TITLE CONSTRUCTION AND ANALYSIS OF ADENOVIRUS/HIV-1 REV RECOMBINANTS

AUTHOR Richard Dafydd WILLIAMS

DEGREE Ph.D

AWARDING BODY Warwick University

DATE 1993

THESIS NUMBER DX182473

Construction and Analysis of Adenovirus/HIV-1 *rev* recombinants

Richard Dafydd Williams

BSc.(Hons.) (Lancaster)

A thesis presented for
the degree of PhD in the
University of Warwick

Department of Biological Sciences

September 1993

Table of contents

Table of contents	i
List of figures	viii
List of tables	x
Acknowledgements	xi
Declaration	xii
Abbreviations	xiii
Summary	xvii
Chapter 1 Introduction	1
1.1. Overview	2
1.2. Replication cycle and gene expression of HIV-1	3
1.2.1. Introduction	3
1.2.2. Structure and life cycle of HIV-1	4
1.2.3. Organisation and expression of the HIV-1 genome	6
1.3. Activity, properties, and functional domains of the Rev protein	9
1.3.1. Discovery of rev function	9
1.3.2. Properties of Rev	10
1.3.3. RRE binding domain of the Rev protein	12
1.3.4. Multimerisation domain	13
1.3.5. Subcellular localisation domain	14
1.3.6. Phosphorylation	15
1.3.7. Activation domain	16
1.4. The Rev responsive element (RRE)	19
1.4.1. RRE function and location	19
1.4.2. Rev binding ability of the RRE	20
1.4.3. RRE secondary structure and functional domains	20
1.4.4. Rev binding site	24
1.5. Stoichiometry of Rev-RRE binding and Rev oligomerisation	26

1.6. *cis*-acting repression sequences 29
1.7. Mechanisms of Rev action 32
1.7.1. Effects on mRNA transport 32
1.7.2. Effects on mRNA splicing 34
1.7.3. Effects on mRNA stability 36
1.7.4. Effects on mRNA utilisation 36
1.7.5. Combined mechanisms 38
1.8. Replication cycle and gene expression of adenovirus 39
1.8.1. Introduction 39
1.8.2. Structure and life cycle of Ad5 40
1.8.3. Organisation and expression of the adenovirus genome 41
1.8.4. Regulation of adenovirus gene expression 45
1.9. E1B 55K 46
1.9.1. Expression 46
1.9.2. Association with E4 34K protein 47
1.9.3. Subcellular localisation 48
1.9.4. Phosphorylation 48
1.9.5. Functional domains 49
1.9.6. Role of E1B 55K in lytic infection 49
1.9.7. Mechanism of action of E1B 55K 55
1.9.8. Interactions of E1B 55K and E4 proteins 58
1.10. Comparison of HIV-1 Rev and Ad5 E1B 55K function - project objectives 60

Chapter 2 Materials and methods 62
2.1. Materials 63
2.1.1 General buffers and solutions 63
2.1.2 Suppliers 63
2.2. Nucleic acid purification 64

2.2.1. Ethanol precipitation of nucleic acids 64
2.2.2. Phenol/chloroform extraction 64
2.2.3. Quantification of nucleic acids 64
2.3. DNA manipulation techniques 65
2.3.1. Digestion by restriction enzymes 65
2.3.2. Partial digestion by restriction enzymes 65
2.3.3. End-fill reaction and 3' overhang removal 66
2.3.4. Dephosphorylation 66
2.3.5. Linker kinase reaction 66
2.3.6. Ligations 67
2.4 Agarose gel electrophoresis 67
2.4.1. Separation of restriction fragments 67
2.4.2. Purification of fragments from agarose gels 67
2.5. Bacteriological techniques 68
2.5.1. Growth and maintenance of bacterial stocks 68
2.5.2. Transformation of bacteria with plasmid or M13 replicative form DNA (calcium chloride method) 69
2.5.3 Transformation of bacteria with plasmid DNA (electroporation method) 70
2.5.4. Rapid preparation of plasmid or M13 DNA from bacteria 71
2.5.5. Large-scale preparation of plasmid / M13 DNA 72
2.6. Tissue culture and virus techniques 73
2.6.1. Maintenance and passage of human cell lines 73
2.6.2. Freezing and recovery of cell stocks 73
2.6.3. Calcium phosphate transfection for adenovirus reconstruction by in vivo recombination 74
2.6.4. Calcium phosphate transfection for transient assays of gene expression 75
2.6.5. Virus infection of cell monolayers 75

2.6.6. Rapid preparation of adenovirus DNA 75
2.6.7. Large-scale preparation of adenovirus DNA 76
2.6.8. Preparation of virus stocks 77
2.6.9. Titration of virus stocks by plaque assay 77
2.7. Analysis of viral protein expression 78
2.7.1. ^{35}S-methionine pulse-labelling of proteins 78
2.7.2. Quantification of ^{35}S-methionine incorporation into protein 78
2.7.3. SDS polyacrylamide gel electrophoresis (SDS-PAGE) 78
2.8. Analysis of viral RNA expression 79
2.8.1. Preparation of cytoplasmic and nuclear RNA 79
2.8.2. Preparation of antisense RNA probes 80
2.8.3. Ribonuclease protection analysis 81
2.9. Autoradiography 83
2.10. Densitometry 83
2.11. Chloramphenicol Acetyl Transferase (CAT) assays 83
2.11.1. Preparation of cytoplasmic extracts 83
2.11.2. Assay for protein concentration 84
2.11.3. CAT assay procedure 84

Chapter 3 Construction and isolation of recombinant adenoviruses 86
3.1. Introduction 87
3.2. Insertion of the RRE into adenovirus - Construction of dl309/RRE 94
3.2.1 Subcloning 94
3.2.2. Virus reconstruction by ligation 97
3.3. Insertion of rev expression cassette into adenovirus - Construction of dl309/rev 97
3.3.1 Subcloning 97
3.3.2 Virus reconstruction by overlap recombination in vivo 100
3.4. Construction of the other recombinants 100

3.5. Restriction analysis of viral recombinants 101
3.5.1. Restriction analysis of Ad5 recombinants with Hind III 101
3.5.2. Restriction analysis of Ad5 recombinants with Sal I 106
3.5.3. Restriction analysis of Ad5 recombinants with Bsu36 I 106
3.6. Summary 108

Chapter 4 Detection of functional Rev expression 109
4.1. Introduction 110
4.2. CAT reporter constructs 111
4.3. CAT assay 111
4.4. Demonstration of Rev expression 113
4.5. Results 113
4.6. Discussion 116

Chapter 5 Analysis of the effects of RRE insertion on gene expression from recombinant adenoviruses 117
5.1. Introduction 118
5.2. Analysis of viral protein expression in infected HeLa cells 119
5.3. Analysis of viral RNA expression in infected HeLa cells 121
5.3.1. Analysis of L3 RNA species from infected HeLa cells 3' terminus 124
5.3.2. Analysis of specific L3 RNA species from infected HeLa cells 124
5.3.3. Analysis of L4 and E3 species from infected HeLa cells 126
5.4. Discussion 128

Chapter 6 Analysis of the effects of a rev expression cassette on Ad5 gene expression 129
6.1. Introduction 130
6.2. Analysis of protein expression in infected HEK-293 cells 130

6.3. Analysis of RNA expression in infected HEK-293 cells 133

6.3.1. Analysis of L3 RNA species from infected HEK-293 cells: 3' terminus 135

6.3.2. Analysis of specific L3 RNA species from infected HEK-293 cells 135

6.3.3. Analysis of L1, L2 and L4 RNA species from infected HEK-293 cells 138

6.4. Discussion 140

Chapter 7 Analysis of the effects of a rev expression cassette and the RRE on Ad5 gene expression 141

7.1. Introduction 142

7.2. Analysis of protein expression in infected KB8a cells 142

7.2.1. Protein analysis at 12, 16.5 and 19 hr post-infection 142

7.2.2. Protein analysis at 49 hr and 72 hr post-infection 142

7.3. Analysis of RNA expression in infected KB8a cells 146

7.3.1. Analysis of L3 RNA species from infected KB8a cells: 3' terminus 146

7.3.2. Analysis of specific L3 RNA species from infected KB8a cells 148

7.3.3. Analysis of L1 and L2 RNA species from infected KB8a cells 151

7.3.4 Analysis of L4 and E3 species from infected KB8a cells 153

7.3.5 Analysis of L5 RNA species from infected KB8a cells 155

7.3.6 Analysis of IVa2 RNA species from infected KB8a cells 157

7.4. Discussion 157

Chapter 8 General discussion 161

8.1. Summary and interpretation of results 162

8.2. Implications for the Rev / E1B 55K analogy 165

8.3. Further experiments 167

References 168

List of figures

Fig. 1.1. Organisation and RNA expression of the HIV-1 genome 5

Fig. 1.2. Domain structure of the Rev protein 11

Fig. 1.3. Predicted secondary structure of the HIV-1 Rev-responsive element (RRE) 21

Fig. 1.4. Transcription and translation map of Ad2 42

Fig. 3.1. rev and RRE insertion sites in Ad5 88

Fig. 3.2. Viral reconstruction plasmids 90

Fig. 3.3. 3-fragment ligation (e.g. dl309/RRE) 91

Fig. 3.4. Overlap recombination in vivo (e.g. dl309/rev) 92

Fig. 3.5. Isolation of adenovirus recombinants 93

Fig. 3.6. Construction of pE2RREx 95

Fig. 3.7. Construction of pwtE1rev 98

Fig. 3.8. Restriction sites in Ad5 and Ad5 recombinants 102

Fig. 3.9. Restriction analysis of Ad5 recombinants with Hind III 103

Fig. 3.10. Restriction analysis of Ad5 recombinants with Sal I 105

Fig. 3.11. Restriction analysis of Ad5 recombinants with Bsu36 I 107

Fig. 4.1. pTK34 CAT reporter construct 112

Fig. 4.2. CAT assay of cytoplasmic extracts (virus + pTK34) 115

Fig. 5.1. Analysis of protein expression in infected HeLa cells 120

Fig. 5.2. Map of regions that protect L3 riboprobes 122

Fig. 5.3. Analysis of L3 RNA species from infected HeLa cells 3' terminus 123

Fig. 5.4. Analysis of specific L3 RNA species from infected HeLa cells 125

Fig. 5.5. Analysis of L4 and E3 species from infected HeLa cells 127

Fig. 6.1. Analysis of protein expression in infected HEK-293 cells 131

Fig. 6.2. Analysis of L3 RNA species from infected HEK-293 cells 3' terminus 134

Fig. 6.3. Analysis of specific L3 RNA species from infected HEK-293 cells 137

Fig. 6.4. Analysis of L1, L2 and L4 RNA species from infected HEK-293 cells 139

Fig. 7.1. Analysis of protein expression in infected KB8a cells at 12, 16.5 and 19 hr post-infection 143

Fig. 7.2. Analysis of protein expression in infected KB8a cells at 49 hr and 72 hr post-infection 145

Fig. 7.3. Analysis of L3 RNA species from infected KB8a cells 3' terminus 147

Fig. 7.4. Analysis of specific L3 and L2 RNA species from infected KB8a cells 149

Fig. 7.5. Analysis of L1 and L2 RNA species from infected KB8a cells 152

Fig. 7.6. Analysis of L4 and E3 species from infected KB8a cells 154

Fig. 7.7. Analysis of L5 RNA species from infected KB8a cells 156

Fig. 7.8. Analysis of IVa2 RNA species from infected KB8a cells 158

List of tables

Table 2.1. pGEM plasmids used to transcribe probes for RNase protection assays 80

Table 3.1. Recombinant adenoviruses 94

Table 4.1 CAT assay scintillation counts 113

Table 6.1a Quantitation of viral proteins from HEK-293 cell infections 133

Table 6.1b Ratios of band densities in recombinant infections to equivalent band densities in *dl*309 infection 133

Acknowledgements

I wish to thank the past and present members of the Adenovirus and Rotavirus Groups of the Biological Sciences Department for their help and advice during the last four years at Warwick. In particular, I am grateful for the advice of Cally Caravokyri and Sean Mason on many of the techniques used in this thesis. I would also like to thank Simon Swingler for advice on CAT assays, Mrs. B.D. Williams for checking references, and Wanda M. Waterworth for proof-reading. Finally, I wish to thank my supervisor, Dr. Keith Leppard, for his invaluable guidance, helpful suggestions, and constructive criticism at all stages of the project.

This work was supported by a grant from the Medical Research Council AIDS Directed Programme.

Declaration

All the work presented in this thesis was performed by the author in the Biological Sciences Department of the University of Warwick, under the supervision of Dr. K.N. Leppard. None of this material has been previously published or submitted for examination at another institution.

Abbreviations

General abbreviations

AAV	adeno-associated virus
AIDS	acquired immunodeficiency syndrome
Ad	adenovirus
Ad(5) (2)	human adenovirus serotype (5) (2)
ATF	activating transcription factor
ATP	adenosine triphosphate
bp	base pairs
cAMP	cyclic adenosine monophosphate
CAT	chloramphenicol acetyl transferase
CoA	coenzyme A
CR1 (2) (3)	conserved region 1 (2) (3) of Ad E1A
CRS	*cis*-acting repression sequence
CS	calf serum
CTP	cytidine triphosphate
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
DEAE	diethylaminoethyl
DEPC	diethylpyrocarbonate
dGTP	deoxyguanosine triphosphate
DMEM	Dulbecco's modification of Eagle's medium
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DOC	deoxycholate
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
EIAE	adenovirus E1A enhancer

EDTA	ethylene diamine tetraacetic acid (disodium salt)
FCS	foetal calf serum
GTP	guanosine triphosphate
HEK	human embryonal kidney cells
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
HIV(-1) (2)	human immunodeficiency virus (type 1) (type 2)
hnRNP	heterogeneous nuclear ribonucleoprotein
HTLV(-I)	human T-cell lymphotropic virus (type I)
LB	Luria-Bertani medium
LTR	retroviral long terminal repeat
M_r	relative molecular mass
MHC	major histocompatibility complex
MLP	Ad major late promoter
NP40	Nonidet P40 detergent
NRE	negative regulatory element in HIV LTR
nt	nucleotides
ORF	open reading frame (e.g. Ad5 E4 ORF6)
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
pfu	plaque forming units
RNA	ribonucleic acid (prefixes: m = messenger, r = ribosomal, sn = small

	nuclear, t=transfer)
RNase	ribonuclease
rpm	revolutions per minute
RRE	Rev-responsive element
SDS	sodium dodecyl sulphate
SV40	simian virus 40
TAR	HIV *trans*-activator response element
TBP	TATA binding protein
TCA	trichloroacetic acid
TEMED	N, N, N',N'-tetramethylethylenediamine
TPL	tripartite leader of Ad major late transcription unit
tris	tris-(hydroxymethyl)-methylamine
UTP	uridine triphosphate

The compositions of individual buffer solutions given abbreviations (TE, TNE, TBE, TD, TS, etc.) are described in chapter 2. The names given to recombinant viruses constructed in the course of the project are listed in table 3.1. Standard abbreviations are used for SI units and amino acids.

Virus gene nomenclature

For HIV-1 the standard retroviral structural gene nomenclature, and the regulatory gene nomenclature of Gallo *et al.* (1988) are used throughout the thesis - i.e., *gag* (group-specific antigen gene), *pol* (polymerase), *env* (envelope), *tat* (*trans*-activator), *rev* (regulator of expression of virion proteins), *vif* (virion infectivity factor), *nef* (negative factor), *vpr* (viral protein R), and *vpu* (viral protein U). The corresponding gene products are denoted by Roman script abbreviations (Gag, Pol, Env, etc.) or protein/glycoprotein molecular masses (e.g. p17, gp41). For adenovirus, the genes and proteins are named according to the early region (E1A,

E1B, E2A, E2B, E3, E4) major late region (L1, L2, L3, L4, L5), or minor late region (IVa2, IX) of the genome from which they are expressed (fig. 1.4), together with the open reading frame number (e.g. E4 ORF6), number of amino acids of the predicted gene product (e.g. E1B 495R), molecular mass of the gene product (e.g. L3-23K), or a Roman numeral identifying a virion component (e.g. L3 II = hexon protein).

Summary

The human immunodeficiency virus Rev protein is required for the cytoplasmic accumulation and probably the translational utilisation of mRNAs encoding the late viral structural proteins. Rev function is mediated by direct binding of the protein to a *cis*-acting RNA sequence, the Rev-responsive element (RRE), carried by these mRNAs. Human adenovirus type 5 (Ad5) also encodes a protein, E1B 55K, required for the cytoplasmic accumulation of late viral mRNAs. Rev and E1B 55K both regulate gene expression post-transcriptionally and directly or indirectly act on the mRNA transport pathway. To explore the potential functional analogy between these two regulatory proteins, the components of the Rev/RRE system were introduced into Ad5. By inserting the RRE into Ad5 late region L3 (expression of which is normally dependent on E1B 55K), and a *rev* expression cassette into early region E1A, a system was set up where the action of Rev could be directly compared with that of E1B 55K.

A series of six recombinant adenoviruses was constructed which, together with two viruses already available, contained the *rev* gene and/or the RRE and/or the Ad5 E1B 55K gene in all possible combinations. Expression of functional Rev from the appropriate recombinants was confirmed by a CAT reporter gene assay. The eight viruses were used to study the effects of Rev and the RRE on the expression of Ad5 late RNAs and proteins. It was shown that the Rev/RRE system can detectably increase the cytoplasmic accumulation of certain Ad5 mRNAs in the absence of E1B 55K. Surprisingly, these included some mRNAs in which the RRE, although present in the primary transcripts, was removed from the mature species. A mechanism was proposed in which Rev/RRE action in this system may commit an RNA to a pathway of facilitated nuclear export before excision of the RRE during processing.

Chapter 1

Introduction

1.1. Overview

The aim of this study was to investigate and compare two viral systems that regulate gene expression - the Rev/RRE system of human immunodeficiency virus type 1 (HIV-1), and the E1B 55K/E4 34K complex of human adenovirus type 5 (Ad5).

The HIV-1 *rev* gene encodes a 13K protein, Rev, required for the synthesis of certain viral proteins. Specifically, Rev upregulates the expression of the late proteins Gag, Pol, Env, Vif, Vpr and Vpu at the expense of downregulating expression of the early viral proteins - Tat, Nef and Rev itself. Rev regulation is post-transcriptional, and probably affects more that one level of gene expression. Evidence has been presented for effects on mRNA splicing, transport, stability and efficient translation. These effects are mediated by specific binding of the Rev protein to a *cis*-acting RNA sequence, the Rev-responsive element (RRE) located in the coding region of the *env* gene. The RRE is retained in the unspliced and singly-spliced mRNAs encoding the late proteins, but absent from the multiply-spliced mRNAs coding for the early proteins. Upregulation of late protein expression appears to be a direct consequence of Rev-RRE interaction, while downregulation of early protein expression is thought to be an indirect effect of this process.

Ad5 early region 1B encodes a protein, E1B 55K, that is also required for efficient expression of late viral proteins. E1B 55K has been shown to act in combination with a product of early region 4, E4 34K. Together, the two proteins form a functional complex. Like Rev, the E1B 55K/E4 34K complex acts post-transcriptionally on gene expression, most probably at an intranuclear stage in the mRNA transport pathway, and/or on cytoplasmic stabilisation of recently transported RNA. Its action results in the selective facilitation of viral late mRNA accumulation in the cytoplasm.

There are clear similarities between the roles of Rev and E1B 55K/E4 34K in their respective virus life cycles. This thesis presents the results of experiments designed to explore the extent of this functional analogy. The functions and possible mechanisms of action of the Rev/RRE system and the E1B 55K/E4 34K complex within the life cycles of HIV-1 and adenovirus respectively are discussed in detail below.

1.2. Replication cycle and gene expression of HIV-1

1.2.1. Introduction

Human immunodeficiency virus type 1 (HIV-1) is the primary actiological agent of the global pandemic of acquired immunodeficiency syndrome (AIDS), a fatal disease of the immune system. In the decade since the discovery of the virus (Barré-Sinoussi *et al.*, 1983), extensive research has amassed an enormous amount of information on its structure, immunology, pathogenesis, and gene expression, although as yet no effective treatment is available for AIDS. This section is limited to a brief description of the life cycle of the virus and the regulation of its gene expression.

HIV-1 belongs to the family *Retroviridae* - i.e., its replication cycle involves reverse transcription of the RNA virion genome to a DNA provirus which integrates in the host cell genome. More specifically, HIV-1 is classed as a member of the subfamily *Lentivirinae*, which also includes such viruses as maedi-visna virus of sheep, the simian immunodeficiency viruses (SIVs) and human immunodeficiency virus type 2 (HIV-2), a virus closely related to the SIVs that is genetically distinct from HIV-1. All the viruses in this group share a similar genetic organisation that identifies them as 'complex' retroviruses - i.e., they encode regulatory proteins (in addition to the structural proteins found in all retroviruses) that mediate a biphasic temporal regulation of gene expression (see review by Cullen, 1991).

1.2.2. Structure and life cycle of HIV-1

The structure, replication cycle, and pattern of gene expression of HIV-1 have been well characterised (see reviews by Wong-Staal, 1990; Vaishnav and Wong-Staal, 1991; Cullen, 1991; Kräusslich, 1992, and references therein). The HIV-1 virion consists of an outer capsid of Gag (group-specific antigen) matrix protein p17 enveloped by a lipid membrane derived from the host cell. On the membrane, knobs of Env (envelope) surface glycoprotein gp120 are noncovalently attached to the Env transmembrane glycoprotein gp41. Within the outer shell of p17 is a cylindrical core of Gag capsid protein p24. The core contains two copies of the single-stranded RNA genome bound to Gag protein p9, and the viral enzymes required for replication - the Pol (polymerase) proteins p66/p51 (reverse transcriptase/RNase H heterodimer), p11 (protease), and p32 (integrase).

The CD4 cell surface antigen, carried by cells such as T4 helper lymphocytes, has been identified as the primary HIV-1 receptor (Maddon *et al.*, 1986). The virus binds to CD4 via its gp120 envelope glycoprotein (McDougal *et al.*, 1986). Subsequent features of the establishment of HIV-1 infection are similar to those of classical retroviruses. Internalisation of the viral core occurs by membrane fusion with the target cell, apparently in a pH-independent process mediated by the gp41 transmembrane glycoprotein. In the cytoplasm, reverse transcriptase catalyses the synthesis of proviral DNA from the RNA template, a process that utilises both the DNA polymerase and RNase H activities of the enzyme, and requires a $tRNA^{Lys}$ molecule bound to genomic RNA. The 10 kb proviral DNA, probably as a nucleoprotein complex, then migrates to the nucleus where it is integrated into the host cell genome by the p32 integrase. Gene expression, regulated by both cellular factors and viral nonstructural proteins, may then occur (as described below, 1.2.3). This culminates in the expression of genomic viral RNA and viral structural proteins that are packaged together into new virions - a process dependent on packaging signals in the genomic RNA. The structural proteins are expressed initially as

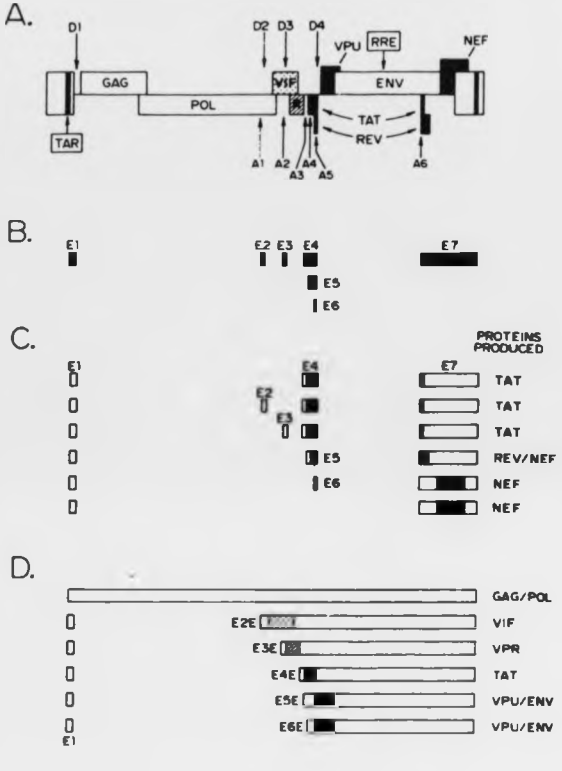

A.
D1
D2
D3
D4
VPU
RRE
NEF
GAG
VIF
ENV
POL
TAT
REV
TAR
A1
A2
A3
A4
A5
A6
B.
E1
E2
E3
E4
E7
E5
E6
C.
PROTEINS PRODUCED
E1
E4
E7
E2
TAT
TAT
E3
TAT
E5
REV/NEF
E6
NEF
NEF
D.
GAG/POL
E2E
VIF
E3E
VPR
E4E
TAT
E5E
VPU/ENV
E6E
VPU/ENV
E1

Fig. 1.1. Organisation and RNA expression of the HIV-1 genome (reproduced from Cullen, 1991). **(A)** Genomic organisation, showing the locations of the *gag*, *pol*, *env*, *tat*, *rev*, *nef*, *vif*, *vpu* and *vpr* (R) genes, the Rev-responsive element (RRE), *trans*-activator response element (TAR), splice donors (D1-D4), splice acceptors (A1-A6), and long terminal repeats (LTRs: large rectangles at each end of the genome). **(B)** Exons (E1-E6) found in multiply-spliced mRNAs (E1-E3 are noncoding). **(C)** Early mRNAs (additional species can also be formed by inclusion of E2 and E3 in *rev* and *nef* mRNAs - not shown). **(D)** Late mRNAs.

precursor polyproteins that are cleaved to the mature forms by the p11 protease. Cleavage of the Gag and Pol polyproteins appears to occur at a late stage in virion maturation, during or after budding of newly formed virus particles at the cell membrane. The Env precursor, gp160, is cleaved intracellularly to form gp120 and gp41 which are then targeted to the plasma membrane where they become associated with budding capsids, presumably by gp41-p17 interaction. Virions are then released from the cell complete with lipid envelopes derived from the host plasma membrane.

1.2.3. Organisation and expression of the HIV-1 genome

The organisation of the HIV-1 genome (fig. 1.1) and its pattern of gene expression have been reviewed in the references cited in the previous section. All the known functional genes of HIV-1 are expressed from a single active promoter located in the 5' long terminal repeat (LTR) region. The LTR is a multifunctional unit duplicated at each end of the linear proviral form of the HIV-1 genome. The LTR contains a variety of signals, some of which are utilised at the 3' end of the genome, while others are active at the 5' end. At the 3' end, the signals responsible for polyadenylation of all HIV-1 mRNAs are active, and sequences encoding the 3' end of the *nef* gene are utilised. At the 5' end, sequences in use include the promoter, and enhancer, a *cis*-acting negative regulatory element (NRE), the response element (TAR) for the viral Tat transactivator protein, and binding sites for several cellular constitutive and inducible transcription factors including AP1, NFAT-1, USF (upstream binding factor), NF-kB, Sp1, TFIID, LBP (leader binding protein) and CTF/NF-1.

HIV-1 primary transcripts extend from an RNA start site in the 5' LTR to a termination site in the 3' LTR. However, differential splicing gives rise to three major classes of mRNA from which viral proteins are translated. The regulatory proteins Tat, Rev and Nef are translated from small ($\simeq$2 kb) multiply-spliced

mRNAs, the Env structural proteins and the Vif, Vpr, and Vpu accessory proteins are translated from ≈4 kb singly-spliced mRNAs, and the Gag and Pol structural proteins are translated from unspliced, ≈9 kb mRNAs. The *gag* and *pol* genes overlap by 241 nt (fig. 1.1) with *pol* in the -1 translational reading frame with respect to *gag*; production of the Pol protein is dependent on a low frequency ribosomal frame shift that creates a Gag-pol fusion.

The absolute levels of primary transcripts and relative levels of the three classes of mRNA in the cytoplasm are linked to the phase of infection. In latently infected cells or in the early phase of a lytic infection, transcription levels are low and the mRNAs in the cytoplasm are predominantly 2 kb multiply-spliced species. In the late phase of lytic infection, the singly-spliced and unspliced species appear in the cytoplasm, while levels of the multiply-spliced species fall. Therefore, as in the case of adenovirus and other DNA tumour viruses, nonstructural regulatory proteins are expressed in the early phase of infection, whereas structural virion proteins are synthesised in the late phase. A viral RNA binding protein, Rev, is implicated in controlling the cytoplasmic accumulation of viral mRNAs; its activity will be reviewed in detail below. Control of viral transcription levels is effected by a second RNA binding protein, Tat. Expression of the Tat protein results in a powerful transactivation effect on transcription from the HIV-1 promoter.

The possible mechanisms of Tat action are reviewed in Vaishnav and Wong-Staal (1991), Cullen (1991), and Krausslich (1992). Tat has been shown to bind the RNA form of the TAR element in the LTR directly; this region of the LTR is incorporated into the 5' end of all HIV-1 transcripts. Binding is mediated by an arginine-rich domain of the Tat protein and is dependent on the secondary structure of TAR; the RNA element forms a helical stem-loop with a three nucleotide bulge that constitutes the Tat binding site. Binding of Tat to HIV RNA is essential for Tat function, but activity is retained if the arginine-rich domain and bulged RNA stem-

loop are replaced by heterologous binding elements. Tat appears to act directly on transcription by stabilising inefficiently elongating transcription complexes (anti-termination), and possibly by increasing the levels of transcription initiation. Tat may also be active at the level of translation.

An additional viral protein, Nef, may also regulate the level of viral gene expression, although the data presently available are conflicting (reviewed in Cullen, 1991; Vaishnav and Wong-Staal, 1991; Krausslich, 1992). While initial experiments suggested that Nef acted as a negative regulator of gene expression, this was not confirmed in subsequent studies. Moreover, experiments with SIV mutated in the equivalent *nef* region showed that a functional gene was selected for *in vivo* and was required for normal viral replication. Sequence similarities with G-proteins and the location of at least some Nef molecules on the inner surface of the plasma membrane suggested that Nef might belong to this class of signal-transducing molecules. However, early reports of detectable Nef GTP-binding, GTPase and autokinase activities (characteristic of G-proteins) were not confirmed in later experiments. Thus the function of Nef is at present enigmatic.

In addition to the proteins discussed above, HIV-1 encodes a number of gene products that have been termed accessory proteins (reviewed in Vaishnav and Wong-Staal, 1991) since many of them are dispensable for maintenance of virus in tissue culture. However, their conservation in natural isolates indicates that they have important functions in infections *in vivo*. The 23K Vif protein appears to function as a virion infectivity factor - i.e., it is required for cell-free transmission of virus, but not for cell-to-cell infection by membrane fusion. However, Vif is apparently not a virion component, and can function in *trans* when supplied to *vif*-negative mutants from a complementing expression vector. Its mechanism of action is at present unclear.

The 15K Vpr protein is a virion component that upregulates gene expression synergistically with Tat, possibly at the levels of transcription initiation and transcript stabilisation. The Vpu protein, of similar size to Vpr, is expressed in the cytoplasm, and is required for extracellular accumulation of virus particles, and may therefore have a role in virion assembly or release. Several other polypeptides, produced by alternative mRNA splicing and comprising sequence elements from one or more of the proteins discussed above, have also been described.

1.3. Activity, properties, and functional domains of the Rev protein

1.3.1. Discovery of *rev* function

The *rev* open reading frame was first identified as a functional gene by Sodroski *et al.* (1986) and Feinberg *et al.* (1986). These groups showed that mutations distinct from those of the *tat* gene (i.e., mutations that did not affect viral Tat expression, or could not be complemented by Tat supplied in *trans*) altered the pattern of HIV-1 gene expression. Specifically, *rev* was found to be required intact for the expression of Gag and Env proteins and was therefore necessary for viral replication. Conversely, *rev* was shown to be unnecessary for Tat expression; indeed Feinberg *et al.* (1986) found that the levels of small multiply-spliced RNA species (which would include *tat* and *rev* RNAs) were reduced in the presence of *rev*. Although both groups suggested that the *rev* gene functioned post-transcriptionally, different mechanisms were proposed for its activity. Sodroski *et al.* (1986) postulated that *rev* acted to counter the effects of *cis*-acting repression sequences that prevented the expression of structural genes in the absence of *rev* - hence their name *art* (antirepression *trans*-activator) for the *rev* gene. This theory was based on the observation that *tat* (and heterologous genes) under the control of the HIV-1 LTR did not require *rev* for expression, whereas the structural genes were *rev*-dependent. However, Feinberg *et al.* (1986) proposed that the *rev* gene product functioned as a *trans*-acting regulator of splicing ('*trs*') since cytoplasmic structural gene mRNA was only detected in the presence of *rev*, whereas predominantly small multiply-

spliced species were detected in its absence. Since these initial experiments considerable progress has been made in clarifying the mechanism of action of the *rev* product, although as yet no single mechanism can unambiguously account for all the available data. However, the apparently conflicting suggestions of Sodroski *et al.* (1986) and Feinberg *et al.* (1986) may represent two aspects of a single mechanism (see discussion in 1.7, below).

