDEPMPAEPPGEGAASGDGGPAEAPPARRARVRPRRPRRRPRRRQPAGEQRSR
PRV         MEDSGNSSGSEASRSGSEERRPVRERLGSRPPERRPVRARLGAIRRRRGGRGGRAARQALRQRRRQQQQQQRQQQHQR
VZV       PDVSTICEDFMNLLPDEPSDDFALEVTDWANDEAIGSTPGEDSTTSRTVYVERTADTAYNPRYSKRRHGRRESYHHNRPK
MDV       TANNKNEVLFAPPCTQELLTERPSPDSKNSQGDDDSNSIYGNVIRDAQHSASRYATRCLDNAIPRKRLRLANLTVDSACI

HSV-1     SDPRRRAPRTNRNPGGPRPGAGWTDGPGAPHGEAWRGSEQPDPPGGQRTRGVRQAPPPLMTLAIAPPPADPRAPAPERKA 239
HSV-2     PKPRRRVSRNAHNQGGRHPASARTDGPGATHGEARRGGEQLDVSGGPRPRGTRQAPPPLMALSLTPPHADGRAPVPERKA
EHV-1     WEKPEPENVPARQSAGSWPKRSSLPVHMRLGQRGGDSSSADSGHGGAGPSDRWRFKTRTQSVARVHRNRRRGNANHGSNT
EHV-4     WEKQSQEAAPANQGSRNWPKRSSMPVHMRLGHRSGDFQSADAGHCTAGPSGGWRFKTRTHSASRVYHNRQRGNTNKSGNA
BHV-1     GPAAKREAALATSSHGGGGAAARSIGSSLRLARSLAEAAQ RATAERVTAVFAGARLDLMRPVQNGGFRAAGV.......
PRV       RRQEADRPDGGPD APPDRLSESARAAVSATHA RVGATRVNELFASARHDLSRPVFNDGFRAAGS..............
VZV       TLVVVLPDSNHHGGRDVETGYARIERGHRRSSRSYNTQSSRKHRDRSLSNRRRRPTTPPAMTTGERNDQTHDESYRLRFS
MDV       SQTKRPHGTGNRKQYHRRNFPMSPTSQEKIHLRLHNRLGSRSEKQQRSLNYDRRLQEGHHRRRFYSERRIYDQNHSHHRT

HSV-1     PAADTIDATTRLVLRSISERAAVDRISESFGRSAQVMHDPFGGQPFPAAN.........SPWAPVLAGQGG.PFDAETR 308
HSV-2     PSADTIDPAVRAVLRSISERAAVERISESFGRSALVMQDPFGGMPFPAAN.........SPWAPVLATQAG.GFDAETR
EHV-1     PGRSAGDRLNAAAASSIADVCRRVTSSRIGEMFHGARETLTTPVKNGGFRAENS......SPWAPVLGFGSD.QFNPEAR
EHV-4     SSSRSGDRLNAAAANAIADVSKRVTSSRISDMFHGARETLTSPVKNGGFRAEHS......SPWSPVLNFGLE.QFNPEGR
BHV-1     ..........................................................SPWAAVLDFGAE.QFVPEGR
PRV       ..........................................................SPWAAVLEFGAE.QFTPDGR
VZV       KRDARRERIRKEYDIPVDRITGRAIEVVSTAGASVTIDSVRHLDETIEKLVVRYATIQEGDSWASGGCFPGI......KQ
MDV       HDIRVPLEKYRVSRQHDLPVHEELNEILQREKHRLASISNECDFRVSSK.........NRWAAVLTFSSNAESTLCGP

HSV-1     RVSWETLVAHGPSLYRTFAGNPRAASTAKAMRDCVLRQENFIEALASADETLAWCKMCIHHNLPLRPQDPIIGTTAAVLD 388
HSV-2     RVSWETLVAHGPSLYRTFAANPRAASTAKAMRDCVLRQENLIEALASADETLAWCKMCIHHNLPLRPQDPIIGTAAAVLE
EHV-1     RITWDTLVEHGVNLYKLFEVRSHAAEAARSLRDAVMRGENLLEALASADETLSWCKMIVTKNLPMRTRDPIISSSVALLD
EHV-4     RITWDTLVTHGENLYKLFEVRSHAAEAARSLRDLVMRGENLLEALASADETISWCKMIITKNLPMRTRDPIIHSSIALLE
BHV-1     RVTWETLMFHGRDLYRMFEVRPHAAQAARALRDLVLRSANLVDALASADECLTWCKFIATKNLRLRTKDPIVATAGAVLE
PRV       RVTWETLMFHGADLHRLFEVRPHATEAARVLREMVLLNEGLTESLASADETLTWVKLILTKGLTLRTLDPIVATAGAVLQ
VZV       NTSWPELMLYGHELYRTFESYKMDSRIARALRERVIRGESLIEALESADELLTWIKMLAAKNLPIYTNNPIVATSKSLLE
MDV       QITWEYLLHAGPELRNTFEIRPRISLQASAAREAVLRGESFIAALGSAEETLSWLKLHAVLKLRLVNHDPIFKTAGAVLD

HSV-1     NLATRLRPFLQCYLKAR....GLCGLDELCSRRRLADIKDIASFVFVILARLANRVERGVAEIDYATLGVGVGEKMHFYL 464
HSV-2     NLATRLRPFLQCYLKAR....GLCGLDDLCSRRRLSDIKDIASFVLVILARLANRVERGVSEIDYTTVGVGAGETMHFYI
EHV-1     NLRLKLEPFMRCYLSSS....GSPTLAELCDHQRLSDVACVPTFMFVMLARIARAVGSGAETVSRDALGPD.GRVLADYV
EHV-4     NLRLKLEPFMRCYLSSS....GSPTLAELCDHQRLSDVACVPTFMFVTLARIARAVGSGAEAVSPDALGPA.GHALANYV
BHV-1     NLRLKLAPFLRCYLRGR....GLPSLEELCAARRLSLATCPASYMFVMLARLSRAVRSGAECVPLLEVTVG.DAPFEEYI
PRV       NLRLKLGPFLRCYLRDT.......PVDELVRRRRLRDVRCIVTYTLVMLARIARVVERGSSCVLPEDLGDS.PVPLEEYV
VZV       NLKLKLGPFVRCLLLNRDNDLGSRTLPELLRQQRFSDITCITTYMFVMIARIANIVVRGSKFVEYDDISCN.VQVLQEYT
MDV       NLRLKLAPIMMCKYGTE.....KRSMGDMLRRSAPEDINDSLTLCLILLSRIRRVMHRTSGSKYSYMIDPR..GCMIDYV

HSV-1     PGACMAGLIEILDTHRQECSSRVCEL.TASHIVAPPYVHGKYFYCNSLF                                512
HSV-2     PGACMAGLIEILDTHRQECSSRVCEL.TASHTIAPLYVHGKYFYCNSLF
EHV-1     PGACLAGTLEAIDAHKRRCKADTCSL.VSAYTLVPVYLHGKYFYCNQIF
EHV-4     PGTCLAGTLEAIDLHKRRCKESTCSL.VSSYTLVPVYLHGKYFYCNQIF
BHV-1     PGTCVAGLIDALDTHKQACDSMTCKL.VANFTLVPVYMHGKYFYCNEIF
PRV       PGACLGGIMDALDSHKTGCDAPTCRL.TCSYTLVPVYMHGKYFYCNHLF
VZV       PGSCLAGVLEALITHQRECGRVECTLSTWAGHLSDARPYGKYFKCSTFNC
MDV       PGECMTNILRYVDAHTRRCSDPACNL.YISCTLMPIYIHGRYFYCNTLFGM


HSV       Herpes simplex virus
EHV       Equine herpesvirus
BHV       Bovine herpesvirus
PRV       Pseudorabies virus
VZV       Varicella zoster virus
MDV       Marek's disease virus
```

(B)Amino acid conservation of HHV8-ORF 57 and homologues in the gamma herpesviruses.

```
HHV-8        MVQAMIDMDIM.KGILE/DSVSSSEFDESRDDETDAPTLEDEQLSEPAEPPADERIRGTQSAQGIPPPLGRIPKKSQ
HVS          MAQAMVTNCQM.EDIIE/GISSDDDFDSSDSSSDEEESDTSPQIMKSDVTMASPPSTPEPSPDVSASTSNLKRERQR
BHV-4        MAQAMLTMDCM.REIIE/DLSSDIDSFSGGESIDMESELEEGEIESDTNSSKPPPPQDLSKPPMMRIPRKRVASPDN
AHV-1        MAQQAIVTMSALRRTME/VSDSGDVSIDISAEDSNDSFHLEESVDDCMDDCKPNNRPNPISMKPAKRRVFMVPKRER
MHV-68       MAQQMLEAGAL.DQMME/GLPSDFDFDTSDEEGELSDSPPVEEPTGPVRDVVYEPDPLFDDPPPTPSPDVKPPSPKA
EBV       MVPSQRLSRTSSISSNEDPA/ESHILELEAVSDTNTDCDLDPMEGSEEHSTDGEISSSEEEDEDPTPAHAIPARPSSVVI

HHV-8     GRSQLRSEIQFCSPLSRPRSPSPVNRYGKKIKFGTAGQNTRPPPEKRPRRRPRDRLQYGRTTRGGQCRAAPKRATRRPQV
HVS       SPITWEHQSPLSRVYRSPSPMRFGKRPRISSNSTSRSCKTSWADRVREAAAQRRPSRPFRKPYSHPRNGPLRNGPPRAPP
BHV-4     ERMEYRSPLNRTYPPPFTERYGKRRRLTAGRPNWSGRVNEDKGRYRRRGLSDNKTIRHTQASIKDEVAVSLRKMKIPTGM
AHV-1     SKTPVQHTSPLNRLYPNVVLGKQHGYKQRPAPSARSRRPQPYSARKDSAAKPQSTPSNQNPLTELLKNVDPAIASRITEM
MHV-68    RKRALSPEIVHNSPLLRDTTKYEPAPKRSYSYHPRRSPQRENANQKQKRGPDSRRPNRWNQKSQKQYWSPKPLLDYSKIP
EBV       TPTSASFVIPRKKWDLQDKTVTLHRSPLCRDEDEKEETGNSSYTRGHKRRRGEVHGCTDESYGKRRHLPPGARAPRAPRA

HHV-8     NCQRQDDDVRQGVSDAVKKLRLPASMIIDGESPRFDDSIIPRHHGACFNVFIPAPPSHVP.................EVF
HVS       LLKLFDISILPKSGEPKLFLPVPSLPCQEA.....................................................EKT
BHV-4     IRRAGEKPFDETLLSSGGPGRYSVFLPRAPEFKL.................................................ERY
AHV-1     RIPRSMLRTPSGQPFAHWLMPSAEDSSKFINVNPVNMEV...........................................EEH
MHV-68    RAEYKNAKLLVPTTGKLRP.................................................................EFY
EBV       PRVPRAPRSPRAPRSNRATRGPRSESRGAGRSTRKQARQERSQRPLPNKPWFDMSLVKPVSKITFVTLPSPLASLTLEPI

HHV-8     TDRDITALIRAGGKDDELIN.KKISAKKIDHLHRQMLSFVTSRHNQAYWVSCRRETAAAGGLQTLGAFVEEQMTWAQTVV
HVS       NDKYVLAMAQRAMHDVPISS.KQLTANLLPVKFKPLLSIVRYTPNYYYWVSMRKETIASANLCTVAAFLDESLCWGQQYL
BHV-4     TDKLVSSLVEKGGENGAGIS.KKLSHLKLSSNFSVIHSFLNKSINYHYWVCLRKETMGSCGLTSLMLFLEETCCWAQLCT
AHV-1     VNVVVRRCTEWALISSRLQD.KSISTKYLAENFYDLRDFAQRSINKSAWINLRREAIANAGFVNLCAFADEMMMWLQLNL
MHV-68    TDRFVDAIIQNAARNCPVSE.KAVSLKNIEESFKLLNSFFNSGINKDHWLSTRYFAIFNNGLVVLTHMLDEQLAWAYACL
EBV       QDPFLQSMLAVAAHPEIGAWQKVQPRHELRRSYKTLREFFTKSTNKDTWLDARMQAIQNAGLCTLVAMLEETIFWLQEIT

HHV-8     RHGGWFDEKDIDILDTAIFVCNAFVTRFRLL.HLSCVFDKQSELALIKQVAYLVAMGNRLVEACNLLGEVKLNFRGGLL
HVS       KNDFIFSENGKDIILDTSSALLSQLVHKIKMLPFCHCLMQTTPQDHIVKQVCYLIASNNRILDAVRYLQTSVIKSPIVLL
BHV-4     SNDVSINGFSNDIILNSANFLSVQIMFKLRSL.VMPCFAREAHNISLVKQLGYLVSTTNKIQTAASLIRELKLDTKLCLL
AHV-1     NNQGSWKACREDIILTGAPDMCFHALQKVRAF..IKCFLRERHQRALVNALCHIICFEGGIKQAATLCQELFFDFKVGLM
MHV-68    KHGRELP..TDDILMSTSEKLSQQLVIKLIEV..IKCIEKDGIFSRILKGVADAVCLKAQFLRGMITLKRTPCSLPMYTL
EBV       YHGDLPLAPAEDILLACAMSLSKVILTKLKEL..APCFLPNTRDYNFVKQLFYITCATARQNKVVETLSSSYVKQPLCLL