An important initial stage in the elucidation of *rev* function was the discovery of its gene product, Rev, in HIV-1-infected cells, strongly suggesting that the effects of the gene were protein-mediated. Goh *et al.* (1987) expressed the *rev* gene in bacteria, and demonstrated that the protein expressed was recognised by serum from an HIV-1-infected individual. This serum also reacted with a protein found in infected cells that, like the bacterially-synthesised Rev protein, had an apparent (SDS-PAGE) molecular mass (M_r) of 20K (the disparity with the predicted M_r of 13K was attributed to the charge or structure of the protein). The two proteins were shown to compete for antibody binding, and the protein from infected cells was also detected by an antibody to a synthetic Rev oligopeptide. It is clear from these results that the *rev* gene is expressed as a protein *in vivo*. A second genetic element, the Rev-responsive element (RRE), is also crucial for *rev* function; its presence in an mRNA is required for Rev-facilitated cytoplasmic accumulation (see 1.4.1, below, and references therein). This sequence is located within the viral *env* gene and is therefore incorporated into the unspliced and singly-spliced viral mRNAs that encode the late proteins, but is absent from the multiply-spliced mRNAs that encode the early proteins (fig. 1.1). The evidence supporting the current definition of the RRE as an RNA structural element is discussed further in section 1.4.

1.3.2. Properties of Rev

The Rev protein has several distinct biochemical properties, some of which are crucial for functional activity. Rev binds RRE RNA directly, localises to the

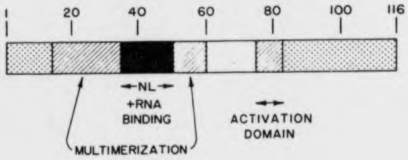

1
20
40
60
80
100
116
NL
+RNA
BINDING
ACTIVATION
DOMAIN
MULTIMERIZATION

Fig. 1.2. Domain structure of the Rev protein (reproduced from Malim and Cullen, 1991). The three distinct functional domains located within the Rev primary structure are indicated. The arginine-rich domain (amino acids 35-50) is required for RRE binding and nuclear/nucleolar localisation. Sequences flanking this domain between amino acids 56 and 75 (approximately) are required for multimerisation of Rev on the RRE. The leucine-rich activation domain (amino acids 75-83) is essential for Rev function *in vivo* and may interact with cellular factors. The N- and C-terminal domains are dispensable for Rev function. It should be noted that some aspects of this proposed domain structure are still controversial (see text).

nucleus and nucleolus, is phosphorylated, and has the ability to self-associate into multimers (see 1.3.3-1.3.6 and references therein). Numerous attempts have been made to map these properties to specific amino acid domains and to correlate them with the biological function of Rev. Most of these studies have utilised mutational analysis or domain exchange with heterologous proteins to dissect the structure-function relationship of Rev. A provisional functional domain map of Rev established by these studies is shown in fig. 1.2. The data relevant to the mapping of each function are summarised in subsequent sections. One notable and rather surprising observation is that the 25 C-terminal residues of the 116 amino acid protein are dispensable for Rev function - deletion of these residues has no effect on Rev activity (Cochrane *et al.*, 1989b).

1.3.3. RRE binding domain of the Rev protein

The Rev protein has been shown to bind RRE RNA directly by a variety of methods (see 1.4.2., below, and references therein). The initial evidence for the location of the RRE-binding domain in Rev was indirect. Several groups (Malim *et al.*, 1989b; Perkins *et al.*, 1989; Hope *et al.*, 1990b; Venkatesh *et al.*, 1990) described mutations in the region of residues 25-30 that disrupted or abolished Rev function. These mutations were centred around an arginine-rich sequence typical of nucleic acid binding domains. The direct correlation between binding ability and functional activity for mutants in this region was first noted by Olsen *et al.* (1990b), who described a series of mutations between residues 28 and 65 that disrupted RNA binding and Rev function. Data consistent with these results were also obtained by Berger *et al.* (1991), Malim and Cullen (1991) and Zapp *et al.* (1991). Subsequently, Kjems *et al.* (1992) demonstrated that a peptide containing Rev residues 34-50 can interact specifically with the RRE. Very recently, Tan *et al.* (1993) have shown that this peptide adopts an α-helical conformation *in vivo* that is required for specific binding to the RRE. The arginine-rich region of Rev has been recognised as one of a family of similar motifs (Lazinski *et al.*, 1989) which

includes the RNA binding domains of bacteriophage antiterminator proteins and HIV-1 *tat* - both of which, like Rev, have been shown to bind to RNA hairpin or stem-loop structures. Similar domains are also present in histones, protamines and the Adenovirus E4 ORF 4 13K protein (Lazinski *et al.*, 1989).

1.3.4. Multimerisation domain

The ability of Rev to multimerise in the presence or absence of bound RRE is well-documented (see 1.5, below, for references and discussion). The domains required for multimerisation have been mapped by mutational analysis, but the results from different groups have not been in complete agreement. Malim and Cullen (1991) found that substitution of residues 23, 25, & 26, or 54-56, prevented the binding of multiple Rev molecules to the RRE but did not affect the binding of a single Rev molecule. Olsen *et al.* (1990b) demonstrated that mutations of residues 28-31 or 55-57 prevented Rev multimerisation, but these mutations also abolished Rev-RRE binding. In contrast, deletion of residues 45-51 or substitution of residues 38-39 in the arginine-rich domain was found to abolish RRE binding, but not to affect Rev multimerisation. Finally, Zapp *et al.* (1991) found that substitutions of residues 14-16 or 23, 25 & 26, or 27-29, disrupted both binding and multimerisation, whereas mutations of residues 9-11 or 17-18 had no effect on either property. These authors also found that multimerisation and binding were severely disrupted by mutations in the arginine rich domain, including substitutions of 38-39, 41-43, 41-44, 48 & 50, or 54 & 56. Moreover, an individual substitution of Arg 44 by either Trp or Gly, or substitution of Trp 45 with Arg, had no effect on oligomerisation but abolished binding, whereas substitution of Trp 45 with Gly abolished both oligomerisation and binding.

The results of these three studies therefore broadly define the multimerisation domain as a sequence between residues 14 and 56, but differ in the importance of

residues in the arginine-rich domain (35-50) for oligomerisation, and in the correlation between multimerisation and RRE-binding ability.

1.3.5. Subcellular localisation domain

Subcellular fractionation and immunofluorescence have demonstrated that Rev is located predominantly in the nucleus, where it is localised within the nucleoli (Cullen *et al.*, 1988; Rosen and Haseltine, 1988; Felber *et al.*, 1989). Several studies have identified nuclear and nucleolar localisation signals, although conflicting data on their positions have been reported. According to several groups, a conserved arginine-rich sequence located between Asn 40 and Trp 45 is both necessary and sufficient for nuclear accumulation of Rev (Malim *et al.*, 1989b; Perkins *et al.*, 1989; Venkatesh *et al.*, 1990). However, Hope *et al.* (1990b) have found that introduction of acidic residues in positions 14-16 or 18-20 abolishes specific nuclear localisation and inactivates Rev function. The sequences required for nucleolar localisation have not been defined unequivocally. Hope *et al.* (1990a), using a Rev protein targeted to the nucleus by fusion with a steroid receptor, have found that Rev residues 41-43 are required for specific localisation in the nucleolus. Cochrane *et al.* (1990a) have shown that mutant proteins with amino acid residues 45-51 deleted, or with Arg 38 and Arg 39 replaced by His and Gly, accumulate in the nucleus but do not localise in the nucleolus, implicating these residues in specific nucleolar localisation. However, in their experiments a fusion protein of Rev residues 38-51 and ß-galactosidase failed to localise in the nucleolus, whereas a Rev 1-59 - ß-galactosidase fusion product did exhibit nucleolar targeting. These data suggest that residues in the amino terminal region of the Rev protein are also involved in nucleolar localisation, but in a similar experiment, Venkatesh *et al.* (1990) have shown that a Rev 38-52 - ß-galactosidase fusion protein can localise in the nucleolus (albeit with increased levels in the nucleoplasm compared to wild-type Rev), indicating that the residues of the central basic domain may indeed be sufficient to confer nucleolar targeting. Moreover, Rev residues 35-50 can

functionally replace a nucleolar targeting signal located at the amino terminus of the analogous Rex protein of HTLV-I (Kubota *et al.*, 1991). It is therefore likely that a functional nucleolar localisation signal resides in the 35-50 region of Rev.

Mutations which affect localisation in the nucleolus disrupt Rev function, suggesting that nucleolar targeting may be a functional requirement for Rev activity (Malim *et al.*, 1989b; Perkins *et al.*, 1989; Cochrane *et al.*, 1990a; Venkatesh *et al.*, 1990). However, due to the overlap of the RRE binding and localisation domains, it is difficult to draw definite conclusions from these experiments. Recently, McDonald *et al.* (1992) have shown that, in experiments where Rev is bound to RNA via a heterologous binding domain/RNA target sequence (phage MS2 coat protein/MS2 operator sequence), nucleolar localisation is dispensable for Rev activity. A fusion protein found in the nucleus, but excluded from the nucleoli, retained Rev function. It is therefore possible that nucleolar localisation may simply reflect the ability of Rev to bind a nucleolar factor, such as the B23 protein described by Fankhauser *et al.* (1991). However, Kubota *et al.* (1992) have shown that a mutant Rev protein with residues 38-44 deleted (thus abolishing nucleolar accumulation) functionally inhibits wild type Rev (possibly by forming heterodimers which cannot target the nucleolus). This suggests that targeting to this site is important. The precise significance of nucleolar localisation of Rev therefore remains to be determined.

1.3.6. Phosphorylation

Rev is phosphorylated *in vivo* by a specific nuclear serine kinase (Hauber *et al.*, 1988; Cochrane *et al.*, 1989a). However, mutant proteins with deletions of Ser 5 and 8 or Ser 99, or substitutions of Ser 92 or Ser 99, are subject to little or no phosphorylation, yet retain approximately wild-type levels of Rev activity, indicating that phosphorylation is not a functional requirement (Malim *et al.*, 1989b; Cochrane *et al.*, 1989b). Moreover, only a small proportion of Rev molecules appears to undergo phosphorylation (Cochrane *et al.*, 1989b).

1.3.7. Activation domain

The existence of an 'activation' domain towards the C-terminal end of Rev was first noted by Malim *et al.* (1989b), who described the effects of several clustered point mutations of the *rev* sequence designed to produce codon substitutions. One mutant, M10, in which Leu 78 was substituted with Asp, and Glu 79 was substituted with Leu, was found to be phosphorylated and to localise in the nucleus normally, but to have a *trans*-dominant negative phenotype for Rev function - i.e., the protein lacked Rev activity and was able to inhibit the function of wild-type Rev in *trans*. This observation was later confirmed by Hope *et al.* (1990b) and Olsen *et al.* (1990b). Malim *et al.* (1989b) suggested that M10 might compete with Rev for its RNA substrate, or alternatively form mixed inactive multimers with wild-type protein. Subsequently, Mermer *et al.* (1990) showed that substitutions or insertions in the Leu 81-Asp 88 region produced dominant negative Rev mutants. Similarly, Olsen *et al.* (1990b) found that substitution of Leu 75 with Arg, or deletion of Arg 80-Leu 81-Thr 82 created mutants that retained the ability to oligomerise and bind RNA, but had *trans*-dominant negative phenotypes. Further confirmation of the effects of mutations in this region was provided by Venkatesh and co-workers (Venkatesh *et al.*, 1990; Venkatesh and Chinnadurai, 1990). They found that a Rev protein truncated at residue 73 was nonfunctional, whereas Rev truncated at residue 98 had wild-type activity. Subsequent experiments demonstrated that the Leu 73 truncated protein had a *trans*-dominant negative phenotype, as did individual substitutions of Leu 75 or Leu 81 (conserved residues in all known HIV-1 strains) with Asp; substitution of Leu 81 had a particularly strong dominant negative effect and completely inhibited the Rev response in cotransfection assays.

A more extensive analysis of the putative activation domain was carried out by Malim *et al.* (1991). Using a series of substitution mutants in the residue 67-92 region of Rev, they showed that leucines at positions 78, 81 and 83 were

functionally critical - substitutions of any or all of these residues with alanine induced a *trans*-dominant negative phenotype. Taken together with the results of Venkatesh and co-workers, these results defined a nine amino acid domain (residues 75-83) containing four leucines (residues 75, 78, 81, 83). Malim *et al.* (1991) also found that positioning of the leucine residues was crucial - a mutant in which Leu 83 and Glu 84 were reversed was dominant negative. Sequence comparisons between HIV-1 and six other lentiviruses showed that a nine amino acid domain containing four hydrophobic (usually leucine) residues is a conserved feature of lentivirus Rev proteins. Furthermore, Tiley *et al.* (1991) showed that the equivalent Rev activation domain of Visna (a lentivirus only distantly related to HIV) could functionally substitute for the HIV-1 Rev domain in a chimaeric protein. Similarly, Hope *et al.* (1991) and Weichselbraun *et al.* (1992) identified a corresponding region in the HTLV-I Rex protein, and demonstrated functional substitution of the Rev domain by the Rex domain.

Although it is clear that the leucine-rich activation domain is essential for Rev activity, the mechanism of action of the sequence is unknown. Presumably (as first noted by Malim *et al.*, 1989b), the activation domain is required for interaction with a cellular factor, since the sequence is dispensable for RNA binding, multimerisation, and nuclear localisation. Fankhauser *et al.* (1991) have shown that Rev specifically binds to the nucleolar protein B23, but did not establish which domain of the Rev protein was responsible for binding. The results of Trono and Baltimore (1990) suggest that a human cell factor (which might conceivably act by binding to Rev) is required for Rev function. In their experiments, Rev was nonfunctional in murine cell lines such as NIH 3T3, or in Chinese hamster ovary cells, but this defect could be complemented by fusion with uninfected human cells. However, in contrast to these results, Malim *et al.* (1991) have found that Rev is functional in avian (quail QCI-3), primate (COS and CV-1), and murine (Ltk$^-$ and NIH 3T3) cell lines, albeit with reduced efficiency in NIH 3T3 cells. Similarly,

Ivey-Hoyle and Rosenberg (1990) have shown that Rev-dependent expression of *env* gp160 can be observed in *Drosophila melanogaster* cells. These results suggest that any Rev cofactors, including potential Rev binding proteins, are likely to be found in a wide range of eukaryotic cells.

The identification of *trans*-dominant negative mutants of the Rev activation domain has led to proposals (first made by Malim *et al.*, 1989b) that such mutants could be used as specific and potent antiviral agents. The possible applications of this concept of 'intracellular immunisation', using *trans*-dominant negative mutants of Rev and other HIV proteins, have been reviewed by Feinberg and Trono (1992). Initial studies of the feasibility of this approach to somatic gene therapy have already been made. Bevec *et al.* (1992) have shown that HIV-1 replication is inhibited in cells of the human T-lymphocyte line CEM previously transduced with a retroviral vector expressing a *trans*-dominant negative Rev mutant.

The failure of *trans*-dominant negative Rev mutants to facilitate the cytoplasmic expression of Rev-responsive RNA species is presumably due to the failure of the activation domain to interact with a Rev cofactor. However the mechanism of inhibition of coexpressed wild-type Rev is less clear. As noted above, inhibition could conceivably occur by competition for the RNA substrate, or by the formation of inactive mixed multimers between mutant and wild-type proteins. The results of Hope *et al.* (1992) support the latter hypothesis. This group demonstrated that a *trans*-dominant negative mutant can inhibit wild-type Rev function even when confined to the cytoplasm by glucocorticoid receptor localisation signals - an effect which they attributed to the formation of inactive mixed multimers in the cytoplasm. However, very recently Bogerd and Greene (1993) found that *trans*-dominant negative mutants of the activation domain fail to multimerise *in vivo*, suggesting that inhibition is mediated by nonfunctional monomers. These authors suggested that an intact activation domain may be required for interaction with a

bridging cellular cofactor required (in addition to the intrinsic multimerisation domain) for the formation of functional and stable multimers *in vivo*.

It is becoming increasingly clear that the Rev protein can act at several levels of gene expression (1.6 and 1.7.5, below), but it is not yet known at what level(s) the function of the activation domain is important. Interestingly, Arrigo *et al.* (1992) have found that a Rev protein mutated in the leucine-rich domain can be co-immunoprecipitated with RRE-containing RNAs in both nuclear and cytoplasmic extracts, whereas wild-type Rev can only be coprecipitated with RNA in the cytoplasm. This suggests that the activation domain is directly involved in a Rev-mediated nuclear export mechanism. Kjems *et al.* (1991b) have shown that a peptide containing Rev residues 34-50 (and therefore lacking the activation domain) can inhibit splicing at the level of spliceosome formation in an *in vitro* assay. Thus the interaction of the activation domain with cellular factors may not be required for inhibition of splicing, but may instead be necessary for other activities such as the apparent effect on transport noted by Arrigo *et al.* (1992).

1.4. The Rev responsive element (RRE)

1.4.1. RRE function and location

Deletion analyses of subgenomic proviral DNA cloned into CAT reporter constructs (Dayton *et al.*, 1988; Rosen *et al.*, 1988), of infectious molecular clones (Hadzopoulou-Cladaras *et al.*, 1989; Arrigo *et al.*, 1989), and of *env* and *tat* gene expression vectors (Malim *et al.*, 1989a, Hammarskjöld *et al.*, 1989) defined a *cis*-acting intragenic sequence, the Rev responsive element (RRE), required in conjunction with Rev for the expression of viral structural proteins. Complementary experiments involving the insertion of sections of the *env* gene into *gag*, *tat*, and *env*-gp120 expression vectors produced similar results, and also demonstrated that RRE function was retained when the element was moved, but not when it was inverted (Emerman *et al.*, 1989; Hadzopoulou-Cladaras *et al.*, 1989; Hammarskjöld

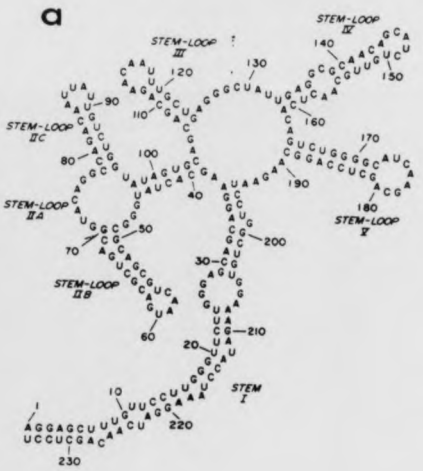
et al., 1989; Malim *et al.*, 1989a). The results of the deletion and insertion experiments, combined with RNA secondary structure predictions (Malim *et al.*, 1989a) localised the RRE to a 234 nt element in the gp41 region (transmembrane protein coding sequence) of the *env* gene, immediately 3' to the gp120/gp41 cleavage site.

1.4.2. Rev binding ability of the RRE

Rev protein has been shown to bind the RRE RNA directly by filter binding assays (Daly *et al.*, 1989; Heaphy *et al.*, 1990; Holland *et al.*, 1990; Nalin *et al.*, 1990), gel mobility shift assays (Daly *et al.*, 1989; Daefler *et al.*, 1990; Heaphy *et al.*, 1990; Holland *et al.*, 1990; Malim *et al.*, 1990), gel mobility shift following ribonuclease protection by Rev (Zapp and Green, 1989; Cochrane *et al.*, 1990b; Heaphy *et al.*, 1990), and immunoprecipitation with Rev antisera (Daefler *et al.*, 1990; Heaphy *et al.*, 1990). The dissociation constant of the Rev-RRE interaction has been variously estimated as 3×10^{-10} M (Daly *et al.*, 1989), $1\text{-}3 \times 10^{-9}$ M (Heaphy *et al.*, 1990), 3.5×10^{-9} M (Nalin *et al.*, 1990), and 5×10^{-9} M (Holland *et al.*, 1990) from filter binding and immunoprecipitation data. These relatively consistent results indicated that Rev-RRE binding is a high affinity interaction.

1.4.3. RRE secondary structure and functional domains

RNA folding algorithms predict that the RRE has a complex secondary structure of stem-loops (Malim *et al.*, 1989a; Dayton *et al.*, 1989). Statistical methods (Le *et al.*, 1990) have demonstrated that the secondary structure of this region is highly stable and conserved in 25 different HIV-1 sequence isolates (with similar structures found in HIV-2 and SIV). One predicted structure for the RRE (from Malim *et al.*, 1990) is reproduced in fig. 1.3. The domain notation and numbering of the RRE used in this diagram is used throughout the following discussion for ease of comparison between models. Several studies have attempted to refine the RRE model by correlating the predicted structure with information from ribonuclease


a
STEM-LOOP III
STEM-LOOP IV
STEM-LOOP IIC
STEM-LOOP IIA
STEM-LOOP IIB
STEM-LOOP V
STEM I
1
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230

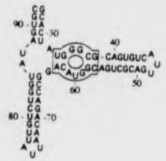

b
CAGUGUC
GCAGUCGCAG

Fig. 1.3. Predicted secondary structure of the HIV-1 Rev-responsive element (RRE) (a) Reference structure (reproduced from Malim *et al.*, 1990). By convention the structure is divided into five RNA domains termed stem I (1-37; 195-234), stem-loop II (39-104), stem-loop III (107-124), stem-loop IV (133-161) and stem-loop V (164-189). Stem-loop II is subdivided into stem-loop IIA (39-48; 71-77; 98-104), stem-loop IIB (49-70) and stem-loop IIC (78-96). **(b) Amended structure of the stem-loop II region** (reproduced from Heaphy *et al.*, 1991). This structure differs from that of Malim *et al.* (1990) in containing a purine-rich 'bubble' (outlined) now known to be required for high-affinity Rev-RRE binding (see 1.4.4, below). Nucleotide 30 in the Heaphy structure corresponds to nucleotide 42 in the Malim structure.

digestion and chemical probing experiments, and by examining the effects of specific mutations on RRE function. These studies have generally focused on determining the precise location and secondary structure of the primary Rev binding site and other potential functional domains of the RRE. The results of Cochrane *et al.* (1990b) indicated that formation of the major stem I domain is required for Rev binding. They suggested, however, that stem I need not itself bind Rev, but could rather function to allow the formation of the other stem-loops, one of which could contain the primary Rev binding site. Heaphy *et al.* (1990) mapped the Rev binding region to a 71 nt section of the RRE including stem-loops IIA, IIB and IIC of the structure of Malim *et al.* (1990); however, they proposed a significantly different secondary structure for the domain II region. Holland *et al.* (1990) and Malim *et al.* (1990) came to similar conclusions on the location of the Rev binding site from deletion mapping experiments. Holland *et al.* (1990) also noted the importance of a 9 nt sequence (CACUAUGGG) at the 5' end of stem-loop IIA, mutations in which were found to abolish the Rev response *in vivo* and Rev binding *in vitro*.

Holland *et al.* (1990) also demonstrated that Rev binding could be divorced from Rev response by analysis of an unresponsive Rev-binding RRE truncated at the 3' end of stem-loop III. This implied that the RRE contains separate Rev-binding and other functional domains. However, Huang *et al.* (1991) showed that an 88 nt truncated RRE, comprising stem-loop II fused to an 8 bp sequence from the base of stem I, was capable of mediating the *in vivo* Rev response. Furthermore, they found that two tandem copies of this minimal RRE functioned as efficiently as the full-length element, indicating that the other domains (stem-loops III-V, and most of stem I) were dispensable for the Rev response. Other investigators (McDonald *et al.*, 1992; Venkatesan *et al.*, 1992) have shown that the RRE can be completely dispensed with if Rev is bound to HIV RNA via a heterologous binding protein/RNA target site: A fusion protein of Rev and phage MS2 coat protein binds an HIV RNA in which the MS2 operator sequence replaces the RRE, and

upregulates cytoplasmic expression of incompletely spliced mRNA (albeit with reduced activity compared to the normal system). Nevertheless, the RRE domains (III-V) which appear to be dispensable for Rev binding and response may have important functions *in vivo*. For example, Constantoulakis *et al.* (1993) have isolated a cellular interferon-inducible protein, RBP9-27, that specifically binds the RRE in the domain III-V region and inhibits Rev-mediated gene expression. Intriguingly, they have suggested that HIV may utilise this cellular factor to downregulate its own gene expression (possibly maintaining a latent infection).

Olsen *et al.* (1990a) have mapped a region required for Rev binding *in vitro* and Rev response *in vivo* to stem-loop IIB, but have suggested that Rev probably has several contact points on the RRE, including the major stem I. They also found that limited mutations of the RRE which were not predicted to cause changes in the secondary structure were generally well tolerated for Rev binding and Rev response - indicating that secondary, rather than purely primary structure, is the major determinant for Rev-RRE interaction. However, Dayton *et al.* (1992) found that an entirely 'base-switched' RRE was inactive. (i.e., they prepared and analysed a construct in which the primary structure was completely changed - generally by replacing bases with their respective Watson-Crick pairing partners - but the predicted secondary structure was maintained). More limited base switching demonstrated that primary structure was particularly important for RRE activity in the stem-loops of domain II. Cook *et al.* (1991) have also produced results which identify domain II as the Rev binding region, from experiments which studied Rev binding to truncated RRE fragments. They identified a 40 nt minimal Rev binding substrate including stem-loop IIB and flanking nucleotides in stem-loop IIA which was capable of binding Rev independently (and was contained within the 71 nt Rev binding fragment identified by Heaphy *et al.*, 1990).

1.4.4. Rev binding site

Several groups have attempted to identify precisely the Rev binding site within the RRE. Kjems *et al.* (1991a) probed the RRE with reagents specific to single or double-stranded regions in the presence or absence of bound Rev. While largely confirming the structure of Malim *et al.* (1989a), their experiments suggested the presence of five potential Rev binding sites in the double-stranded regions of stem-loops I, IIA, IIB, IIC, and III. These sites were characterised by the ability of bound Rev to protect nucleotides from RNase cleavage or chemical modification, and contained a conserved G-C, A-U 2 bp sequence within a 6 bp stem flanked by a single-stranded region of 3 nt. Kjems *et al.* (1991a) also noted a lack of sensitivity to double-stranded specific nuclease activity in stem I, and to single-stranded specific reagents in the loop of IIA between IIB and IIC. This was attributed to tertiary RRE structure - increased accessibility of these regions on Rev binding suggested a conformational change of the RRE RNA.

In contrast to Kjems *et al.* (1991a), most investigators have proposed a single binding site in the RRE for Rev. Heaphy *et al.* (1991), using data from an alternative (Zuker) RNA folding prediction algorithm from that used by Malim *et al.* (1989a), combined with a reinterpretation of the results of Kjems *et al.* (1991a), proposed a novel secondary structure for the RRE (fig. 1.3b). This differed markedly in the stem-loop II domain from the predicted structure of Malim *et al.* (1989a) and the modified structure originally proposed by Heaphy *et al.* (1990). In the new structure, the Malim stem-loop IIB and part of the single-stranded region of stem-loop IIA were replaced by an extended stem-loop starting with paired U45-A75 residues (Malim notation), and containing unpaired GG and GUA residues in a 'purine-rich bubble'. Heaphy *et al.* (1991) found that the 13 nts (8 paired, 5 'bulged') of the purine-rich bubble were sufficient to bind Rev with wild-type affinity when stabilised within a heterologous RNA hairpin structure. They showed that mutations of the bubble nucleotides reduced or abolished Rev binding and,

noting the stability and resistance of the binding element to chemical modification discovered by Kjems *et al.* (1991a), suggested possible stabilisation of the structure (and local helix distortion) by a non-Watson-Crick G47-A73 base pair. Using different techniques, Bartel *et al.* (1991) predicted a similar structure for the RRE, including a similar bulged stem-loop domain. This group used Rev binding ability to identify residues critical for binding by *in vitro* genetic selection from a pool of 10^{13} artificially generated random variants of the stem-loop II Rev binding domain. They obtained results consistent with a bulged stem-loop model, and also noted the covariation of conserved residues G48 and G71, which could be functionally replaced by adenine at these positions, but not by other bases. This suggested the functional requirement for a noncanonical purine base pair in the Rev binding domain which, like the G47-A73 base pair suggested by Heaphy *et al.* (1991), Bartel *et al.* (1991) and Dayton *et al.* (1992), could act by locally distorting the RNA backbone - a possible requirement for protein binding.

The crucial role of the purine residues of the RNA bulge was independently confirmed by Tiley *et al.* (1992). They tested the effects of diethylpyrocarbonate (DEPC) modification of purine residues on Rev binding, and showed that the residues of the bulge were crucial for Rev binding. This information, combined with analysis of the secondary structure of a truncated 29 nt Rev binding fragment of the RRE (including the essential purine residues) led them to propose a secondary structure for the Rev binding domain essentially identical to that of Heaphy *et al.* (1991). Moreover, the application of phylogenetic analysis and Zuker RNA secondary structure prediction to the RREs of 25 different HIV-1 isolates has demonstrated that a base pairing scheme consistent with an RNA bubble containing a putative A-G base pair is a conserved feature (Dayton *et al.*, 1992).

Recently, Iwai *et al.* (1992) have found by functional group modification of a 29 nt RNA duplex containing the Rev binding bubble, that the G48-G71 base pair exists

in an *anti : syn* configuration and is essential for Rev binding; similarly the G47-A73 pair exists in either an *anti : syn* or *anti : anti* configuration and is also required for binding. Both Iwai *et al.* (1992) and Kjems *et al.* (1992) have suggested three dimensional structure is important in the Rev-RRE interaction. DEPC modification (Kjems *et al.*, 1992) or substitution (Iwai *et al.*, 1992) at the N^7 position of purines in the bubble region interfered with Rev binding - a characteristic of protein interaction with the major groove of a nucleic acid helix. Thus the RRE bubble could function by distorting the helix to expose the normally inaccessible groove for Rev binding. However, Holland *et al.* (1992) have concluded that the secondary and three dimensional structures of the Rev binding element are irrelevant, provided that the CACUAUGGG (nt 40-48) sequence (originally identified by Holland *et al.*, 1990 - see above) is preserved, and at least one of the G residues remains unpaired. They found that this sequence, optimally presented in the context of a stem-bulge-stem structure, is sufficient for Rev recognition, and is required for the Rev response.

1.5. Stoichiometry of Rev-RRE binding and Rev oligomerisation

There is now clear evidence that multiple Rev molecules bind the RRE. Several groups have demonstrated the existence of multiple Rev-RRE complexes of differing electrophoretic mobilities, which were interpreted as containing a varying number of Rev molecules bound to each response element (Daly *et al.*, 1989; Heaphy *et al.*, 1990; Olsen *et al.*, 1990b; Cook *et al.*, 1991). Scintillation counting of labelled protein and RNA from gel-isolated Rev-RRE complexes (Cook *et al.*, 1991) and filter binding assays (Daly *et al.*, 1989; Heaphy *et al.*, 1991) have confirmed this interpretation - Daly *et al.* (1989) have estimated that up to 8 Rev molecules bind the RRE. Although it is possible that the RRE contains a number of discrete Rev binding elements (Kjems *et al.*, 1991a), gel mobility shift assays of Rev binding to the stem-loop IIB high affinity site inserted in heterologous RNA constructs have shown that the high affinity site is necessary and sufficient for

binding multiple Rev molecules (Iwai *et al.*, 1992). Similarly, chemical interference mapping with diethylpyrocarbonate (Tiley *et al.*, 1992) has shown that binding of more than one Rev molecule requires only the stem-loop IIB site - suggesting that binding of multiple Rev molecules is mediated by protein-protein interactions.

The intrinsic ability of Rev molecules to multimerise even in the absence of RNA has been demonstrated by a variety of methods, and (as noted above, 1.3.4) the domains responsible for multimerisation have been mapped. Experiments using gel filtration chromatography, chemical cross-linking, and density gradient centrifugation (Nalin *et al.*, 1990; Wingfield *et al.*, 1991; Zapp *et al.*, 1991) have shown that Rev can exist as stable multimers is solution; Rev species with molecular masses consistent with the protein existing as a dimer, trimer and tetramer were reported by these groups. Heaphy *et al.* (1991), using electron microscopy, reported the formation of hollow filaments with a diameter of about 14 nm and a length of up to 1500 nm in preparations of purified Rev at concentrations greater than 100 μg/ml; in the presence of RRE RNA, short rod-like ribonucleoprotein complexes were formed with a preferred length of 60 nm. Similarly, Wingfield *et al.* (1991) described the formation of hollow filaments with a diameter of 20 nm at Rev concentrations above 80 μg/ml; in the presence of RRE RNA, filaments 8 nm in diameter and up to 0.12 μm long could be formed at Rev concentrations as low as 40 μg/ml. Both groups have suggested that the ability of Rev to polymerise and form ribonucleoprotein complexes could be important in the mechanism of Rev action. A 'coating' of Rev molecules on the RNA could directly block splicing factor access to splice sites, or alternatively channel the transcripts into a ribonucleoprotein transport pathway. In support of this theory, Heaphy *et al.* (1991) found that Rev filaments can extend over hundreds of nucleotides of template RNA and coat the entire molecule, a process facilitated by 'nucleation' at an RRE sequence and probably by low affinity nonspecific binding to RNA. However, observations of filament formation was restricted to *in vitro* systems.