HHV-8     LAFVLTIPGM.QSRRSISARGQELFRTLLEYYRPGDVMGLLNVIVMEHHSLCRNSECAAATRAAMGSAKFNKGLFFYPLS
HVS       LAYAVCLPAA.IICTKNETQLYSHCMRILKEYRPGDVMNILHESLTQHLNKCPSSTCAYTTRAIVGTKANTTGLFFLPTQ
BHV-4     AAFAIVVPTL.LETDKTEHGTYAFFMQYINRYRPGCIMSLYNDVISSHSRECTSRLCIANTRALAGTKDKTKGLFFCPI
AHV-1     VLYFLTPYAFLYSHTIPQCNFGGYFSKCVAQYTPGAVTGLLNSAIEDHYKDCTSQDCTNLITAIVSPETSNKGLLFFPLPM
MHV-68    FVYVLTIPTL..RTRVIRDPLLTQCKDVVLKYQPGDCITLLKAALNCHQCNKDCDKCKYILDPLLGQTHRTKGVFFVCE
EBV       AAYAAVAPAYINANCRRRHDEVEFLGHYIKNYNPGTLSSLLTEAVETHTRDCRSASCSRLVRAILSPGTGSLGLFFVPGLNQ
```

HHV Human herpesvirus
HVS Herpesvirus saimiri
BHV Bovine herpesvirus
AHV Avian herpesvirus
MHV Murine herpesvirus
EBV Epstein-Barr virus

Thanks to Dr Andrew Davison for help with the amino acid alignments using Pileup programme.

(reviewed by Choo & Klug, 1997) and in interactions with RNA and DNA-RNA hybrids. In addition, zinc fingers can mediate protein-protein interactions, potentiating interactions with other proteins as well as self interaction (reviewed by Mackay & Crossley, 1998). The yeast genome contains some 500 zinc finger proteins and estimates are that some 1% of all mammalian genes encode zinc fingers.

These findings suggest that other alphaherpesvius homologues of IE63, such as varicella-zoster virus (VZV) ORF4 and bovine herpesvirus-1 (BHV-1) BICP27, will have a similar function to IE63, and that these proteins will engage some cellular partner proteins in common. Indeed the VZV ORF4 homologue is an immediate-early protein with transcriptional and posttranscriptional regulatory properties (Defechereux *et al.*, 1997) although unlike IE63, ORF4 is a component of the VZV virion (Kinchington *et al.*, 1995). The bovine herpesvirus BICP27 homologue, while expressed with early kinetics, like IE63 acts post-transcriptionally to increase RNA 3' processing efficiency at weak poly(A) sites (Singh *et al.*, 1996).

The data presented here, show that sequences within IE63 aa 166-242 contributed to the interaction with hnRNP K and CKIIβ and this was the only region required for the interaction with p32. This region is not conserved among the alphaherpesviruses but in HSV-1 and HSV-2 it contains runs of prolines similar to the hnRNP K SH3 binding domain, a region of hnRNP K known to potentiate interactions with partner proteins (Bomsztyk *et al.*, 1997).

3.4 IE63 INTERACTS WITH A NUCLEOPORIN-LIKE PROTEIN IN YEAST CELLS

3.4.1 IE63 interacts with the nucleoporin-like RIP1 protein in the yeast two-hybrid system but not the nucleoporin CRM1

Although the yeast two-hybrid library screen should theoretically identify all of the cellular proteins which interacts with IE63, other workers frequently

report that screens fail to identify interactions with proteins that are known to interact with their bait protein. In other words although a screen is an excellent method for identifying protein-protein interactions, it does not guarantee that all proteins which interact with the bait protein *in vivo* will be identified. The reason for this is not as yet not known but it may be due to the fact that the portion of the library screened is not always 100% representative. For this reason and considering that a protein directly implicated in RNA transport had not been identified in the yeast two-hybrid assay, we decided to test for an interaction between IE63 and; (i) the cellular nuclear protein RIP1 (Bogerd *et al.*, 1995, Fritz *et al.*, 1995) which contains a series of phenylalanine-glycine repeats (FG repeats) present in the FG nucleoporin family of proteins and is implicated in the process of nucleocytoplasmic transport (Davis, 1995) (ii) the nucleoporin export factor hCRM1 (Neville *et al.*, 1997, Stutz *et al.*, 1997).

The expression plasmids 502CBD and 502CAD were simultaneously transformed into the yeast strain Y187 with an expression plasmid encoding either RIP1 fused to the GAL4 AD or CRM1 fused to the GAL4-BD respectively. Following three days of incubation, a beta-galactosidase assay was performed. Experiments showed that beta-galactosidase activity was present in the presence of RIP1 (++) but not in the presence of CRM1 (-), indicating an interaction with RIP1 but not with CRM1.

3.4.2 The significance of the interaction between IE63 and RIP1 and the lack of interaction with CRM1

Several lines of investigation have demonstrated that CRM1 is the HIV-1 Rev leucine-rich NES receptor (reviewed in Stutz & Robash, 1998). The FG repeat region of RIP1, initially identified as a Rev NES interacting region in a two-hybrid screen, was shown to interact via the bridging activity of CRM1. Since IE63 contains a leucine-rich NES, it seems puzzling that an interaction with CRM1 was not found. However, the cytotoxin leptomycin B, shown to target the CRM1 protein and prevent shuttling of Rev, does not efficiently inhibit

cytoplasmic shuttling of IE63 (T F Soliman and S J Silverstein, Columbia University, New York, unpublished observations), indicating that IE63 may be exported by an alternative pathway.

RIP1 is absolutely required for the selective export of heat shock RNAs (Saavedra *et al.*, 1997) with no substantial contribution from the FG repeat regions (Stutz *et al.*, 1997) and is a nuclear pore component. It is possible that both IE63 and hnRNP K could directly recognise a common nuclear pore component (such as RIP1), via the SADET and KNS sequences respectively, to facilitate nucleocytoplasmic transport. The influenza virus NS2 protein mediates the nuclear export of virion RNAs and interacts with several cellular nucleoporins including RIP1; however the interaction of NS2 with CRM1 was not examined (O'Neill *et al.*, 1998).

3.5 hnRNP K, SAP145, CKII AND p32 CLONES IDENTIFIED IN THE YEAST TWO-HYBRID SCREEN DO NOT INTERACT WITH THE KAPOSI'S SARCOMA IE63 HOMOLOGUE

Human herpesvirus-8 (HHV-8), or Kaposi's sarcoma (KS) associated herpesvirus is a gammaherpesvirus (reviewed in Schuly, 1998). KS is a neoplasm found predominantly in AIDS patients. HHV-8 expresses an IE63 homologue from the ORF57 gene, and this gene product, like IE63, can regulate viral gene expression at the post-transcriptional level (Gupta *et al.*, 2000), exhibits a punctate nuclear distribution that co-localises with the cellular splicing factor SC35 and shuttles from a transfected monkey nucleus to a recipient mouse nucleus in an interspecies heterokaryon assay (Bello *et al.*, 1999). ORF57 of the gammaherpesvirus herpesvirus saimiri acts post-transcriptionally to repress gene expression, dependent on the presence of an intron in the reporter gene (Whitehouse *et al.*, 1998) and causes the clumping of splicing factor SC35 and spliceosome components and co-localises with these redistributed components (Cooper *et al.*, 1999). CRM1 appears to mediate the function and the shuttling of the Epstein-Barr virus IE63 homologue (Boyle *et al.*, 1999). These findings suggested that HHV-8 ORF57

may function in a similar manner to IE63. For this reason, a two-hybrid screen was set up to test for interactions of HHV-8 ORF578 with the proteins expressed from the hnRNP K, CKIIβ, SAP145 and p32 clones which had been shown to interact with IE63 in the library screen.

The expression plasmid encoding the full ORF57 coding sequence of 455 amino acids fused to the GAL4 BD was constructed from ORF57 cDNA (a kind gift of Dr L Bello) as the ORF57 gene contains a single intron (Bello *et al.*, 1999). This plasmid was simultaneously transformed into the yeast strain Y187 with an expression plasmid encoding either hnRNP K, CKIIβ, SAP145 or p32 fused to the GAL4 AD. The ORF57-GAL4 BD construct was not sequenced to ensure that it lay in the same frame as the GAL4 BD or tested for expressed in yeast cells due to time limitations. Following three days of incubation, a β-galactosidase assay was carried out. The results of the experiment demonstrated that there was no β-galactosidase activity (-) detected in any of the assays compared to the positive control of pVA3-1 and pTD-1 (++) indicating that HHV-8 ORF57 does not interact with the cellular proteins identified in this study to interact with IE63. However, since the construct had not been sequenced or tested for expression it is not possible to say conclusively that HHV-8 ORF57 does not interact with these proteins.

3.5.1 Interpretation of the finding that the Kaposi's Sarcoma IE63 homologue does not interact with hnRNP K, CKII, SAP145 or p32

HHV-8 is a gammaherpesvirus and the ORF 57 gene, unlike IE63, contains a single intron and shuttles in a heterokaryon assay but not after actinomycin D treatment of transfected cells suggesting that protein re-import into the nucleus is not transcription dependent (Bello *et al.*, 1999). Figure 3.10 (B) shows a Pileup analysis of the ORF57 amino acid sequence together with the sequences of the other gammaherpesviruses. ORF57, like other gammaherpesviruses, contains an N-terminal exon of 16-17 amino acids (with the exception of EBV which is 20 amino acids in length). A feature of the ORF57 extreme N terminus is the absence of an acidic region as compared to IE63 and this is a

feature of all gammaherpesvirus homologues. Full length ORF57 did not cause transcriptional activation of the two-hybrid system's reporter genes and this may reflect the absence of this acidic region which in IE63 did cause transcriptional activation in this assay system.

Comparison of the HHV-8 ORF57 sequence with that of HSV-1 IE63 showed little amino acid conservation. The arrangements of cysteine and histidine residues in the putative zinc finger of the gammaherpesvirus homologues are positionally conserved, with the exception of one cysteine in MHV-68, but the arrangement (H-X_3-C-X_4-C) located between HHV-8 ORF57 aa 423-432 is very different from the zinc finger arrangement of the alphaherpesviruses. These findings suggest that although IE63 and HHV-8 ORF57 proteins share a number of functional similarities, these distantly related homologues are likely to possess certain different properties. Studies of molecular evolutionary history indicate that branches giving rise to the alpha- and gammaherpesviruses diverged around 200 million years ago (McGeoch *et al.*, 1995) while lines giving rise to HSV-1 (α-1) and VZV (α-2) have been separate for some 75 million years.

3.6 IE63 DIMERISES IN YEAST CELLS

A number of viral regulatory proteins such as hnRNP K , the HIV-1 Rev and HSV-1 IE175 and IE110 form dimers or multimers (Ciufo *et al.*, 1994, Gallinari *et al.*, 1994, Hope *et al.*, 1992). Therefore the expression plasmid 502CBD was simultaneously transformed into the yeast strain Y187 with the expression plasmid 502CAD to determine whether or not IE63 could also oligomerise. After three days a filter lift assay showed that β-galactosidase was present (++) indicating that IE63 is able to dimerise *in vivo*. These findings have recently been confirmed by Zhi *et al.* using coimmunoprecipitation and the yeast two-hybrid system (Zhi *et al.*, 1999) and these workers postulate that IE63 may function *in vivo* as a multimer.

3.6.1 The putative zinc finger of IE63 mediates its dimerisation

In order to map the region of IE63 required for its dimerisation a series of IE63 truncations were expressed as hybrids with the GAL4 DNA-binding domain. These were mated into host cells transformed with the IE63 truncation 502C AD. Figure 3.11 shows the region of IE63 present in each construct and the degree of activation of the two reporter genes HIS3 and LacZ in the presence of IE63 fused to the activation domain relative to the positive control pVA3-1 and pTD-1 and the negative controls pAS2-1 and pLam5'1.

These results show that it is the zinc finger region of IE63 which is essential for its ability to form oligomers, a result recently confirmed by Zhi *et al.* (Zhi *et al.*, 1999) who showed that an insertion within but not outside the zinc finger resulted in a loss of interaction. However, the data presented here show that interaction of the mutants 346CBD and 270CBD is weaker than the interaction between 502CBD and 502CAD. This suggests that amino acids 72-166 also contribute to the self-interaction of IE63.

3.7 SUMMARY

This project has found that IE63 interacts with the cellular proteins hnRNP K, CK II, SAP145, p32 and ALY. Identification of the regions of IE63 which interact with these proteins and comparison with previously identified functional domains of IE63 has enabled putative functions to be attributed to each interaction but these will need to be investigated further. Disruption of these interactions using truncated forms of IE63 has confirmed the authenticity of these interactions in yeast cells.

In addition, sequence analysis has shown that certain regions of IE63 which mediate the interaction between hnRNPK, CKII, SAP145 and p32 are well conserved in IE63 homologues throughout the alphaherpesviruses indicating that these proteins may share common molecular partners. The HHV-8 ORF57 protein which has some similar properties to IE63 has been shown

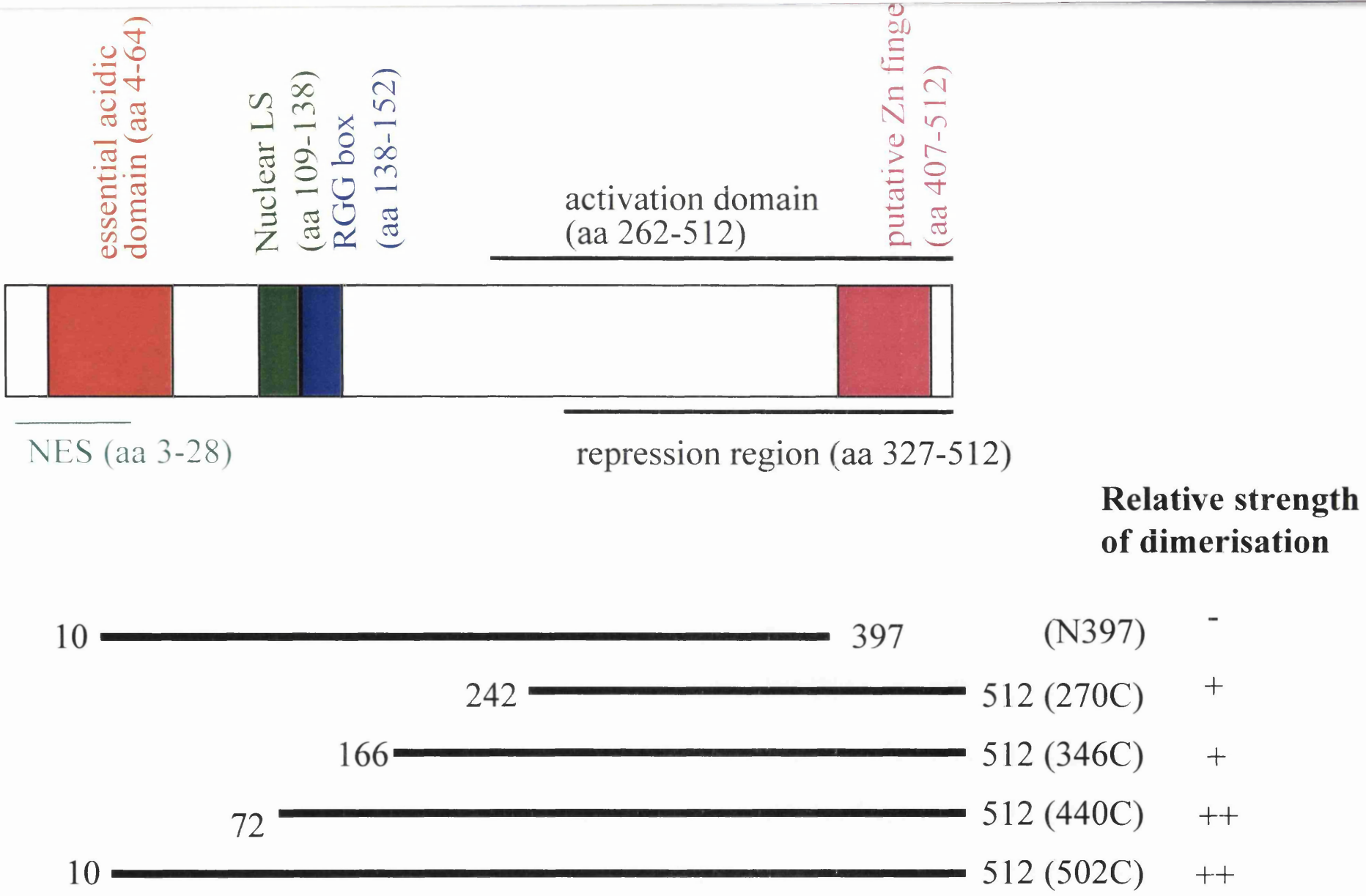


Figure 3.11