At present, it is unclear if prior Rev multimerisation is a prerequisite for RRE binding. Cook *et al.* (1991) have shown by scintillation counting of complexes between ^{125}I-labelled Rev, and a ^{32}P-labelled RRE fragment containing the (high affinity) stem-loop IIB binding site, that monomeric Rev can bind the RRE. Multiple Rev molecules were shown to bind the complete RRE sequence, but the results from the subfragment binding experiment indicated that multimerisation is not essential for binding. Furthermore, Malim and Cullen (1991) isolated Rev mutants which were able to bind the RRE, but not multimerise. However, Olsen *et al.* (1990b) found that mutants which were unable to multimerise (including one mutant with a deletion outside the putative RNA binding domain of Rev) did not bind RNA. Similarly, Zapp *et al.* (1991), in experiments which included the mutants isolated by Malim and Cullen, found that Rev molecules which were unable to multimerise did not bind RNA. However, both Olsen *et al.* (1990b) and Zapp *et al.* (1991) assayed Rev-RRE binding by an RNase protection gel mobility shift assay, rather than the simple mobility shift assay with undigested RNA employed by Malim and Cullen; the RNase protection method may be a more stringent assay for efficient Rev binding.

Whether or not multimerisation is necessary for binding, the ability of Rev molecules to multimerise is clearly essential for Rev function, since the multimerisation-deficient mutants which Malim and Cullen (1991) reported to bind the RRE were deficient in Rev activity. It is also probable that Rev multimerisation occurs before nuclear import, since Rev mutants of the activation domain with a *trans*-dominant negative phenotype were able to inhibit wild-type Rev function even when the mutant protein was restricted to the cytoplasm by fusion with glucocorticoid receptor localisation signals (Hope *et al.*, 1992) - presumably by forming inactive protein complexes in the cytoplasm. Similarly, dominant negative mutants of the Rev nucleolar localisation signal which retained the ability to

multimerise were able to inhibit Rev function while remaining in the cytoplasm (Kubota *et al.*, 1992).

Although there is some conflict in the results summarised above, it is possible to suggest a mechanism which is relatively consistent with the published data. After translation, Rev molecules associate to form small oligomers which are transported to the nucleus. A single molecule is then released from an oligomer on binding the RRE at the high affinity site; alternatively the unit RRE binding species may be a small Rev oligomer. Additional molecules then bind by protein-protein interactions with the initial 'nucleating' molecule at the high affinity site, in the type of ordered assembly process proposed by Heaphy *et al.* (1991), eventually forming a multimeric ribonucleoprotein complex. Presumably, this process is facilitated by nonspecific RNA binding, and possibly by interaction with the additional binding sites described by Kjems *et al.* (1991a).

1.6. *Cis*-acting repression sequences

Although the RRE is necessary for Rev-facilitated cytoplasmic expression of Rev-responsive genes, it is not responsible for repressing expression of these RNAs in the absence of Rev. Other less well-defined RNA sequences are responsible for this phenomenon. According to one model of the mechanism of Rev action (see 1.7.2, below, and reviews by Chang and Sharp, 1990; Cullen and Malim, 1991), 'inefficient' HIV splice sites are necessary and sufficient to make RNA species containing the RRE dependent on Rev for expression. It is envisaged that such inefficient splice sites can interact with spliceosome components, but are only processed comparatively slowly. This allows intervention by Rev-RRE interaction to take place before splicing can occur - Rev binding may lead to premature dissociation of spliceosome components, releasing the unspliced mRNA for transport. In support of this model, Chang and Sharp (1989) have shown that constitutive cytoplasmic expression of spliced ß-globin mRNA from a construct

containing the RRE can be altered to Rev-responsive cytoplasmic expression of unspliced mRNA by replacing the ß-globin splice sites with HIV-1 *tat* splice sites, or by introducing specific point mutations in either the 5' or 3' ß-globin splice site. Similarly, Itoh *et al.* (1989) have demonstrated that Rev can overcome splice donor mediated repression of cytoplasmic expression from a chimaeric HTLV-I construct containing the RRE.

Several groups have identified other HIV-1 sequences (termed *cis*-acting repression sequences (CRS) by Dayton *et al.*, 1988) which apparently prevent cytoplasmic expression of mRNAs in the absence of Rev/RRE interaction. The existence of such sequences was first suggested by Sodroski *et al.* (1986). Subsequently, it was shown by deletion analysis of *env*-CAT/human growth hormone reporter gene constructs (Rosen *et al.*, 1988) and of proviral expression plasmids (Dayton *et al.*, 1988) that such sequences appeared to be present in the *env* and *gag* genes. The *gag* gene sequence and at least one of the *env* gene sequences clearly did not overlap with the Rev-response element, and the effect of the *env* gene sequences was retained when known splice donor and acceptor sites were deleted. However, Malim *et al.* (1989a) failed to detect truncated Tat expression - the expected product of cytoplasmic expression of unspliced RNA in their assay - when the *env* sequences defined by Rosen *et al.* (1988) were deleted from an expression vector, suggesting that other sequences were involved in nuclear RNA retention. Emerman *et al.* (1989) identified an additional *cis*-acting negative regulatory sequence at the 3' end of the *env* gene by adding *env* sequences 3' to the stop codon of a truncated *env* (gp120 only) expression vector. Again, the effect was observed in the assumed absence of splicing, as the *tat*/*rev* 5' splice site was removed. The presence of a novel *cis*-acting repression sequence in *gag*, and of a sequence in *env* overlapping those defined by Rosen *et al.* (1988) and Emerman *et al.* (1989) has been reported by Hadzopoulou-Cladaras *et al.* (1989). Maldarelli *et al.* (1991) have located *cis*-acting regions in a 1295 nt section of *gag* and a 1932 nt section of *pol* which can inhibit

the cytoplasmic expression of a heterologous CAT reporter gene. Cochrane *et al.* (1991) have identified *cis*-acting repression sequences in both *gag* and the 3' end of *pol*. Their experiments, based on adding HIV-1 sequences downstream of the CAT gene in a reporter construct, showed that the 260 nt *pol* element had a particularly strong repressive effect which could, however, be reversed by the addition of the RRE in *cis* and Rev in *trans*. In contrast to the other groups, they attributed the mechanism of repression to an effect on mRNA utilisation by cellular translation machinery, rather than mRNA transport, since the subcellular distribution of viral mRNA appeared to be unaffected. Another mechanism of *cis*-acting inhibition of expression has been proposed by Schwartz *et al.* (1992a, 1992b). These workers have found a 218 nt AU-rich inhibitory sequence (INS-1) at the 5' end of *gag* which appears to reduce the stability of *gag* RNA. They showed that the inhibitory effect of the sequence could be reversed either by the presence of Rev and the RRE, or by a series of 28 silent point mutations over a 270 nt region of *gag*. An additional element in the *gag-pol* region was also detected.

It is not yet clear if the specific repression sequence model is compatible with the inefficient splice site recognition model. Known splice sites were not excluded by all groups which favoured specific repression sequences. For example, although Emerman *et al.* (1989) removed the 5' splice site to prevent splicing, the 3' splice site (as shown by Chang and Sharp, 1989) could still mediate nuclear retention of RNA. In experiments where all known functional splice sites were removed, it is possible that the repression sequences represented cryptic or unused splice sites capable of interacting with splicing components. Alternatively, sequences which are not themselves splice sites may function by interacting with components of the splicing machinery. Olsen *et al.* (1992) have shown that the *pol* repressive element defined by Cochrane *et al.* (1991) interacts strongly and specifically with at least three cellular factors, including one which cross-reacts with a monoclonal antibody to the hnRNP C protein, a factor believed to be involved in the splicing pathway.

Moreover, this interaction was blocked by the addition of a heterologous intron sequence. Possibly a number of elements, including splice sites and more specific *cis*-acting repression sequences, have inhibitory activities which can be countered by the Rev/RRE system.

It is more difficult to accommodate the results of Cochrane *et al.* (1991) and Schwartz *et al.* (1992a, b) within this framework, since the repression sequences they defined appeared to act at the levels of mRNA utilisation in the cytoplasm and mRNA stabilisation, respectively. However, repression and its Rev-dependent relief may act at more than one level of mRNA metabolism. It is possible to envisage a situation where the Rev/RRE system counters multiple levels of repression exerted by a variety of elements responsible for nuclear retention, destabilisation, and poor translatability of RNA in the absence of Rev. There is some evidence that different mechanisms predominate in different cell types. For example, Arrigo and Chen (1991) have shown that in lymphoid cells, the principal effect of Rev is to allow the association of exported unspliced mRNAs with polysomes and their subsequent translation. Possibly, an adaptable system able to function by any of several different mechanisms is required for HIV tropism in a variety of cells.

1.7. Mechanisms of Rev action

1.7.1. Effects on mRNA transport

A mechanism for Rev action based on a direct effect on the mRNA transport pathway has been proposed by several groups. Felber *et al.* (1989), while noting a Rev-mediated effect on mRNA stability, concluded that Rev also functioned at the level of mRNA transport. Their results showed that Rev was required for the cytoplasmic accumulation of unspliced RNAs carrying the RRE. Rev did not appear to affect nuclear levels of these mRNAs. The effect on cytoplasmic accumulation was presumed to be independent of splicing since it persisted with proviral constructs from which all known functional splice sites were removed. Felber *et al.*

(1989) proposed a mechanism in which Rev action channelled unspliced RNAs away from splicing (or degradation in the case of unspliceable mutant RNAs) and towards nuclear export. Similar conclusions were reached by Malim *et al.* (1989a), who showed that a Rev response was maintained with a non-spliceable *env* construct, and when the RRE was moved to various exonic and intronic locations within *env*. They suggested that Rev functioned to allow the interaction of incompletely spliced RNAs with nuclear export machinery. Comparable results were also obtained by Arrigo *et al.* (1989), who noted that the transport facilitation effect might be indirect - i.e., Rev might allow transport by blocking spliceosome formation, permitting nuclear export (spliceosomes might still be formed even in constructs where known splice donors were deleted by the activation of cryptic splice sites). A similar mechanism has been suggested by Emerman *et al.* (1989). Direct evidence for this type of mechanism is discussed below (1.7.2). Results broadly consistent with those of Felber *et al.* (1989), Malim *et al.* (1989a) and Arrigo *et al.* (1989) have been obtained by Hammarskjold *et al.* (1989) and Emerman *et al.* (1989). However, Hammarskjold *et al.* (1989) found that incompletely spliced RNA accumulates in the nucleus in the absence of Rev, rather than being unaffected. Either result could be consistent with a facilitation of transport mechanism - incompletely spliced RNA could either be further spliced, degraded in the nucleus, or retained to accumulate. The differing results may reflect the different expression systems and methods of analysis employed by these groups. Recently, Arrigo *et al.* (1992) have demonstrated that RNA containing the RRE coimmunoprecipitates with Rev protein in the cytoplasm, but cannot be detected in the nucleus unless Rev is replaced by a nonfunctional protein mutated in the activation domain. This suggests that Rev is involved directly in rapid nuclear export of incompletely spliced RNA; an indirect (splicing inhibition) mechanism might be expected to show relatively high levels of incompletely spliced RNA bound to Rev in the nucleus.

1.7.2. Effects on mRNA splicing

A direct effect on pre-mRNA splicing which served to regulate the relative proportions of viral transcripts was first proposed as a mechanism for Rev action by Feinberg *et al.* (1986). Subsequently, Gutman and Goldenberg (1988) showed that an extract from HIV-1 infected cells specifically inhibited splicing of an Sp6/HIV-1 *env* transcript by preventing spliceosome assembly. However, the factor responsible for this effect was not identified. As noted above (1.6) Chang and Sharp (1989) demonstrated that Rev/RRE-mediated regulation depended on splice site recognition. Their results showed that relatively inefficient splice sites that are processed comparatively slowly are required for dependence of an mRNA on Rev-facilitated cytoplasmic accumulation mediated by the RRE. Itoh *et al.* (1989) have shown that Rev in *trans* and the RRE in *cis* can allow cytoplasmic expression of incompletely spliced RNA from a chimaeric HTLV/HIV construct; in the absence of Rev, expression is prevented by an HTLV splice donor. Thus Rev is capable of overcoming splice-site mediated repression of gene expression in a heterologous system, suggesting that the protein may have a similar role in HIV. Moreover, Lu *et al.* (1990) have demonstrated that Rev-mediated expression of Env proteins from an SV40-based vector requires an intact splice donor.

Kjems *et al.* (1991b) have shown that Rev can specifically inhibit the splicing of RNA containing the RRE by 3- to 4-fold *in vitro*. A 17 amino acid peptide containing the Rev RNA binding domain (residues 34-50) was also functional in the splicing inhibition assay - surprisingly with greater activity than the whole protein. Rev requires an activation domain (1.3.7) for functional activity *in vivo* which would not be present in the 34-50 peptide, suggesting either that the *in vitro* experiments do not adequately reflect the situation *in vivo*, or that Rev has a number of important activities, one of which can be isolated in the splicing inhibition assay. This assay also showed that Rev could inhibit splicing from constructs containing the RRE and an adenovirus major late intron, or a ß-globin gene with wild-type

splice sites, a result which contrasts with the *in vivo* findings of Chang and Sharp (1989). Again, these results probably reflected the artificial situation of the *in vitro* system, in which splicing occurs much less efficiently than *in vivo*.

The precise mechanism by which Rev-RRE interaction might inhibit the splicing reaction is not yet clear. Kjems *et al.* (1991b) found that the Rev 34-50 peptide was only an active inhibitor if added before or immediately after adding nuclear extract to the splicing reaction. This suggested a role in preventing spliceosome assembly. Furthermore, sucrose gradient fractionation demonstrated that complete 60S spliceosomes were not formed on transcripts containing the RRE in the presence of Rev 34-50. Gutman and Goldenberg (1988) also found that spliceosome formation is inhibited in reactions using extracts from HIV1-infected cells. It is possible that Rev blocks spliceosome formation and hence the splicing reaction by directly interfering with the binding of splicing components at the splice sites. Indeed, Wingfield *et al.* (1991) and Heaphy *et al.* (1991) have proposed that an RRE-nucleated coating of Rev molecules on incompletely spliced RNA (forming the ribonucleoprotein filaments observed by these groups *in vitro*) could directly block the splice sites and protect the RNA from splicing. Equally, the formation of a specific ribonucleoprotein might be important in directing the RNA to a facilitated nuclear export pathway, rather than (or as well as) blocking splicing. It should be noted that Kjems *et al.* (1991b) saw no decrease in accessibility to RNase T1 of RNA bound to the Rev 34-50 peptide, despite significant inhibition of splicing, suggesting that the effect observed *in vitro* was independent of protein coating. An alternative possibility is that conformational changes caused by Rev binding may directly inhibit splicing or form binding sites for inhibitory factors. Differences in the circular dichroism spectra of free RRE RNA and bound Rev-RRE complexes (Daly *et al.*, 1990) suggest that such changes do occur.

1.7.3. Effects on mRNA stability

Evidence for an effect of Rev on HIV mRNA stability has been described by Felber *et al.* (1989). This group found that Rev both promoted the transport (1.7.1, above) and increased the half-life of unspliced viral mRNAs carrying the RRE. The effect on stability, studied with constructs from which transcription was arrested with actinomycin D, was found to be independent of splicing and RNA encapsidation by Gag proteins. Schwartz *et al.* (1992a, b) isolated a 218 nt inhibitory sequence at the 5' end of the *gag* gene which appears to act by destabilising *gag* RNA in the absence of Rev. This sequence was found to have an unusually high AU content (61.5% AU) compared to *tat* and most cellular RNAs. Within the inhibitory sequence, two very AU-rich subregions were identified; deletion of one of these abolished the inhibitory effect, suggesting that AU content may be an important determinant of the stability of these RNA species. AU-rich elements have been described as determinants of instability in several cellular mRNAs, one example being granulocyte-macrophage colony stimulating factor (Shaw and Kamen, 1986). If such elements are present in the other incompletely spliced mRNAs, then RNA stabilisation may be a mechanism of Rev action.

1.7.4. Effects on mRNA utilisation

A mechanism involving an effect on mRNA utilisation was first proposed for Rev action by Sodroski *et al.* (1986). Subsequently, Knight *et al.* (1987) concluded that Rev acted at a translational level from experiments in which viral RNA was quantitated in the presence and absence of Rev expression. However, Knight *et al.* (1987) measured total cellular rather than fractionated nuclear and cytoplasmic RNA levels, and therefore did not distinguish translational from transport or splicing mechanisms. More recently, Arrigo and Chen (1991), using quantitative PCR techniques, showed that in lymphoid 729 cells, Rev had no significant effect on the cytoplasmic accumulation of singly spliced (*env*/*vpu*, *vif*, *vpr*) mRNA species, and an effect on cytoplasmic accumulation of unspliced (*gag*/*pol*) mRNA was not

sufficient to explain observed Rev-mediated increases in Gag protein expression. Rev was, however, required for the association of these singly-spliced and unspliced RNA species with polysomes. In the absence of Rev, such RNAs were found to be associated with a 40S-80S complex which it was speculated might interact with CRS elements in HIV RNA, preventing polysome loading unless released by Rev action; alternatively, the complexes could simply be single ribosomes or their subunits. Consistent with the observations, Cochrane *et al.* (1991) defined a CRS element in the *pol* gene that acted at the level of mRNA utilisation (1.6, above). D'Agostino *et al.* (1992) have also reported that Rev acts at the level of mRNA utilisation. In their experiments, carried out in Tat-expressing HeLa cells, Rev expression caused an 800-fold increase in the levels of Gag protein expression, but only a 4-16-fold increase in the cytoplasmic level of *gag* mRNA. The *gag* mRNA produced in the absence of Rev was found to be fully translatable in a cell-free system but was not associated with polysomes in the absence of Rev. Even in the presence of Rev, a notably lower proportion of the total *gag* mRNA was polysome-associated than was the case for a control ß-globin RNA. This suggested that the HIV RNA might have reached the cytoplasm by a (possibly Rev-mediated) transport pathway distinct from that followed by cellular RNAs, leading to slower association with translation machinery. These authors also examined the subcellular distribution of *gag* mRNA by *in situ* hybridisation and found that, in the absence of Rev, this species was restricted to the nucleus and the area of the cytoplasm surrounding it. This indicated that the *gag* mRNA might be sequestered in a specific cytoplasmic subcompartment, perhaps by binding to a cellular factor via a CRS sequence; release from this compartment might therefore be mediated by Rev-RRE binding.

The conclusion that Rev is active at the level of mRNA utilisation implies that Rev protein remains bound to the RRE in the cytoplasm. This has been confirmed by coimmunoprecipitation of Rev with the RRE from cytoplasmic cell extracts (Arrigo *et al.*, 1992).

1.7.5. Combined mechanisms

It is clear from the preceding sections that the mechanism of action of the Rev protein is still a controversial subject. It should, however, be noted that the mechanisms discussed above are not necessarily mutually exclusive. As noted in 1.6, Rev may be a multifunctional protein active at several different levels of expression. Indeed, the apparently distinct processes of splicing, nuclear export, mRNA stabilisation and translation may be artificially divided stages of a continuous, unified process of gene expression.

Given the data discussed above, it is possible to suggest a plausible combined mechanism in which Rev functions at more than one level of expression: In the absence of Rev, incompletely-spliced transcripts are either further spliced or retained in the nucleus by interaction of their splice sites and/or *cis*-acting repressive sequences with cellular factors (which may be components of the splicing machinery). However, when the *rev* gene is expressed, multiple Rev molecules bind the RRE and may then cause premature spliceosome dissociation and/or block splice sites directly by polymerisation along the transcripts. Alternatively, Rev binding may allow transcript release from some other mechanism that prevents nuclear export of incompletely-spliced transcripts (possibly involving a distinct cellular factor or compartmentalisation within the nucleus). A third (and not exclusive) possibility is that Rev packages transcripts for interaction with a facilitated nuclear export mechanism (Rev has the potential to interact with splicing components, other nuclear factors, or transport machinery via its activation domain). Rev continues to be associated with exported RNAs in the cytoplasm where it may stabilise the transcripts and/or form part of a translation complex and/or facilitate release from a cytoplasmic subcompartment, allowing polysome loading. Alternatively, by allowing transport into the cytoplasm through a facilitated nuclear export pathway, Rev may locate the RNA in a particularly favourable site for translation.

1.8. Replication cycle and gene expression of adenovirus

1.8.1. Introduction

Adenoviruses are the aetiological agents of a variety of diseases in mammals and birds. In humans, conditions caused by infection with particular adenovirus serotypes include a number of acute respiratory illnesses, epidemic keratoconjunctivitis, and acute haemorrhagic cystitis (reviewed in Horwitz, 1990b). Some human adenoviruses have also been shown to cause malignant tumours in rodents, but these viruses do not appear to be oncogenic in humans.

The known serotypes of human adenoviruses have been classified into six subgroups (A-F) based on their haemagglutination properties, oncogenic potential in animals, ability to transform cells in tissue culture, and GC base content (reviewed in Horwitz, 1990a). Human adenovirus serotype 5 (Ad5), the virus used in this study, belongs to subgroup C. This group is characterised by partial agglutination of rat erythrocytes, a relatively high (57-59%) GC content, the ability to transform cells in tissue culture, and the inability to induce tumours in animals. Of the four serotypes in this subgroup, two - Ad2 and Ad5 - have been completely sequenced. The 35938 bp Ad5 sequence (completed by Chroboczek *et al.*, 1992, and amended by Dix and Leppard, 1992) shows 94.7% homology with the sequence of Ad2. Detailed sequence comparison has suggested that the Ad2 genome may have evolved from the highly related Ad5 sequence, primarily by nucleotide transitions, and probably by recombination of the fibre gene (Chroboczek *et al.*, 1992). As these two serotypes are so closely related, it is generally assumed in this thesis, as elsewhere, that the results of experiments conducted with Ad2 can be applied to Ad5; no distinction between the serotypes is usually made in the discussion of these results.

1.8.2. Structure and life cycle of Ad5

The structure and replication cycle of Ad2/5 have been well characterised (see review by Horwitz, 1990a, and references therein). The double-stranded DNA genome is bound to a 55K terminal protein at each 5' terminus. The genome is associated with viral core proteins V, VII and μ, and the core is in turn encased in a regular nonenveloped icosahedral capsid, 65-80 nm in diameter, which carries 12 projecting fibres. The major capsid protein is hexon (protein II); the fibres consist of 3 molecules of protein IV, and are linked to the capsid via penton base (protein III) pentamers at each vertex of the icosahedron. Other virion components include the hexon-associated proteins (VI, VIII and IX) and the penton-associated protein (IIIa).

Attachment of adenovirus to permissive cells is believed to be mediated by binding of the fibre proteins to a cell surface receptor. The virus is then internalised at clathrin-coated pits by endocytosis, forming a receptosome. The receptosome is ruptured, and penton protein lost, in a pH-dependent process. The virus is then transported to the nucleus, probably by a mechanism involving hexon interaction with microtubules. The remaining capsid proteins appear to be shed as the viral core enters the nucleus through the nuclear pores. In the nucleus, viral genomes undergoing active DNA replication and transcription appear to exist as distinct inclusion bodies associated with the nuclear matrix (described in Ornelles and Shenk, 1991).

The expression of viral genes depends extensively on host cell mechanisms. A number of viral gene products serve to regulate gene expression at transcriptional, post-transcriptional and translational levels (see 1.8.4). Viral gene expression can be broadly divided into two phases - early (before the onset of viral DNA replication) and late (after the onset of viral DNA replication). In general, the genes expressed in the early phase encode regulatory and other nonstructural viral proteins, while

most of the genes expressed in the late phase encode virion components. However, the early/late distinction is not rigid - most early proteins continue to be expressed in the late phase and limited expression of some late proteins occurs in the early phase. Replication requires three viral gene products directly - the DNA polymerase, a 72K DNA binding protein, and the 55K terminal protein, the latter as an 80K precursor form. Cellular proteins, including nuclear factors (NF) I and III which bind the Ad origin of replication, and topoisomerase I activity are also required. A replication model involving single strand displacement by the nascent chain (Lechner and Kelly, 1977) is now widely accepted.

During the late phase of infection, host cell protein synthesis is shut down and biosynthetic activity is diverted to the synthesis of viral structural proteins. Expression of these late proteins is followed by virion assembly, a process dependent on viral scaffold proteins that is initiated in the cytoplasm by the formation of multimeric penton and hexon capsomeres and completed in the nucleus. Replicated DNA genomes appear to enter hollow preformed capsids in one of the late stages of virion maturation. The mechanism of virus release from the dying cell is not well understood.

1.8.3. Organisation and expression of the adenovirus genome

The pattern of gene expression from the adenovirus genome is highly complex (see reviews by Flint, 1986, and Horwitz, 1990a, and references therein). This complexity is illustrated in fig. 1.4, a transcription and translation map of Ad2.

Early region E1A encodes two major proteins of 289 and 243 amino acid residues (R), synthesised from 13S and 12S mRNAs respectively. Multiple functional domains have been characterised in the 289R protein, and are responsible for induction of DNA synthesis, induction of mitosis/transformation, transactivation (transcriptional activation) of viral early genes (including E1A itself), and nuclear

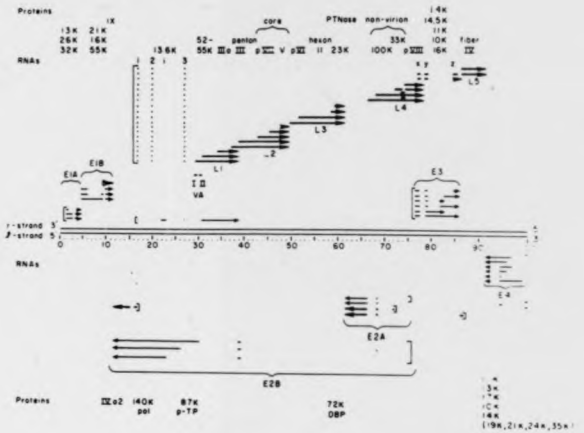

Proteins
IX
13K 21K
26K 16K
32K 55K
13.6K
core
52- penton
55K IIIa III pVI V pVII
hexon
II 23K
PTNase
non-virion
33K
100K pVIII
14K
14.5K
11K
10K fiber
16K IV
RNAs
1 2 3
L1
L2
L3
L4
L5
E1A
E1B
IVa
E3
r-strand 5'
l-strand 3'
0 10 20 30 40 50 60 70 80
RNAs
E4
E2A
E2B
Proteins
IIIa2 140K 87K 72K
pol p-TP DBP
3K
7K
12K
14K
(19K,21K,24K,35K)

Fig. 1.4. Transcription and translation map of Ad2 (reproduced from Horwitz, 1990a). Exons are denoted by bold lines, polyadenylation sites by arrowheads, nonstructural gene products by their molecular masses (e.g. 55K), and virion components by Roman numerals (e.g. II). Some minor RNA species and corresponding peptides are not shown at this scale.

localisation (Moran and Mathews, 1987). The 243R protein which, by differential splicing of the RNA encoding it, lacks one of the domains (CR3) involved in transactivation, has only a weak effect on transcription but retains the other activities (Winberg and Shenk, 1984).

Early region E1B encodes two major proteins of molecular mass 21K and 55K (as discussed below, 1.9.1). The role of the 55K protein, a principal focus of this thesis, is discussed in detail below (1.9). The 21K protein appears to inhibit the apoptosis responses (degradation of host cell DNA and cytolysis) to Ad E1A and cellular tumour necrosis factor-α proteins (White *et al.*, 1991; 1992), and can transactivate the expression of viral early genes (Yoshida *et al.*, 1987).

The proteins encoded by both E1A and E1B are involved in the induction and maintenance of cell transformation and are therefore classed as the adenovirus tumour antigens. Significantly, both E1B 55K and the E1A 289R protein bind the products of tumour suppressor genes and modulate their activities. Binding to E1B 55K has recently been shown to inhibit the transactivation activity of p53 (Yew and Berk, 1992), while binding to E1A dissociates the p105-retinoblastoma protein from complexes with cellular transcription factors (reviewed in Dyson and Harlow, 1992). A detailed review of the extensive literature on the role of these adenovirus early proteins in cell transformation and tumourigenesis is outside the scope of this thesis.

Regions E2A and E2B encode proteins required for DNA replication (reviewed in Stillman, 1989). E2A encodes a 72K single-stranded DNA binding protein necessary for DNA replication whilst E2B encodes the 80K precursor of the 55K genome terminal protein and the 140K adenovirus DNA polymerase. The E3 region encodes several proteins that modulate the host response to adenovirus infection. These include a 19K glycoprotein that blocks cell surface expression of MHC class I

polypeptides, a 14.7K protein that inhibits tumour necrosis factor-mediated lysis of infected cells, and a 10.4K protein that binds the epidermal growth factor receptor (reviewed by Wold and Gooding, 1991). The E4 region encodes a number of small proteins (Virtanen *et al.*, 1984), at least two of which (the products of ORFs 3 and 6) are involved in the post-transcriptional regulation of gene expression (as discussed below, 1.9.8).

All the genes discussed above - E1A, E1B, E2A, E2B, E3 and E4 - are classed as early - i.e., are expressed before the onset of viral DNA replication. Early gene expression can be further divided into pre-early (E1A only) and delayed early (E1B, E2A, E3 and E4) phases. The delayed early genes are dependent on E1A-mediated transactivation for their expression (reviewed in Flint, 1986). In contrast to E1A, no significant expression of the delayed early gene products can be detected in the first two hours after infection. Expression of the early genes continues at least into the early part of the late phase of infection.

The viral late genes comprise the major late transcription unit (L1, 2, 3, 4, 5) together with genes for polypeptides IX and IVa2. A distinct late promoter for the E2 gene (E2-L) is also active in the late phase. The major late transcription unit encodes all the virion components or their precursors except IX and the terminal protein (E2B), as well as nonstructural proteins including L1 55K, L4 100K, L4 33K and L3 23K (the protease responsible for processing precursor proteins to their mature forms). The 50K IVa2 protein is also a nonstructural scaffolding protein. Expression of the late genes, with the exception of L1, is limited to the period after DNA replication; transcription of these genes probably only occurs from a replicated DNA template. In the early phase, major late transcription stops in the L3 region and only L1 RNA reaches the cytoplasm. The minor late genes (IX, IVa2) are expressed before the major late transcription unit becomes fully activated. The IX gene, embedded in the E1B transcription unit but expressed from an autonomous

promoter, encodes a hexon-associated virion component (as noted above). A third minor late promoter, E2-L, is responsible for the expression of E2 proteins in the late phase of infection. It is distinct from the E2-E control region in position and expression kinetics, and controls about a third of E2 expression in the late phase.

1.8.4. Regulation of adenovirus gene expression

Adenovirus gene expression is regulated at several levels (see review by Flint, 1986, and references therein). A number of viral proteins are involved in regulation of transcription (reviewed in Akusjärvi, 1993). The E1A 289R protein activates transcription from several viral and cellular promoters by mechanisms involving its conserved amino acid domains CR1, CR2 and CR3. The CR3 domain is involved in phosphorylation of transcription factors such as E2F and is required for activation by association with TBP (TATA binding protein) and with upstream binding factors such as the ATFs (activating transcription factors). CR1 and CR2 (also carried by and active in the E1A 243R protein) are required for activation by dissociating transcription factors and inhibitory proteins (e.g. dissociating the p105-retinoblastoma protein from E2F, reviewed in Dyson and Harlow, 1992). The E4 ORF6/7 protein is also a transcriptional activator, specifically of E2A transcription. It appears to act by inducing co-operative binding of E2F to the E2A promoter (reviewed in Akusjärvi, 1993). Specific mechanisms also exist for repression of transcription from particular genes. The E2A 72K DNA binding protein has been shown to repress transcription of the E4 region (Handa *et al.*, 1983) while both the 289R and 243R E1A proteins have an autorepressive effect on their own transcription that may be involved in the general decline in transcription rates of the early genes in the latter part of the early phase of infection (Smith *et al.*, 1985).

Regulatory control is also exerted at the level of transcriptional termination. Transcription from the major late promoter terminates in L3 in the early phase of infection, whereas regions L1-L5 are transcribed in the late phase. This control of

termination is temporally regulated with respect to the phase of infection. Moreover, although L2 and L3 sequences are transcribed in the nucleus in the early phase, they do not appear to undergo significant polyadenylation and subsequent nuclear export. This contrasts with the situation in the late phase, when all the major late mRNAs are both transcribed and polyadenylated. Thus control of polyadenylation also appears to be an important mechanism in the regulation of adenovirus gene expression.

Differential splicing is used extensively by Ad5 to produce multiple RNA species from a single primary transcript. The 5' ends of major late mRNAs are generated by splicing of tripartite leader sequences to one of several alternative 3' splice acceptor sites. e.g., the L3 region forms a template for a family of three major RNA species, where the longest (coding for pVI) utilises the most 5' of the 3' acceptor sites, the intermediate size RNA (coding for hexon protein II) utilises the middle 3' acceptor site, and the shortest (coding for the 23K protease) utilises the most 3' of the acceptor sites (fig. 5.2). In some cases, splice site usage is temporally regulated - i.e., depends on the phase of infection. This type of regulation has been shown to occur in the formation of L1, E1A and E1B RNA species (see for example Akusjärvi and Persson, 1981; Montell *et al.*, 1984). A further mechanism of regulation exists at the level of mRNA transport - i.e., between release of mature mRNAs from their processing sites and their appearance in the cytoplasm (a multi-step process that could potentially be regulated at several points). In adenovirus infections, these processes are regulated by a functional complex of the E1B 55K and E4 34K proteins as discussed below.

1.9. E1B 55K

1.9.1. Expression

Adenovirus 5 early region E1B encodes two major protein products - a 495 amino acid (495R) protein with a molecular mass of 55K, and a 175 amino acid (175R)

protein with a molecular mass of 21K (Bos *et al.*, 1981; Anderson *et al.*, 1984). At early times in infection both E1B 55K and E1B 21K are translated from overlapping reading frames in a single 2.2 kb (22S) RNA. Subsequently, the 21K protein is also expressed from a 1 kb (13S) RNA which is the predominant E1B species produced in the late phase of infection (Bos *et al.*, 1981). Although differential stabilisation may be a factor in controlling relative levels of the E1B mRNAs, studies of splice site mutations (Montell *et al.*, 1984) and staggered superinfection experiments (Adami and Babiss, 1991) have shown that the changes in 22S and 13S message abundances are due to a *cis*-effect on splice site usage.

A 155 amino acid (155R) 16.6K E1B protein expressed from a minor (1.2 kb) RNA species has also been identified (Anderson *et al.*, 1984). This protein is translated in the 55K reading frame, with the middle section removed by alternative splicing. The 16.6K protein primary structure therefore corresponds to the N- and C-termini of E1B 55K. Other minor E1B polypeptides, translated from mRNAs that have undergone additional splicing events that truncate the 55K or 21K reading frames, have also been detected. (Anderson *et al.*, 1984; Virtanen and Pettersson, 1985).

1.9.2. Association with E4 34K protein

Sarnow *et al.* (1984) have shown that the E1B 55K protein is physically associated with a protein with an apparent molecular mass of 25K. This protein was mapped to open reading frame 6 (ORF 6) of early region E4 in the adenovirus genome. E4 ORF 6 was predicted to encode a 34K protein, indicating that the molecular mass of the detected protein could not be determined accurately by SDS-PAGE (or that a larger precursor was processed post-translationally). The E1B 55K/E4 34K complex is believed to be the functional unit of 55K activity. (Cutt *et al.*, 1987) (see 1.9.8 for detailed discussion). Thus, in the following discussion of the role of E1B 55K it should be noted that E4 34K is also required for the functions described.

1.9.3. Subcellular localisation

Two recent studies have attempted to define precisely the intracellular localisation of the E1B 55K protein in lytically infected cells. Smiley *et al.* (1990) used subcellular fractionation and indirect immunofluorescence techniques to examine the subnuclear distribution of the 55K protein by probing with a monoclonal antibody. E1B 55K was detected in RNase A-releasable and in pore complex-lamina (nuclear envelope) fractions of the nucleus. However, comparisons of the distributions and elution profiles of 55K and cellular pore complex proteins indicated that 55K was not stably or intimately associated with the pore complex. Free and E1B 55K-bound forms of the E4 34K protein were also concentrated in the pore complex-lamina fraction, but the presence of the 34K protein was not required for the localisation of 55K in this fraction. However, Ornelles and Shenk (1991) showed that E4 34K was required for localisation of E1B 55K in and around the periphery of the nuclear viral inclusion bodies in which DNA replication and transcription are believed to occur. Using immunoelectronmicroscopy and immunofluorescence, this group demonstrated localisation of 55K in cytoplasmic fibrous bodies adjacent to the nucleus. The 55K protein was also detected in a diffuse reticular distribution in the cytoplasm, in a granular non-nucleolar distribution in the nucleolus, and in discrete nuclear spicules. However, only the localisation to viral inclusion bodies required the E4 34K protein. Since 55K requires binding of the 34K protein to form an active complex, then the 34K-dependent localisation to the inclusion bodies is likely to be functionally significant. The possible relevance of this localisation to the mechanism of action of the 55K/34K complex is discussed below (1.9.7).

1.9.4. Phosphorylation

The E1B 55K protein is phosphorylated *in vivo* by a cAMP-independent protein kinase. Malette *et al.* (1983) found that each 55K molecule contains 2-3 phosphates on average. The ratio of phosphoserine : phosphothreonine was estimated as 2 : 1 (with essentially no detectable phosphotyrosine). Taken together, these results

suggested that 55K is phosphorylated at two serine residues and one threonine. However, the exact positions of the phosphorylation sites were not determined and the importance of phosphorylation to 55K function was not assessed.

1.9.5. Functional domains

Yew *et al.* (1990) have attempted to identify the functional domains of the Ad2 E1B 55K protein. By constructing a series of recombinant viruses with insertion mutations in the E1B 55K gene, they were able to show that a mutant could be obtained (with an insertion at 55K amino acid 380) that efficiently expressed late viral proteins, caused shutoff of host cell synthesis, and replicated in noncomplementing HeLa cells, but was deficient in transformation of cloned embryo fibroblast (CREF) cells following infection. However, distinct domains in the primary sequence required for individual 55K functions were not identified. Several regions (at amino acids 180, 262-326 and 443-474) were found to be necessary for viral replication. Overlapping with two of these regions, amino acids 180-354 were shown to be required for efficient shutoff of cellular protein synthesis. Numerous regions throughout the sequence were required for late viral protein synthesis, and transformation ability in CREF cells was impaired in all the mutants studied. These results suggest that E1B 55K may not have a modular domain structure of the type found for the E1A proteins.

1.9.6. Role of E1B 55K in lytic infection

While earlier studies had focused on the role of E1B 55K in cell transformation, or on the effects of the whole E1B region on adenovirus and cellular gene expression, Babiss and Ginsberg (1984) were the first workers to study directly the effects of specific mutations of the 55K sequence on the phenotype of viral lytic infection. They constructed Ad5 insertion or deletion mutants that expressed truncated 55K protein. These mutants exhibited greatly reduced yield (10% of wild type) when grown in HeLa cells. DNA replication, early and late transcription, cytoplasmic

accumulation of early mRNAs and early protein synthesis were comparable to wild type levels. However, cytoplasmic accumulation of late mRNAs (L3 and L5) and late protein expression were greatly reduced in the mutants with respect to the wild type. Moreover, the mutants, unlike the wild type, could not efficiently shut off host cell protein synthesis. In further studies with these mutants (Babiss *et al.*, 1985) results were obtained which suggested that the defect in cytoplasmic accumulation of late viral mRNA in the presence of truncated 55K was due to a failure in both cytoplasmic RNA stabilisation and RNA transport. The rate of nuclear export of late RNA, monitored by cellular ^{3}H-uridine labelling and filter hybridisation of extracted poly(A)$^{+}$ RNA to viral DNA sequences, was decreased by 30-50% in the absence of full-length 55K, but the cytoplasmic accumulation of late RNAs measured by Northern analysis was reduced 5- to 10-fold in mutant infections. Thus, the effect could not be explained by a reduction in the rate of transport alone, indicating that another process, presumably RNA stabilisation, was involved. Furthermore these authors found that an intact E1B region was required to block both transport and translation of cellular RNAs.

In a separate study, Williams *et al.* (1986) investigated the phenotypes of the previously isolated Ad5 host-range mutants *hr*6 and *hr*^{CS}13. These viruses possess mutations that disrupt E1B 55K by truncating the protein coding sequence, but do not affect E1B 21K or the minor E1B 55K-related proteins. The mutants were cold-sensitive for growth in HeLa cells - i.e., they showed significantly reduced yield at 32.5 °C, compared with the yield at 38.5 °C - but grew normally in complementing HEK-293 cells at both temperatures. The only significant differences between 32.5 °C and 38.5 °C infections of HeLa cells by these viruses that were observed were in the transport of late RNA from the nucleus to the cytoplasm, and in late RNA stability. Transport and stability were measured and compared by pulse-chase methods for mutant and wild type at the restrictive temperature. Transport was reduced by approximately 5-fold and stability by about 4-fold in the absence of

intact 55K. Moreover, the transport of cellular RNA (ß-actin mRNA and 28S rRNA) into the cytoplasm, normally blocked by wild type infection, continued in the mutant-infected cells. It was therefore concluded that 55K acted primarily at the levels of RNA stabilisation and transport.

An extensive analysis of the properties of a mutant virus deficient in E1B 55K was also made by Pilder *et al.* (1986). This group utilised Ad5 deletion variant *dl*338 - a virus with a 524 bp deletion in E1B that removes a large section of the E1B 55K gene while leaving the 21K gene intact. Results broadly consistent with those of Babiss *et al.* (1985) and Williams *et al.* (1986) were obtained. They found that E1B 55K was required for normal growth in HeLa cells (yield was reduced 100-fold in the mutant). Viral DNA replication and early gene expression were normal in *dl*338 infections, but accumulation of late viral RNAs in the cytoplasm measured by Northern analysis was greatly reduced. This effect correlated with a decrease in the rate of appearance of late RNAs in the cytoplasm (or increased cytoplasmic stabilisation) measured by ^{3}H-uridine continuous labelling. Transcription rates were normal for all transcripts at 12 and 16 hours post infection, discounting transcriptional effects as an explanation of the reduced late mRNA accumulation seen at these times. A reduction in the levels of late transcription was however noted at 20 and 24 hr in *dl*338 infections. Neither transcription rates nor cytoplasmic mRNA levels from early genes were affected by the *dl*338 deletion even at late times; in fact transcription from E2A was actually increased twofold at 20 hr in the mutant infection.

Pilder *et al.* (1986) also looked at the effect of the *dl*338 deletion on RNA processing. They found by Northern analysis of poly(A)$^{+}$ nuclear RNA that neither the size nor the quantity of a late (L5) RNA species in this fraction were altered in the mutant infection, indicating that both splicing and polyadenylation were independent of 55K. Moreover, an Ad5 variant carrying a non-spliceable Ad2 late

(L5) construct was shown to require intact 55K for efficient RNA accumulation in the cytoplasm, indicating that the 55K response was not related to splicing. L5 mRNA reaching the cytoplasm was also shown to have the same half-life in wild type and *dl*338-infected cells at 12 hr post infection, suggesting that 55K did not primarily affect this aspect of RNA metabolism.

Studies of cellular RNA metabolism in wild type or *dl*338-infected cells showed that host gene (actin) transcription rates were not affected by 55K, but cytoplasmic accumulation and the transport rate / cytoplasmic stabilisation of host-specific RNA was severely reduced only in the wild type infection (Pilder *et al.*, 1986). From these results, it was concluded that the primary effect of 55K is on RNA transport or the stabilisation of RNA immediately after it enters the cytoplasm. However, it was not possible to determine whether the reciprocal effects on late viral and cellular RNA accumulation were due to the participation of 55K in two distinct processes, or whether 55K acts directly on one process and hence indirectly affects the other (e.g. efficient facilitation of viral RNA transport might saturate a limited capacity for nuclear export, preventing transport of cellular RNA).

The analysis of RNA metabolism in *dl*338-infected cells was further extended by Leppard and Shenk (1989). They showed that *dl*338 (like the *hr* mutants analysed by Williams *et al.*, 1986) was cold-sensitive for growth - i.e., the yield of *dl*338 compared to wild type was further reduced by a factor of 4 at 32 °C, compared to the yield at 37 °C, and late (L3, L4 and L5) viral RNA accumulation was more severely inhibited at the lower temperature. Qualitatively, however, the effect observed on RNA metabolism was similar to that described by Pilder *et al.* (1986) at 37 °C. These results were reflected at the protein level by reduced synthesis of late proteins in the mutant infection (exacerbated at the restrictive temperature) whilst early (E1A and E2A) protein synthesis in the early phase of infection was comparable to wild type; synthesis of these proteins continued at high levels into the

late phase in the *dl*338 infection only. Consistent with the results of Pilder *et al.* (1986) at 37 °C, Leppard and Shenk (1989) showed that the effect on RNA accumulation at 32 °C could not be attributed to differences in transcription rates. Rates were similar for early genes (and the cellular actin gene) at all infection times analysed, while the rates for L3, L4 and L5 were also comparable up until 40 hr post infection, after which time there was a widening differential in rates between mutant and wild type. This as yet unexplained observation could possibly be due to the failure of the mutant to produce sufficient quantities of a late protein required for stimulation of late transcription.

The increased severity of the *dl*338 phenotype at 32 °C, coupled to the generally lower rate of all molecular processes allowed Leppard and Shenk (1989) to examine its basis at the RNA level in more detail than previous studies. One cytoplasmic and four nuclear (nuclear membrane, DNase I-soluble, salt-soluble, and nuclear matrix) fractions were defined and the levels of RNA species and their kinetics of movement between the defined compartments determined. In both wild type and *dl*338 infections, transcription of viral genes, like cellular genes, was found to occur in association with the nuclear matrix. In a wild type infection, late RNA then moved through the DNase I-soluble and salt-soluble fractions (in an order that was not determined) into the nuclear membrane fraction, before arriving and accumulating in the cytoplasm. Late viral transcripts synthesised by *dl*338 were found to move out of the nuclear matrix fraction more slowly and less efficiently than transcripts synthesised by the wild type. These transcripts also failed to accumulate in the DNase I-soluble fraction, indicating that they were degraded before or immediately after reaching it. In contrast, early *dl*338 transcripts from both viruses were able to pass through all the nuclear fractions and accumulate efficiently in the cytoplasm. Thus the function of the 55K protein appeared to be to prevent degradation of late viral transcripts in a nuclear compartment, or to

facilitate transport between nuclear compartments, thus allowing nuclear export and cytoplasmic RNA accumulation.

Several studies have demonstrated that the RNA transport facilitation function of E1B 55K is not restricted to products of the adenovirus major late transcription unit. Samulski and Shenk (1988) showed that E1B 55K and E4 34K together facilitate cytoplasmic accumulation of mRNA encoded by adeno-associated virus (AAV) - a parvovirus whose replication is considerably enhanced by coinfection with adenovirus. Leppard (1993) demonstrated that the cytoplasmic accumulation of mRNAs expressed from minor late promoters (i.e., promoters for IVa2, polypeptide IX and late E2-L expression) was dependent on 55K at late times. Conversely, accumulation of mRNA expressed from the major late promoter at early times was independent of 55K. Similarly, Dix and Leppard (1993) showed that E4 mRNA A, one of a number of E4 mRNAs produced by differential splicing and showing late expression kinetics, was dependent on E1B 55K and E4 34K for efficient accumulation in the cytoplasm. Significantly, this RNA species contains an intron and unused flanking splice sites that are utilised to produce other E4 mRNAs expressed at late times. It was suggested that these intronic sequences or unused splice sites could be acting as signals for RNA retention in the nucleus (presumably on the nuclear matrix in association with splicing factors or matrix proteins) in the absence of 55K/34K. This would be consistent with the putative role of matrix releasing factor ascribed to 55K by Leppard and Shenk (1989). Furthermore, detailed analysis of individual mRNAs within the major late transcription unit families L1, L2 and L3 (Leppard, 1993) indicated that longer mRNAs were more dependent on 55K than the shorter species. Since the longer species contain unused splice sites and sequences spliced out of the shorter mRNAs as introns, these results, like those of Dix and Leppard (1993), suggest that splice sites or intronic sequences may function as nuclear retention signals in these RNAs, to be overcome by E1B 55K activity. Thus both the phase of infection and the splicing potential of

RNA species appear to be important in determining their dependence on 55K function.

The results of the studies described above give a consistent picture of E1B 55K function in lytically infected cells. The 55K protein (in combination with E4 34K) acts post-transcriptionally on RNA metabolism, probably at the level of RNA transport or stabilisation, to allow the cytoplasmic accumulation and subsequent translation of late viral mRNAs, including the products of major and minor late transcription units. mRNA expressed by adeno-associated virus is similarly dependent on 55K for cytoplasmic accumulation. However, mRNA expressed by adenovirus early transcription units in the early phase does not require 55K for cytoplasmic accumulation. Finally, the E1B 55K/E4 34K complex blocks the cytoplasmic accumulation of cellular RNA species.

1.9.7. Mechanism of action of E1B 55K

As noted above, current evidence indicates that the E1B 55K/E4 34K complex is active at the levels of RNA transport and possibly stabilisation. To arrive at a more precise mechanism of activity it is necessary to consider the individual post-processing stages of mRNA metabolism. RNA transport, as reviewed in Schröder (1987), is a multi-step process that involves release of mature mRNAs from their processing sites on the nuclear matrix, movement through nuclear subcompartments, translocation through the nuclear pore complex, and binding to the cytoskeleton. The 55K/34K complex could potentially act at any of these individual steps. However, as noted above, localisation studies have demonstrated that 55K is not stably or intimately associated with nuclear pore complexes, but significant concentrations are found in and around the matrix-associated viral inclusion bodies (in which transcription occurs) only in the presence of the functionally essential 34K protein. Moreover, Leppard and Shenk (1989) have shown that E1B 55K is required for the efficient movement of late viral mRNAs

from the nuclear matrix to other nuclear subcompartments and the cytoplasm. These experiments suggest that the 55K protein acts at an early stage of the mRNA transport pathway - most probably a matrix-releasing event. Any viable model of 55K function has to take account of the ability of this protein to block nuclear export of cellular RNAs, and the 55K-independent expression of viral genes in the early phase of infection. Such a model has been proposed (Moore *et al.*, 1987; Leppard and Shenk, 1989; Ornelles and Shenk, 1991; Leppard, 1993), based on elements of the gene gating hypothesis of Blobel (1985). Blobel suggested that the eukaryotic genome is organised into a specific three-dimensional structure that is reflected in the distribution of pore complexes in the nuclear membrane. Actively transcribed regions of the genome, or regions potentially active in that cell type, are structurally expanded and linked, or 'gated', to specific pore complexes. Nuclear export of RNA can only occur efficiently if the corresponding gene is gated to a pore complex. Blobel envisaged that the linkage between the interior of the nucleus and the pore complex would take the form of a channel in which transcription and RNA processing could take place. Direct microscopic evidence of the existence of RNA 'tracks' in which splicing and intranuclear transport occurs has in fact been obtained for cellular and Epstein-Barr viral transcripts by Lawrence and co-workers (Lawrence *et al.*, 1989; Carter *et al.*, 1993; Xing *et al.*, 1993).

The Blobel model implies that the cell has a limited capacity for efficient nuclear export of RNA. Active cellular genes would be expected to be located in favourable sites in the nuclear architecture - facilitated microenvironments linked to efficient RNA transport pathways. In the early phase of an adenovirus infection, the unreplicated genomes from which early gene expression occurs could potentially colonise such favourable sites and make use of the facilitated transport pathways. In the late phase the number of replicated viral genomes would be predicted to exceed the number of available favourable sites, forcing the new DNA templates to occupy microenvironments without efficient transport mechanisms. According to this

model, E1B 55K/E4 34K function would allow the virus to circumvent this situation. The 55K/34K complex might act by recruiting cellular transport factors to viral RNA, by displacing cellular genes from favourable sites, or by taking part in a virus-specific transport mechanism (in each case allowing release of transcripts from the nuclear matrix). By competing for limited supplies of transport factors, favourable sites, or some other aspect of transport capacity, the E1B 55K/E4 34K complex could facilitate late viral RNA transport at the expense of cellular RNA transport. This scenario is consistent with the experimental data, including the recent observation that RNAs expressed from both the major and minor late promoters in the late phase are 55K-dependent, whereas RNAs expressed from the major late promoter in the early phase are 55K-independent (Leppard, 1993).

An implication of this type of mechanism is that 55K dependence might be expected to be independent of RNA composition - i.e., no *cis*-acting sequences would necessarily be required. However, as noted above, 55K dependence varies and can be correlated with the presence of unused splice sites and intervening sequences in the RNA (Dix and Leppard, 1993; Leppard, 1993). Although it has been shown that both cellular (Ciejeck *et al.*, 1982) and viral (Mariman *et al.*, 1982) RNAs are normally associated with the nuclear matrix, it is reasonable to suppose that unused splice sites or intronic signals could act as nuclear retention signals that would attach RNA to the matrix more securely by binding to components of the splicing machinery. Thus the matrix-releasing activity of the 55K/34K complex may have a proportionally greater effect on incompletely spliced RNA species than on fully-spliced RNAs.

The mechanism discussed above is also consistent with the results of experiments which showed that the export of cellular RNAs whose synthesis is induced by or after adenovirus infection is independent of 55K (Moore *et al.*, 1987). Recently induced genes could occupy particularly favourable positions in the nuclear

architecture close to pore complexes and would therefore be more able to compete effectively with 55K-mediated export for limited transport resources.

The above discussion has focused on the mechanisms that act at the level of RNA transport. As noted previously, 55K may also act at the level of RNA stabilisation. However, there is as yet no direct evidence for any specific 55K-mediated mechanism of stabilisation so this has not been addressed in any detail.

1.9.8. Interactions of E1B 55K and E4 proteins

As noted above (1.9.2) the E1B 55K protein is physically associated with E4 34K, the product of E4 ORF6, and functions in combination with this protein. The existence of an E1B 55K/E4 34K complex was first noted by Sarnow *et al.* (1984), who demonstrated specific co-immunoprecipitation of the ORF6 protein in the presence of E1B 55K with antibodies to the latter. Subsequently, mutants with deletions of E4 ORF6 were shown to have similar phenotypes to mutants defective in E1B 55K expression (Halbert *et al.*, 1985; Weinberg and Ketner, 1986). Cutt *et al.* (1987) demonstrated that a recombinant with mutations designed to disrupt the expression of both E1B 55K and E4 34K had an identical phenotype to equivalent mutants disrupted in only one of the genes. This indicated that the E1B 55K and E4 34K function in combination, rather than by parallel pathways - consistent with the physical association of the two proteins. Conversely, the E1B 55K/E4 34K complex appears to act in parallel with another product of early region 4, the E4 ORF3 116 amino acid protein, since mutants deficient in both ORF3 and ORF6 have more severely defective phenotypes than either of the equivalent single mutants (Bridge and Ketner, 1989, 1990; Huang and Hearing, 1989). Both the ORF 3 and ORF6 proteins have been implicated in viral mRNA stabilisation in the nucleus, but only the ORF 6 protein is involved in the parallel mechanism of facilitating transport of late mRNAs into the cytoplasm (Bridge and Ketner, 1990; Bridge *et al.*, 1991). Moreover, the ORF3 protein is ineffective in providing the adeno-associated virus

helper function conferred by ORF6 in combination with E1B 55K (Samulski and Shenk, 1988; Huang and Hearing, 1989). According to Bridge *et al.* (1991), the effect on nuclear RNA stabilisation exerted by both ORF3 and ORF6 affects only products of the major late transcription unit, whereas the ORF6/E1B 55K-mediated effect is also active on mRNA from the minor late promoter, IVa2 (consistent with the results of Leppard, 1993, for E1B 55K). However, Bridge *et al.* (1991) also found that polypeptide IX and E2-L minor late mRNAs accumulated efficiently in the cytoplasm in the absence of ORF6. This contrasts with the results of Leppard (1993), who showed that cytoplasmic accumulation of these RNA species was facilitated by E1B 55K. These results may reflect differences in the phenotypes of E1B 55K and E4 ORF6 deletion mutants, or differences in the sensitivities of the RNase protection assay used by Leppard (1993) and the S1 nuclease assay / primer extension analysis employed by Bridge *et al.* (1991). Recently, Ohman *et al.* (1993) have found that in transient cotransfection experiments (in the absence of E1B 55K) the ORF3 protein and the ORF6 protein individually stimulate accumulation of mRNAs carrying the tripartite leader. Moreover, in experiments using an adenovirus mutant in which all E4 ORFs except ORF1 were deleted, tripartite leader assembly was found to be impaired by inefficient splicing out of a fourth 5' leader segment, the *i*-leader. This result suggested that the E1B 55K-independent effects of ORF6 and ORF3 may be exerted at the level of alternative RNA splicing. It is at present unclear if the apparently separate E1B-independent and E1B-linked effects of the E4 ORF6 protein are unrelated activities. In a lytic infection, E4 function may be a prerequisite for E1B 55K action, or the two proteins may act on interrelated steps of the RNA processing pathway rather than co-operating to facilitate a single step.

1.10. Comparison of HIV-1 Rev and Ad5 E1B 55K function - project objectives

In the previous sections of this introduction I have discussed the roles of the HIV-1 Rev and Ad5 E1B 55K proteins in the regulation of viral gene expression. From this

discussion, certain functional similarities between the two proteins are apparent. Both proteins clearly act at a post-transcriptional stage of gene expression, and are required for the synthesis of late (generally structural) but not early (generally nonstructural) proteins. In both cases modes of action have been proposed that involve selective facilitation of RNA transport. Results consistent with a mechanism in which transport is facilitated at an intranuclear stage, by catalysing the release of transcripts otherwise bound by nuclear factors or retained in a nuclear subcompartment, have been obtained for E1B 55K and for Rev. In both systems, the presence of intronic sequences or unused splice sites is a factor in making an RNA dependent on the appropriate regulatory protein for nuclear export. Finally, an additional mechanism involving cytoplasmic stabilisation of late mRNAs has been proposed for both regulatory proteins. Thus the activities of Rev and E1B 55K are, at least to some extent, analogous.

It is also clear that there are some distinct differences between the Rev and E1B 55K regulatory systems. Multiple activities (including direct effects on splicing and mRNA utilisation in the cytoplasm) have been attributed to Rev, but not to E1B 55K. Moreover, no specific response element analogous to the RRE has been defined for 55K; current evidence indicates that the microenvironment of the template, rather than a particular *cis*-acting sequence, is probably the major determinant of the 55K dependence of an RNA species. Indeed, it is not known whether 55K binds RNA directly, or functions by interacting with a cellular factor. Nevertheless, the functional analogy between the two systems is sufficiently strong to make a comparative study worthwhile.

It is the purpose of the work described in this thesis to further explore this analogy by constructing recombinant adenoviruses in which the effects of Rev/RRE on Ad gene expression can be tested and compared with those of E1B 55K. The construction and isolation of recombinant viruses containing a *rev* expression

system, the RRE, and/or the E1B 55K gene in all possible combinations will be described in subsequent sections. Data will be presented from experiments in which gene expression from these constructs was analysed at the protein and RNA levels, and the implications for the potential Rev/E1B 55K analogy will be discussed.

Chapter 2

Materials and methods

2.1. Materials

2.1.1 General buffers and solutions

The following buffers and solutions were in general use:

TE: 10 mM Tris.HCl pH 8, 1 mM EDTA pH 8

TNE: 10 mM Tris.HCl pH 8, 100 mM NaCl, 1 mM EDTA pH 8

TBE: 89 mM Tris base, 89 mM boric acid, 2 mM EDTA pH 8

PBS: 137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4.7H_2O$, 1.4 mM KH_2PO_4

TD: 8 g/l NaCl, 0.38 g/l KCl, 0.1 g/l Na_2PO_4

TS: 8 g/l NaCl, 0.38 g/l KCl, 0.1 g/l Na_2PO_4, 0.1 g/l $CaCl_2$, 0.1 g/l $MgCl_2$

Chloroform/amyl alcohol: 96% (v/v) chloroform, 4 % (v/v) amyl alcohol

Phenol/TNE: liquefied phenol (Fisons Tris-washed) equilibrated with excess TNE

Phenol/chloroform: 50% (v/v) phenol/TNE, 50% (v/v) chloroform/amyl alcohol

2.1.2 Suppliers

Sterile PBS, versene, and tissue culture media including DMEM and 2x DMEM were supplied by the media preparation staff of the Biological Sciences Department, University of Warwick. Calf serum and foetal calf serum were supplied by Northumbria Biologicals Ltd. (Cramlington, Northumbria). Calf intestinal phosphatase was supplied by Boehringer Mannheim UK (Lewes, E. Sussex). Other DNA modification enzymes, restriction enzymes, T4 DNA ligase and nucleases were obtained from Gibco-BRL Life Technologies Ltd. (Renfrewshire, Scotland), New England Biolabs (CP Laboratories, Bishops Stortford, Herts.) and Pharmacia Ltd (Milton Keynes, Bucks.). Econofluor and ^{3}H-acetyl CoA were supplied by DuPont (UK) Ltd (NEN products, Stevenage, Herts.). ^{32}P-UTP and ^{35}S-Met were supplied by Amersham International PLC (Bucks.). Nuctrap columns were obtained from Stratagene Ltd. (Cambridge). Noble agar, bacto-agar, bacto-tryptone, minimal agar and yeast extract were supplied by Difco Laboratories (Basingstoke, Hants.). Acrylamide, N,N'-methylene-bis-acrylamide, ammonium persulphate and TEMED were obtained from Bio-Rad (Hemel Hempstead, Herts.). Other chemicals, unless

stated otherwise, were supplied by Sigma Chemical Company Ltd. (Poole, Dorset), BDH Laboratory Supplies (Merck Ltd, Poole, Dorset) and Fisons Scientific Equipment (Loughborough, Leics.) and were of molecular biology or analytical grade.

2.2. Nucleic acid purification

2.2.1. Ethanol precipitation of nucleic acids

DNA and RNA were precipitated from aqueous solutions containing 0.3 M sodium acetate or 0.1 M sodium chloride by the addition of 95% (v/v) ethanol (2 volumes for DNA, 3 volumes for RNA). The nucleic acid was allowed to precipitate, either for 30 min at -20 °C or 10 min at -80 °C, then pelleted by centrifugation in a microcentrifuge for 15 min at 12 000 rpm (variations from these conditions used in individual procedures are indicated where applicable). The pellet was washed, if necessary, in 70% (v/v) ethanol, dried under vacuum, and resuspended in TE, TNE or water.

2.2.2. Phenol/chloroform extraction

An equal volume of phenol/chloroform was added to the DNA solution and the phases were mixed by vortexing (or inversion for high molecular weight viral DNA). The tube was centrifuged for 2 min at 12000 rpm in a microcentrifuge, and the aqueous phase was removed to a fresh tube (variations from this procedure are indicated where applicable).

2.2.3. Quantification of nucleic acids

DNA and RNA were quantified in solution by spectrophotometry. Absorbance readings were taken at 260 nm (A_{260}) where an absorbance of 1.0 indicates a concentration of 50 μg/ml of double-stranded DNA, or 40 μg/ml of RNA. Absorbance readings were also taken at 280 nm (A_{280}) to assess the purity of the samples (A_{260}/A_{280} = 1.8 for pure DNA solutions, or 2.0 for pure RNA solutions

- significantly lower values indicate contamination, preventing accurate quantitation by this method)

2.3. DNA manipulation techniques

2.3.1. Digestion by restriction enzymes

Restriction enzyme digests were carried out under the conditions recommended by the suppliers. Reaction buffers supplied with the enzymes (usually as 10x concentrates) were used where available; otherwise buffers were prepared according to the suppliers' instructions. Most reactions were carried out at 37 °C - where higher temperatures were required, reactions were carried out under liquid paraffin to minimise evaporation. Multiple enzyme digests were done simultaneously, where possible, in a buffer which allowed both enzymes to function with adequate efficiencies.