Schematic diagram (not to scale) of the IE63 truncations which form a dimer in a yeast two-hybrid assay and their binding relative to the positive control pVA3-1 and pTD-1 (++) and the negative controls pAS2-1 and pLam5'1 (-).

here not to interact with the same cellular proteins. This is likely due to the fact that HHV-8 is a gammaherpesvirus, with ORF57 exhibiting little sequence homology to IE63 and probably possessing certain different functions in viral infection.

Finally, this study has indicated that IE63 may use the nucleoporin RIP1 to cross the nuclear pore. In addition, IE63 appears to form dimers in yeast cells and this function may enable the protein to interact with multiple partners *in vivo*.

It should be noted that although this thesis describes a screen for IE63-protein interactions in an uninfected-cell library, a screen of an infected-cell library would confirm that these interactions exist in HSV-1 infection where cellular protein levels can be increased or decreased. This approach would make an interesting future study.

GENERAL DISCUSSION

4.1 THE YEAST TWO-HYBRID SYSTEM HAS BEEN EFFECTIVE IN IDENTIFYING IE63-PROTEIN INTERACTIONS

The HSV-1 protein IE63 has several, previously defined effects on the cell in which it replicates and on HSV-1 lytic infection itself. IE63-deficient viral mutants show alterations in transcription, polyadenylation, splicing, mRNA transport and viral DNA replication. That IE63 has such diverse effects on the cell and on viral replication led us to propose IE63 functions by interacting with and modifying the cellular proteins involved in these processes. Consequently, the aim of this study was to identify the cellular proteins that interact with IE63. It was anticipated that this would, at least in part, elucidate the mode of action of this protein.

Evidence presented here shows that IE63 interacts with the cellular proteins hnRNP K, CKIIβ, SAP145, p32, ALY and RIP1. The interactions between IE63 and hnRNP K, CKIIβ, SAP145 and p32 have been confirmed independently using biochemical methods by colleagues. Therefore, the supposition that IE63 interacts with cellular proteins has been proven and the aim of this study has undoubtedly been satisfied by the results obtained.

On the whole, the yeast two-hybrid system has been an excellent 'first step' towards determining IE63 function. Indeed, this system has shown that IE63 interacts with SAP145, p32 and ALY and these interactions may be the mechanism by which IE63 inhibits splicing. Testing for an interaction with the hundred or so proteins required for splicing using traditional biochemical methods would of course have been enormously time consuming. Other methods which can screen for interactions of a single protein with the entire set of cellular genes do not give immediate access to part of a gene and its cDNA for ready sequencing and identification and are not performed *in vivo*. Nevertheless, this study has highlighted a complication of the yeast two-hybrid system, ie how can biologically relevant interactions be distinguished from irrelevant ones and what criteria should be adopted to identify a protein which warrants further investigation.

In this study, no less than eight proteins, SAP145, p32, hnRNP K, CKII, RIP1, Gallus Gallus Glutamine Rich Protein, human JTV-1, and transcriptional activator ALY interacted with IE63 and survived false-positive screening experiments. Why then were the latter three proteins reserved for future studies? In short, to fully investigate all of these interactions at any one time would require an enormous amount of manpower. Each interaction would have to be confirmed using an independent biochemical method and this requires antibodies to be obtained or generated for each protein. Once any interaction was confirmed independently, the lengthy process of determining the function, if any, of that interaction *in vivo* would need to be initiated. Therefore the investigation of the interactions identified in this study only in part was unavoidable. SAP145, p32, hnRNP K and CKII were chosen as the first batch of proteins to be examined primarily because it was evident that by interaction with these proteins IE63 could, conceivably mediate some of its previously identified roles in viral infection.

However, many critical scientific discoveries would not have been made had the investigators accepted results that 'made sense' and discounted the unexpected. Indeed the interaction of IE63 with 'transcriptional' activator ALY identified in this project did not initially appear to be of interest but recent work has suggested that ALY is a splicing factor (Neubauer *et al.*, 1998) and therefore the interaction may have a role in IE63 function. For these reasons the proteins *Gallus gallus* Glutamine Rich Protein and human JTV-1 are discussed below and will be investigated at a later date.

4.1.1 The *Gallus Gallus* Glutamine Rich Protein

Gallus gallus is a species of chicken and as the library being screened was a human cell cDNA library the protein that interacts with IE63 cannot be *Gallus gallus* Glutamine Rich Protein (sequenced by S Niu, C A Adamson, E Morkin, University of Arizona, 1997, Accession Number U90567). However, the similarity between the clone identified in the library screen and *Gallus gallus*

Glutamine Rich Protein was 80.6% over 453 bp and two such clones were identified in the library screen. This suggests that IE63 may interact with a human homologue of this protein. Glutamine/Glutamic acid-rich proteins are frequently transcription factors and it is plausible that this clone is a novel transcription factor. Extensive sequencing of this clone is to be done.

4.1.2 Human JTV-1

Finally human JTV-1 is a protein which has been sequenced (Nicholaides *et al.*, 1995) but no function or characteristics have been attributed to this protein at the present time. One clone was identified in this study which had a 85.8% homology over 387 bp of the JTV-1 gene. In these circumstances, no putative function can be assigned to an interaction between IE63 and JTV-1 until the function of this protein is further investigated.

4.2 THE FINDINGS OF THIS STUDY IN CONTEXT

How do the findings of other workers accord with the findings of this study and how do the interpretations made here integrate with existing knowledge/beliefs. This study was novel in that it used the two-hybrid system to screen a library of cellular proteins with IE63 to identify protein-protein interactions. The results in general integrated very well with existing knowledge of the IE63 protein. We anticipated that IE63 would interact with multiple cellular partners and that one or more of these partners would likely play a role in the regulation of splicing, transcription, mRNA transport and phosphorylation. However, since IE63 also plays a role polyadenylation it was anticipated that IE63 might interact with a polyadenylation factor, a likely candidate being the 64 kDa CstF component, a protein which exhibits increased binding to weak viral poly A sites at late times during infection only in the presence of IE63. It may of course be that the effect of IE63 on CstF is an indirect one. Alternatively, Huang *et al.* (Huang *et al.*, 1999) have shown that cis-acting RNA elements can promote histone 2a transcript polyadenylation and it is possible that such elements exist in virus intronless

RNAs, allowing IE63 to polyadenylate viral RNAs with no interaction with cellular polyadenylation factors. In addition, since the phosphorylation state of the U1 snRNP 70 kDa protein is thought to be altered in the presence of IE63 why was this interaction not identified? It is possible that this interaction exists in the cell but that that it has not been demonstrated in the two-hybrid system or that IE63 up-regulates/modulates CKII activity (as is the case of HIV-1 Rev, Ohtsuki *et al.*, 1998) facilitating phosphorylation of the U1 70 kDa protein. In this respect it is interesting that Dr Margy Koffa in our laboratory has shown that the activity of CKII α subunit declines whilst the α1 subunit is activated following HSV-1 infection.

Since this work was carried out, Sciabica and Sandri-Goldin have also used IE63 to screen a cellular cDNA library using the yeast two-hybrid system (K S Sciabica, R M Sandri-Goldin, University of California, Irvine, unpublished work). These workers found that IE63 interacts with the cellular splicing factor SRp20 and propose that IE63 impairs splicing by affecting the phosphorylation of this protein. Sciabica and Sandri-Goldin have also shown that ALY and p32 interact with IE63 in a two-hybrid screen. We did not identify SRp20 as interacting with IE63. This demonstrates that whilst the two-hybrid system is a powerful tool for identifying protein-protein interactions, a library screen is not exhaustive and not all proteins that interact with the bait will necessarily be identified in any one screen. Zhi *et al.* have also recently confirmed the findings of this study by showing that IE63 can form oligomers and also that it is the putative zinc finger of IE63 which is required for its multimerisation (Zhi *et al.*, 1999). Finally Helen Bryant in our laboratory has confirmed that IE63 interacts with CKII, hnRNP K, SAP145 and p32 using independent biochemical methods. Clearly this work is now supported by others working in this field and has allowed important new insights into how this key herpesvirus protein exerts its various activities.

4.3 FUTURE DIRECTIONS

Work in our group was initially focused on using independent biochemical methods to confirm the results of this study, and has also demonstrated that CKII can phosphorylate hnRNP K in the presence of IE63 and that CKII can phosphorylate IE63 itself (Wadd *et al.*, 1999). Future studies will include:

(a) Examining the effect of phosphorylation on IE63/ hnRNP K functions in HSV-1 infected cells using kinase inhibitors and IE63 viral mutants in CKII phosphorylation sites. These studies will examine nucleocytoplasmic shuttling using the heterokaryon assay, the cellular partner proteins present in the IE63 protein complex, and whether or not virus transcripts are exported and intron-containing RNAs still accumulate in the nucleus using *in situ* hybridisation. Effects on the phosphorylation/shuttling of hnRNP A1 will be examined in the presence of expressed IE63. Studies will be performed on the upregulation of CKII activity in HSV-1 infected cells, determining the viral gene products required and examining for phosphorylation of CKII sites on hnRNP K.

(b) The possible significance of interactions between IE63 and the Glutamine Rich Protein and protein JTV-1 will be examined.

(c) Immunofluorescence will be carried out to determine whether or not hnRNP K co-localises with IE63 at sites of viral transcription and effects of phosphorylation on the cellular distribution of IE63/hnRNP K.

(d) The identity of the export receptor which IE63 uses to cross the nuclear pore will be investigated. Mutations in the CKII consensus site present in the IE63 SADET sequence will be introduced into viruses and effects on IE63/hnRNP K shuttling examined.

(e) The ability of SAP145 to participate in splicing complex formation *in vitro* will be examined using splicing extracts from infected cells and from uninfected cells with added recombinant IE63.

(f) Various IE63 mutant viruses located throughout the protein will be used to infect cells and examine for effects on the partner proteins in the IE63 complex by immunoprecipitation.

4.4 IE63 IS A HIGHLY PLEIOTROPIC PROTEIN

To conclude, the evidence presented here suggests that IE63 is a protein with many diverse functions which it carries out via interactions with host cell proteins. There is a close inter-relationship between the processes of transcription and RNA processing as reflected in interactions between RNA polymerase II and cellular factors involved in pre-mRNA splicing and RNA 3' processing (Hirose & Manley, 1998). Some proteins such as WT1 protein (Davies *et al.*, 1999) and hnRNP K appear to co-ordinate transcription and RNA processing/export (reviewed in Ladomery, 1997) by interacting with cellular proteins and nucleic acid. Similarly, IE63 is a multifunctional protein which enables HSV-1 to efficiently utilise host cell mechanisms and replicate efficiently. The proposed role of IE63 on the HSV-1 lytic cycle is illustrated diagrammatically in Figure 4.1. This study has not only greatly improved understanding of the functions of the essential HSV-1 protein IE63, but as is so often the case, has opened many new avenues of research for further investigators to explore. This promises to be an exciting era in the study of IE63.

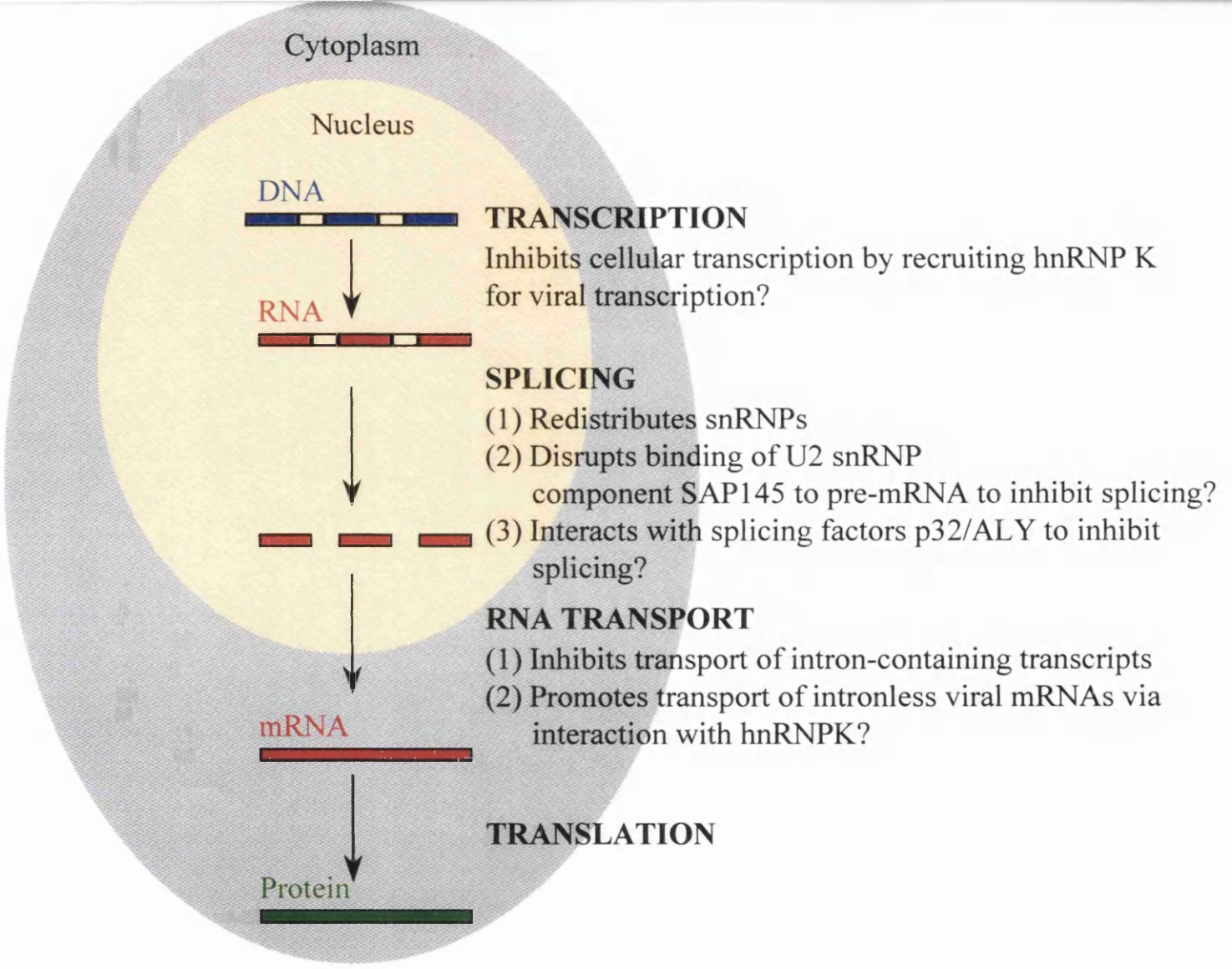


Figure 4.1
Diagram illustrating the stages of cellular protein synthesis at which IE63 acts. The interaction of IE63 with CKII may modulate its interactions with other cellular and viral proteins.

Appendix 1

Sequence comparison of clones identified in library screen with genes that they have homology with

Homo sapiens pre-mRNA splicing factor p32 subunit

```
Sw97.Seq
Em_Hum2:Hssf2p32a

ID   HSSF2P32A  standard; RNA; HUM; 1163 BP.
AC   L04636;
NI   g338044
DT   21-OCT-1992 (Rel. 33, Created)
DT   19-FEB-1994 (Rel. 38, Last updated, Version 2)
DE   Homo sapiens pre-mRNA splicing factor 2 p32 subunit (SF2p32) mRNA, . . .


SCORES      Init1: 1910  Initn: 1988  Opt: 2042  z-score: 1734.5  E(): 0
  98.8% identity in 421 bp overlap

                50        60        70        80        90       100       110
Sw97.Seq        ATCNCCCGGGCTGCAGGAATTCGGCACGAGGCGCTGCCGCGCCCGCCTCGCCTTTCCGGC
                                                  ||||||||||||||||||||||||||||||
Hssf2p32a       GTGTGCTGGGCTCCTCCGTCGCCGGCTCCGCGCTGCCGCGCCCGCCTCGCCTTTCCGGC
                 110       120       130       140       150       160

              110       120       130       140       150       160       170
Sw97.Seq        AGCTCCTGCAGCCGGCACCCCGGCTGTGCACCCGGCCCTTCGGGCTGCTCAGCGTGCCGC
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hssf2p32a       AGCTCCTGCAGCCGGCACCCCGGCTGTGCACCCGGCCCTTCGGGCTGCTCAGCGTGCCGC
                 170       180       190       200       210       220

              170       180       190       200       210       220       230
Sw97.Seq        CAGGTTCCGAGCGGCGGCCGGGCCTCCTGCGGCCTCGCGGACCCTGCGCCTGTGGCTGTG
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hssf2p32a       CAGGTTCCGAGCGGCGGCCGGGCCTCCTGCGGCCTCGCGGACCCTGCGCCTGTGGCTGTG
                 230       240       250       260       270       280

              230       240       250       260       270       280       290
Sw97.Seq        GCTGCGGCTCGCTGCACACCGACGGAGACAAAGCTTTTGTTGATTTCCTGAGTGATGAAA
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hssf2p32a       GCTGCGGCTCGCTGCACACCGACGGAGACAAAGCTTTTGTTGATTTCCTGAGTGATGAAA
                 290       300       310       320       330       340

              290       300       310       320       330       340       350
Sw97.Seq        TTAAGGAGGAAAGAAAAATTCAGAAGCATAAAACCCTCCCTAAGATGTCTGGAGGTTGGG
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hssf2p32a       TTAAGGAGGAAAGAAAAATTCAGAAGCATAAAACCCTCCCTAAGATGTCTGGAGGTTGGG
                 350       360       370       380       390       400

              350       360       370       380       390       400       410
Sw97.Seq        AGCTGGAACTGAATGGGACAGAAGCGAAATTAGTGCGGAAAGTTGCCGGGGAAAAAATCA
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hssf2p32a       AGCTGGAACTGAATGGGACAGAAGCGAAATTAGTGCGGAAAGTTGCCGGGGAAAAAATCA
                 410       420       430       440       450       460

              410       420       430       440       450       460       469
Sw97.Seq        CGGTCACTTTCAACATTAACAACAGCATCCCACCAACATTTGATGGTGAGGAAGGAACCC
                ||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Hssf2p32a       CGGTCACTTTCAACATTAACAACAGCATCCCACCAACATTTGATGGTGAGG-AGGAACCC
                 470       480       490       500       510       520

              470       480       490       500
Sw97.Seq        TCGCAAAGGGCANAANGTTGAAGAACAAGAG
                |||| |||||| || |||||||||||| |||
Hssf2p32a       TCGC-AAGGGCAGAAGGTTGAAGAACAGGAGCCTGAACTGACATCAACTCCCAATTTCGT
```

Casein Kinase II beta subunit

```
ID   HSCKII      standard; RNA; HUM; 921 BP.
AC   X16312;
NI   g29964
DT   23-NOV-1989 (Rel. 21, Created)
DT   12-SEP-1993 (Rel. 36, Last updated, Version 2)
DE   Human mRNA for phosvitin/casein kinase II beta subunit . . .

SCORES      Init1:  1950  Initn:  1950  Opt:  1953  z-score: 1852.6  E():
  99.2% identity in 396 bp overlap

                    80        90        100       110       120       130
21.Seq        TAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGCCCACTTCGCCTGCCGCGGTCGGG
                                                  | | | ||||||||||||||||||||
Hsckii                                          ACTCCCCCCACCCCACTTCGCCTGCCGCGGTCGGG
                                                      10        20        30

                   140       150       160       170       180       190
21.Seq        TCCGCGGCCTGCGCTGTAGCGGTCGCCGCCGTTCCCTGGAAGTAGCAACTTCCCTACCCC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hsckii        TCCGCGGCCTGCGCTGTAGCGGTCGCCGCCGTTCCCTGGAAGTAGCAACTTCCCTACCCC
                  40        50        60        70        80        90

                   200       210       220       230       240       250
21.Seq        ACCCCAGTCCTGGTCCCCGTCCAGCCGCTGACGTGAAGATGAGCAGCTCAGAGGAGGTGT
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hsckii        ACCCCAGTCCTGGTCCCCGTCCAGCCGCTGACGTGAAGATGAGCAGCTCAGAGGAGGTGT
                 100       110       120       130       140       150

                   260       270       280       290       300       310
21.Seq        CCTGGATTTCCTGGTTCTGTGGGCTCCGTGGCAATGAATTCTTCTGTGAAGTGGATGAAG
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hsckii        CCTGGATTTCCTGGTTCTGTGGGCTCCGTGGCAATGAATTCTTCTGTGAAGTGGATGAAG
                 160       170       180       190       200       210

                   320       330       340       350       360       370
21.Seq        ACTACATCCAGGACAAATTTAATCTTACTGGACTCAATGAGCAGGTCCCTCACTATCGAC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hsckii        ACTACATCCAGGACAAATTTAATCTTACTGGACTCAATGAGCAGGTCCCTCACTATCGAC
                 220       230       240       250       260       270

                   380       390       400       410       420       430
21.Seq        AAGCTCTAGACATGATCTTGGACCTGGAGCCTGATGAAGAACTGGAAGACAACCCCAACC
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hsckii        AAGCTCTAGACATGATCTTGGACCTGGAGCCTGATGAAGAACTGGAAGACAACCCCAACC
                 280       290       300       310       320       330

                   440       450       460       470       480       490
21.Seq        AGAGTGACCTGATTGAGCAGGCAGCCGAGATGCTTTATGGATTGATCCACGCCCGCTACA
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Hsckii        AGAGTGACCTGATTGAGCAGGCAGCCGAGATGCTTTATGGATTGATCCACGCCCGCTACA
                 340       350       360       370       380       390

                   500
21.Seq        TCCTTA
              ||||||
Hsckii        TCCTTACCAACCGTGGCATCGCCCAGATGTTGGAAAAGTACCAGCAAGGAGACTTTGGTT
                 400       410       420       430       440       450
```

Human Spliceosome Associated Protein 145

```
ID   Hs413711 standard; RNA; HUM; 2839 BP.
AC   U41371;
NI   g1173904
DT   03-FEB-1996 (Rel. 46, Created)
DT   19-APR-1996 (Rel. 47, Last updated, Version 2)
DE   Human spliceosome associated protein (SAP 145) mRNA, complete cds. . . .


SCORES      Initl:   879  Initn:  1144  Opt:  1907 z-score: 1943.2 E():      0
  89.4% identity in 499 bp overlap

                                      10        20        30        40
Sw34.Seq                      CCCAAAAAAAAGAGATCCTAAG-AACTAGTGGNTC-CCCCGGG
                                      :: : ::: : :: :::: ::  ::::
Hs413711     TTGCGCCGAATGAACCGCTTCACTGTGGCTGAACTCAAGCAGCTGGTGGCTCGGCCCGAT
           1340      1350      1360      1370      1380      1390

                    50        60        70        80        90        99
Sw34.Seq     CTGCAGGAATTCGGCACGA-GTGACAGCGCAGGACCCTAAGCTCTTGGTTCACCTCAAGG
              :  ::: : ::::::: ::::::::::::::::::::::::::::::::::::::::
Hs413711     GTCGTGGAGAT--GCACGATGTGACAGCGCAGGACCCTAAGCTCTTGGTTCACCTCAAGG
           1400      1410      1420      1430      1440      1450

           100       110       120       130       140       150       159
Sw34.Seq     CCACTCGGAACTCTGTGCCTGTGCCACGCCACTGGTGTTTTAAGCGCAAATACCTGCANG
             ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::: :
Hs413711     CCACTCGGAACTCTGTGCCTGTGCCACGCCACTGGTGTTTTAAGCGCAAATACCTGCAGG
             1460      1470      1480      1490      1500      1510

           160       170       180       190       200       210       219
Sw34.Seq     GCAAACGGGGCATTGAGAAGCCCCCCTTCGAGCTGCCAGACTTCATCAAACGCACANGCA
             :::::::::::::::::::::::::::::::::::::::::::::::::::::::::: :::
Hs413711     GCAAACGGGGCATTGAGAAGCCCCCCTTCGAGCTGCCAGACTTCATCAAACGCACAGGCA
             1520      1530      1540      1550      1560      1570

           220       230       240       250       260       270
Sw34.Seq     TCCAAGAGATGCGAGAGG-CCTGCAGGAGAAGGAAGAACAGAAGACCATGAAGTCAAAAA
             :::: :::::::::::::: ::::::::::::::::::::::::::::::::::::::::
Hs413711     TCCAGGAGATGCGAGAGGCCCTGCAGGAGAAGGAAGAACAGAAGACCATGAAGTCAAAAA
             1580      1590      1600      1610      1620      1630

           280       290       300       310       320       330
Sw34.Seq     TGCGAGAGAAAGTTCGGCCTAAGATNGGCAAAATTGACATCGACTACCGGAAACTGCATG
             :::::::::::::::::::::::::: :::::::::::::::::::::: ::::::::::
Hs413711     TGCGAGAGAAAGTTCGGCCTAAGATGGGCAAAATTGACATCGACTACCAGAAACTGCATG
             1640      1650      1660      1670      1680      1690

           340       350       360       370       380       390
Sw34.Seq     ATGCCTTCTTCAANTGGCAAGACCAAGCCAAAGCTGA-CATCCATGGGGA-GTGT-CTAT
             :::::::::::::: :::: :::::::::::::::::: :::::::::: :::: ::::
Hs413711     ATGCCTTCTTCAAGTGGC-AGACCAAGCCAAAGCTGACCATCCATGGGGACCTGTACTAT
             1700      1710      1720      1730      1740      1750

             400       410       420       430       440       450
Sw34.Seq     GAGGGGAAGG-TTTCCANACACNANTGAAGGANAANAACCCAGGAAATCTNTCTGATNAC
             :::::::::: ::: : :::: : ::::::: :: :: ::::::: :::: ::::::: :
Hs413711     GAGGGGAAGGAGTTCGAGACACGACTGAAGGAGAAGAAGCCAGGAGATCTGTCTGATGAG
             1760      1770      1780      1790      1800      1810

             460       470       480       490       500
Sw34.Seq     CTAAGGATT--CTTGGGGATNCCCATTNGGGCCAAATG-CCANAAGGTC
             :::::::::  :::::::::  ::: : :: ::::::: ::: ::::::
             ----------------------------------------------------
```

Heterogenous nuclear ribonucleoprotein K

```
AC   S74678;
NI   gi241477
DT   20-MAY-1993 (Rel. 36, Created)
DT   20-MAY-1993 (Rel. 36, Last updated, Version 1)
DE   heterogeneous nuclear ribonucleoprotein complex K (human, mRNA, . . .


SCORES      Init1: 1663  Initn: 1738  Opt: 2007  z-score: 1949.9  E():      0
  96.3% identity in 432 bp overlap

              40        50        60        70        80        90
Sw73.Seq      GGATCCCCCGGGCTGCAGGAATTCGGCACGAGCTGAGTCTCCCATCAAAGGACGTGCACA
                                        ::::::::::::::::::::::::::::::
S74678        AGAGTGCATAAAGATCATCCTTGATCTTATATCTGAGTCTCCCATCAAAGGACGTGCACA
                720       730       740       750       760       770

             100       110       120       130       140       150
Sw73.Seq      GCCTTATGATCCCAATTTTTACGATGAAACCTATGATTATGGTGGTTTTACAATGATGTT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
S74678        GCCTTATGATCCCAATTTTTACGATGAAACCTATGATTATGGTGGTTTTACAATGATGTT
                780       790       800       810       820       830

             160       170       180       190       200       210
Sw73.Seq      TGATGACCGTCGCGGACGCCCAGTGGGATTTCCCATGCGGGGAAGAGGTGGTTTTGACAG
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
S74678        TGATGACCGTCGCGGACGCCCAGTGGGATTTCCCATGCGGGGAAGAGGTGGTTTTGACAG
                840       850       860       870       880       890

             220       230       240       250       260       270
Sw73.Seq      AATGCCTCCTGGTCGGGGTGGGCGTCCCATGCCTCCATCTAGAAGAGATTATGATGATAT
              ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
S74678        AATGCCTCCTGGTCGGGGTGGGCGTCCCATGCCTCCATCTAGAAGAGATTATGATGATAT
                900       910       920       930       940       950

             280       290       300       310       320       330
Sw73.Seq      GAGCCCTCGTCGANGACCACCTCCCCCTCCTCCCGGGCNAAGCGCGCCGGGGTGGTAACAA
              :::::::::::::: :::::::::::::::::::::::: : :::::::::::::: ::
S74678        GAGCCCTCGTCGAGGACCACCTCCCCCTCCTCCCGGACGAGGCGCGCCGGGGTGGTAGCAG
                960       970       980       990      1000      1010

             340       350       360       370       380       390
Sw73.Seq      AGCTCGGAATCTTCCTCTTCCTCCAACAACAACACCTANANGGGGAGACCTCATGGCCTA
              :::::::::::::::::::::::::: :: :: :::::: : ::::::::::::::::
S74678        AGCTCGGAATCTTCCTCTTCCTCCACCACCACCACCTAGAGGGGGAGACCTCATGGCCTA
               1020      1030      1040      1050      1060      1070

             400       410       420       430       440       450
Sw73.Seq      TGACAGAAGAAGGAGACCTGGAAGACCGTTACGACGGCATGGTTGGTTTCAATGCTGATG
              :::::::::: :::::::::: :::::::::::::::::::::::::::: ::::::::
S74678        TGACAGAAGAGGGAGACCTGG-AGACCGTTACGACGGCATGGTTGGTTTCAGTGCTGATG
               1080      1090      1100      1110      1120      1130

             460       470       480       490       500
Sw73.Seq      AAACTTGGGACTCTGCAATAAGATACATGGAGCCCATCANAA
              :::::::::::::::::::: :::::::::::::::::::: ::
S74678        AAACTTGGGACTCTGCAAT-AGATACATGGAGCCCATCAGAATGGCAGATGGCTTATGAA
               1140      1150      1160      1170      1180      1190
```

Transcriptional Coactivator ALY

```
Sw118.Seq
Em_Hum1:Af047002

ID   AF047002   standard; RNA; HUM; 993 BP.
AC   AF047002;
NI   g2896145
DT   20-FEB-1998 (Rel. 54, Created)
DT   20-FEB-1998 (Rel. 54, Last updated, Version 1)
DE   Homo sapiens transcriptional coactivator ALY mRNA, partial cds. . . .