2.3.2. Partial digestion by restriction enzymes

Partial digests were used in subcloning when the required enzyme recognition sequence occurred more than once in the (circular) DNA to be digested. The conditions necessary to linearise the DNA by cleavage at a single site were determined in a preliminary reaction: 1 μg of DNA was digested with 1 unit of restriction enzyme in a volume of 10 μl for 10 min; 1 μl aliquots were removed at 1 min intervals and added to 2 μl of 0.25 M EDTA pH 8.0 to stop the reaction. Each aliquot was analysed by agarose gel electrophoresis, and the optimum reaction time (with the highest proportion of linearised product) was noted. The reaction was scaled up to a preparative level by increasing the amounts of enzyme and DNA and the overall volume proportionally, but retaining the optimum reaction time. The linearised fragments were then isolated by agarose gel electrophoresis. The target DNA, either uncut or linearised by complete digestion with an enzyme having a single recognition site, provided molecular size markers.

2.3.3. End-fill reaction and 3' overhang removal

T4 DNA polymerase (BRL) was used to fill 5' protruding cohesive termini when a blunt terminus was necessary for further cloning. Reactions contained 33 mM Tris.acetate pH 7.9, 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT, 0.1 mg/ml BSA, 0.2 mM dATP, 0.2 mM dCTP, 0.2 mM dGTP, 0.2 mM dTTP and 2 units of enzyme per μg of DNA. The reaction (typical volume 25 μl) was incubated at 37 °C for 30 min, then extracted with phenol/chloroform.

Conversely, where blunt ends were produced by removing 3' protruding cohesive termini, the deoxyribonucleotides were omitted from the reaction.

2.3.4. Dephosphorylation

To prevent self ligation of cloning vectors cleaved at a single restriction site (or at two sites with compatible termini), the 5' phosphate groups were removed using calf intestinal alkaline phosphatase (Boehringer) in a buffer containing 50 mM Tris.HCl pH 9.0, 1 mM $MgCl_2$, 1 mM $ZnCl_2$, and 1mM spermidine. 1 μg of cleaved plasmid DNA was dephosphorylated by 1 unit of enzyme in a 10 μl reaction. For protruding 5' termini, the reaction was incubated at 37 °C for 30 min. For blunt or recessed termini, the reaction was incubated at 37 °C for 15 min, a second aliquot of enzyme was added, and the reaction was incubated at 55 °C for a further 45 min. The enzyme was removed by phenol/chloroform extraction.

2.3.5. Linker kinase reaction

Oligonucleotide linkers were prepared for ligation by treatment with T4 polynucleotide kinase (BRL). The kinase reaction contained 2 μg of DNA, 10 units of enzyme, 1 mM ATP, 70 mM Tris.HCl pH 7.6, 10 mM $MgCl_2$, and 5 mM DTT in a volume of 10 μl. The reaction was incubated for 30 min at 37 °C, then the enzyme was removed by extraction with phenol/chloroform.

2.3.6. Ligations

DNA fragments were ligated with BRL T4 DNA ligase in a volume of 15-30 μl using the manufacturer's buffer. Typically, a fivefold molar excess of insert DNA was ligated into a cloning vector (200-500 ng) previously cleaved at one or more restriction sites. Blunt-end ligation reactions were incubated at room temperature for at least 5 hr; ligations between cohesive ends were incubated overnight (12-16 hr) at 15 °C.

2.4 Agarose gel electrophoresis

2.4.1. Separation of restriction fragments

Restriction fragments were separated on horizontal agarose gels containing between 0.4 and 2.0% agarose and 0.5 μg/ml ethidium bromide, in 0.5x or 1x TBE. Samples were prepared for loading by adding 1/5 the volume of 6x loading buffer (50% glycerol, 5x TBE, 0.02% bromophenol blue, 0.02% xylene cyanol). Electrophoresis was carried out at 150 V in 0.5x or 1x TBE running buffer. Restriction fragments were visualised on an ultraviolet transilluminator and photographed with Polaroid 665 or 667 film.

2.4.2. Purification of fragments from agarose gels

The required fragment was sliced out of the gel with a scalpel and placed in a short section of dialysis tubing filled with 0.5x TBE buffer. The tubing was clip-sealed at both ends and the DNA was electroeluted from the gel fragment into the buffer for 30-45 min at 150 V in a minigel tank filled with 0.5x TBE. The polarity of the current was reversed for 30 s immediately before removing the tubing from the apparatus to ensure that the DNA did not stick to the wall of the tubing. The effectiveness of the elution was checked under ultraviolet illumination.

A DEAE-Sephacel or DEAE-cellulose column was prepared (using column matrix pre-equilibrated in TNE buffer) in a Pasteur pipette plugged with autoclaved glass

wool. The column was rinsed with 3 ml of TNE buffer, and the electroeluted DNA was transferred to the column. The column was rinsed with 6 ml of TNE, and the DNA was eluted into a microcentrifuge tube with 1 ml of 1 M NaCl in TNE. The eluate was extracted with phenol/chloroform and the DNA was precipitated with ethanol, dried, and resuspended in water or TE.

2.5. Bacteriological techniques

2.5.1. Growth and maintenance of bacterial stocks

(All growth media were autoclaved at 15 lb/in^2 before use)

E. coli strain DH1 (host strain for plasmid vectors, Hanahan, 1983) was routinely grown in LB (Luria-Bertani) medium containing 10 g of bacto-tryptone, 5 g of bacto-yeast extract, and 10 g of NaCl per litre. Cultures were maintained in the short term on LB-agar plates (stored at 4 °C) containing 15 g of Bacto-agar per litre of LB medium. Where ampicillin selection of transformed bacteria was required, LB medium was supplemented with 50 μg/ml of ampicillin and LB-agar was supplemented with 60 μg/ml of ampicillin. Bacterial stocks were stored in the long term by mixing 0.5 ml of a stationary phase culture with 0.5 ml of 80% (v/v) sterile glycerol and placing at -20 °C.

E. coli strain TG1 (Gibson, 1984, and described in Sambrook *et al.*, 1989) was used as a host strain for M13 phage vectors (Messing *et al.*, 1977). Cells were grown in 2YT medium, containing 16 g of bacto-tryptone, 10 g of bacto-yeast extract, and 5 g of NaCl per litre. 2YT plates for phage plating were prepared using 2YT soft agar, which contains 15 g of bacto-agar per litre of 2YT medium. TG1 cells were maintained in the short term on minimal agar plates to select for retention of the F' factor. The minimal plating medium was prepared by autoclaving a solution of Difco minimal agar, then adding sterile salt and sugar solutions to give final concentrations of 15 g/l minimal agar, 10.5 g/l K_2HPO_4, 4.5 g/l KH_2PO_4,

1 g/l $(NH_4)_2SO_4$, 0.5 g/l sodium citrate dihydrate, 0.2 g/l $MgSO_4$, 0.5 g/l thiamine hydrochloride, and 2 g/l glucose.

2.5.2. Transformation of bacteria with plasmid or M13 replicative form DNA (calcium chloride method)

Preparation of competent cells

A single colony of bacterial cells (DH1 for plasmid transformations, TG1 for M13) was grown overnight in 2 ml of the appropriate medium at 37 °C with continuous shaking. The overnight culture was added to 100 ml of fresh medium the following day and shaken at 37 °C until the absorbance of the suspension at 600 nm was approximately 0.4 (1.5-2 hr). The cells were transferred to 50 ml Oak Ridge tubes and pelleted in the JA-20 rotor of a Beckman J2-21M centrifuge for 5 min at 5000 rpm (4 °C). The supernatants were discarded and the tubes were kept on ice for the remainder of the procedure. The pellets were gently resuspended (without vortexing or hard shaking) in a total of 50 ml of $CaCl_2$ solution (0.1 M for DH1, 50 mM for TG1), and kept on ice for 20-30 min. The pellets were recentrifuged (as above) and resuspended in a total of 1-2 ml of $CaCl_2$ solution. The cells were stored on ice for up to 24 hr until required.

Transformation procedure (plasmids)

The plasmid DNA for transformation (typically a ligation product) was diluted to 100 µl with sterile distilled water and mixed with 100 µl of competent DH1 cells. The mixture was incubated on ice for 40 min, then transferred to a 42 °C water bath for exactly 3 min. Immediately afterwards, the cells were mixed with 0.8 ml of LB medium and shaken at 37 °C for 30 min. The bacterial suspension was plated out by spreading 10% and 25% of the total volume onto separate ampicillin-supplemented LB-agar plates. The plates were incubated overnight at 37 °C. Individual colonies were picked the next day for small-scale preparation of DNA.

Transformation procedure (M13)

The M13 DNA to be transformed was resuspended in a small volume (typically less than 10 μl) of water or TE and mixed with 200 μl of competent TG1 cells. The cells were incubated for 40 min on ice, then transferred to a 42 °C water bath for exactly 3 min. Immediately afterwards, 89%, 10% and 1% volume aliquots of the suspension were each mixed with 300 μl of an exponential phase culture of TG1 cells, then 3 ml of molten 2YT soft agar (45 °C) were added. Each agar mix was plated out and incubated overnight at 37 °C. Plaques were picked with a Pasteur pipette the next day for small-scale DNA preparation.

2.5.3 Transformation of bacteria with plasmid DNA (electroporation method)

Preparation of frozen cell stocks

A single colony of DH1 bacterial cells was grown in 5 ml of LB medium overnight at 37 °C with continuous shaking. The overnight culture was added to 500 ml of LB the following day and shaken at 37 °C until the absorbance of the suspension at 600 nm was in the range 0.5-1.0 (approximately 4 hr). The culture was chilled on ice for 15 min, then the cells were pelleted in a Beckman GPR centrifuge (GH 3.7 rotor) for 10 min at 3000 rpm (4 °C). The supernatants were discarded and the pellets were resuspended in a total of 500 ml of sterile distilled water. The suspensions were recentrifuged and the pellets were resuspended in a total of 250 ml of sterile distilled water. The supernatants were again discarded, the pellets were resuspended in a total of 10 ml of sterile 10% (v/v) glycerol, pooled, and centrifuged for 10 min at 3000 rpm as before. The final pellet was resuspended in 500-1000 μl of 10% glycerol and stored at -70 °C in 40 μl aliquots.

Electroporation

DNA (typically 100-500 ng) was prepared for electroporation by phenol/chloroform extraction, ethanol precipitation, and repeated washing with 70% (v/v) ethanol to eliminate salt. After drying, the DNA was resuspended in 10 μl of water, and

mixed with a 40 μl aliquot of the frozen DH1 stock (allowed to thaw briefly on ice before use). The DNA was electroporated into the cells on ice using a Bio-Rad Gene Pulser with a voltage pulse of 2.5 kV at a capacitance of 25 μF (pulse controller resistance 200 Ω). Immediately after electroporation, the cells were mixed with 700-1000 μl of SOC medium (2% bacto-tryptone, 0.5% bacto-yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) and incubated at 37 °C for 30-45 min (with continuous shaking). The bacterial suspension was plated out by spreading 25%, 10% and 5% onto separate ampicillin-supplemented LB-agar plates. The plates were incubated at 37 °C overnight. Individual colonies were picked the next day for small-scale preparation of DNA.

2.5.4. Rapid preparation of plasmid or M13 DNA from bacteria

For plasmid DNA, a bacterial culture was prepared by picking a single colony into 5 ml of ampicillin-supplemented LB medium, and incubating overnight at 37 °C (with continuous shaking). A 1.5 ml aliquot of the culture was transferred to a microcentrifuge tube and the remainder was stored at 4 °C to initiate a large-scale DNA preparation if required. The bacteria were pelleted in a microcentrifuge for 1 min and the supernatant was removed by aspiration.

For M13 DNA, a single phage plaque was picked with a Pasteur pipette into 1 ml of 2YT medium, vortexed, and heated at 65 °C for 10 min to kill the bacteria. The medium was allowed to cool for 10 min and mixed with 1 ml of an early log phase culture of TG1 cells (previously prepared by diluting an overnight culture of these bacteria 1:50 in 2YT medium and shaking at 37 °C for 1 hr). The culture was shaken for 5 hr at 37 °C, then the bacteria from a 1.5 ml aliquot were pelleted in a microcentrifuge for 2 min. The supernatant was removed and stored at 4 °C (for large scale DNA preparation).

All subsequent manipulations were carried out on ice, using the method of Birnboim and Doly (1979), as described in Sambrook *et al.* (1989). The final nucleic acid pellet was resuspended in 50 μl of TE or TNE (pH 8.0) containing 20 μg/ml DNase-free pancreatic RNase.

2.5.5. Large-scale preparation of plasmid / M13 DNA

A 500 ml culture was initiated and incubated for 18-20 hr (cultures were supplemented with 50 μg/ml chloramphenicol 4-5 hr after initiation in the case of plasmid preparations). The cells were pelleted by centrifuging for 15 min at 5000 rpm in the JA14 rotor of a Beckman J2-21M centrifuge (4 °C). The pellets were resuspended in a total of 5 ml of ice-cold 25% (w/v) sucrose, 50 mM Tris.HCl pH 8.0. Ice-cold 0.25 M EDTA pH 8.0 (2 ml) was added and the mixture was swirled and stored on ice for 5 min, then mixed with 1 ml of lysozyme solution (10 mg/ml) in TNE buffer. The mixture was swirled, incubated on ice for 15 min, then 8 ml of Triton lytic mixture (50 mM Tris.HCl pH 8, 62.5 mM EDTA, 0.1% Triton X-100) were added. The mixture was swirled and stored on ice for a further 5 min, centrifuged at 17 000 rpm for 40 min (4 °C) in the JA-20 rotor of a Beckman J2-21M centrifuge, and the supernatant collected. Nucleic acid was purified from the supernatant by the sequential addition of 7 ml of phenol/TNE and 1 ml of chloroform/amyl alcohol with vigorous shaking. Phases were separated by centrifugation (3000 rpm, 10 min, 4 °C, Beckman GPR, GH 4.7 rotor) and the aqueous phase was further extracted with 25 ml of chloroform/amyl alcohol as before. The nucleic acid in the supernatant was precipitated with ethanol (30 min at -70 °C), then pelleted by centrifugation (3000 rpm, 10 min, 4 °C, Beckman GPR, GH 4.7 rotor). The pellet was dried and resuspended in 4.5 ml of TNE buffer, to which was added 4.72 g of caesium chloride and 0.32 ml of 10 mg/ml ethidium bromide. The solution was mixed and centrifuged to equilibrium (60 000 rpm, 18-20 hr, 20 °C, Beckman L8-70M, VTi 65.2 rotor) and the purified plasmid or M13

DNA band was recovered; when necessary faint bands were first located under ultraviolet illumination.

Ethidium bromide was removed from the recovered fraction by repeated extractions with water-NaCl-saturated propan-2-ol. The final aqueous phase was diluted with three volumes of TNE and the DNA was precipitated with ethanol (-70 °C, 1 hr), pelleted by centrifuging at 10 000 rpm for 30 min (4 °C) in the Beckman J2-21M (JS-13.1 rotor), and washed with 70% (v/v) ethanol. The pellet was dried and resuspended in 300-500 μl of TE buffer.

2.6. Tissue culture and virus techniques

2.6.1. Maintenance and passage of human cell lines

HeLa, HEK-293 (Graham *et al.*, 1977), and KB8a (Babiss *et al.*, 1983) human cell lines were maintained in 90 mm tissue culture dishes containing 10 ml of medium (other culture vessels used in experiments contained volumes of medium in proportion to their surface areas). HeLa and HEK-293 cells were grown in DMEM/10% CS (or occasionally in DMEM/10% FCS); KB8a cell were grown in DMEM/10% FCS. All cells were grown at 37 °C in a 5% CO_2 humidified atmosphere.

KB8a and HEK-293 cells were passaged at an area ratio of 1:3 or 1:4 and HeLa cells were passaged at a ratio of 1:8 or 1:10 every 3-4 days. HEK-293 and HeLa cell monolayers were washed with TD and passaged with TDE/trypsin (TD supplemented with final concentrations of 0.5 mM EDTA and 0.25 g/l trypsin); KB8a monolayers were washed and passaged with versene only.

2.6.2. Freezing and recovery of cell stocks

Long-term cell stocks were prepared by suspending the cells from one 90 mm dish in 0.5 ml of 8% (v/v) dimethylsulphoxide in FCS. The freezing vials were slowly

cooled to -70 °C, then immersed in liquid nitrogen. Cells were recovered by rapidly thawing the contents of one vial and transferring the cells to a dish of pre-warmed medium. The medium was changed, or the cells passaged, the following day.

2.6.3. Calcium phosphate transfection for adenovirus reconstruction by *in vivo* recombination

The HEK-293 cells from one 90 mm dish were passaged and divided between three 60 mm dishes 24 hr before transfection. Three hours before transfection the medium was changed to DMEM/10% FCS. DNA/calcium phosphate precipitates were prepared by mixing 50 μl of 2x HEPES buffer (16 g/l NaCl, 0.74 g/l KCl, 2 g/l glucose, 10 g/l HEPES), 3 μl of 100x phosphate solution (9.9 g/l Na_2HPO_4), and DNA solutions + water to 90 μl, then adding 10 μl of 1.25 M calcium chloride and blowing air through the solution. The precipitates were allowed to form for 30 min at room temperature, then the medium was removed from the dishes and the cells were overlaid with the precipitate suspensions diluted to 1 ml with DMEM/5% FCS. The cells were then incubated for 2.5-3 hr at 37 °C (5% CO_2), with occasional rocking. The medium was removed from each dish, and 1 ml of 20% (v/v) glycerol in TS buffer was carefully added to each monolayer. This was removed after exactly 1 min, and the cells were washed twice with TS, then overlaid with DMEM/10% CS. The cells were incubated at 37 °C for 30-60 min, then overlaid with semi-solid medium (1) (4% (v/v) CS, 0.375% (w/v) $NaHCO_3$, 1% (w/v) Noble agar, in 1x DMEM). After 3-4 days this was supplemented with semi-solid medium (2) (1% (v/v) CS, 0.1875% (w/v) $NaHCO_3$, 1% (w/v) Noble agar, in 1x DMEM). When plaques became visible (6-9 days) they were stained with neutral red semi-solid medium (1% (v/v) CS, 0.1875% (w/v) $NaHCO_3$, 0.01% (w/v) neutral red, 1% (w/v) Noble agar, in 1x DMEM) and picked the next day with a Pasteur pipette.

2.6.4. Calcium phosphate transfection for transient assays of gene expression

This procedure was similar to 2.6.3, above, with the following exceptions. The medium was changed to DMEM/10% FCS 4 hr before transfection. The DNA/calcium phosphate precipitate was prepared by slowly adding 500 µl of 0.25 M $CaCl_2$ containing 5 µg of plasmid DNA and 10 µg of salmon sperm DNA to 500 µl to a freshly prepared solution containing 490 µl of 2x HEPES buffer and 10 µl of 100x phosphate solution, blowing air through the solution, and allowing to stand for 30 min at room temperature. The precipitate was added directly to the medium in a 60 mm dish, and the cells were incubated for 4 hr. After glycerol shock, the cells were retained in DMEM/10% CS.

2.6.5. Virus infection of cell monolayers

Cell monolayers at 70-100% confluence were gently overlaid with virus suspension (diluted, if necessary, in TS/2% CS), and the plates rocked to ensure even distribution of virus across the monolayers. Volumes of the suspensions used were 200 µl for 35 mm or 60 mm dishes, 500 µl for 90 mm dishes, or 2.5 ml for 150 mm dishes. The plates were incubated for 30 min (37 °C), rocked again, and incubated for a further 30 min. The virus suspension was removed by aspiration and replaced with fresh medium.

2.6.6. Rapid preparation of adenovirus DNA

HEK-293 cell monolayers in 90 mm dishes were harvested when full cytopathic effect became visible (usually in 3-4 days) and cells were pelleted by centrifuging for 5 min at 1500 rpm in a Beckman GPR (GH 3.7 rotor). The supernatant was removed and the pellet was resuspended in 0.4 ml of TE (pH 9)/10 mM spermine tetrahydrochloride (Sigma). The suspension was mixed with 0.4 ml DOC lysis buffer (20% (v/v) ethanol, 100 mM Tris.HCl pH 9, 0.4% (w/v) sodium deoxycholate) and cell debris was pelleted by centrifugation (15 min, 12000 rpm, microcentrifuge). The supernatant was collected and 60 µl 10% (w/v) SDS, 40 µl

0.25 M EDTA, and 20 μl 20 mg/ml proteinase K (Boehringer) were added. The solution was mixed several times by inversion and incubated at 37 °C for 1 hr.

Residual protein was then extracted with phenol/chloroform and the DNA precipitated by adding 30 μl 5 M NaCl and 0.6 ml of propan-2-ol at room temperature. The solution was mixed several times by inversion and the DNA was pelleted by centrifuging for 15 min at 12 000 rpm in a microcentrifuge (room temperature). The pellet was dried and resuspended in 50 μl distilled water or TE.

2.6.7. Large-scale preparation of adenovirus DNA

Infected cells were harvested as described above. A typical preparation used cells from ten 90 mm monolayer cultures. The cells were resuspended in 10 ml of 0.1 M Tris.HCl (pH 8) and disrupted by sonication on ice (two sets of 10 x 1 s pulses separated by 30 s using a Jencons GE375 ultrasonic processor and 3 mm tapered microtip). The sonicates were clarified by centrifugation at 6000 rpm for 10 min (4 °C) in a Beckman J2-21M centrifuge, (JS-13.1 rotor). The supernatants were layered over CsCl step gradients consisting of 2 ml of CsCl in TD (density 1.40 g/ml) with an overlay of 3 ml of CsCl in TD (density 1.25 g/ml). The gradients were centrifuged at 35 000 rpm for 1 hr (15 °C) in a Beckman L8-70M ultracentrifuge (SW 41 rotor). The virus formed a clearly visible band (below cellular debris) which was recovered by puncturing the bottom of the tube and collecting drip fractions. Collected virus was then centrifuged to equilibrium (at least 15 hr) in CsCl/TD of density 1.35 g/ml at 40 000 rpm (15 °C, Beckman L8-70M ultracentrifuge, SW 50.1 rotor).

The virus band was collected (as above), diluted with two volumes of water and precipitated with six band volumes of ethanol at -70 °C for 15 min. The virus was pelleted by centrifuging for 20 min at 7000 rpm (4 °C) in a Beckman J2-21M centrifuge (JS-13.1 rotor). The pellet was dried and gently resuspended in 2 ml of

TNE, then 120 µl 10% (w/v) SDS, 40 µl 0.25 M EDTA, and 20 µl 10 mg/ml proteinase K were added. The solution was incubated at 37 °C for 1-2 hr then extracted with 2 ml of phenol/TNE at room temperature. The tube was gently swirled for 1 min, 1 ml chloroform/amyl alcohol was added, and the solution was again gently swirled to mix. Phases were separated by centrifuging for 10 min at 3000 rpm (4 °C) in a Beckman GPR centrifuge, and the organic extraction and centrifugation were repeated with the aqueous phase. DNA was precipitated from the final aqueous phase by adding sodium acetate solution to 0.3 M and two volumes of ethanol. The DNA was allowed to precipitate at -70 °C for 30 min, and then pelleted at 7000 rpm for 30 min (4 °C) in a Beckman J2-21M centrifuge (JS-13.1 rotor). The DNA pellet was washed with 70% ethanol, dried, and allowed to redissolve in 0.5 ml TE at 4 °C overnight.

2.6.8. Preparation of virus stocks

Infected cells were harvested when cytopathic effect was visible (usually in 3 days) and pelleted by centrifuging (Beckman GPR centrifuge, GH 3.7 rotor, 1500 rpm, 3 min). To concentrate the stocks, half the medium was removed by aspiration, and the cells were lysed in the remainder by three freeze-thaw cycles. Cell debris was pelleted by centrifugation (as before). The supernatant (virus stock) was removed and stored in 2-3 ml aliquots at -70 °C.

2.6.9. Titration of virus stocks by plaque assay

Tenfold serial dilutions with TS/2% CS in the range 10^{-1} to 10^{-8} were prepared from the virus stock; 200 µl aliquots of the appropriate dilutions were used to infect duplicate cultures of HEK-293 cells in 60 mm dishes or 35 mm wells in 6-well trays. Cells were then overlaid with semi-solid medium (1); after 3-4 days a second overlay of semi-solid medium (2) was added (the compositions of semi-solid media are given in 2.6.3). When plaques became visible, they were stained by the addition of neutral red semi-solid medium and counted the next day.

2.7. Analysis of viral protein expression

2.7.1. ^{35}S-methionine pulse-labelling of proteins

Cell monolayers on 35 mm dishes were incubated with 0.5 ml of methionine-free DMEM for 30 min. This was then replaced with 0.25 ml of similar medium supplemented with 25 μCi of ^{35}S-methionine (specific activity >800Ci/mmol, Amersham). The dish was incubated for a further 30 min, then placed on ice. The labelling medium was removed, the cells were washed with ice-cold PBS, and lysed by the addition of 0.2 ml of RIPA buffer (150 mM NaCl, 10 mM Tris.HCl pH 7.6, 1% (v/v) Triton X-100, 1% (w/v) sodium deoxycholate, 0.1% (w/v) SDS). The dish was incubated for 15 min on ice, then the lysate was harvested by pipetting and stored at -70 °C.

2.7.2. Quantification of ^{35}S-methionine incorporation into protein

A small aliquot (5-10 μl) of labelled cell extract was diluted with 400 μl of distilled water and mixed with 100 μl of 100% (w/v) trichloroacetic acid (TCA) in water. The solution was allowed to stand on ice for 15 min, and the precipitate was collected on a 2.5 cm Whatman GF/C glass microfibre filter. The filter was washed with 1 ml of 5% (w/v) TCA, then with 1 ml of ethanol, and allowed to dry. The filter was placed in 5 ml of Optiphase Safe scintillation fluid and the radioactivity measured in an LKB Wallac 1219 Rackbeta scintillation counter.

2.7.3. SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples in 100 mM DTT, 25 mM Tris.HCl pH 6.8, 10% (v/v) glycerol, 2% (w/v) SDS, and 0.005% (w/v) bromophenol blue were heated to 95 °C for 3 min. Discontinuous SDS-polyacrylamide gels were prepared and run in a BRL vertical gel electrophoresis system according to the method of Laemmli (1970). The resolving gel contained 0.43 M Tris.HCl pH 8.8, 9.98% (w/v) acrylamide, 0.27% (w/v) N,N'-methylene-bis-acrylamide, 0.095% (w/v) SDS, 0.67 mg/ml ammonium

persulphate and 0.20 µl/ml TEMED. The stacking gel contained 0.12 M Tris.HCl pH 6.8, 4.57% (w/v) acrylamide, 0.12% (w/v) N,N'-methylene-bis-acrylamide, 0.10% (w/v) SDS, 0.95 mg/ml ammonium persulphate and 0.48 µl/ml TEMED. Gels were run at variable voltage and current (40 mA, 250 V set maxima) for 2-3 hr in a buffer containing 303 g/l Tris, 1442 g/l glycine, and 100 g/l SDS (buffer pH 8.3). Gels were fixed for at least 1 hr in 25% (v/v) methanol, 7% (v/v) acetic acid, treated for 15 min with Amplify (Amersham) for fluorographic enhancement of radioactive emissions, then dried for 1-2 hr at 70 °C under vacuum.

2.8. Analysis of viral RNA expression

2.8.1. Preparation of cytoplasmic and nuclear RNA

(All manipulations were carried out on ice or at 4 °C)

The cells from a 150 mm dish were washed twice in 10 ml of complete PBS buffer (PBS supplemented with final concentrations of 0.1 g/l $CaCl_2$ and 0.1 g/l $MgCl_2$) and pelleted by centrifuging at 1500 rpm for 3 min (4 °C) in a Beckman GPR centrifuge (GH 3.7 rotor). The cells were resuspended in 2 ml of isotonic buffer (150 mM NaCl, 10 mM Tris.HCl pH 7.6, 1.5 mM $MgCl_2$), then 200 µl of 10% (v/v) Nonidet P40 (NP40) were added. The suspension was vortexed, allowed to stand for 10 min on ice, then pelleted at 1500 rpm for 3 min (4 °C) in the Beckman GPR centrifuge. The supernatant was removed and the pellet was resuspended in 1 ml of isotonic buffer and pelleted as before. Both supernatants (containing cytoplasmic RNA) were combined, then 60 µl of 10% (w/v) SDS and 120 µl of 0.25 M EDTA were added. The pellet (containing nuclear RNA) was resuspended in 1 ml of RSB (10 mM Tris.HCl pH 7.6, 3 mM $MgCl_2$, 10 mM NaCl), and 2 µl of 10 mg/ml DNase I (Gibco BRL, RNase-free) were added. The suspension was incubated for 30 min on ice, then 20 µl of 10% SDS and 40 µl of 0.25 M EDTA were added to lyse the nuclei. The solutions of both cytoplasmic and nuclear RNA were extracted with one volume of phenol/TNE and one volume of chloroform/amyl alcohol, then centrifuged for 10 min at 3000 rpm (4 °C)

Table 2.1. pGEM plasmids used to transcribe probes for RNase protection analyses

Name	Promoter	Linearised with	Ad5 fragment	pGEM sites
pGEM3.L3-hexon	T7	*Hind* III	*Hind* III 18310 -*Sma* I 18922	*Hind* III *Sma* I
pGEM3.L3-23K	Sp6	*EcoR* I	*BamH* I 21562 -*Bgl* II 21775	*BamH* I
pGEM4.L2-3'/L3-5'	T7	*Nru* I	*Kpn* I 17068 -*Hind* III 18323	*Kpn* I *Hind* III
pGEM4.L3/E2A	T7	*EcoR* I	*Kpn* I 22188 -*Dra* I 22444	*Kpn* I *Hinc* II
pGEM3.L4-5'	T7	*Hind* III	*Pst* I 23286 -*Pst* I 24369	*Pst* I
pGEM3.L1-3'/L2-5'	T7	*Hind* III	*Hind* III 13721 -*Kpn* I 14291	*Hind* III *Kpn* I
pGEM3.E1A	Sp6	*EcoR* I	*Pvu* II 455 -*Pst* I 1838	*Sma* I *Pst* I
pGEM4.E3/L4	T7	*EcoR* I	*EcoR* I 27331 -*Bgl* II 28137	*EcoR* I *BamH* I
pGEM2.L5	Sp6	*Hind* III	*Hind* III 31993 -*Sma* I 33096	*Hind* III *Sma* I
pGEM3.IVa2	T7	*Hind* III	*Bst* E2 5186 *Xho* I 5792	*Sma* I *Sal* I

Notes:

For each clone the enzyme required to linearise the plasmid for transcription and the promoter from which an antisense probe was transcribed are indicated. Details of the adenovirus DNA inserted into each vector, and the cloning sites utilised in pGEM2, pGEM3, or pGEM4 (Promega) are also shown. Plasmids pGEM4.L2-3'/L3-5 and pGEM4.L3/E2A were prepared in the course of the project. pGEM3.L3-hexon and pGEM3.L3-23K were prepared by P. Dancey; pGEM3.L1-3'/L2-5' was prepared by S. Khandke; pGEM3.L4-5' was prepared by K. Leppard. All other plasmids listed were constructed and provided by J. Schaack (Dept Microbiol. & Immunol., University of Colorado, Denver, CO, USA).

(Beckman GPR). The extraction of the aqueous phase was repeated with one volume of phenol/TNE and two volumes of chloroform/amyl alcohol, then with two volumes of chloroform/amyl alcohol. The final aqueous phase was precipitated with ethanol (-70 °C, 1 hr) and the RNA was pelleted by centrifuging for 30 min at 10 000 rpm (4 °C) in a Beckman J2-21M centrifuge (JS-13.1 rotor). The pellet was dried and resuspended in 300-350 μl of water.

2.8.2. Preparation of antisense RNA probes

Preparation of template DNA

A 2 μg stock of each linearised probe template (table 2.1) was prepared by appropriate digestion of the relevant pGEM vector subclone (i.e. the vector was linearised at an enzyme site between the 3' end of the sequence to be transcribed and the unused promoter). The template DNA stock was purified by phenol/chloroform extraction and ethanol precipitation and dissolved in water at a concentration of 0.32 μg/μl.

In vitro *transcription in the pGEM Sp6 / T7 system*

The components of the *in vitro* transcription reaction were mixed at room temperature, and in the order listed below, to avoid precipitation of the DNA by spermidine (a constituent of the 5x transcription buffer)

	Volume/μl
10 μCi/μl α-^{32}P UTP (Amersham)	2.5
5x Sp6 or T7 buffer (Gibco BRL)	1.25
10x nucleotide mix	0.625
200 mM DTT (Sigma)	0.3125
Water	0.3125
Human placental RNase inhibitor (Gibco BRL)	0.25
Template DNA (0.32 μg/μl)	0.625

Sp6 or T7 polymerase (Gibco BRL) 0.40

(where 10x nucleotide mix contained 5 mM ATP, 5 mM CTP, 5 mM GTP, 1 mM UTP)

The reaction was incubated at 37 °C (for T7 polymerase) or 40 °C (for Sp6 polymerase) for 1 hr.