SCORES    Init1:   640  Initn: 1407  Opt:  1764  z-score: 1559.1  E():      0
  92.5% identity in 429 bp overlap

                50        60        70        80        90        100
Sw118.Seq    GACCAGATCCTAGAACTAGTGGATCCCCCGGGCTGCAGG--AATTCGGCACGAGCGGCAC
                                      ::: :: ::  :: ::: ::   :::: :
Af047002     GGGGCGCAGGCCGCCGCGCGAGTGAATCGAGGCGGCGGGCCCATCCGGAAC---CGGC-C
             50        60        70        80        90          100

                110       120       130       140       150       160
Sw118.Seq    GAGCCATCGCCCGCGGCGCGGCCGGCGGAGGCGGCAGGAACCGAACCGGCGCCCTACAGCA
             : ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Af047002     G-GCCATCGCCCGCGGCGCGGCCGGCGGAGGCGGCAGGAACCGAACCGGCGCCCTACAGCA
                110       120       130       140       150       160

                170       180       190       200       210       220
Sw118.Seq    GGCCAAAACAACTTCCCGACAAGTGGCAGCACGATCTTTTCGACAGTGGCTTCGGCGGTG
             ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Af047002     GGCCAAAACAACTTCCCGACAAGTGGCAGCACGATCTTTTCGACAGTGGCTTCGGCGGTG
                170       180       190       200       210       220

                230       240       250       260       270       280
Sw118.Seq    GTGCCGGCGTTGGAGACAGGTGGGAACTGCTGGTGTCCAATCTGGATTTTGGAGTCTCA
             ::::::::: ::::::::::::::::::::::::::::::::::::::::::::::::::
Af047002     GTGCCGGCG-TGGAGACAGGTGGGAAACTGCTGGTGTCCAATCTGGATTTTGGAGTCTCA
                230       240       250       260       270       280

                290       300       310       320       330       340
Sw118.Seq    GACGCCGATATTCAAGGAACTCTTTGCTGAATTTGGAACGCTGAAGAAGCGGCTGTGCA
             :::::::::::::: :::::::::::::::::::::::::::::::: ::::::::::::
Af047002     GACGCCGATATTC-AGGAACTCTTTGCTGAATTTGGAACGCTGAAGAAGCGGCTGTGCA
                290       300       310       320       330       340

                350       360       370       380       390       400
Sw118.Seq    CTATGATCGCTCTGGTCGCCAACTTAAGAACAACANACGTGCACTTTGAACCGGAAAGGCA
             :::::::::::::::: :: :::: ::::: :: ::::::::::::: :  : :::::::
Af047002     CTATGATCGCTCTGGTCGCAGCTTAGGAAGAGCAAAGCGTGCACTTTGAGC--GGAAGGCA
                350       360       370       380       390

                410       420       430       440       450       460
Sw118.Seq    AATGCCCTGAANGCCATGAANCAANTACAACCGGCGTCCCTCTGGATGGCCGCCCCATGA
             :::::::::: :::::::: : : ::::: :::: ::::::::::::::::::::::::
Af047002     GATGCCCTGAAGGCCATGAAGC-AGTACAA-CGCCTTCCCTCTGGATGGCCGCCCCATGA
           400       410       420       430       440       450

                470       480       490       500
Sw118.Seq    ACATTCAGCTTGTCACGTCACAGATTGACGCACAACGGA
             :::::::::::::::::::::::::::::::::: ::::
Af047002     ACATTCAGCTTGTCACGTCACAGATTGACGCACAGCGGAGGCCTGCACAGAGCGTAAACA
```

BIBLIOGRAPHY

Ackerman, M., Braun, D., Pereira, K. & Roizman, B. (1984). Characterisation of Herpes Simplex 1 α proteins 0, 4 and 27 with monoclonal antibodies. *Journal of Virology* **52**, 108-118.

Allende, J. E. & Allende, C. C. (1995). Protein kinase CK2: An enzyme with multiple substrates and a puzzling regulation. *FASEB J* **9**, 313-323.

Atwater, J. A., Wilson, R. & Verma, I. M. (1990). Regulated mRNA stability. *Annual Review Genetics* **24**, 519-541.

Bachvarova, R. F. (1992). A maternal tail of poly(A). The long and short of it. *Cell* **69**, 895-897.

Bagga, P. S., Arhin, G. K. & Wilusz, J. (1998). DSEF-1 is a member of the hnRNP H family of RNA-binding proteins and stimulates pre-mRNA cleavage and polyadenylation in vitro. *Nucleic Acids Research* **26**, 5343-50.

Bartel, P. L., Chien, C.-T., Sternglanz, R. & Fields, S. (1993). Using the two-hybrid system to detect protein-protein interactions. In *Cellular Interactions in Development: A Practical Approach*, pp. 153-179. Edited by D. A. Hartley. Oxford: Oxford University Press.

Beier, D. R. & Young, E. T. (1982). Characterization of a regulatory region upstream of the ADR2 locus of S. cerevisiae. *Nature* **300**, 724-728.

Bello, L. J., Davison, A. J., Glenn, M. A., Whitehouse, A., Rethmeier, N., Schulz, T. F. & Clements, J. B. (1999). The human herpesvirus-8 ORF 57 gene and its properties. *Journal of General Virology* **80**, 3207-3215.

Blaho, J., Mitchell, C. & Roizman, B. (1993). Guanylylation and adenylation of the α regulatory proteins of herpes simplex virus require a viral β or γ function. *Journal of Virology* **67**, 3891-3900.

Block, T. M., Deshmane, S., Masonis, J., Maggioncaldo, J., Valyi-Nagi, T. & Fraser, N. (1993). A herpes simplex virus type-1 latency-associated transcript mutant reactivates with normal kinetics from latent infection. *Journal of Virology* **64**, 3714-3426.

Block, T. M. & Hill, J. M. (1997). The latency associated transcripts (LAT) of herpes simplex virus: still no end in sight. *Journal of Neurovirology* **3**, 313-321.

Bogerd, H. P., Fridell, R. A., Madore, S. & Cullen, B. R. (1995). Identification of a novel cellular cofactor for the Rev/Rex class of retroviral regulatory proteins. *Cell* **82**, 485-494.

Bomsztyk, K., Seuningen, I. V., Suzuki, H., Denisenko, O. & Ostrowski, J. (1997). Diverse molecular interactions of the hnRNP K protein. *FEBS Letters* **403**, 113-115.

Boyle, S. M., Ruvolo, V., Gupta, A. K. & Swaminathan, S. (1999). Association with the cellular export receptor CRM1 mediates function and intracellular localisation of Epstein-Barr virus SM protein, a regulator of gene expression. *Journal of Virology* **73**, 6872-6881.

Brent, R. & Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**, 729-736.

Bridge, E., Carmo-Fonseca, M., Lamond, A. I. & Petersson, U. (1993). Nuclear organisation of splicing snRNPs in adenovirus-infected cells. *Journal of Virology* **67**, 5792-5802.

Brosi, R., Groning, K., Behrens, S.-E., Luhrmann, R. & Kramer, A. (1993a). Interaction of mammalian splicing factor SF3a with U2 snRNP and relationship of its 60-kD subunit to yeast PRP9. *Science* **262**, 102-105.

Brosi, R., Hauri, H. P. & Kramer, A. (1993b). Separation of splicing factor SF3 into two compartments and purification of SF3a activity. *Journal of Biological Chemistry* **1993**, 17640-17646.

Brown, C. R., Nakamura, M. S., Mosca, J., Hayward, G. S., Straus, S. E. & Perera, L. P. (1995). Herpes simplex virus trans-regulatory protein ICP27 stabilizes and binds to 3' ends of labile mRNA. *Journal of Virology* **69**, 7187-7195.

Bruhn, L., Munnerlyn, A. & Grosschedl, R. (1997). ALY, a context-dependent coactivator of LEF-1 and AML-1 is required for TCR alpha enhancer function. *Genes and Development* **11**, 640-653.

Bruni, R. & Roizman, B. (1996). Open reading frame P - a herpes simplex virus gene repressed during productive infection encodes a protein that binds a splicing factor and reduces synthesis of viral proteins made from spliced mRNA. *Proceedings of National Academy of Sciences. USA* **93**, 10423-10427.

Buchenau, P., Saumweber, H. & Arndt-Jovin, D. J. (1997). The dynamic nuclear redistribution of an hnRNP K-homologous protein during Drosphila embryo development and heat shock. Flexibility of transcription sites in vivo. *Journal of Cell Biology* **137**, 291-303.

Buisson, M., Hans, F., Kusters, I., Duran, N. & Sergeant, A. (1999). The C-terminal region but not the arg-x-pro repeat of Epstein Barr virus is required for its effect on RNA splicing and transport. *Journal of Virology* **73**, 4090-4100.

Bustello, X. R., Suen, K., Michael, W. M., Dreyfuss, G. & Barbacid, M. (1995). Association of the *vav* proto-oncogene product with poly(rC)- specific RNA-binding proteins. *Molecular Cell Biology* **15**, 1324-1332.

Byrappa, S., Pan, Y. B. & Gupta, K. C. (1996). Sendai virus P protein is constitutively phosphorylated at serine 249: High phosphorylation potential of the P protein. *Virology* **216**, 228-234.

Caceres, J. F., Screaton, G. R. & Krainer, A. R. (1998). A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes and Development* **12**, 55-66.

Cai, W., Astor, T. L., Liptak, M., Cho, C., Coen, D. M. & Schaffer, P. A. (1993). The herpes simplex virus type 1 ICP0 enhances virus replication during acute infection and reactivation from latency. *Journal of Virology* **67**, 7501-7512.

Campbell, M. E. M., Palfreyman, J. W. & Preston, C. M. (1984). Identification of herpes simplex virus sequences which encode a trans-acting polypeptide responsible for stimulation of immediate-early gene transcription. *Journal of Molecular Biology* **180**, 1-19.

Cao, W., Jamison, S. F. & Garcia-Blanco, M. A. (1997). Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing *in vitro*. *RNA* **3**, 1456-1467.

Challberg, M. D. (1991). Herpes simplex virus DNA replication. *Seminars in Virology* **2**, 247-256.

Champion-Arnaud, P. & Reed, R. (1996). The prespliceosome components SAP49 and SAP145 interact in a complex implicated in tethering U2 snRNP to the branch site. *Genes and Development* **8**, 1974-1983.

Chang, D. D. & Sharp, P. A. (1990). Regulation of HIV depends upon recognition of splice sites. *Cell* **59**, 789-795.

Chen, F., MacDonald, C. & Wilusz, J. (1995). Cleavage site determinants in a mammalian polyadenylation signal. *Nucleic Acids Research* **23**, 2614-2620.

Chen, J., Zhu, X. & Silverstein, S. (1991). Mutational analysis of the sequence encoding ICP0 from herpes simplex virus type 1. *Virology* **180**, 207-220.

Choo, Y. & Klug, A. (1997). Physical basis of a protein-DNA recognition code. *Current Opinion in Structural Biology* **7**, 117-125.

Ciufo, D. M., Mullen, M.-A. & Hayward, G. S. (1994). Identification of a dimizeration domain in the C-terminal segment of the IE110 trans-activator protein from herpes simplex virus. *Journal of Virology* **68**, 3267-3282.

Clements, J., Watson, R. & Wilkie, N. (1977). Temporal regulation of herpes simplex virus type 1 transcription : location of transcripts on the viral genome. *Cell* **12**, 275-285.

Colas, P., Cohen, B., Jessen, T., Grishina, I., McCoy, J. & Brent, R. (1996). Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2. *Nature* **380**, 548-550.

Collier, B., Goobar-Larsson, L., Sokolowski, M. & Schwartz, S. (1998). Translational inhibition in vitro of human papillomavirus type 16 L2 mRNA mediated through interaction with herogenous ribonucleoprotein K and poly(rC)-binding proteins 1 and 2. *Journal of Biological Chemistry* **273**, 22648-22656.

Constanzo, F., Campadelli-Fiume, G., Foa-Thomas, L. & Cassai, E. (1977). Evidence that herpes simplex virus DNA is transcribed by cellular RNA polymerase II. *Journal of Virology* **21**, 996-1001.

Cooper, M., Goodwin, D., Hall, K., Stevenson, A., Meredith, D., Markham, A. & Whitehouse, A. (1999). The gene product encoded by ORF57 of herpesvirus saimiri regulates the redistribution of the splicing factor SC-35. *Journal of General Virology* **80**, 1311-1316.

Curtin, K. D. & Knipe, D. M. (1993). Altered properties of the herpes simplex virus ICP8 binding protein in cells infected with ICP27 mutant viruses. *Virology* **196**, 1-14.

Dargan, D. J., Patel, A. H. & Subak-Sharpe, J. H. (1995). PREPS - Herpes simplex virus type 1 specific particles produced by infected cells when viral-DNA replication is blocked. *Journal of Virology* **69**, 4924-4932.

Davies, R. C., Calvio, C., Bratt, E., Larsonn, S. H., Lamond, A. I. & Hastie, N. D. (1999). WT1 interacts with the splicing factor U2AF65 in an isoform-dependent manner and can be incorporated into spliceosomes. *Genes and Development* **20**, 3217-3225.

Davis, L. I. (1995). The nuclear pore complex. *Annual Review Biochemistry* **64**, 865-896.

Davison, A. J. & Wilkie, N. M. (1983). Nucleotide sequences of the join between the L and S segments of herpes simplex virus type 1 and 2. *Journal of General Virology* **55**, 315-331.

Defechereux, P., Debrus, S., Baudoux, L., Rentier, B. & Piette, J. (1997). Varicella-zoster virus open reading frame 4 encodes an immediate early protein with posttranscriptional regulatory properties. *Journal of Virology* **71**, 7073-7079.

Dejgaard, K., Leffers, H., Rasmussen, H. H., Madsen, P., Kruse, T. A., Gesser, B., Nilesen, H. & Celis, J. E. (1994). Identification, molecular cloning, expression and chromosome mapping of a family of tranformation upregulated hnRNP-K

proteins derived by altnerative splicing. *Journal of Molecular Biology* **236**, 33-48.

Delius, H. & Clements, J. B. (1976). A partial denaturation map of herpes simplex virus type 1 DNA: evidence for inversions of the unique DNA regions. *Journal of General Virology* **33**, 125-133.

Denisenko, O. N., O'Neill, B., Ostrowski, J., Seuningen, I. V. & Bomsztyk, K. (1996). Zik1, a transcriptional repressor that interacts with the heterogenous nuclear ribonucleoprotein particle K protein. *Journal of Biological Chemistry* **271**, 27701-27706.

Deshmane, S. L., Nicosia, M., Valyi-Nagy, T., Feldman, L. T., Dillner, A. & Fraser, N. W. (1993). An HSV-1 mutant lacking the LAT TATA element reactivates normally in explant cocultivation. *Virology* **196**, 868-872.

Dixon, R. & Schaffer, P. (1980). Fine structure mapping and functional analysis of temperature sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early protein VP175. *Journal of Virology* **36**, 189-203.

Dobbelstein, M., Roth, J., Kimberly, W. T., Levine, A. J. & Shenk, T. (1997). Nuclear export of the E1B-kDa and E4 34-kDa adenoviral oncoproteins mediated by a Rev-like signal sequence. *EMBO J* **16**, 4276-4284.

Dobson, A. T., Sederati, F., Devi-Rao, G., Flanagan, J., Farrell, M. J., Stevens, J. G., Wagner, E. K. & Feldman, L. T. (1989). Identification of the latency-associated transcript promoter by expression of rabbit beta-globin mRNA in mouse sensory nerve ganglia latently infected with a recombinant herpes simplex virus. *Journal of Virology* **63**, 3844-3851.

Eggleton, P., Ghebrehiwet, B., Sastry, K. N., Coburn, J. P., Zaner, K. S., Reid, K. B. & Tauber, A. I. (1995). Identification of a gC1q-binding protein (gc1q-R) on

the surface of human neutrophils. Subcellular localization and binding properties in comparison with the cC1q-R. *Journal of Clinical Investigation* **95**, 1569-1578.

Estojak, J., Brent, R. & Golemis, E. (1995). Correlation of two-hybrid affinity data with in vitro measurements. *Mollecular and Cellular Biology* **15**, 5820-5829.

Everett, R. D. (1984). Trans activation of transcription by herpes virus products: requirement for two HSV-1 immediate-early polypeptides for maximum activity. *EMBO J* **13**, 3135-3141.

Fearon, E. R., Finkel, T., Gillison, M. L., Kennedy, S. P., Casella, I., Tomaselli, G. F., Morrow, J. S. & Van, D. C. (1992). Karyoplasmic interaction selection strategy: A general strategy to detect protein-protein interactions in animal cells. *Proceedings of the National Academy of Science. USA.* **89**, 7958-7962.

Fields, S. & Song, O. (1989). A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246.

Fischer, U., Darzynkiewicz, E., Tahara, S. M., Dathan, N. A., Luhrmann, R. & Mattaj, I. W. (1991). Diversity in the signals required for nuclear accumulation of U snRNPs and vareity in the pathways of nuclear transport. *Journal of Cell Biology* **113**, 705-714.

Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W. & Luhrmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475-483.

Flint, J. & Shenk, T. (1997). Viral transactivating proteins. *Annual Review of Genetics* **31**, 177-212.

Fortes, P., Lamond, A. I. & Ortin, J. (1995). Influenza virus NS1 protein alters the subnuclear localisation of cellular splicing components. *Journal of General Virology* **76**, 1001-1007.

Fridell, R. A., Fischer, U., Luhrmann, R., Meyer, B. E., Meinkoth, J. L., Malim, M. H. & Cullen, B. R. (1996). Amphibian transcription factor IIIA proteins contain a sequence element functionally equivalent to the nuclear export signal of human immunodeficiency virus type 1 Rev. *Proceedings of the National Academy of Science USA* **93**, 2936-2940.

Frink, R. J., Eisenberg, R., Cohen, G. & Wagner, E. K. (1983). Detailed analysis of the portion of the herpes simplex virus type 1 genome encoding glycoprotein C. *Journal of Virology* **45**, 634-647.

Fritz, C. C., Zapp, M. L. & Green, M. R. (1995). A human nucleoporin-like protein that specifically interacs with HIV Rev. *Nature* **376**, 530-533.

Fu, X. D. (1995). The superfamily of arginine/serine-rich splicing factors. *RNA* **1**, 663-680.

Fuller, A. O. & Lee, W. C. (1992). Herpes simplex virus type-1 entry through a cascade of virus-cell interactions requires different roles of gD and gH in penetration. *Journal of Virology* **66**, 5002-5012.

Gaffney, D. F., McLauchlan, J., Whitton, J. L. & Clements, J. B. (1985). A modular system for the assay of transcription regulatory signals: the sequence TAATGARAT is required for herpes simplex virus immediate-early gene activation. *Nucleic Acids Research* **13**, 7874-7882.

Gaillard, C., Weber, M. & Strauss, F. (1988). A sequence specific single-strand binding protein for the late coding strand of the simian virus-40 control region. *Molecular Cell Biology* **62**, 2380-2385.

Gallinari, P., Wiebauer, K., Nardi, M. C. & Jiricny, J. (1994). Localization of a 34-amino-acid segment implicated in dimerization of the herpes simplex virus type 1 ICP4 polypeptide by a dimerization trap. *Journal of Virology* **68**, 3809-3820.

Ghebrehiwet, B., Lu, P. D., Zhang, W., Keilbaugh, S. A., Leigh, L. E., Eggleton, P., Reid, K. B. & Peerschke, E. I. (1997). Evidence that the two C1q binding membrane proteins, gC1q-R and cC1q-R, associate to form a complex. *Journal of Immunology* **159**, 1429-1436.

Gozani, O., Feld, R. & Reed, R. (1996). Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of splicesomal complex A. *Genes and Development* **8**, 1974-1983.

Guerra, B. & Issinger, O. (1999). Protein Kinase CK2 and its role in cellular proliferation, development and pathology. *Electrophoresis* **20**, 391-408.

Guo, N. (1997). Assignment of C1QBP encoding the C1q globular domain binding protein (gC1q-R) to human chrosome 17 band p13.3 by *in situ* hybridization. *Cytogenetics and Cell Genetics* **77**, 283-284.

Gupta, A. K., Ruvolo, V., Patterson, C. & Swaminathan, S. (2000). The human herpesvirus 8 homologue of Epstein-Barr virus SM protein is a posttranscriptional activator of gene expression. *Journal of Virology* **74**, 1038-1044.

Hamm, J. & Mattaj, I. W. (1990). Monomethylated cap structures facilitate RNA export from the nucleus. *Cell* **63**, 109-118.

Hardwicke, M. A. & Sandri-Goldin, R. M. (1994). The herpes simpex virus regulatory protein ICP27 contributes to the decrease in cellular mRNA levels during infection. *Journal of Virology* **68**, 4797-4810.

Hardy, W. R. & Sandri-Goldin, R. M. (1994). Herpes simplex virus inhibits host cell splicing, and regulatory protein ICP27 is required for this effect. *Journal of Virology* **68**, 7790-7799.