Purification of probe

The reaction was diluted to 50 µl with water, then 2 µl of vanadyl ribonucleoside complex (200 mM Gibco BRL) and 1 µl of RNase-free DNase I were added. The reaction was incubated at 37 °C for 10 min, SDS (0.5 µl 10% w/v) was added, and the solution was extracted with phenol/chloroform. The probe was finally purified on a Stratagene Nuctrap push column (according to the manufacturer's instructions) and stored at -20 °C.

2.8.3. Ribonuclease protection analysis

Ribonuclease protection assays were carried out according to the method of Melton *et al.* (1984) :

Hybridisation

5 µg of RNA and 5 µl of antisense riboprobe were co-precipitated in three volumes of ethanol for 30 min at -20 °C. The dried RNA pellet was resuspended in 20 µl of hybridisation buffer (80% (v/v) deionised formamide; 40 mM PIPES pH 6.7, 0.4 M NaCl, 1 mM EDTA), heated to 80 °C for 10 min, allowed to cool slowly (over several hours) to 56 °C, and incubated at this temperature overnight.

Ribonuclease digestion

The hybridisation reactions were diluted to 200 μl in RNase solution (10 mM Tris.HCl pH 7.5, 5 mM EDTA, 300 mM NaCl, 100 units/ml RNase T1, 5 μg/ml pancreatic RNase A) and incubated for 30 min at room temperature. SDS to 0.5% and 1 μl of 20 mg/ml proteinase K were added, the solutions were vortexed vigorously, and the reactions were incubated for a further 15 min at room temperature. The solutions were extracted once with phenol/chloroform, and RNA precipitated with three volumes of ethanol (in the presence of 1 μg of carrier tRNA) for 30 min at -20 °C. Dried RNA pellets were resuspended in 6 μl of formamide dye mix (80% (v/v) deionised formamide, 10 mM EDTA pH 8.0, 1 mg/ml xylene cyanol, 1 mg/ml bromomphenol blue) and denatured by heating for 2 min at 90 °C, then cooled rapidly on ice.

Separation on denaturing acrylamide gel

Using a Bio-Rad SequiGen apparatus, a 5% polyacrylamide denaturing gel was prepared in 1x TBE buffer, containing 4.75% (w/v) acrylamide, 0.25% (w/v) N,N'-methylene-bis-acrylamide, 46% (w/v) urea, 0.8 mg/ml ammonium persulphate and 0.83 μl/ml TEMED. A 2 μl aliquot of each sample electrophoresed on the gel for 2-3 hr in 1x TBE buffer at 45 W (running temperature 50 °C). The gel was dried for 90 min at 60 °C under vacuum, and the RNA bands were visualised by autoradiography.

Preparation of molecular size markers for denaturing gel

Labelled markers were prepared by digesting 1 μg of pBR322 DNA with restriction enzyme Hpa II, then adding 1 unit of Klenow DNA polymerase and 5 μCi of α-^{32}P-dCTP to the reaction mix and incubating for a further 10 min at room temperature. The mix was diluted with 30 μl of formamide dye mix, and the amount required to give clearly visible markers was determined empirically (typically 1-2 μl when

freshly prepared, but usable at increased volumes for several weeks when stored at -20 °C).

2.9. Autoradiography

Preflashed Fuji RX X-ray film was exposed to dried radioactive gels at -70 °C (initially overnight, then for longer periods as required). The preflash distance was calibrated by spectrophotometry - the absorbance at 545 nm of preflashed film was measured and compared to the A_{545} for unflashed film. The preflash distance was adjusted until the absorbance ratio of preflashed to unflashed was in the range 0.1-0.2, corresponding to a linear response of the film to radioactivity.

2.10. Densitometry

Film band intensities were quantified on a Molecular Dynamics scanning laser densitometer, using Image Quant 3 analysis software.

2.11. Chloramphenicol Acetyl Transferase (CAT) assays

2.11.1. Preparation of cytoplasmic extracts

(All manipulations were carried out on ice or at 4 °C).

Cytoplasmic extracts were prepared from virus-infected or mock-infected HEK-293 cell monolayers in 60 mm dishes. Each monolayer was washed with 3 ml of ice-cold phosphate buffered saline (PBS). The cells were resuspended in 6 ml of PBS and pelleted by centrifugation at 1500 rpm for 3 min (4 °C) in a Beckman GPR centrifuge (GH 3.7 rotor). Each cell pellet was resuspended in 150 µl ice-cold CAT lysis buffer (250 mM Tris.HCl pH 7.8, 5 mM DTT, 10% (v/v) glycerol, 0.25% (v/v) Nonidet-P40) and kept on ice for 10 min. The cells were further broken by three freeze-thaw cycles. Alternatively, the extracts were stored at -70 °C prior to freeze-thaw until required. Cell debris was pelleted in a microcentrifuge at 15 000 rpm for 12 min and the supernatant (i.e. the cytoplasmic extract) was kept on ice until required.

2.11.2. Assay for protein concentration

The protein content of each cytoplasmic extract was determined using the Bio-Rad protein assay reagent (following the manufacturer's microassay procedure derived from Bradford, 1976). A small quantity (1-2 μl) of extract was diluted to 800 μl with distilled water, mixed with 200 μl of assay reagent and vortexed briefly. The colour change was allowed to develop for 5 min, then the absorbance at 595 nm was measured against a reagent blank (200 μl reagent + 800 μl water). Where necessary, the assay was repeated with diluted samples to give absorbance readings in the linear range of the assay (A_{595} = 0.1-0.6, corresponding to 1-10 μg protein).

2.11.3. CAT assay procedure

CAT activity was assayed by the direct scintillation diffusion method (Neumann *et al.*, 1987; Eastman, 1987), which relies on the differential solubilities of acetyl CoA and acetylated chloramphenicol in a nonaqueous scintillant (4.3, below). Each cytoplasmic extract was heated to 65 °C for 15 min to reduce background signal by inactivating endogenous acetyltransferase activity (CAT itself is comparatively resistant to heat inactivation). Five absorbance units of cytoplasmic extract (where one unit is defined as the quantity of extract required to produce an absorbance of 1.0 in the protein assay) were mixed with 50 μl 5 mM chloramphenicol (aq), 25 μl 1 M Tris.HCl (pH 7.8), 124 μl distilled water, and 100 μl 0.1 M Tris.HCl in a plastic scintillation vial. The assay reaction was started by adding 1 μl (0.1 μCi) ^{3}H-acetyl coenzyme A (DuPont NEN); immediately afterwards the assay mixture was covered with 3 ml of a water-immiscible scintillation fluid (Econofluor, DuPont NEN) and the vial was placed at 37 °C. The formation of ^{3}H-acetyl chloramphenicol (and hence the activity of CAT) was measured in a scintillation counter (LKB Rackbeta) for 10 s per reaction every 30 min over a period of 2 hr.

The reaction was allowed to continue at 37°C overnight and final readings were taken the next day.

Chapter 3

Construction and isolation of recombinant adenoviruses

3.1. Introduction

In this chapter I will describe the strategies employed to construct and isolate the HIV-adenovirus recombinants used in all subsequent experiments.

As noted in the general introduction, the aim of this study was to investigate the action of the Rev/RRE system in adenovirus, and more specifically to compare the effects of this system with those of Ad E1B 55K. The approach taken was to insert a *rev* expression construct and the RRE into suitable locations in the adenovirus genome, and analyse their effects on gene expression in either the presence or absence of E1B 55K. Several factors were considered before selecting the cloning strategy used to insert the foreign DNA into the Ad genome: First, it was clearly vital to place the *rev* gene in a context where it could be expressed at adequate levels, and to insert the RRE in a position where it could be incorporated into otherwise normal viral transcripts - i.e., to provide the components of a functional Rev/RRE system that could be compared with E1B 55K. Second, it was important that insertion of heterologous DNA should cause minimal disruption to adenovirus gene expression. Third, to facilitate cloning, the availability of suitable restriction sites in the genome was evaluated.

In the case of the RRE, examination of the Ad5 sequence showed that suitable sites for insertion were strictly limited. In particular, the complexity of the pattern of gene expression from both strands of the genome made it difficult to find locations that did not disrupt at least one coding sequence. However, a 32 nt sequence was located in the L3 region of the major late transcription unit between the translation stop signal for the L3 hexon (major capsid) protein, and the ATG start codon of the L3 23K protease. This sequence was also non-coding on the opposite strand of the genome. Inserting the RRE in this position results in its incorporation into the L3 RNAs coding for pVI (hexon-associated protein precursor), hexon protein, and 23K protease (fig. 3.1). Each L3 RNA normally requires the E1B 55K protein to

Fig. 3.1. rev and RRE insertion sites in Ad5

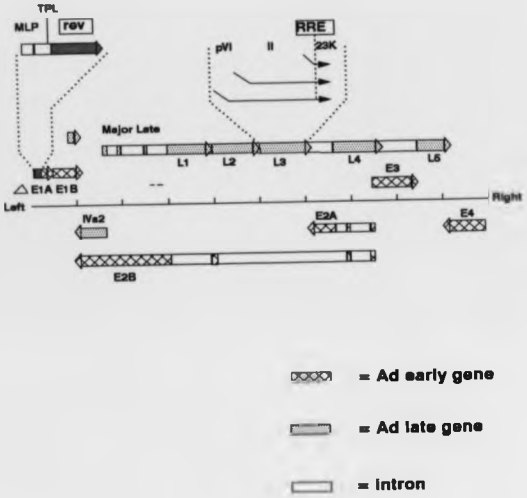


Fig. 3.1. *rev* and RRE insertion sites in Ad5. The diagram is a simplified version of the adenovirus transcription map in chapter 1, with the insertion positions of the *rev* expression cassette (in place of E1A) and the RRE (in the 5' untranslated region of the L3-23K 3' exon) indicated. MLP=major late promoter; TPL=tripartite leader sequence.

facilitate its expression, so insertion of the RRE in this region makes a direct comparison of the two regulatory systems possible. Moreover, this intergenic sequence contains a convenient Pvu II restriction site for cloning. Therefore this region was chosen as the target for RRE insertion.

In the case of *rev*, the gene was cloned into an expression construct (based on pNL3C - a kind gift of Dr. R Schneider, New York University) containing the adenovirus E1A enhancer, major late promoter, and tripartite leader, which enhances the translatability of mRNA expressed from this promoter (Logan and Shenk, 1984). This construct was designed for insertion into the adenovirus genome in place of most of early region E1A (fig. 3.1). In this context, all the necessary elements for efficient expression of the cloned gene are present - upstream sequences are supplied by the construct, and E1A RNA processing signals from the virus can be utilised. The absence of a large section of the E1A frame in such a recombinant makes the virus defective in cell lines such as HeLa, since E1A products are required to transactivate other viral genes. However, if E1A is supplied in *trans* from a complementing cell line such as HEK-293 or KB-8a the virus can go through a successful infectious cycle. Moreover, for safety reasons, the failure of such a virus to infect noncomplementing cells productively is an advantage, since it is unclear whether Rev expression has any deleterious effects on host cells which might make a replication competent virus hazardous. Therefore, the *rev* gene was inserted in a major late promoter expression construct in this region, in place of E1A.

Although adenovirus DNA can be manipulated directly, it is generally more convenient to work on smaller fragments of the genome subcloned into plasmid vectors. Constructs containing appropriate regions of the Ad5 genome and the RRE or *rev* expression cassette were therefore prepared (diagrams, fig 3.2). Complete viral DNA can be reconstructed either by ligation of appropriate cloned and viral

Fig. 3.2. Viral reconstruction plasmids

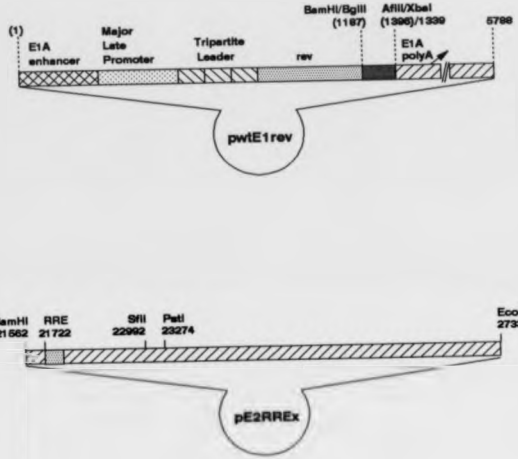


Fig. 3.2. Viral reconstruction plasmids. The diagram illustrates the essential features of the plasmids used to insert the *rev* expression cassette and the RRE into adenovirus. The *rev* cassette plasmids pwtE1*rev* and pdl338E1*rev* are identical except for deletion of the 55K gene (Ad5 2804-3328 *Hin*d III-*Bgl* II fragment) in the latter. Nucleotide numbers in parentheses are relative to the start of the *rev* expression cassette in pwtE1*rev*. All other numbering refers to the unmodified Ad5 sequence.

Fig. 3.3. 3-fragment ligation (e.g. dl309/RRE)

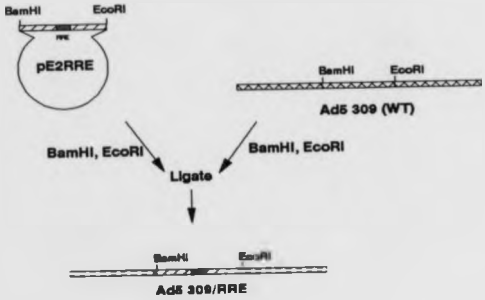


Fig. 3.3. 3-fragment ligation (e.g. *dl*309/RRE). The diagram shows the method of virus reconstruction by 3-fragment ligation, as exemplified by *dl*309/RRE.

Fig. 3.4. Overlap recombination in vivo (e.g. dl309/rev)

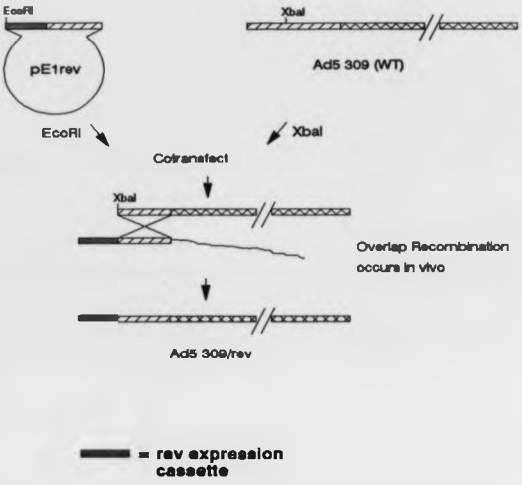


Fig. 3.4. Overlap recombination *in vivo* (e.g. *dl*309/*rev*). The diagram shows the method of virus reconstruction by overlap recombination *in vivo*, as exemplified by *dl*309/*rev*.

Fig. 3.5 Isolation of adenovirus recombinants

Fig. 3.5. Isolation of adenovirus recombinants. The diagram lists the stages involved in isolating recombinant adenoviruses by the methods detailed in chapter 2.

Table 3.1 : Recombinant Adenoviruses

	E1B 55K	RRE	rev	name
1	+	-	-	309 (WT)
2	-	-	-	338
3	+	+	-	309/RRE
4	-	+	-	338/RRE
5	+	+	+	309/RRErev
6	-	+	+	338/RRErev
7	+	-	+	309/rev
8	-	-	+	338/rev

genomic fragments, or by overlap recombination *in vivo* between such fragments (Chinnadurai *et al.*, 1979). These two methods were used to reconstitute complete genomes from the RRE and *rev* cassettes respectively, as illustrated in figs. 3.3 and 3.4. Since naked viral genomes are infectious, stocks of viable virus can be recovered from cells transfected with recombinant viral DNA. In practice, this involves an extended process of plaque purification, screening, growth, and titration by plaque assay, summarised in fig 3.5. The details of subcloning, virus reconstruction, screening and titration are given in the following sections of this chapter.

In order to make comparisons between the Rev/RRE and E1B 55K systems, and to differentiate any individual effects of the RRE or Rev alone from those of the complete Rev/RRE system, it was necessary to construct a series of recombinants - i.e., viruses either containing or not containing the RRE, *rev* construct, and E1B 55K gene in all possible combinations. The genotypes of the eight possible viruses are listed in table 3.1. Two of these - *dl*309 (effectively wild-type for this purpose) and *dl*338 (lacking E1B 55K) - were already available (Jones and Shenk, 1979; Pilder *et al.*, 1986). The other viruses were constructed and isolated in the course of this project as described below.

3.2. Insertion of the RRE into adenovirus - Construction of *dl*309/RRE

3.2.1 Subcloning

(All adenovirus numbering refers to the published complete Ad5 sequence of Chroboczek *et al.*, 1992, GenBank accession number M73260, as modified by Dix and Leppard, 1992, GenBank accession number D12587).

The following cloning strategy is summarised in fig. 3.6. The starting materials for subcloning were plasmids pE2, pGEMRRE and pBR322. Plasmid pE2 (originated in the laboratory of Prof. T. Shenk, Princeton, NJ, and obtained from K. Leppard) contains a 5769 bp fragment of Ad5 DNA, running from the *Bam*H I site at

Fig. 3.6. Construction of pE2RREx

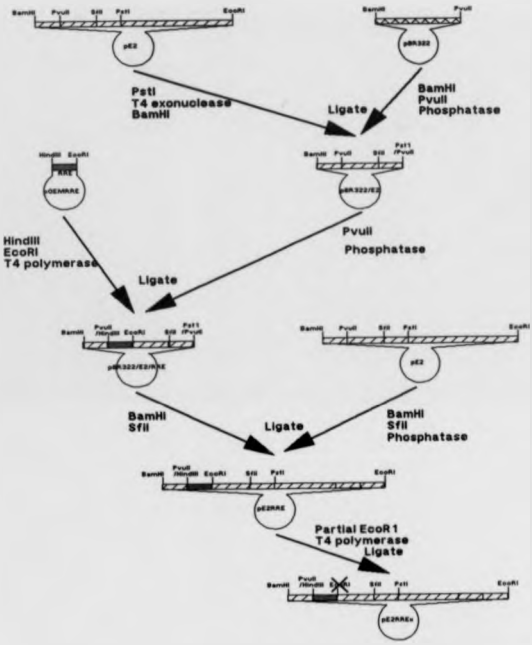


Fig. 3.6. Construction of pE2RREx. The diagram summarises the subcloning strategy used to construct pE2RREx, the plasmid used to insert the RRE in the Ad5 genome.

21562 bp to the *Eco*R I site at 27331 bp, cloned into pBR322 between the *Bam*H I and *Eco*R I sites. Plasmid pGEMRRE (kindly supplied by Dr. Shaun Heaphy, Dept. of Microbiology, University of Leicester) contains a *Sty* I fragment (7787-8009 bp) covering the RRE region from HIV-1 strain ARV2, cloned into pGEM1 between the *Pst* I and *Sac* I sites (Heaphy *et al.*, 1990).

The pE2 plasmid was digested with *Pst* I, and 3' overhangs were converted to blunt ends with T4 polymerase (exonuclease activity). The DNA was further digested with *Bam*H I, and the *Bam*H I-*Pst* I fragment (Ad5 sites 21562-23274 bp) was isolated and cloned into pBR322 between the *Bam*H I and *Pvu* II sites, to make plasmid pBR322/E2. The RRE was excised from the polylinker of pGEMRRE as a *Hin*d III-*Eco*R I fragment, and 5' overhangs were filled in with T4 polymerase (polymerase activity). The blunted RRE fragment was cloned into pBR322/E2 at the *Pvu* II site (Ad5 site 21722 bp). Recombinants were selected with the RRE in the correct 5'-3' orientation relative to the Ad5 sequence by restriction analysis (plasmid pBR322/E2/RRE). A *Bam*H I-*Sfi* I (Ad5 21562-22992 bp) fragment was excised from pBR322/E2/RRE and cloned back into pE2 in place of the equivalent unmodified fragment. The resulting plasmid, pE2RRE, was equivalent to pE2 with the RRE inserted at *Pvu* II site 21722 bp (Ad5 numbering). By inserting the RRE at a *Pvu* II site, the blunted *Eco*R I site at the 3' end of the RRE cassette was regenerated. However, since it was necessary to have only a single *Eco*R I site in the plasmid for use in reconstruction of the virus (see below), the regenerated site was removed by partially digesting the plasmid with *Eco*R I, selecting singly-cut DNA, end-filling with T4 polymerase, and religating. The recombinants were screened by restriction analysis for clones end-filled at the regenerated site (destroying the recognition sequence), but intact at the original site in pE2 (Ad5 site 27331 bp). The required clone, pE2RREx (fig. 3.2), was used in the reconstruction of complete viral DNA (see below).

3.2.2. Virus reconstruction by ligation

The reconstruction procedure is summarised in fig. 3.3. Full-length adenovirus DNA containing the RRE inserted in the *Pvu* II site at 21722 bp was reconstructed from pE2RREx and cleaved genomic DNA from Ad5 deletion variant *dl*309. The *dl*309 variant (Jones and Shenk, 1979) has an essentially wild-type phenotype, but contains unique *Eco*R I (27331 bp) and *Xba* I (1339 bp) sites, facilitating virus reconstruction by ligation or overlap recombination. Plasmid pE2RREx (5 μg) was cleaved at the unique *Bam*H I and *Eco*R I sites (Ad5 sites 21562 and 27331 bp), and the 6 kbp adenovirus-RRE fragment was isolated. Ad5 *dl*309 DNA (3 μg) was cleaved at the equivalent genomic *Bam*H I and *Eco*R I sites (generating three fragments), then dephosphorylated with alkaline phosphatase. The 6 kbp fragment from pE2RREx was ligated with the cleaved viral DNA (dephosphorylation prevented religation of *dl*309 'wild-type' DNA - a complete viral genome could only be generated by incorporation of the phosphorylated RRE-recombinant central fragment). The ligation products were extracted with phenol/chloroform/amyl alcohol, ethanol precipitated, and resuspended in sterile water for transfection. Recombinant virus and viral DNA were isolated by the methods outlined in fig. 3.5 (see chapter 2 for detailed methodology).

3.3. Insertion of *rev* expression cassette into adenovirus - Construction of *dl*309/*rev*

3.3.1 Subcloning

The following cloning strategy is summarised in fig. 3.7. The starting materials for cloning the *rev* expression construct were M13mp19*rev*#209 (constructed by Dr. Shaun Heaphy and kindly supplied by Dr. Ian Jones, NERC Institute of Virology, Oxford), pNL3C, and pwtXho1-C, (originated in the laboratory of Prof. T. Shenk, Princeton, NJ, and obtained from K. Leppard). M13mp19*rev*#209 contains a synthetic *rev* gene, designed to be translated as a protein identical to the Rev of HIV-1 strain BRU, cloned between the *Hin*d III and *Bam*H I sites of the M13mp19

Fig. 3.7. Construction of pwtE1rev

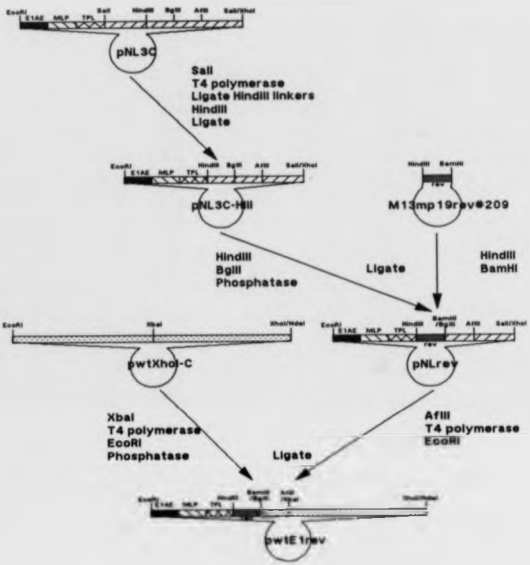


Fig. 3.7. Construction of pwtE1*rev*. The diagram summarises the subcloning strategy used to construct pwtE1*rev*, a plasmid used to insert the *rev* expression cassette into the Ad genome.

polylinker (Heaphy *et al.*, 1990). Plasmid pNL3C contains Ad5 sequences between the *Eco*R I and *Sal* I sites of the pML2 vector. The Ad5 sequences begin with the E1A enhancer, located upstream of the major late promoter and the elements of the tripartite leader. The leader sequences are fused to the *Sma* I site at 1008 bp (in E1A) with *Sal* I linkers, making a *Sal* I site available for cloning. The Ad5 sequence continues up to the *Xho* I site at 5788 bp, which is joined via complementary cohesive ends to the *Sal* I site of the vector. Plasmid pwtXho1-C contains the *Xho* I-C fragment of Ad5 (1-5788 bp) cloned into a pBR322 derivative between the *Eco*R I and *Nde* I (modified with a *Xho* I linker) sites.

Plasmid pNL3C was linearised by cleavage at the *Sal* I site (immediately 3' to the tripartite leader) and ligated to *Hind* III linkers. The DNA was digested with *Hind* III, which removed linker multimers and Ad5 sequence in the vector from the fused *Sma* I site at 1007 bp to the *Hind* III site at 2804 bp (Ad 5 numbering). The DNA was religated to form a plasmid, pNL3C-HIII, with a *Hind* III site immediately 3' to the tripartite leader. The pNL3C-HIII DNA was cleaved with *Hind* III and *Bgl* II, removing a fragment running from the new *Hind* III site to the unique *Bgl* II site (Ad5 site 3328 bp). The *rev* gene was excised from M13mp19*rev*#209 as a *Hind* III-*Bam*H I fragment and cloned into pNL3C-HIII between the opened *Hind* III and *Bgl* II sites, creating plasmid pNL*rev*. The promoter-leader-*rev* gene expression cassette was excised from pNL*rev* by cleaving with *Afl* II (Ad5 site 3533 bp), end-filling with T4 polymerase, then cleaving with *Eco*R I to release the fragment. Plasmid pwtXho1-C was cleaved with *Xba* I (Ad5 site 1339), end filled, and cleaved with *Eco*R I. The *rev* cassette was cloned into the opened *Eco*R I and blunted *Xba* I sites of pNL*rev*, to form the virus reconstruction plasmid pwtE1*rev* (fig. 3.2) - suitable for insertion of the *rev* cassette in place of the E1A region of adenovirus.

3.3.2 Virus reconstruction by overlap recombination *in vivo*

The reconstruction procedure is summarised in fig. 3.4. In practice, the method is similar to that used for the construction of the RRE virus. It was not however necessary to ligate the DNA before transfection, as the *rev* vector was constructed to allow overlap recombination to occur between the 3' end of the linearised construct and the homologous 5' end of viral genomic DNA, if the first 1340 bp of the genome were removed by cleavage with *Xba* I.

Ad5 *dl*309 DNA (1 μg) was prepared for overlap recombination by cleavage at the unique *Xba* I site (1339 bp) to destroy its infectivity; pwtE1*rev* (5 μg) DNA was linearised at the *Eco*R I site. Both restriction products were extracted with phenol/chloroform, co-precipitated with ethanol and resuspended in sterile water for transfection. Recombinant virus and viral DNA were isolated by the same methods used for the RRE-virus (above and fig. 3.5).

3.4. Construction of the other recombinants

The methods used to construct the remaining viral recombinants described in table 3.1 followed those used to prepare *dl*309/*rev*. Each virus containing the RRE used *dl*309/RRE DNA as a starting point - the other viruses utilised *dl*309 DNA in their construction. Recombinant *dl*338/RRE was constructed by overlap recombination of *dl*309/RRE with pdl338XhoI-C, an analogue of pwtXhoI-C that has a *Hin*d III-*Bgl* II deletion of the E1B 55K gene (Ad5 sites 2804-3328 bp) equivalent to that found in Ad5 *dl*338 (Pilder *et al.*, 1986). Recombinant *dl*309/RRE/*rev* was prepared by overlap recombination of *dl*309/RRE with pwtE1*rev*. The construction of *dl*338/*rev* and *dl*338/RRE/*rev* required an additional subcloning stage - the *rev* expression cassette from pNL*rev* (see above) was cloned into pdl338XhoI-C, rather than pwtXhoI-C, making plasmid pdl338E1*rev*. Recombinant *dl*338/*rev* was constructed by overlap recombination of pdl338E1*rev* with *dl*309; similarly, *dl*338/RRE/*rev* was prepared by overlap recombination of pdl338E1*rev* with

*dl*309/RRE. In every case the viral DNA was cleaved at the *Xba* I site prior to recombination and the plasmid was linearised at the *Eco*R I site.

3.5. Restriction analysis of viral recombinants

As outlined in fig. 3.5, new recombinants were screened for the expected characteristics by restriction analysis. Restriction enzymes were selected that allowed discrimination between recombinant viruses carrying or lacking the *rev* expression cassette, the RRE, and the *dl*338 deletion of the 55K gene. The positions of the recognition sequences of three representative enzymes in wild type and recombinant viruses are denoted in fig. 3.8. The restriction patterns produced from viral DNAs by these enzymes are shown in figs. 3.9-3.11.

3.5.1. Restriction analysis of Ad5 recombinants with *Hin*d III

*Hin*d III was selected for analysis of the recombinants since its restriction profile allows *dl*309-derivatives to be readily distinguished from *dl*338-derivatives and permits *rev*-recombinants to be differentiated from viruses that do not carry the *rev* cassette. More specifically, deletion of the 55K gene in *dl*338 destroys the *Hin*d III site at 2804 bp and removes 524 bp 3' to this site (fig. 3.8). Therefore the *dl*309 *Hin*d III restriction fragments of 2804 bp (1-2804 bp) and 3437 bp (2805-6241 bp) are replaced by a single fragment of 5717 bp in *dl*338 and *dl*338/RRE. Insertion of the *rev* cassette increases the sequence length 5' to the *dl*309 *Xba* I site from 1339 bp to 1396 bp and introduces a new *Hin*d III site at 821 bp from the new 5' end of the genome. Thus the *dl*309 2804 bp *Hin*d III fragment is replaced by 820 bp and 2041 bp fragments in *dl*309/RRE*rev* and *dl*309/*rev*. Similarly, in *dl*338/RRE*rev* and *dl*338/*rev*, which have the 55K deletion and carry the *rev* cassette, the *dl*309 2804 bp and 3437 bp *Hin*d III fragments are replaced by 820 bp and 4954 bp fragments. Furthermore, insertion of the RRE sequence in the *Pvu* II site at 21720 bp increases the size of the 8010 bp (18319-26328 bp) fragment to 8246 bp

Fig. 3.8. Restriction sites in Ad5 and Ad5 recombinants

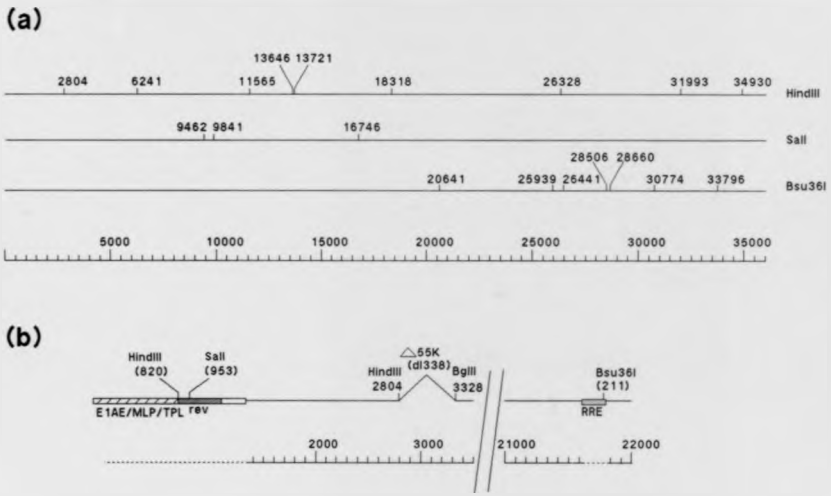


Fig. 3.8. Restriction sites in Ad5 and Ad5 recombinants. (a) *Hind* III, *Sal* I and *Bsu*36 I recognition sites in Ad5. **(b)** Enlarged map showing regions of the Ad genome affected by insertion of the *rev* cassette and RRE, and by the *dl*338 deletion of the 55K gene. All *rev*-recombinants lack Ad5 sequences 5' of the *Xba* I site at 1339 bp, but have an additional 1396 bp of *rev* cassette sequence 5' of this position. All *dl*338 derivatives lack 524 bp between the *Hind* III site at 2804 bp and the *Bgl* II site at 3328 bp (both sites are abolished by the deletion). All RRE-recombinants have an additional 236 bp of RRE sequence inserted in the *Pvu* II site at 21720 bp. Numbers in parentheses are measured from the start of the relevant insert (*rev* cassette or RRE). All other numbering refers to the standard Ad5 sequence. E1AE=E1A enhancer; MLP=major late promoter; TPL=tripartite leader.