Harris, R. A., Everett, R. D., Zhu, X., Silverstein, S. & Preston, C. M. (1989). Herpes simplex virus type 1 immediate-early protein Vmw110 reactivates latent herpes simplex virus type 2 in an in vitro latency system. *Journal of Virology* **63**, 3513-3515.

Hayward, S. D., Jacob, R. J., Wadsworth, S. C. & Roizman, B. (1975). Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components. *Proceedings of the National Academy of Science, USA* **72**, 4243-4247.

Hermann, H., Frabrizio, P., Raker, V. A., Foulaki, K., Hornig, H., Brahms, H. & Luhrmann, R. (1995). snRNP Sm proteins share two evolutionary conserved sequence motifs which are involved in Sm protein-protein interactions. *EMBO J* **14**, 2076-2088.

Hibbard, M. & Sandri-Goldin, R. M. (1995). Arginine-rich regions succeeding the nuclear localization region of herpes simplex virus type 1 regulator protein ICP27 are required for efficient nuclear localization and late gene expression. *Journal of Virology* **69**, 4656-4667.

Hill, J. M., Sedarati, F., Javier, R. T., Wagner, E. K. & Stevens, J. G. (1990). Herpes simplex virus latent phase transcription facilitates *in vivo* reactivation. *Virology* **174**, 117-125.

Hirose, Y. & Manley, J. L. (1998). RNA polymerase II is an essential mRNA polyadenylation factor. *Nature* **395**, 93-96.

Hodges, P. E. & Beggs, J. D. (1994). U2 fills a committment. *Current Biology* **4**, 264-267.

Honess, R. W. & Roizman, B. (1974). Regulation of herpesvirus macromolecular synthesis. Cascade regulation of the synthesis of three groups of viral proteins. *Journal of Virology* **14**, 8-19.

Hope, I. A. & Struhl, K. (1986). Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. *Cell* **46**, 885-894.

Hope, T. J., Klein, N. P., Elder, M. E. & Parslow, T. G. (1992). Trans-dominant inhibition of human immunodeficiency virus type 1 Rev occurs through formation of inactive protein complexes. *Journal of Virology* **66**, 1849-1855.

Hsieh, T., Matsumoto, M., Chou, H., Schneider, R., Hwang, S., Lee, A. & Lai, M. (1998). Hepatitis C virus core protein interacts with heterogenous ribonucleoprotein K. *Journal of Biological Chemistry* **273**, 17651-17659.

Huang, Y., Wimler, K. M. & Carmichael, G. G. (1999). Intronless mRNA transport elements may affect multiple sites of pre-mRNA processing. *EMBO J* **18**, 1642-1652.

Ingram, A., Phelan, A., Dunlop, J. & Clements, J. B. (1996). Immediate early protein IE63 of herpes simplex virus type 1 binds RNA directly. *Journal of General Virology* **77**.

Jackson, R. J. & Standart, N. (1990). Do the poly(A) tail and 3' untranslated region control mRNA translation. *Cell* **62**, 15-24.

Jang, K. & Latchman, D. (1992). The herpes simplex virus immediate early protein ICP27 stimulates the transcription of cellular Alu repeated sequences by

increasing the activity of transcription factor TFIIIC. *Biochemical Journal* **284**, 667-673.

Javier, R. T., Stevens, J. G., Dissette, V. B. & Wagner, E. K. (1988). A herpes simplex virus transcript abundant in latently infected neurons is dispensable for establishment of the latent state. *Virology* **166**, 254-257.

Kalderon, D., Richardson, W. D., Markham, A. F. & Smith, A. E. (1984). Sequence requirements for nuclear localisation of SV40 large T antigen. *Nature* **311**, 33-38.

Kalland, K. H., Szilvay, M., Brokstad, K. A., Saetrevik, W. & Haukenes, G. (1994). The human immunodeficiency virus type 1 Rev protein shuttles between the cytoplasm and nuclear compartments. *Molecular Cell Biology* **14**, 7436-7444.

Kanopka, A., Muhlemann, O., Petersen-Mahrt, S., Estmer, C., Ohrmalm, C. & Akusjarvi, G. (1998). Regulation of adenovirus alternative RNA splicing by dephosphorylation of SR proteins. *Nature* **393**, 185-187.

Kim, F. J., Beeche, A. A., Hunter, J. J., Chin, D. J. & Hope, T. J. (1996). Characterization of the nuclear export signal of human T-cell lymphotrophic virus type 1 Rex reveals that nuclear export is mediated by position-variable hydrophobic interactions. *Molecular Cell Biology* **1996**, 5147-5155.

Kinchington, P. R., Bookey, D. & Turse, S. E. (1995). The transcriptional regulatory proteins encoded by varicella-zoster virus open reading frames (ORFs) 4 and 63 but not ORF61 are associated with purified virus particles. *Journal of Virology* **69**, 4274-4282.

Krainer, A. R., Conway, G. & Kozak, D. (1990). Purification and characterisation of SF2, a human pre-mRNA splicing factor. *Genes and Development* **4**, 1158-1171.

Krainer, A. R., Mayeda, A., Kozak, D. & Binns, G. (1991). Functional expression of cloned human splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell* **62**, 35-42.

Kramer, A. (1996). The structure and function of proteins involved in mammalian pre-mRNA splicing. *Annual Review of Biochemistry* **65**, 367-409.

Kristensson, K., Lycke, E., Ryotta, M., Svennerholm, B. & Vahlne, A. (1986). Neuritic transport of herpes simplex virus in rat sensory neurons in vitro. Effects of substances interacting with microtubular function and axonal flow. *Journal of General Virology* **67**, 2023-2028.

Ladomery, M. (1997). Multifunctional proteins suggest connections between transcriptional and post-transcriptional processes. *Bioessays* **19**, 903-909.

Leib, D. A., Bogard, C. L., Kosz-Vnenchak, M., Hicks, K. A., Coen, D. M., Knipe, D. M. & Schaffer, P. A. (1989). A deletion mutant of the latency associated transcript of herpes simplex virus type-1 reactivates from the latent infection. *Journal of Virology* **63**, 2893-2900.

Lieb, D. A., Coen, D. M., Bogard, C. L., Hicks, K. A., Yager, D. R., Knipe, D. M., Tyler, K. L. & Schaffer, P. A. (1989). Immediate-early regulatory gene mutants define different stages in the establishment and reactivation of herpes simplex virus latency. *Journal of Virology* **63**, 759-768.

Lin, A., Frost, J., Deng, T., Smeal, T., al-Alawi, N., Kikkawa, U., Hunter, T., Brenner, D. & Karin, M. (1992). Casein kinase-II is a negative regulator of C-Jun DNA-binding and AP-1 activity. *Cell* **70**, 777-789.

Liu, Q. & Dreyfuss, G. (1995). In vivo and in vitro arginine methylation of RNA-binding proteins. *Molecular Cell Biology* **15**, 2800-2808.

Luhrman, R., Kastner, B. & Bach, M. (1990). Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. *Biochim Biophys Acta* **1087**, 265-292.

Luo, Y., Yu, H. & Peterlin, B. M. (1994). Cellular protein modulates effects of human immunodeficiency virus type 1 Rev *Journal of Virology* **68**, 3850-3856.

Ma, J. & Ptashne, M. (1988). Converting a eukaryotic transcriptional inhibitor into an activator. *Cell* **55**, 443-446.

MacDonald, C., Wilusz, J. & Shenk, T. (1994). The 64 kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNA downstream of the cleavage site and influences cleavage site location. *Molecular and Cellular Biology* **14**, 6647-6654.

Mackay, J. P. & Crossley, M. (1998). Zinc fingers are sticking together. *Trends in Biochemical Science* **23**, 1-4.

Manley, J. & Tacke, R. (1996). SR proteins and splicing control. *Genes and Development* **10**, 1569-1579.

Marais, R. M., Hsuan, J., McGuigan, C., Wynne, J. & Treisman, R. (1992). Casein Kinase-II phosphorylation increases the rate of serum response factor-binding site exchange. *EMBO J* **11**, 97-105.

Martin, T. E., Barghusen, S. C., Leser, G. P. & Spear, P. G. (1987). Redistribution of the nuclear ribonucleoprotein antigens during herpes simplex virus infection. *Journal of Cell Biology* **105**.

Massimi, P., Pim, D., Storey, A. & Banks, I. (1996). HPV-16 E7 and adenovirus E1a complex formation with a TATA box binding protein is enhanced by casein kinase II phosphorylation. *Oncogene* **12**, 2325-2330.

Matthews, D. A. & Russell, W. C. (1998). Adenovirus core protein V interacts with p32 - a protein which is associated with both the mitochondria and the nucleus. *Journal of General Virology* **79**, 1677-1685.

Matunis, M. J., Michael, W. M. & Dreyfuss, G. (1992). Characterization and primary structure of the poly(c)-binding heterogenous nuclear ribonucleoprotein complex K-protein. *Molecular Cell Biology* **12**, 164-171.

McCarthy, A. M., McMahan, L. & Schaffer, P. A. (1989). Herpes simplex virus type 1 ICP27 deletion mutants exhibit altered patterns of transcription and are DNA defficient. *Journal of Virology* **63**, 18-27.

McGeoch, D. J., Cook, S., Dolan, A., Jamieson, F. E. & Telford, E. A. R. (1995). Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *Journal of Molecular Biology* **257**, 443-458.

McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., McNab, D., Perry, L. J., Scott, J. E. & Taylor, P. (1988). The complete DNA sequence of the long unique region in the genome of HSV-1. *Journal of General Virology* **69**, 1531-1574.

McGregor, F., Phelan, A., Dunlop, J. & Clements, J. B. (1996). Regulation of herpes simplex virus type 1 poly(A) site usage and the action of the immediate-early protein IE63 in the early-late switch. *Journal of Virology* **70**, 1931-1940.

McLauchlan, J., Simpson, S. & Clements, J. B. (1989). Herpes simplex virus induces a processing factor that stimulates poly(A) site usage. *Cell* **59**, 1093-1105.

Mears, W. E., Lam, V. & Rice, S. A. (1995). identification of nuclear and nucleolar localization signals in the herpes simplex virus regulatory protein ICP27. *Journal of Virology* **69**, 935-947.

Mears, W. E. & Rice, S. A. (1996). The RGG motif of the herpes simplex virus ICP27 protein mediates an RNA-binding activity and determines *in vivo* methylation. *Journal of Virology* **70**, 7445-7453.

Mears, W. E. & Rice, S. A. (1998). The herpes simplex virus immediate-early protein ICP27 shuttles between nucleus and cytoplasm. *Virology* **242**, 128-137.

Meggio, F., Dagostino, D. M., Ciminale, V., Bianchi, L. C. & Pinna, L. A. (1996). Phosphorylation of HIV-1 Rev protein: Implication of protein kinase CK2 and pro-directed kinases. *Biochemical and Biophysical Research Communications* **226**, 547-554.

Mellerick, D. M. & Fraser, N. W. (1987). Physical state of the latent herpes simplex virus genome in a mouse model system : evidence suggesting an episomal state. *Virology* **158**, 265-275.

Mendelsohn, A. & Brent, R. (1994). Applications of interaction traps/two-hybrid systems to biotechnology research. *Current Opinion in Biotechnology* **5**, 482-486.

Mendelsohn, A. & Brent, R. (1999). Protein interaction methods - towards an endgame. *Science* **284**, 1948-1950.

Mermoud, J. E., Cohen, P. T. W. & Lamond, A. I. (1994). Regulation of mammalian spliceosome assembly by protein phoshporylation mechanism. *EMBO J* **13**, 5679-5688.

Meyer, B. E. & Halim, M. H. (1994). The HIV-1 Rev *trans*-activator shuttles between the nucleus and the cytoplasm. *Genes and Development* **8**, 1538-1547.

Meyer, B. E., Meinkoth, J. L. & Malim, M. H. (1996). Nuclear transport of human immunodeficieny virus type 1, visna virus, and equine infectious agent virus

Rev proteins: Identification of a family of nuclear export signals. *Journal of Virology* **70**, 2350-2359.

Miau, L., Chang, C., Shen, B., Tsai, W. & Lee, S. (1998). Identification of heterogenous nuclear ribonucleoprotein K (hnRNP K) as a repressor of C/EBPβ -mediated gene activation. *Journal of Biological Chemistry* **273**, 10784-10791.

Michael, N., Spector, D., Mavromara-Nazos, P., Kristie, T. M. & Roizman, B. (1988). The DNA-binding properties of the major regulatory protein α4 of herpes simplex virus. *Science* **239**, 1531-1534.

Michael, W. M., Eder, P. S. & Dreyfuss, G. (1997). The K nuclear shuttling domain: A novel signal for nuclear import and nuclear export in the hnRNP K protein. *EMBO J* **16**, 3587-3598.

Michelotti, E. F., Michelotti, G. A., Aronsohn, A. I. & Levens, D. (1996). Multiple single-stranded cis elements are associated with activated chromatin of the human c-myc gene in vivo. *Moelcular Cell Biology* **16**, 2350-2360.

Misteli, T. & Spector, D. L. (1997). Protein phosphorylation and the nuclear organisation of pre-mRNA splicing. *Trends in Cell Biology* **7**, 135-138.

Mitchell, C., Blaho, J., McCormick, L. & Roizman, B. (1997). The nucleotidylation of herpes simplex virus I regulatory protein α22 by human casein kinase II. *Journal of Biological Chemistry* **272**, 25394-25398.

Muller, M. T. (1987). Binding of the herpes simplex virus immediate-early gene product ICP4 to its own transcriptional start site. *Journal of Virology* **61**, 858-865.

Murphy, R. & Wente, S. R. (1996). An RNA-export mediator with an essential nuclear export signal. *Nature* **383**, 357-360.

Muta, T., Kang, D., Kitajima, S., Fujiwara, T. & Hamaski, N. (1997). p32 protein, a splicing factor 2-associated protein, is localized in mitochondrial matrix and is functionally important in maintaining oxidative phosphorylation. *Journal of Biological Chemistry* **272**, 24363-24370.

Neubauer, G., King, A., Rappsilber, J., Calvio, C., Watson, M., Ajuh, P., Sleeman, J., Lamond, A. & Mann, M. (1998). Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nature Genetics* **20**, 46-50.

Neville, M., Stutz, F., Lee, L., Davis, l. & Rosbash, M. (1997). The importin-beta family member Crm 1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Current Biology* **7**, 767-775.

Nicholades, N., Kinzer, K., Vogelstein, B. (1995). Analysis of the 5' region of PMS2 reveals heterogenous transcripts and a novel overlapping gene. *Genomics* 29, 329-334.

O'Hare, P. (1993). The virion transcription activator of herpes simplex virus. *Seminars in Virology* **4**, 145-155.

O'Hare, P. & Hayward, G. S. (1985). Evidence for a direct role for both the 175,000- and 110,000-molecular-weight immediate-early proteins of herpes simplex virus in the transactivation of delayed-early promoters. *Journal of Virology* **53**, 751-760.

Ohtsuki, K., Maekawa, T., Harada, S., Karino, A., Morikawa, Y. & Ito, M. (1998). Biochemical characterization of HIV-1 Rev as a potent activator of casein kinase II in vitro. *FEBS Letters* **428**, 235-240.

O'Neill, R. E., Talon, J. & Palese, P. (1998). The influenza virus NEP (NS2 protein) mediates the nuclear export of viral nucleoproteins. *EMBO J* **17**, 288-296.

Ostareck, D. H., Ostareck-Lederer, A., Wilm, M., Thiele, B. J., Mann, M. & Hentze, M. W. (1997). mRNA silencing in erythroid differentiation: hnRNP K and hnRNP E1 regulate 15-Lipoxygenase Translation from the 3' end. *Cell* **89**, 597-606.

Ostrowski, J., Seuningen, I. V., Seger, R., Rouch, C. T., Sleath, P. R., McMullen, B. A. & Bomsztyk, K. (1994). Purification, cloning, and expression of a murine phosphoprotein that binds the kappa-B motif in vitro identifies it as the homolog of the human heterogenous nuclear ribonucleoprotein K-protein - description of a novel DNA-dependent phosphorylation process. *Journal of Biological Chemistry* **269**, 17626-17634.

Ostrowski, J., Sims, J. E., Sibley, C. H., Valentine, M. A., Dower, S. K., Meier, K. E. & Bomsztyk, K. (1991). A serine threonine kinase acitvity is closely associated with a 65 kDa phosphoprotein specifically recognized by the kappa-B enhancer element. *Journal of Biological Chemistry* **266**, 12722-12733.

Palmer, D. & Malim, M. H. (1996). The human T-cell leukemia virus type 1 pottranscriptional trans-activator Rex contains a nuclear export signal. *Journal of Virology* **70**, 6442-6445.

Pangiotidis, C. A., Lium, E. K. & Silverstein, S. J. (1997). Physical and functional interactions between herpes simplex virus immediate-early proteins ICP4 and ICP27. *Journal of Virology* **71**, 1547-1557.

Pereira, L., Wolff, M. H., Fenwick, M. & Roizman, B. (1977). Regulation of herpesvirus macromolecular synthesis. V. Properties of alpha polypeptides made in HSV-1 and HSV-2 infected cells. *Virology* **77**, 733-749.

Petersen-Mahrt, S. K., Estmer, C., Ohrmalm, C., Matthews, D. A., Russell, W. C. & Akusjarvi, G. (1999). The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. *EMBO J* **18**, 1014-1024.

Peterson, K. L., Zhang, W., Lu, P. D., Keilbaugh, S. A., Peerschke, E. I. & Ghebrehiwet, B. (1997). The C1q-binding cell membrane proteins cC1q-R and gC1q-R are released from activated cells: subcellular distribution and immunochemical characterisation. *Clinical Immunology and Immunopathology* **84**, 17-26.

Phelan, A., Carmo-Fonseca, M., McLauchlan, J., Lamond, A. I. & Clements, J. B. (1993). A herpes simplex virus type 1 immediate-early gene product, IE63, regulates small nuclear ribonucleoprotein distribution. *Proceedings of National Academy of Science. USA* **90**, 9056-9060.

Phelan, A. & Clements, J. B. (1997). Herpes simplex virus type 1 immediate early protein IE63 shuttles between nuclear compartments and the cytoplasm. *Journal of General Virology.*

Phelan, A., Dunlop, J. & Clements, J. B. (1996). Herpes simplex virus type 1 protein IE63 affects the nuclear export of virus intron-containing transcripts. *Journal of Virology* **70**, 5255-5265.

Phillips, A. C. & Vousden, K. H. (1997). Analysis of the interaction between human papillomavirus type 16 E7 and the TATA binding protein TBP. *Journal of General Virology* **78**, 905-909.

Picard, D., Khursheed, B., Garabedian, M. J., Fortin, M. G., Lindquist, S. & Yamamoto, K. R. (1990). Reduced levels of HSP90 compromise steroid-receptor action in vivo. *Nature* **313**, 166-168.

Post, L. E. & Roizman, B. (1981). A generalized technique for deletion of specific genes in large genomes: alpha gene 22 of herpes simplex virus 1 is not essential for growth. *Cell* **1981**, 227-32.

Preston, C. M., Frame, M. C. & Campbell, M. E. M. (1988). A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA-sequence. *Cell* **425-434**.

Pugh, B. F. & Tijian, B. (1990). Mechanism of transcriptional activation by SP1 - evidence for coactivators. *Cell* **61**, 777-789.

Purves, F. C., Ogle, W. O. & Roizman, B. (1993). Processing of the herpes virus regulatory protein alpha 22 mediated by the UL13 protein kinase determines the accumulation of a subset of alpha and gamma mRNAs and proteins in infected cells. *Proceedings of the National Academy of Sciences USA.* **90**, 6701-6705.

Quinlan, M. P. & Knipe, D. M. (1985). Stimulation of expression of a herpes simplex virus DNA-binding protein by two viral functions. *Molecular Cell Biology* **5**, 957-963.

Quinn, J. P., McGregor, R. A., Fiskerstand, C. E., Davey, C., Allan, J. & Dalzel, R. G. (1998). Identification of a novel multifunctional structural domain in herpes simplex virus 1 genome: implications for viral latency. *Journal of General Virology* **79**, 2529-2532.

Rice, S. A. & Knipe, D. M. (1988). Gene-specific transactivation by herpes simplex virus type 1 alpha protein ICP27. *Journal of Virology* **62**, 3814-3823.

Rice, S. A. & Knipe, D. M. (1990). Genetic evidence for two distinct transactivation functions of the herpes simplex virus α protein ICP27. *Journal of Virology* **64**, 1704-1715.

Rice, S. A. & Lam, V. (1994). Amino acid substitution mutations in the herpes simplex virus ICP27 protein define an essential gene regulation function. *Journal of Virology* **68**, 823-833.

Rice, S. A., Lam, V. & Knipe, D. M. (1993). The acidic amino-terminal region of herpes simplex virus type 1 alpha protein ICP27 is required for an essential lytic function. *Journal of Virology* **67**, 1778-1787.

Rice, S. A., Long, M. C., Lam, V., Schaffer, P. A. & Spencer, C. A. (1995). HSV ICP22 protein is required for viral modification of RNA polymerase II and establishment of the normal viral transcription program. *Journal of Virology* **69**, 5550-5559.

Richard, N., Iacampo, S. & Cochrane, A. (1994). HIV-1 Rev is capable of shuttling between the nucleus and the cytoplasm. *Virology* **204**, 123-131.

Rock, D. (1993). The molecular basis of latent infections by alphaherpes viruses. *Seminars in Virology* **4**, 157-165.

Roizman, B. (1979). The structure and isomerization of herpes simplex virus genomes. *Cell* **16**, 481-494.

Roizman, B. & Sears, A. (1996). Herpes simplex viruses and their replication. In *Virology*, 3 edn, pp. 2231-2296. Edited by B. Fields, D. Knipe & P. Howley. Philadelphia: Raven Press.

Rout, M. P. & Wente, S. R. (1994). Pores for thought: nuclear pore complex proteins. *Trends in Cell Biology* **89**, 715-725.

Saavedra, C. A., Hammell, C. M., Heath, C. V. & Cole, C. N. (1997). Export of heat shock mRNAs following stress in *S. cerevisiae* employs a distinct pathway defined by Rip1p and also requires a subset of factors essential for export of poly(A)+ RNA. *Genes and Development* **11**, 2845-2856.

Sacks, W. R., Greene, C. C., Aschman, D. P. & Schaffer, P. A. (1985). Herpes simplex virus type 1 ICP27 is an essential regulatory protein. *Journal of Virology* **55**, 796-805.

Sacks, W. R. & Schaffer, P. A. (1987). Deletion mutants in the gene encoding the herpes simplex virus type 1 immediate-early protein ICP0 exhibit impaired growth in cell culture. *Journal of Virology* **61**, 829-839.

Sandri-Goldin, R. M. (1998). ICP27 mediates HSV RNA export by shuttling through a leucine-rich nuclear export signal and bindng viral intronless RNAs through an RGG motif. *Genes and Development* **12**, 868-879.

Sandri-Goldin, R. M. & Hibbard, M. K. (1996). The herpes simplex virus type 1 regulatory protein ICP27 coimmunoprecipitates with anti-Sm antiserum, and the C terminus appears to be required for this interaction. *Jornal of Virology* **70**, 108-118.

Sandri-Goldin, R. M., Hibbard, M. K. & Hardwicke, M. A. (1995). The C-terminal repressor region of herpes simplex virus type 1 ICP27 is requred for the redistribution of small nuclear ribonucleoprotein particles and splicing factor SC35; however, these alterations are not sufficient to inhibit host cell splicing. *Journal of Virology* **69**, 6063-6076.

Sanz-Ezquerro, J., Santaren, J. F., Sierra, T., Aragon, T., Ortega, J., Ortin, J., Smith, G. L. & Nieto, A. (1998). The PA influenza virus polymerase subunit is a phosphorylated protein. *Journal of General Virology* **79**, 471-478.

Schullery, D., Ostrowski, J., Denisenko, O., Stempka, L., Shnyreva, M., Suzuki, H., Gschwendt, M. & Bomsztyk, K. (1999). Regulation interaction of protein kinase Cδ with the heterogenous nuclear ribonucleoprotein K protein. *Journal of Biological Chemistry* **274**, 15101-15109.

Schuly, T. F. (1998). Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8). *Journal of General Virology.* **79**, 1573-1591.

Semmes, O., Chen, L., Sarisky, R., Gao, Z., Zhong, L. & Hayward, S. (1998). Mta has properties of an RNA export protein and increases cytoplasmic accumulation of Epstein-Barr virus replication gene mRNA. *Journal of Virology* **72**, 9526-9534.

Seuningen, I. V., Ostrowski, J. & Bomsztyk, K. (1995). Description of an interleukin-1 responsive kinase that phosphorylates the K protein. Enhancement of phosphorylation by selective DNA and RNA motifs. *Biochemistry* **34**, 5644-5650.

Seuningen, I. V., Ostrowski, J., Bustelo, X., Sleath, P. & Bomxztyk, K. (1996). The K protein domain that recruits the interleukin 1-responsive K-protein-kinase lies adjacent to a cluster of c-src and Vav SH3-binding sites - implications that K-protein acts as a docking platform. *Journal of Biological Chemistry* **270**, 26976-26985.

Sheldrick, P. & Berthelot, N. (1974). Inverted reptitions in the chromsome of herpes simplex virus. *Cold Spring Harbour Symposium on Quantitative Biology* **9**, 471-480.

Simos, G. & Georgatos, S. D. (1994). The lamin B receptor-associated protein p34 shares sequence homology and antigenic determinants with the splicing factor 2-associated protein p32. *FEBS Letters* **346**, 225-228.

Singh, M., Fraefel, C., Bello, L. J., Lawrence, W. C. & Schwyzer, M. (1996). Identification and characterization of BICP27, an early protein of bovine herpesvirus 1 which may stimulate mRNA 3' processing. *Journal of General Virology* **77**, 615-625.

Siomi, H., Choi, M., Siomi, M. C., Nussbaum, R. L. & Dreyfuss, G. (1994). Essential role for KH domains in RNA binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes Fragile X syndrome. *Cell* **77**, 33-39.

Siomi, H. & Dreyfuss, G. (1995). A nuclear localization domain in the hnRNP A1 protein. *Journal of Cell Biology* **129**, 551-560.

Soliman, T. M., Sandri-Goldin, R. M. & Silverstein, S. J. (1997). Shuttling of the herpes simplex virus type 1 regulatory protein ICP27 between the nucleus and the cytoplasm mediates the expression of late proteins. *Journal of Virology* **71**, 9188-9197.

Steveley, W. S., Katan, M., Stirling, V., Smith, G. & Leader, D. P. (1985). Protein-kinase activities associated with the virions of pseudorabies and herpes-simplex virus. *Journal of General Virology* **66**, 661-673.

Stow, N. D. & Stow, E. C. (1986). Isolation and characterization of a herpes simplex virus type 1 mutant containing a deletion within the gene encoding the immediate-early polypeptide Vmw110. *Journal of General Virology* **67**, 2571-1585.

Struhl, K. (1991). Acid connections. *Current Biology* **1**, 188-191.

Stutz, F., Kantor, J., Zhang, D., McCarthy, T., Neville, M. & Rosbash, M. (1997). The yeast nucleoporin Rip1p contributes to multiple export pathways with no essential role for its FG repeat. *Genes and Development* **11**, 2587-2868.

Stutz, F. & Robash, M. (1998). Nuclear RNA export. *Genes and Development* **12**, 3303-3319.

Su, L. & Knipe, D. M. (1989). Herpes simplex virus alpha protein ICP27 can inhibit or augment viral gene transactivation. *Virology* **170**, 496-504.

Szilagyi, J. F. & Cunningham, C. (1991). Identification and characterisation of a novel non-infectious herpes simplex virus related particle. *Journal of General Virology* **62**.

Takagaki, Y., Ryner, L. C. & Manley, J. L. (1989). Four factors are required for 3' end cleavage of pre-mRNAs. *Genes and Development* **3**, 1711-1724.

Takimoto, M., Tomonaga, T., Matunis, M., Avigan, M., Krutzsch, H., Dreyfuss, G. & Levens, D. (1993). Specific binding of heterogeneous ribonucleoprotein particle protein-K to the human c-myc promoter, in vitro. *Journal of Biological Chemistry* **268**, 18249-18258.

Tange, T. O., Jensen, T. H. & Kjems, J. (1996). *In vitro* interaction between human immunodeficiency virus type 1 Rev protein and splicing factor ASF/SF2-associated protein, p32. *Journal of Biological Chemistry* **271**, 10066-10072.

Taylor, S. J. & Shalloway, D. (1994). An RNA-binding protein associated with Src through its SH2 and SH3 binding domains in mitosis. *Nature* **368**, 867-870.

Tomonaga, T. & Levens, D. (1995). Heterogeneous nuclear ribonucleoprotein-K is a DNA-binding transactivator. *Journal of Biological Chemistry* **270**, 4875-4881.

Tomonaga, T., Michelotti, G., Libutti, D., Uy, A., Sauer, B. & Levens, D. (1998). Unrestraining genetic processes with a protein-DNA hinge. *Molecular Cell* **1**, 759-764.

Tornow, J. & Santangelo, G. G. (1990). Efficient expression of Saccharomyces cerevisiae glycolytic gene ADH1 is dependent upon a cis-acting regulator element (UAS_{PRG}) found initially in genes encoding ribosomal proteins. *Gene* **90**, 79-85.

Truant, R., Kang, Y. & Cullen, B. R. (1999). The human Tap nuclear RNA export factor contains a novel transportin-dependent nuclear localization signal that

lacks nuclear export signal function. *Journal of Biological Chemistry* **274**, 32167-32171.

Tsieh, T. Y., Matsumoto, M., Chou, H. C., Schneider, R., Hwang, S., Lee, A. & Lai, M. (1998). Hepatitis C Virus core protein interacts with heterogenous nuclear ribonucleoprotein K. *Journal of Biological Chemistry* **273**, 17651-17659.

Ullman, K. S., Powers, M. A. & Forbes, D. J. (1997). Nuclear export receptors from importin to exportin. *Cell*, 967-970.

Uprichard, S. L. & Knipe, D. M. (1996). Herpes simplex ICP27 mutant viruses exhibit reduced expression of specific DNA replication genes. *Journal of Virology* **70**, 1969-1980.

Vancurova, I., JochovaRupes, J., Lou, W. & Paine, P. L. (1995). Distinct phosphorylation sites differentially influence facilitated transport of an NLS-protein and its subsequent intranuclear binding. *Biochemical and Biophysical Research Communications* **217**, 419-427.

Vasavada, H. A., Ganguly, S., Germino, F. J., Wang, Z. X. & Weissman, S. M. (1991). A contingent replication assay for the detection of protein-protein interactions in animal cells. *Proceedings of the National Academy of Science. USA* **88**, 10686-10690.

Vaughan, P. J., Thibault, K. J., Hardwicke, M. A. & Sandri-Goldin, R. M. (1992). The herpes simplex virus immediate early protein ICP27 encodes a potential metal binding domain and binds zinc *in vitro*. *Virology* **189**, 377-384.

Voit, R., Schnapp, A., Kuhn, A., Rosenbauer, H., Hirschmann, P., Stunnenberg, H. G. & Grummt, I. (1992). The nucleolar transcription factor MUBF is phosphorylated by casein kinase-II in the C-terminal hyperacidic tail which is essential for transactivation. *EMBO Journal* **11**, 2211-2218.

Wadd, S., Bryant, H., Fihol, O., Scott, J. E., Hsieh, T.-Y., Everett, R. D. & Clements, J. B. (1999). The multifunctional herpes simplex virus IE63 protein interacts with heterogenous ribonucleoprotein K and with casein kinase 2. *Journal of Biological Chemistry* **274**, 28991-28998.

Wagner, E. K., Guzowski, J. F. & Singh, J. (1995). Transcription of the herpes simplex virus genome during productive and latent infection. In *Progress in Nucleic Acid Research and Molecular Biology*, pp. 123-165. Edited by W. H. Cohn & K. Moldave. Moldave: Academic Press.

Wang, Y. L., Finan, J. E., Middeldorp, J. M. & Hayward, S. D. (1997). p32/Tap, a cellular protein that interacts with EBNA-1 of Epstein-Barr virus. *Virology* **236**, 18-29.

Watson, R. J. & Clements, J. B. (1980). A herpes simplex virus type 1 function continuously required for early and late virus RNA synthesis. *Nature* **285**, 329-330.

Wen, W., Meinkoth, J. L., Tsien, R. Y. & Taylor, S. S. (1995). Identification of a signal for rapid export of proteins from the nucleus. *Cell* **82**, 463-473.

Weng, Z., Thomas, S. M., Rickles, R., Taylor, J., Brauer, A., Seidel-Dugan, C., Michael, W. M., Dreyfuss, G. & Brugge, J. (1994). Identification of Src, Fyn and Lyn SH3-binding proteins: implications for a function of SH3 domains. *Molecular Cell Biology* **14**, 4509-4521.

Whitehouse, A., Cooper, M. & Meredith, D. (1998). The immediate-early gene product encoded by open reading frame 57 of Herpesvirus saimiri modulates gene expression at the posttranscriptional level. *Journal of Virology* **72**, 857-861.

Wichmann, I., Garcia-Lozano, L., Respaldiza, N., Gonzales-Escribano, M. F. & Nunez-Roldan, A. (1999). Autoantibodies to transcriptional regulation proteins DEK and ALY in a patient with systemic Lupus Erthematosus. *Human Immunology* **60**, 57-62.

Wickens, M. (1990). In the beginning is the end: regulation of poly(A) addition and removal during early development. *Trends in Biological Science* **15**, 320-324.

Wilcox, K. W., Kohn, A., Skylanskaya, E. & Roizman, B. (1980). Herpes simplex virus phosphoproteins. I. Phopshate cycles on and off some viral polypeptides and can alter their affinity for DNA. *Journal of Virology* **33**, 167-182.

Will, C. L. & Luhrmann, R. (1997). Protein functions in pre-mRNA splicing in metazoa. *Current Opinions in Genetics Development* **7**, 205-211.

Winkler, M., Rice, S. & Stammeringer, T. (1994). UL69 of human cytomegalovirus, an open reading frame with homology to ICP27 of herpes simplex virus, encodes a transactivator of gene expression. *Journal of Virology* **68**, 3943-3954.

Xiao, S. H. & Manley, J. L. (1997). Phosphorylation of the ASE/SF2 RS domain affects both protein-protein and protein-RNA interactions and is necessary for splicing. *Genes and Development* **11**, 334-344.

York, I., Roop, C., Andrews, D., Riddell, S., Graham, F. & Johnson, D. (1994). A cytosolic herpes simplex protein inhibits antigen presentation to CD8+ T lymphocytes. *Cell* **77**, 525-535.

Yu, L., Loewenstein, P. M., Zhang, Z. & Green, M. (1995). *In vitro* interaction of the TFIIB with the cellular protein TAP. *Journal of Virology* **69**, 3017-3023.

Zhang, B. L., Kraemer, B., Sengupta, D. & Wickens, M. (1999). Yeast three-hybrid systems to detect and analyse interactions between RNA and protein. *Methods in Enzymology* **306**, 93-113.

Zhi, Y. & Sandri-Goldin, R. (1999). Analysis of the phosphorylation sites of herpes simplex virus type 1 regulatory protein ICP27. *Journal of Virology* **73**, 3246-3257.

Zhi, Y., Sciabica, K. S. & Sandri-Goldin, R. M. (1999). Self-interaction of the herpes simplex virus type 1 regulatory protein ICP27. *Virology* **257**, 341-351.

Zhu, X., Chen, J., Young, C. S. H. & Silverstein, S. (1990). Reactivation of latent herpes simplex virus by adenovirus recombinants encoding mutant IE-0 gene products. *Journal of Virology* **64**, 4489-4498.