Fig. 3.9 Restriction analysis of Ad5 recombinants with Hind III

Fig. 3.9. Restriction analysis of Ad5 recombinants with *Hind* III. 0.5 μg of DNA prepared from *dl*338/*rev* (lane A), *dl*309/*rev* (B), *dl*338/RRE*rev* (C), *dl*309/RRE*rev* (D), *dl*338/RRE (E), *dl*309/RRE (F) and *dl*309 (G) was digested with 5 units of *Hind* III for 1 hr at 37 °C. The restriction fragments were separated by agarose gel electrophoresis and visualised by ethidium bromide staining and ultraviolet illumination.

(although the small proportional difference in fragment size makes resolution of the wild type and recombinant fragments difficult).

As shown in fig. 3.9, digestion of viral DNA from each recombinant with *Hin*d III yielded fragments of 5665 bp, 5324 bp, 4597 bp, 2937 bp, 2081 bp, and 1008 bp identical in size to those produced by digestion of *dl*309 DNA. These fragments were derived from regions of the genome unaffected by the 55K deletion or insertion of the *rev* cassette or RRE, and their sizes corresponded to those predicted from the (sequence-derived) restriction map in fig. 3.8. The wild type 8010 bp fragment was detected in the *dl*309 (G), *dl*309/*rev* (B) and *dl*338/*rev* (A) digests whereas a slightly larger fragment, probably corresponding to the 8246 bp fragment predicted for the RRE-recombinants, was resolved in the *dl*309/RRE (F), *dl*338/RRE (E), *dl*309/RRE*rev* (D) and *dl*338/RRE*rev* (C) digests. The predicted 5717 bp fragment (diagnostic of the 55K deletion) was detected in the *dl*338/RRE digest, while the smaller 4954 bp fragment (indicating 55K deletion and *rev* cassette insertion) was detected in the *dl*338/RRE*rev* and *dl*338/*rev* digests. Conversely, the wild type 3437 bp fragment was detected in the digests of all the *dl*309-derivatives. A 2804 bp wild type fragment, indicative of an intact 55K gene and a normal left end of the genome, was detected in the *dl*309 and *dl*309/RRE digests. The predicted 2041 bp fragment specific for *rev* cassette insertion in a 55K-intact background was detected in the *dl*309/RRE*rev* and *dl*309/*rev* digests. Finally, the 820 bp fragment specific for *rev* cassette insertion was detected in the digests of all four *rev*-recombinants. These results therefore indicate that *dl*309, *dl*309/RRE, *dl*309/RRE*rev* and *dl*309/*rev* have an intact 55K gene, that *dl*338/RRE, *dl*338/RRE*rev* and *dl*338/*rev* have the 55K deletion, and that *dl*309/RRE*rev*, *dl*338/RRE*rev*, *dl*309/*rev* and *dl*338/*rev* carry the *rev* cassette. The data also suggest that *dl*309/RRE, *dl*338/RRE, *dl*309/RRE*rev* and *dl*338/RRE*rev* carry the RRE insert.

Fig. 3.10 Restriction analysis of Ad5 recombinants with Sal I

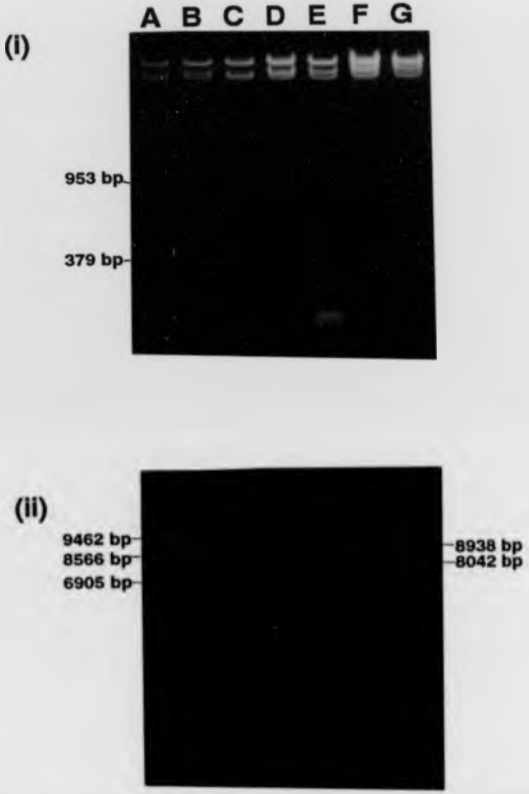


Fig. 3.10. Restriction analysis of Ad5 recombinants with *Sal* I. 0.5 μg of DNA prepared from *dl*338/*rev* (lane A), *dl*309/*rev* (B), *dl*338/RRE*rev* (C), *dl*309/RRE*rev* (D), *dl*338/RRE (E), *dl*309/RRE (F) and *dl*309 (G) was digested with 5 units of *Sal* I for 1 hr at 37 °C. The restriction fragments were separated and visualised as above. (I) Short electrophoresis separation time. (II) Long separation time.

3.5.2. Restriction analysis of Ad5 recombinants with *Sal* I

Sal I was selected to confirm the results obtained by analysis with *Hin*d III. Insertion of the *rev* cassette, as well as increasing the sequence length 5' of the *Xba* I site, introduces a new *Sal* I site 953 bp from the 5' end of the insert (fig. 3.8). This generates 953 bp and 8566 bp fragments in place of the wild type 9462 bp (1-9462 bp) fragment. The size of the 9462 bp fragment is also altered by the 55K deletion in *dl*338-derivatives to 8938 bp, or to 8042 bp in *dl*338-derivatives that carry the *rev* cassette. The size of the 19192 bp (16747-35938 bp) fragment is increased to 19428 bp in the RRE-recombinants, but this small relative difference is difficult to resolve in practice.

As shown in fig. 3.10, the 379 bp and 6905 bp fragments (predicted from the restriction map in fig. 3.8) that are unaffected by the changes made in the genome are found in the digests from all the viral recombinants. The 953 bp *rev* cassette-specific fragment is, as expected, found only in the *dl*309/RRE*rev*, *dl*338/RRE*rev*, *dl*309/*rev* and *dl*338/*rev* digests. The 8566 bp fragment specific to *rev*-recombinants carrying an intact 55K gene was resolved in the *dl*309/RRE*rev* and *dl*309/*rev* digests. The 8042 bp fragment specific to *rev*-recombinants with the 55K deletion was detected in the *dl*338/RRE*rev* and *dl*338/*rev* infections. The 8938 bp fragment specific to *dl*338-derivatives without the *rev* cassette was detected in the *dl*338/RRE digest. The 236 bp difference in the largest *Sal* I fragment between the RRE-recombinants and the other viruses was not resolved. These results therefore confirm the conclusions from the *Hin*d III analysis that the putative *rev*-recombinants carry the *rev* cassette, the *dl*309-derivatives carry an intact 55K gene, and the *dl*338-derivatives have the expected 55K deletion.

3.5.3. Restriction analysis of Ad5 recombinants with *Bsu*36 I.

*Bsu*36 I was selected for analysis of the recombinants as its recognition sequence is positioned asymmetrically within the RRE, generating junction fragments that

Fig. 3.11. Restriction analysis of Ad5 recombinants with Bsu36 I

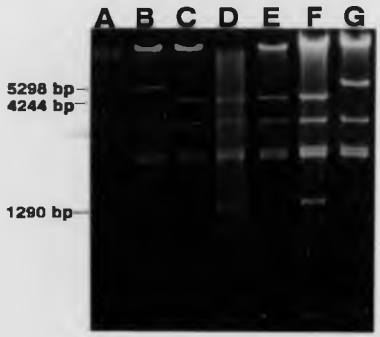


Fig. 3.11. Restriction analysis of Ad5 recombinants with *Bsu*36 I. 0.5 μg of DNA prepared from *dl*338/*rev* (lane A), *dl*309/*rev* (B), *dl*338/RRE*rev* (C), *dl*309/RRE*rev* (D), *dl*338/RRE (E), *dl*309/RRE (F) and *dl*309 (G) was digested with 5 units of *Bsu*36 I for 1 hr at 37 °C. The restriction fragments were separated and visualised as above.

confirm that the response element is inserted in the genome in the correct orientation. Specifically, 1290 bp and 4244 bp fragments are generated in place of the wild type 5298 bp (20643-25940 bp) fragment. As shown in fig. 3.11, 1290 bp and 4244 bp fragments were detected in the digests of *dl*309/RRE, *dl*338/RRE, *dl*309/RRE*rev* and *dl*338/RRE*rev* DNA, confirming the presence of the RRE in the correct orientation in these recombinants.

3.6. Summary

In this chapter I have described the construction and isolation of a series of six adenovirus/HIV recombinants designed for use in investigating the operation of the HIV-1 Rev/RRE regulatory system in an Ad5 context, and in comparing Rev function with that of the Ad5 E1B 55K gene. Restriction analysis data have been presented that confirm the intended presence or absence of a *rev* expression cassette, the RRE, and the Ad5 E1B 55K gene in each virus.

Chapter 4

Detection of functional Rev expression

4.1. Introduction

In this chapter I will describe the approach used to detect the expression of functional Rev protein from the *rev*-recombinant adenoviruses, and present the data obtained in the expression assays.

Several methods have been used to detect Rev expression from HIV and from *rev*-recombinant constructs. These can be divided into methods which directly detect the products of *rev* gene expression, and techniques which depend on the functional effects of the Rev protein. In the former category, Northern analysis and nuclease protection assays have been used to detect expression at the RNA level (see, for example, Emerman *et al.*, 1989; Chanda *et al.*, 1990). Assays based on antibody binding to Rev have similarly been used to demonstrate expression at the protein level (e.g. Goh *et al.*, 1987). Although useful, these methods obviously give no information on the biological activity of the Rev protein expressed. However, a number of functional assays for Rev have been developed. Malim *et al* (1989a) have shown that Rev activity can be detected by its *trans*-regulatory effect on *tat* expression from a reporter plasmid containing the unspliced *tat* exons and intron. In the absence of functional Rev, spliced *tat* RNA appears in the cytoplasm and full-length Tat protein is expressed; in the presence of active Rev, unspliced *tat* RNA enters the cytoplasm and a truncated Tat protein is translated. Other groups have developed assays based on the detection of Gag expression (e.g. by p24 antigen capture assay) from a Rev-responsive reporter plasmid (e.g. Chanda *et al.*, 1990). A third type of assay utilises Rev-responsive chloramphenicol acetyl transferase (CAT) reporter gene constructs (e.g. Rosen *et al.*, 1988). These constructs contain the RRE and a CAT gene in an intron. CAT RNA can only enter the cytoplasm and be expressed in the presence of functional Rev protein. These assays have the advantage that several comparatively simple techniques can be used to detect CAT expression. For its ability to detect functional activity rather than simply Rev

expression, and its simplicity and ease of application, such an assay was selected for use in this study.

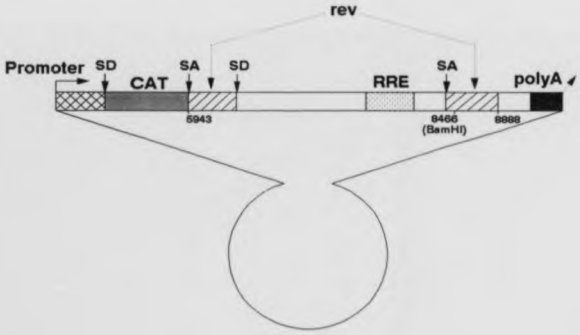
4.2. CAT reporter constructs

The specific constructs used were kindly supplied by Dr Tominori Kimura (now at Kansai Medical University, Osaka, Japan). The descriptions of the plasmids are derived from T. Kimura (personal communication). The Rev-reponsive reporter construct, pTK34, is illustrated in fig. 4.1. It contains a bacterial CAT gene and HIV-1 nucleotides 5943-8888 from molecular clone pNL4-3 (Adachi *et al.*, 1986) inserted downstream of the HIV-1 *gag* 5' splice donor site. Transcription from this construct is controlled by the SV40 early promoter and enhancer. Functional Rev expression is prevented by a frameshift mutation of the *rev* gene at the *Bam*H I site (pNL4-3 nt 8466). A derivative of pTK34, pTK53, was used as a negative control in my assays. This construct has a two nucleotide deletion (G 7805-G 7806) designed to disrupt the Rev-binding 'RNA bubble' of the RRE (fig. 1.3b) described by Heaphy *et al.* (1991). Since the CAT gene of pTK34 is flanked by a *gag* splice donor (SD) and a *rev* splice acceptor (SA) then, as noted above, CAT RNA can only enter the cytoplasm in the presence of Rev protein; in its absence, the CAT sequence will be spliced out as an intron. The deletion in pTK53 makes this plasmid unresponsive to Rev by disrupting its binding site in the RRE RNA.

4.3. CAT assay

The direct scintillation diffusion method was used to assay CAT expression (Neumann *et al.*, 1987; Eastman, 1987). Briefly, in this assay the cytoplasmic extract to be tested for CAT expression (and hence Rev activity) is mixed with [^{3}H]acetyl CoA and chloramphenicol, and placed under a layer of water-immiscible scintillant. CAT catalyses the [^{3}H]acetylation of chloramphenicol. The [^{3}H]acetyl chloramphenicol product is more soluble in the scintillant than in the aqueous extract, and diffuses into the nonaqueous phase where it can be detected by

Fig. 4.1. pTK34 CAT reporter construct



= SV40 promoter/enhancer

= CAT gene

= rev exon

= env exon

= RRE

= SV40 poly A

Fig. 4.1. pTK34 CAT reporter construct. This plasmid carries the bacterial chloramphenicol acetyltransferase gene and HIV-1 nucleotides 5943-8888 (from clone pNL4-3) inserted downstream of the HIV-1 *gag* 5' splice site. CAT expression is controlled by the SV40 early promoter and enhancer. The *rev* gene is disrupted by a frameshift mutation at the *Bam*H I site (8466 nt).

Table 4.1. CAT assay scintillation counts

	Incubation time / hours						
Sample:	0	0.5	1	1.5	2	16	20
309/RRErev + pTK34	97	939	1,822	3,189	4,220	22,825	26,088
338/RRErev + pTK34	131	3,548	8,175	15,451	21,343	72,521	79,034
309/rev + pTK34	159	1,662	4,676	8,540	11,261	54,532	60,002
338/rev + pTK34	97	2,112	5,058	8,911	12,170	53,035	58,616
Mock + pSV2CAT	80	1,053	2,841	4,722	6,656	32,457	35,866
Mock + pTK34	74	137	205	336	376	1,224	1,611
309/rev + Mock	57	85	193	142	216	421	506
309 + pTK34	28	171	205	398	455	2,140	2,323
309/rev + pTK53	68	285	546	1,059	1,298	6,593	7,781
Diffusion control	68	91	85	171	119	319	387

Data are cpm ^{3}H-acetyl chloramphenicol present at the times indicated

Table 4.1. CAT assay scintillation counts

	Incubation time / hours						
Sample:	0	0.5	1	1.5	2	16	20
309/RRErev + pTK34	97	939	1,822	3,189	4,220	22,825	25,088
338/RRErev + pTK34	131	3,548	8,175	15,451	21,343	72,521	79,034
309/rev + pTK34	159	1,662	4,876	8,540	11,281	54,532	60,002
338/rev + pTK34	97	2,112	5,058	8,911	12,170	53,035	58,616
Mock + pSV2CAT	80	1,053	2,841	4,722	6,856	32,457	35,865
Mock + pTK34	74	137	205	336	376	1,224	1,611
309/rev + Mock	57	85	193	142	216	421	506
309 + pTK34	28	171	205	398	455	2,140	2,323
309/rev + pTK53	68	265	548	1,059	1,298	6,593	7,781
Diffusion control	68	91	85	171	119	319	387

Data are cpm ^{3}H-acetyl chloramphenicol present at the times indicated

scintillation counting. This method has the advantage of allowing a single reaction to be monitored over an extended time course (unlike other assay techniques which require multiple reactions or repeated sampling of aliquots at each time point).

4.4. Demonstration of Rev expression

An experimental protocol was devised to assay Rev expression from each of the four recombinant adenoviruses that contain the *rev* construct:

HEK-293 cells were transfected with CAT reporter plasmid or control plasmid. 24 hr after transfection, the cells were infected with control or *rev*-recombinant adenovirus at a multiplicity of 10 pfu/cell. 24 hr after infection, the cells were harvested and a cytoplasmic extract was prepared. CAT activity was assayed for samples of extracts normalised for protein content. pSV2CAT, a construct that constitutively expresses the CAT gene (Gorman *et al.*, 1982), was used as a positive control for CAT expression. The background level of CAT expression from pTK34 in the absence of Rev was assessed by mock-infection of transfected cells. Conversely, mock-transfected cells were infected with *dl*309/*rev* to detect any endogenous acetyltransferase activity. To differentiate any general adenovirus-mediated effects from Rev-specific effects, pTK34-transfected cells were infected with *dl*309 rather than a *rev*-recombinant. As a further negative control, cells transfected with pTK53 (the Rev-unresponsive construct) were infected with *dl*309/*rev*. Finally, a reagent control containing no cytoplasmic extract was included to monitor the diffusion of ^{3}H-acetyl CoA into the nonaqueous scintillant phase.

4.5. Results

The results of the CAT assays for the four *rev*-recombinants and controls are presented in table 4.1, and in graphical form in fig. 4.2 (all controls except *dl*309 + pTK34 omitted in graph for clarity). The reactions were monitored over a 20 hr period. For the first two hours, readings were taken every 30 min; additional

readings were taken at 16 hr and 20 hr. The readings at the zero time point were taken immediately after mixing the reactants, but before incubation at 37 °C. All time point values refer to incubation time at 37 °C - the time required for scintillation counting (at room temperature) was ignored.

From the data, it is clear that the levels of CAT activity in the cytoplasmic extracts from cells infected with each of the four *rev*-recombinants and transfected with the pTK34 reporter construct are significantly higher than the levels in the negative control extracts. At 0.5 hr, the levels of tritiated product accumulating in the scintillant from the reactions using the *rev*-recombinant infected/pTK34 transfected extracts are already 6.9-26 times higher than from the *dl*309 infected/pTK34 transfected negative control extract. By 20 hr, the equivalent levels are 16.2-49.1 times higher for the *rev*-viruses than for the negative controls. The CAT activity of the *rev*-virus + pTK34 extracts is at all time points comparable to that measured in the positive control extract from cells transfected with pSV2CAT. The CAT activity of the extract from cells transfected with pTK34 alone (mock infection negative control) rises from 1.4 times (0.5 hr) to 4.2 times (20 hr) greater than the diffusion control (no cytoplasmic extract), indicating a low level of cytoplasmic CAT expression from the reporter construct in the absence of Rev. The slightly higher levels of CAT expression at later time points from the *dl*309 + pTK34 extract over those from the mock infected + pTK34 extract suggest that adenovirus alone may have a small effect on CAT expression from this construct. The CAT activity of the extract infected with *dl*309/*rev* and transfected with the negative control construct pTK53 (predicted to be unresponsive to Rev) is significantly lower than that of the equivalent *dl*309/*rev* + pTK34 extract (e.g. at 20 hr, the pTK53 reading is nearly eightfold lower than the equivalent pTK34 result). This indicates that specific binding of Rev to an intact RRE is required for high level cytoplasmic expression of CAT. However, the CAT activity in the *dl*309/*rev* + pTK53 extract is significantly

Fig. 4.2. CAT assay of cytoplasmic extracts (virus + pTK34)

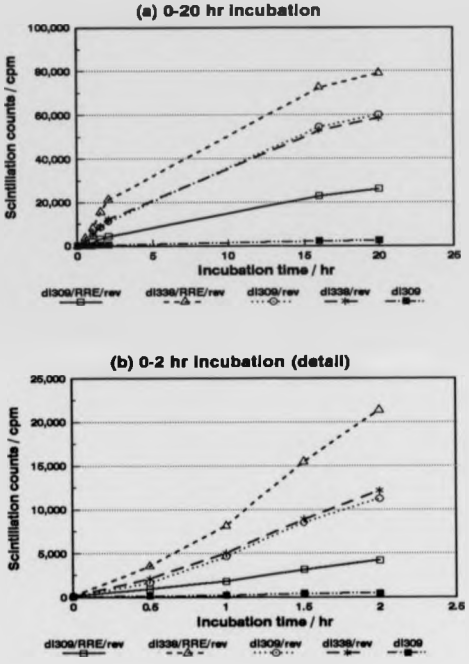


Fig. 4.2. CAT assay of cytoplasmic extracts (virus + pTK34). Cytoplasmic extracts were prepared from cells infected 24 hr previously with *dl*309/RRE*rev*, *dl*338/RRE*rev*, *dl*309/*rev*, *dl*338/*rev* or *dl*309, and transfected 48 hr previously with pTK34. The CAT activities of the extracts were assayed by the direct scintillation diffusion method over a 20 hr period. **(a)** Plot of scintillation counts *versus* incubation time for 0-20 hr incubation. **(b)** Plot for 0-2 hr incubation (enlarged scale).

greater than that in the *dl*309 + pTK34 extract (more than threefold greater at the 20 hr time point). If pTK53 were completely unresponsive to Rev, the levels of CAT from these two controls would be expected to be comparable. Therefore these results suggest that the pTK53 construct retains a residual Rev response. The scintillation counts for the extract from cells infected with *dl*309/*rev* alone (mock transfection) are only slightly higher than those for the diffusion control, showing that virus infected cells have little or no endogenous acetyltransferase activity that survives the heat inactivation step of the CAT assay.

4.6. Discussion

The level of CAT activity from cytoplasmic extracts of cells transfected with pTK34 and infected with *dl*309/RRE*rev*, *dl*338/RRE*rev*, *dl*309/*rev*, or *dl*338/*rev* demonstrates that each *rev*-recombinant expresses functional Rev protein. The 3-5 fold variation in the CAT activities of the different *rev*-virus extracts in this experiment suggests that the level of Rev expression from individual recombinants varies. However, the *dl*338/RRE*rev* > *dl*309/*rev* ≈ *dl*338/*rev* > *dl*309/RRE*rev* order of Rev expression levels suggested by this experiment was not reproducible: In a similar experiment an order of *dl*309/RRE*rev* > *dl*338/RRE*rev* > *dl*309/*rev* ≈ *dl*338/*rev* (within a 2.5 fold range of variation) was obtained. Differences in transfection efficiency of the reporter construct may account for small variations in the CAT activity of individual samples. In contrast, in repeated experiments the larger difference between the *rev*-recombinants as a group and the negative controls was fully reproducible. It was therefore concluded that each of the four *rev*-recombinants expresses functional Rev at similar levels.

Chapter 5

Analysis of the effects of RRE insertion on gene expression from recombinant adenoviruses

5.1. Introduction

In this and the following two chapters I present the results of a series of experiments in which gene expression from the recombinant viruses described in chapter 3 was analysed at the protein and RNA levels. It was originally intended that expression from all the recombinants and wild type virus would be examined in KB8a cells (Babiss *et al.*, 1983). These cells stably express the Ad5 E1A region and should therefore be suitable for complementing the E1A defect in the *rev*-recombinants (see chapter 3). However, initial experiments (7.2.1) indicated that complementation was poor. Subsequently, studies of E1A mRNA levels during infections of these cells suggested that the complement of E1A mRNAs expressed was incomplete, with little or no 13S mRNA detected (data not shown). Thus valid comparisons between the E1A-deficient *rev*-recombinants and the other E1A-intact viruses could not be made in this cell line. To circumvent this difficulty, different aspects of the problem were studied in different cell lines.

The major question to be addressed by this section of the thesis - does the Rev/RRE system affect the pattern of gene expression in adenovirus? - can conveniently be divided into three lines of enquiry:

(a) Does insertion of the RRE alone in L3 affect Ad5 gene expression?

(b) Does insertion of the Rev expression construct affect Ad5 gene expression?

(c) Can Rev and the RRE together regulate Ad5 gene expression, and is this effect similar to regulation by E1B 55K? - i.e., can Rev and the RRE complement deletion of the E1B 55K gene?

(a) and (b) are essentially controls for (c), designed to eliminate the possible *cis*-acting effects on gene expression introduced by RRE insertion, or nonspecific *trans*-acting effects of Rev expressed from a major late promoter on Ad5 genes. Since investigating (a) requires only the RRE single recombinant *dl*309/RRE and wild type *dl*309, no complementation for E1A deficiency is required, and the question

can conveniently be addressed in HeLa cells (this chapter). These cells have the advantage of growing rapidly in tissue culture, are relatively easy to manipulate, and are known to produce high yields of extractable nuclear and cytoplasmic viral RNA.

To answer (b), E1A complementation is required, since gene expression from the *rev* recombinants must be monitored. HEK-293 cells (a human embryonal kidney cell line described in Graham *et al.*, 1977) efficiently express high levels of Ad5 E1A and E1B proteins. This makes them suitable for testing the effect of Rev on adenovirus gene expression using viruses *dl*309, *dl*309/RRE, *dl*309/RRE*rev* and *dl*309/*rev*. However, since HEK-293 cells constitutively express E1B 55K, few useful comparisons can be made with the *dl*338 (55K-deficient) derivatives of these viruses, making these cells unsuitable for answering (c). To address this last question, it is necessary to employ KB8a cells, accepting the limitation of poor E1A complementation. This prevents direct comparisons being made between the *rev*-recombinants as a group and the other viruses, but allows comparisons between the individual *rev* viruses to be made. As discussed in chapter 7, poor complementation also slows the replication cycle of the virus, so that reasonable quantities of viral protein and RNA can only be harvested at relatively late times post-infection.

5.2. Analysis of viral protein expression in infected HeLa cells

The effect of inserting the RRE into region L3 of the adenovirus genome was assessed by comparing viral protein expression in HeLa cells infected with either recombinant *dl*309/RRE or wild type *dl*309 virus. The profile of proteins observed in parallel infections by these viruses is shown in fig. 5.1. Comparable levels of the major viral proteins were detected in the extracts from *dl*309- and *dl*309/RRE-infected cells at each time point. Thus insertion of the RRE does not lead to gross changes in the pattern of viral gene expression. This result is consistent with the

Fig. 5.1 Analysis of viral protein expression in infected HeLa cells

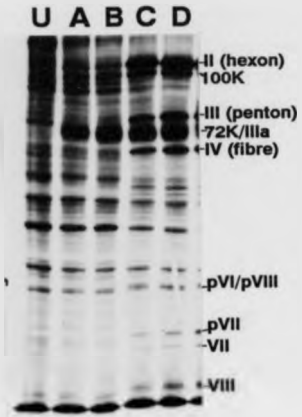


Fig. 5.1. Analysis of protein expression in infected HeLa cells. HeLa cell monolayers were infected at a multiplicity of 10 pfu/cell with *dl*309 (lanes A, C), *dl*309/RRE (B, D), or mock-infected (U). Proteins were labelled *in vivo* with ^{35}S-Met for 30 min immediately before harvesting by RIPA lysis at 12 hr (U, A, B) or 16 hr (C, D) time points post-infection. Samples of each cell extract were standardised to 50 000 counts per minute by scintillation counting of TCA-precipitated aliquots, analysed by SDS-PAGE, and visualised by fluorography.

ability of the recombinant to grow efficiently in tissue culture to a high titre. At a given time point, the level of each of the viral proteins consistently appears slightly higher in the *dl*309/RRE lane. This may be due to small differences in the multiplicities of infection. If RRE insertion in L3 does have any effect on viral gene expression, the most probable observable consequence in this experiment would be an alteration in the levels of proteins encoded by L3. The L3 23K protein is expressed at too low a level to be detectable by this method; L3 pVI levels are also relatively low and it is difficult to resolve the protein from L4 pVIII. However, L3 structural protein II - i.e., hexon, the major capsid protein - is readily detectable. Assessment of hexon levels by inspection at the 16 hr time point does not suggest a significant difference between *dl*309 and *dl*309/RRE hexon expression. - i.e., RRE insertion does not appear to exert a differential effect on gene expression in *cis* that is detectable at this level.

To evaluate the relative levels of viral proteins more rigourously, film band intensities at 16 hr were quantified by densitometry. The hexon (protein II) band density ratio measured between the *dl*309/RRE and *dl*309 lanes was 1.14. Slightly higher ratios were measured for the 100K product of L4 (1.25), the penton (protein III) product of L1 (1.45), and the fibre (protein IV) product of L5 (1.62). Within the limits of accuracy of quantitation, these results indicate that insertion of the RRE in L3 has little or no effect on the levels of L3 hexon expression, since the density ratio for hexon (1.14) differs only slightly from the average density ratio (1.34) calculated for the four late viral proteins quantitated.

5.3. Analysis of viral RNA expression in infected HeLa cells

These experiments were intended to complement and extend the analysis of viral protein expression discussed above. The RNase protection method employed enables specific viral RNA species to be detected with a high degree of sensitivity. Using this method, it was possible to select riboprobes that allowed discrimination between

Fig. 5.2. Map of regions that protect L3 riboprobes

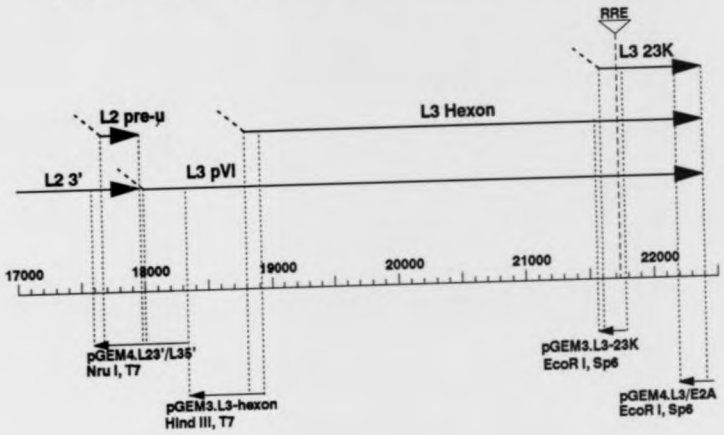


Fig. 5.2. Map of regions that protect L3 riboprobes. Viral and probe transcripts are denoted by solid lines. Directions of transcription are indicated by arrowheads. The positions of the L3 and L2 genes within the adenovirus genome are represented by the scale (in nucleotides from the left end). Polyadenylation of all the L3 RNAs occurs at 22399 nt; polyadenylation of L2 RNAs occurs at 17970 nt.

Fig. 5.3. Analysis of L3 RNA species from infected HeLa cells: 3' terminus

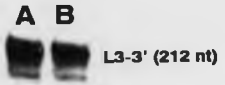


Fig. 5.3. Analysis of L3 RNA species from infected HeLa cells: 3' terminus. HeLa cells were infected with *dl*309 (lane A) or *dl*309/RRE (B) at a multiplicity of 10 pfu/cell. RNA was extracted from the cytoplasm at 16 hr post infection. 5 μg of RNA were hybridised with excess α^{32}P-UTP-labelled antisense riboprobe transcribed from the Ad5 22188-22444 nt *Kpn* I-*Dra* I fragment in pGEM4.L3/E2A (table 2.1). The fragments protected from digestion by single-strand-specific ribonucleases were separated on a denaturing polyacrylamide gel and visualised by autoradiography.

each of the individual L3 RNA species, whose expression might be expected to be affected by RRE insertion. The regions protected by the probes employed and the positions of the processing sites in L3 are illustrated in fig. 5.2. The characteristics of the plasmids encoding all of the probes used in chapters 5-7 are summarised in chapter 2 (table 2.1).

5.3.1. Analysis of L3 RNA species from infected HeLa cells: 3' terminus.

To assess the overall levels of L3 mRNAs, aliquots of cytoplasmic RNA extracted from *dl*309- and *dl*309/RRE-infected cells were analysed using a riboprobe designed to be protected by sequences common to the 3' termini of all the L3 RNA species - i.e., a 212 nt fragment of this probe is protected by pVI, hexon and 23K RNAs (fig. 5.2). As shown in fig. 5.3, the cytoplasmic levels of these mRNAs in the *dl*309 and *dl*309/RRE infections are similar. Thus insertion of the RRE appears to have no gross effect on the overall level of RNA expression from L3.

5.3.2. Analysis of specific L3 RNA species from infected HeLa cells.

Although overall L3 mRNA expression was not found to be affected by RRE insertion, it was possible that levels of individual L3 species were differentially altered in the recombinant. Further riboprobes were therefore selected that would give distinct protection profiles specific to each L3 mRNA (fig. 5.4). The riboprobe transcribed from pGEM3.L3-hexon (table 2.1) spans the hexon splice acceptor at 18805 nt. A 117 nt fragment of the probe is therefore protected by the 5' end of the hexon mRNA 3' exon, and a 600 nt fragment by the body of the pVI mRNA (fig. 5.2). Protection by the hexon RNA gives rise to two pairs of closely spaced bands with sizes of between 110 and 130 nt rather than the single expected band of 117 nt. This pattern is characteristic of hexon mRNA as detected by this probe, and may reflect heterogeneity in splice site usage or instability of the probe : mRNA hybrid. The cytoplasmic RNA levels appear similar for both the *dl*309 and *dl*309/RRE infections, indicating that the appearance of hexon mRNA in the cytoplasm is

Fig. 5.4 Analysis of specific L3 species from infected HeLa cells

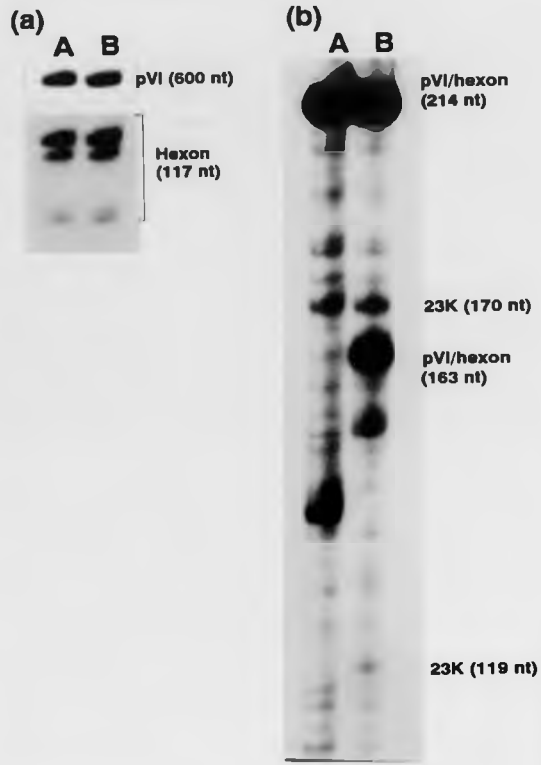


Fig. 5.4. Analysis of specific L3 RNA species from infected HeLa cells. HeLa cells were infected with *dl*309 (lane A) or *dl*309/RRE (B) at a multiplicity of 10 pfu/cell. RNA was extracted from the cytoplasm at 16 hr post infection. **(a) pVI and hexon probe:** 5 μg of RNA were hybridised with an antisense riboprobe transcribed from the Ad5 18318-18922 nt *Hin*d III-*Sma* I fragment in pGEM3.L3-hexon (table 2.1). All other details were as for fig. 5.3. **(b) 23K probe:** 5 μg of RNA were hybridised with an antisense riboprobe transcribed from the Ad5 21562-21775 nt *Bam*H I-*Bgl* II fragment in pGEM3.L3-23K (table 2.1). All other details were as for fig. 5.3.

unaffected by the presence of the RRE in L3. This result is consistent with the SDS-PAGE experiments described above. The cytoplasmic levels of pVI RNA (which will also contain the RRE in *dl*309/RRE) are also similar in both the *dl*309 and *dl*309/RRE infections.

The riboprobe transcribed from pGEM3.L3-23K (table 2.1) spans the 23K splice acceptor at 21605 nt. A 170 nt fragment of the probe is therefore protected by the 5' end of the 23K mRNA 3' exon, and a 214 nt fragment by both the hexon and pVI mRNAs (fig. 5.2). However, the pattern of protection is complicated for *dl*309/RRE, since the RRE was inserted at 21724 nt - i.e., within the sequences that protect the probe. By interrupting these sequences, RRE insertion could split the predicted 170 nt fragment into 119 and 51 nt fragments, and the 214 nt fragment into 163 and 51 nt fragments. Alternatively, the inserted RNA may form an unpaired loop at the insertion site, allowing uninterrupted protection of the probe by base pairing with the RNA flanking the insert. The most prominent species in this analysis (fig. 5.4) is the predicted pVI/hexon 214 nt fragment, which is present in both lanes at comparable intensity. A band at the position predicted for the specific 170 nt 23K fragment is also visible in each of the lanes. Thus, looping out of the RRE allowing protection of the normal length fragment does occur. However, the *dl*309/RRE lane also contains unique species that appear to correspond to the predicted 163 nt, 51 nt (not shown) and (at low levels) the 119 nt protected fragments formed by cleavage at the RRE insertion site. Since both protection and cleavage apparently occur at this site, it is difficult to make direct quantitative comparisons between the *dl*309 and *dl*309/RRE lanes. There is however no gross difference in the levels of 23K mRNA discernible in this assay.

5.3.3. Analysis of L4 and E3 species from infected HeLa cells

To complement the L3 analysis, control experiments were carried out in which the cytoplasmic levels of L4 and E3 species in *dl*309- and *dl*309/RRE-infected cells

Fig. 5.5 Analysis of L4 and E3 species from infected HeLa cells

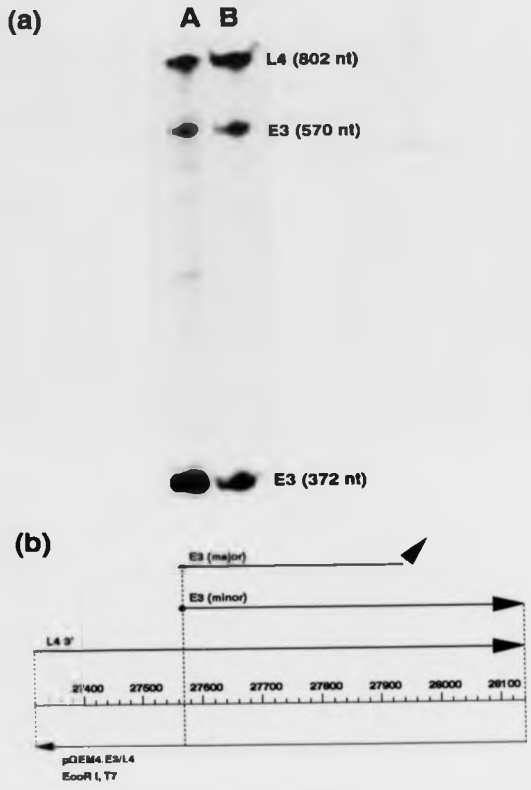


Fig. 5.5. Analysis of L4 and E3 species from infected HeLa cells. (a) HeLa cells were infected with *dl*309 (lane A) or *dl*309/RRE (B) at a multiplicity of 10 pfu/cell. RNA was extracted from the cytoplasm at 16 hr post infection. 5 μg of RNA were hybridised with an antisense riboprobe transcribed from the Ad5 27331-28137 nt *Eco*R I-*Bgl* II fragment in pGEM4.E3/L4 (table 2.1). All other details were as for fig. 5.3. (b) Map of viral mRNA regions that protect pGEM4.E3/L4 riboprobe. Directions of transcription are indicated by arrowheads. Viral and probe transcripts are denoted by solid lines.

were examined. No *cis*-acting effects of RRE insertion on these mRNAs were expected since none of the species detected carry the RRE. The pGEM4.E3/L4 riboprobe (table 2.1) produces a fragment of 802 nt due to protection by L4 mRNAs; 372 nt and 570 nt fragments are generated from protection by E3 mRNAs (fig. 5.5b). As shown in fig. 5.5a, all the predicted species were detected in the protection assay, and the cytoplasmic levels of each were similar in the *dl*309 and *dl*309/RRE infections.

5.4. Discussion

The results presented in this chapter show that RRE insertion has no significant effect on the pattern of viral gene expression detectable by the methods employed. *dl*309 and *dl*309/RRE infections displayed similar profiles of viral protein expression, with comparable relative levels of each of four late proteins quantified by densitometry. This result is consistent with the viability of the recombinant virus, which grows efficiently in tissue culture to a high titre. It was expected that any potential differential effect on gene expression resulting from RRE insertion would most probably be exerted on the L3 genes, since the position of the insertion site results in incorporation of the response element into L3 mRNAs (see chapter 3 and fig. 5.2). Of the three L3 proteins, only hexon (viral protein II) was readily detectable in the SDS-PAGE analysis; the relative level of this protein with respect to the other quantified late proteins was not significantly altered by RRE insertion. Cytoplasmic levels of L3-3', L3-hexon, L3-pVI, L4-3' and E3 mRNA species were similar in *dl*309 and *dl*309/RRE infections, again suggesting that RRE insertion has no significant effect on gene expression from L3 or the other regions probed as controls (L3-23K mRNA could not be accurately quantified but no gross difference was apparent).

Chapter 6

Analysis of the effects of a rev expression cassette on Ad5 gene expression

6.1. Introduction

In this chapter I present the results of experiments in which gene expression from Ad5 recombinants containing the *rev* expression construct was analysed at the protein and RNA levels, to determine whether Rev alone has any effect on this process. For reasons discussed in chapter 5, these experiments were conducted in HEK-293 cells and only included *dl*309/RRE*rev* and *dl*309/*rev* as the viruses to be tested, with *dl*309 and/or *dl*309/RRE as controls. In general, no useful purpose would be served by including *dl*338 and its recombinant derivatives since the 55K deletion in these viruses is complemented by the HEK-293 cells. However, a *dl*338/*rev* infection was employed as a control for an anomalous effect on Ad protein expression resulting from *dl*309/*rev* infection (see below).

6.2. Analysis of protein expression in infected HEK-293 cells

The effects of inserting the *rev* expression construct into the adenovirus genome were analysed by comparing viral protein expression in cells infected with *rev*-recombinants *dl*309/rev or *dl*309/RRE*rev* with expression in cells infected with *dl*309 or *dl*309/RRE. The profile of proteins observed in parallel infection by these viruses is shown in fig. 6.1a. A similar pattern of viral protein expression is visible in all the infections. Although there is some variation in band intensity between lanes, for example between *dl*309 and *dl*309/RRE, this appears to affect all the viral bands in a lane equally - i.e., it is a general effect probably due to slight differences in infection multiplicity, similar to that observed in HeLa cells (chapter 5). In terms of relative band intensities within an individual lane, the *dl*309, *dl*309/RRE and *dl*309/RRE*rev* lanes appear similar, suggesting that Rev expression (even in the presence of the RRE) has no effect on adenovirus gene expression detectable by this type of experiment in HEK-293 cells. However, the pattern of expression in the *dl*309/*rev* lane is slightly anomalous. One viral protein of approximately 35K is reduced in or absent from the *dl*309/*rev* lane but present in the other lanes. This band cannot be identified conclusively from these data, but it may correspond to the

Fig. 6.1 Analysis of viral protein expression in infected HEK-293 cells

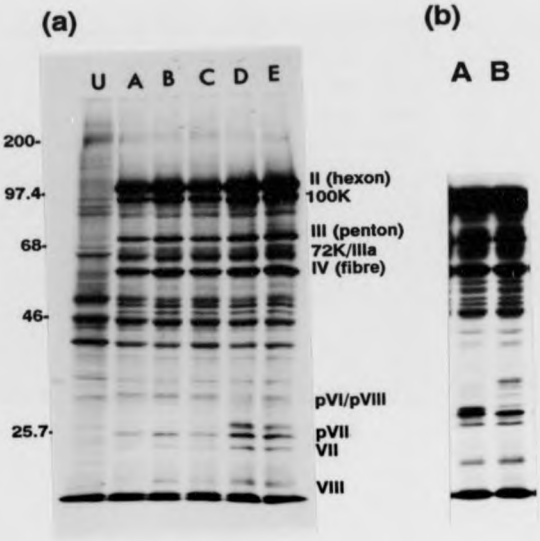


Fig. 6.1. Analysis of protein expression in infected HEK-293 cells. (a) HEK-293 cell monolayers were infected at a multiplicity of 10 pfu/cell with *dl*309 (lane A), *dl*309/RRE (B), *dl*309/RRE*rev* (C), *dl*309/*rev* (D), *dl*309/RRE + *dl*309/*rev* (each at a multiplicity of 10 pfu/cell, lane E) or mock-infected (lane U). Proteins were labelled *in vivo* with ^{35}S-Met and harvested at 16 hr post-infection; preliminary experiments (not shown) had indicated that sufficient viral protein for analysis was synthesised at this time point. Proteins were separated by SDS-PAGE. (b) Conditions and experimental procedures were as for (a), above. Cells were infected with *dl*309/*rev* (lane A) or *dl*338/*rev* (B).

L3 pVI or L4 pVIII protein (precursors of hexon-associated virion components). Furthermore, an additional band of approximately 28K is present in the *dl*309/*rev* lane, but appears to be absent from the other lanes. It is possible that this additional band represents a truncated form of the missing protein. If this is the case, the truncation does not affect the viability of the virus, since it grows efficiently to a high titre in tissue culture with normal cytopathic effect. An alternative explanation is that the additional protein represents the processed form of one of the virion protein precursors (pVI or pVIII). However, there is no evidence in the protein profiles of accelerated processing of other viral protein precursors normally subject to processing by the 23K protease, and it is difficult to envisage a mechanism by which such processing could be facilitated selectively. The *dl*309/*rev* infection is also characterised by unusually high levels of hexon, 100K, penton, fibre and other late proteins. It is unclear if this effect is connected with the anomalous pattern of protein expression in the 25-40K region of the gel. As noted above, some variation in band intensity between the lanes can probably be explained by minor differences in infection multiplicity; this factor may also explain the particularly intense bands in the *dl*309/*rev* lane. Alternatively, these features may reflect an accelerated replication cycle for this virus.

Since neither the effect on late protein expression nor the anomalous pattern of expression in the 25-40K region is evident in the *dl*309/RRE*rev* lane, they are unlikely to be Rev-specific effects. However, it is a formal possibility that the presence of the RRE target sequence in the double recombinant *dl*309/RRE*rev* prevents Rev interacting with adenovirus gene expression to cause these effects, possibly by high-affinity binding of all available Rev molecules. When this possibility was tested by using a *dl*309/RRE + *dl*309/*rev* coinfection to provide Rev in *trans* (fig. 6.1a) the same anomalous pattern of protein expression was observed, indicating that this effect is independent of the presence of RRE transcripts. Moreover, infection of HEK-293 cells with the *dl*338/*rev* single recombinant results

Table 6.1a Quantitation of viral proteins from HEK-293 cell infections (arbitrary units)

	Virus				
Protein:	*dl*309	*dl*309/RRE	*dl*309/RRE*rev*	*dl*309*/rev*	*dl*309/RRE + *dl*309*/rev*
IV (fibre)	19.1	26.9	18.4	30.4	33.9
III (penton)	9.2	14.2	9.5	15.9	19.8
100K	15.5	20.8	17.9	25.7	31.0
II (hexon)	45.1	52.7	37.0	63.5	64.2

Table 6.1b Ratios of band densities in recombinant infections to equivalent band densities in *dl*309 infection

	Virus			
Protein:	*dl*309/RRE	*dl*309/RRE*rev*	*dl*309*/rev*	*dl*309/RRE + *dl*309*/rev*
IV (fibre)	1.41	0.96	1.60	1.78
III (penton)	1.54	1.03	1.73	2.15
100K	1.34	1.15	1.66	2.00
II (hexon)	1.17	0.82	1.41	1.42

in a similar pattern of expression to that obtained with *dl*309, *dl*309/RRE and *dl*309/RRE*rev* at a 16 hr time point (fig. 6.1b). Hence the anomalous pattern of expression is unique to *dl*309/*rev*. The cause of this effect is unclear; uncharacterised changes in the viral genome that were not detected by restriction analysis may have occurred spontaneously during the construction of this virus.

To test more rigourously the observation that the presence of the *rev* construct does not appear to affect the overall pattern of late viral gene expression (with the exception of the probably *rev*-independent effect on proteins in the 25-40K M_r size range from *dl*309/*rev*) the film band densities of four major late proteins were quantified by densitometry (table 6.1a). To facilitate comparisons between lanes, the density ratios between each of the bands in the *dl*309/RRE, *dl*309/RRE*rev*, *dl*309/*rev* and *dl*309/RRE + *dl*309/*rev* lanes and the equivalent bands in the *dl*309 lane were calculated (table 6.1b). From these results, it is evident that the small variation in band intensity between lanes occurs by about the same degree for proteins IV, III, and 100K - i.e., the relative proportions of these proteins in each cell extract are approximately equal. The variation in hexon levels is consistently slightly smaller between the *dl*309 and recombinant lanes than are the variations in the levels of the other proteins. This suggests that relative levels of hexon are lower in the recombinant infections than in the *dl*309 infection (which has the lowest absolute levels of each of the viral proteins). However, this small effect, even if significant, is nonspecific since it occurs in the *dl*309/RRE, *dl*309/*rev*, and *dl*309/RRE*rev* lanes.

6.3. Analysis of RNA expression in infected HEK-293 cells

As in chapter 5, the analysis of gene expression at the protein level was extended by experiments at the RNA level, using the RNase protection assay described previously. Although Rev synthesis in *dl*309/*rev* or *dl*338/*rev* could potentially affect gene expression from any region of the adenovirus genome, possible effects

Fig. 6.2. Analysis of L3 RNA species from infected HEK-293 cells: 3' terminus

Fig. 6.2. Analysis of L3 RNA species from infected HEK-293 cells: 3' terminus

Fig. 6.2. Analysis of L3 RNA species from infected HEK-293 cells: 3' terminus.

HEK-293 cells were infected with *dl*309/RRE (lanes A, D), *dl*309/RRE*rev* (B, E) or *dl*309/*rev* (C, F) at a multiplicity of 10 pfu/cell. RNA was extracted from the cytoplasm (A-C) and nuclei (D-F) at 16 hr post infection. 5 μg of RNA were hybridised with excess α^{32}P-UTP-labelled antisense riboprobe transcribed from the Ad5 22188-22444 nt *Kpn* I-*Dra* I fragment in pGEM4.L3/E2A (table 2.1). The fragments protected from digestion by single-strand-specific ribonucleases were separated on a denaturing polyacrylamide gel and visualised by autoradiography.

on L3 were of particular interest since the RRE was inserted in this region in the *dl*309/RRE*rev* and *dl*338/RRE*rev* double recombinants. Therefore, riboprobes were selected that were protected by several RNA species synthesised from L3. Control probes that were protected by other late viral species were also included in the analysis.

6.3.1. Analysis of L3 RNA species from infected HEK-293 cells: 3' terminus.

To assess the overall levels of L3 mRNAs, aliquots of cytoplasmic RNA from *dl*309/RRE-, *dl*309/RRE*rev*- and *dl*309/*rev*-infected cells were analysed using the riboprobe synthesised from pGEM4.L3/E2A (table 2.1), from which all L3 RNA species protect a 212 nt fragment (fig. 5.2) (*dl*309/RRE was utilised as a control virus since insertion of the RRE was shown not to affect viral gene expression in chapter 5). As shown in fig. 6.2a, the cytoplasmic level of the L3 mRNA varies between the infections. Levels in the *dl*309/RRE and *dl*309/RRE*rev* infections are comparable, suggesting that insertion of the *rev* construct does not affect overall levels of L3 expression. However, the level in the *dl*309/*rev* infection is markedly higher than in the *dl*309/RRE and *dl*309/RRE*rev* infections. This might suggest that insertion of the *rev* construct does have an effect on L3 gene expression that is for some reason limited to *dl*309/*rev*. However, an analysis of the corresponding nuclear RNA samples (fig. 6.2b) shows a similar lane to lane variation in levels of the 212 nt species, implying a transcriptional, rather than post-transcriptional, effect that is unlikely to be due to Rev function. This effect may be the cause of the *dl*309/*rev*-specific aberration of protein synthesis noted in section 6.2, wherein possible explanations were discussed.

6.3.2. Analysis of specific L3 RNA species from infected HEK-293 cells.

Although overall expression of L3 mRNA was not specifically affected by the *rev* construct, it was possible that the levels of individual L3 species were differentially

affected in the *rev*-recombinants. Further assays were therefore performed using riboprobes transcribed from pGEM3.L3-hexon and pGEM3.L3-23K (table 2.1).

As noted in 5.3.2, several fragments of the pGEM3.L3-hexon probe in the 110-130 nt size range are protected by sequences at the 5' end of the hexon mRNA, and all 600 nt are protected by pVI mRNA (fig. 5.2). As shown in fig. 6.3a, approximately equal levels of the pVI species and similar levels of the group of hexon species are present in the *dl*309/RRE and *dl*309/RRE*rev* lanes, but the levels in the *dl*309/*rev* lane are somewhat elevated. The equality of band intensity in the *dl*309/RRE and *dl*309/RRE*rev* lanes suggests that *rev* expression does not affect expression of the pVI and hexon mRNAs from E1B 55K-intact virus in an RRE-dependent manner. The reason for the increase in the levels of hexon and pVI species in the *dl*309/*rev* infection (consistent with the increase in overall L3 levels noted in 6.3.1) is unclear. Since this effect does not occur with *dl*309/RRE*rev* it is unlikely to be Rev-specific. However, as noted above, *dl*309/*rev* has an anomalous phenotype at the protein level which may be a result of uncharacterised genome alterations. Such changes could alter the structure of hexon and pVI mRNAs.

A 170 nt fragment of the pGEM3.L3-23K probe (table 2.1) is protected by the 5' end of the 23K mRNA 3' exon and a second 214 nt fragment is protected by both the pVI and hexon mRNAs (fig. 5.2). When these mRNAs carry the RRE insert, the 170 nt species is partially further digested to 119 and 51 nt fragments and the 214 nt species is partially further digested to 163 and 51 nt fragments, preventing direct comparisons between the RRE recombinants and the other viruses. However,

Fig. 6.3. Analysis of specific L3 RNA species from infected HEK-293 cells.

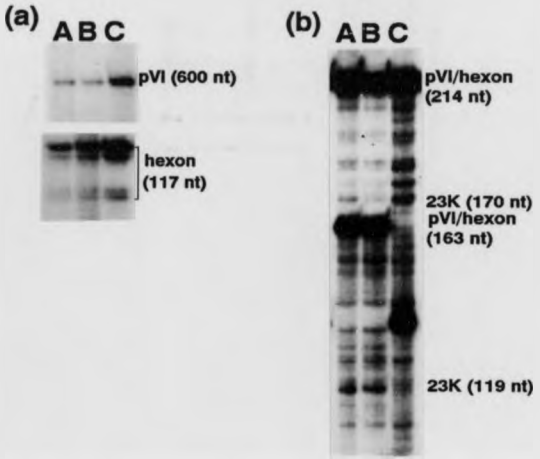


Fig. 6.3. Analysis of specific L3 RNA species from infected HEK-293 cells. HEK-293 cells were infected with *dl*309/RRE (lane A), *dl*309/RRE*rev* (B), or *dl*309/*rev* (C) at a multiplicity of 10 pfu/cell. RNA was extracted from the cytoplasm at 16 hr post infection.

(a) pVI and hexon probe: 5 μg of RNA were hybridised with an antisense riboprobe transcribed from the Ad5 18318-18922 nt *Hin*d III-*Sma* I fragment in pGEM3.L3-hexon (table 2.1). All other details were as for fig. 6.2.

(b) 23K probe: 5 μg of RNA were hybridised with an antisense riboprobe transcribed from the Ad5 21562-21775 nt *Bam*H I-*Bgl* II fragment in pGEM3.L3-23K (table 2.1). All other details were as for fig. 6.2.

comparisons can be made between the *dl*309/RRE and *dl*309/RRE*rev* infections if it is assumed that the proportion of each species that is further digested in the protection assay is constant. For *dl*309/RRE and *dl*309/RRE*rev*, the cytoplasmic levels of 23K mRNA are most easily assessed from the levels of the 119 nt fragment. As shown in fig. 6.3b, these levels are similar in both infections, indicating that Rev expression does not affect 23K mRNA expression in this system.

6.3.3. Analysis of L1, L2 and L4 RNA species from infected HEK-293 cells

To complement the L3 analysis, experiments were carried out in which cytoplasmic levels of L1, L2 and L4 RNA species were examined. The probe transcribed from pGEM3.L1-3'/L2-5' (table 2.1) spans the L1 polyadenylation site at 14118 nt and the L2 splice acceptor at 14154 nt. A 393 nt fragment is therefore protected by L1 sequences, and a 137 nt fragment by L2 sequences (fig. 6.4c(i)). An additional L1 fragment, of approximately 350 nt, is also detected using this probe. This fragment is believed to result from cleavage of the L1 : probe hybrid at an RNase-sensitive site. An oligo A : oligo T region is predicted approximately 40 nt upstream from the L1 polyadenylation site and RNase sensitivity at a similar sequence has been noted within E4 (Dix and Leppard, 1993). As shown in fig. 6.4a, bands corresponding to both L1 species and the L2 species were detected in the assay. Similar levels of each of the three fragments were detected in assays of *dl*309/RRE, *dl*309/RRE*rev* and *dl*309/*rev* cytoplasmic RNA.

A single 312 nt fragment of the probe transcribed from pGEM3.L4-5' (table 2.1) is protected by sequences in L4 extending from the L4 100K splice acceptor at 24057 nt to the *Pst* I recognition sequence at 24369 nt (fig. 6.4c(ii)). Similar levels of this species were detected in the *dl*309/RRE, *dl*309/RRE*rev* and *dl*309/*rev* lanes; (fig. 6.4b). The equivalence of the cytoplasmic levels of L1, L2 and L4 species in

Fig. 6.4. Analysis of L1, L2 and L4 RNA species from infected HEK-293 cells

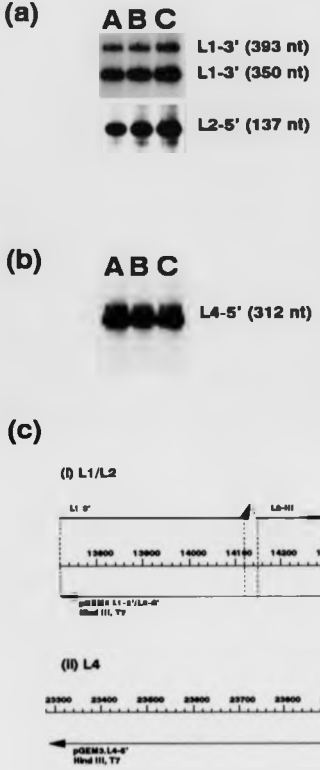


Fig. 6.4. Analysis of L1, L2 and L4 RNA species from infected HEK-293 cells.

HEK-293 cells infected with *dl*309/RRE (lane A), *dl*309/RRE*rev* (B), or *dl*309/*rev* (C) at a multiplicity of 10 pfu/cell. RNA was extracted from the cytoplasm at 16 hr post infection.

(a) L1/L2 probe: 5 μg of RNA were hybridised with an antisense riboprobe transcribed from the Ad5 3721-14291 nt *Hin*d III-*Kpn* I fragment in pGEM3.L1-3'/L2-5' (table 2.1). All other details were as for fig. 6.2.

(b) L4 probe: 5 μg of RNA were hybridised with an antisense riboprobe transcribed from the Ad5 23286-24369 nt *Pst* I-*Pst* I fragment in pGEM3.L4-5'. All other details were as for fig. 6.2.

(c) Map of viral mRNA regions that protect L1/L2 and L4 riboprobes. (i) L1/L2. (ii) L4. Directions of transcription are indicated by arrowheads. Viral and probe transcripts are denoted by solid lines.

the *dl*309/RRE, *dl*309/RRE*rev* and *dl*309/*rev* infections indicates that the presence of Rev does not affect the expression of these mRNAs.

6.4. Discussion

The results presented in this chapter show that expression of Rev from the *rev* cassette inserted in place of E1A has no specific effect on adenovirus gene expression detectable by the methods employed. Although an effect on the expression of proteins in the 25-40K region was noted in infections with recombinant *dl*309*rev*, a similar effect was not observed in *dl*309/RRE*rev* or *dl*338/*rev* infections, indicating that the effect was not Rev-specific. Similarly, experiments conducted at the RNA level did not provide any evidence of Rev-specific effects. The cytoplasmic levels of individual viral RNA species were generally similar in the *dl*309/RRE and *dl*309/RRE*rev* infections. Some differences were noted between the levels of species expressed from *dl*309/*rev* and the other viruses, but these may have been due to other uncharacterised changes introduced into the genome during virus construction (as suggested by the protein experiments), or to small differences in infection multiplicities (such differences would be expected to be minor since the stocks of recombinant viruses were titred repeatedly in parallel). Overall, there was no evidence for any Rev-specific effect.

Chapter 7

Analysis of the effects of a *rev* expression cassette and the RRE on Ad5 gene expression

7.1. Introduction

In this chapter, I describe experiments in which the effects on adenovirus gene expression of inserting the *rev* expression construct into the Ad5 genome together with the RRE were evaluated. These experiments were conducted in KB8a cells which, for reasons discussed in chapter 5, allow direct comparisons to be made between recombinants with an intact or deleted E1B 55K gene (derivatives of *dl*309 and *dl*338 respectively). Thus these experiments address the question of whether the Rev/RRE system can substitute for the E1B 55K system in the facilitation of Ad late gene expression.

7.2. Analysis of protein expression in infected KB8a cells

7.2.1. Protein analysis at 12, 16.5 and 19 hr post-infection

The objective of this experiment was to examine the patterns of protein expression in cells infected with a range of the recombinant viruses described in chapter 3, and to assess the suitability of this cell line for comparisons of expression between the recombinants. As shown in fig. 7.1, little expression of the indicated viral proteins is apparent at 12 hr post-infection. However, by 16.5 hr, several viral proteins are detectable in the *dl*309 and *dl*309/RRE lanes. Lower intensity bands are visible in the *dl*338 and *dl*338/RRE lanes (as would be expected from the 55K$^-$ genotypes of these viruses). However, little or no expression of viral proteins is detectable in the *dl*309/RRE*rev* and *dl*338/RRE*rev* lanes, even at the 19 hr time point. This indicates that the complementation of E1A function required by the *rev* recombinants is at best incomplete in this cell line.

7.2.2. Protein analysis at 49 hr and 72 hr post-infection

The range of data obtained in the previous analysis was extended in a second experiment. This experiment had two objectives: first, to determine if the apparent lack of E1A complementation observed at 19 hr and earlier in infection extended to later time points (i.e., to establish whether the viral replication cycle of the E1A$^-$

Fig. 7.1. Analysis of protein expression in infected KB8a cells at 12, 16.5 and 19 hr post-infection

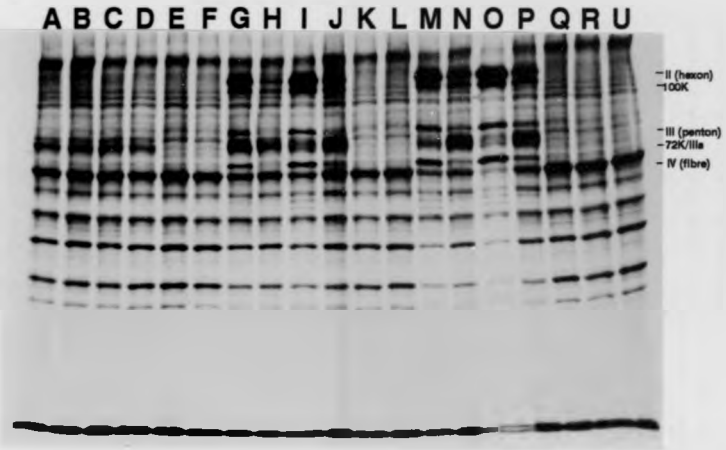


Fig. 7.1. Analysis of protein expression in infected KB8a cells at 12, 16.5 and 19 hr post-infection. KB8a cell monolayers were infected at a multiplicity of 10 pfu/cell with *dl*309 (lanes A, G, M), *dl*338 (B, H, N) *dl*309/RRE (C, I, O), *dl*338/RRE (D, J, P), *dl*309/RRE*rev* (E, K, Q), *dl*338/RRE*rev* (F, L, R), or mock-infected (U). Proteins were labelled *in vivo* with ^{35}S-Met and harvested at 12 hr (A-F, U), 16.5 hr (G-L) and 19 hr (M-R) post-infection, then analysed by SDS-PAGE and visualised by fluorography.

viruses was delayed, or completely halted); second, assuming sufficient complementation, to compare the patterns of gene expression from 55K$^-$ viruses in the presence and absence of the complete Rev/RRE system, using the incomplete system in the *dl*338/*rev* recombinant as the negative control (since Rev alone was not found to affect adenovirus gene expression in chapter 6). In this experiment *dl*312 (an E1A$^-$ virus derived from *dl*309, Jones and Shenk, 1979) was included as a negative control in place of *dl*309; E1A$^+$ viruses were not used as their replication cycles would be much further advanced at these time points, preventing direct comparisons. As shown in fig. 7.2a, viral protein expression is clearly visible from all the E1A-deficient *dl*309 derivatives at both time points tested. This demonstrates that the replication cycles of these viruses are delayed, rather than arrested, and indicates that the KB8a cells provide a degree of complementation sufficient to allow experiments to be conducted in these cells at late time points. Levels of viral protein expression that are reduced compared to their *dl*309 counterparts are visible from the *dl*338 derivatives at both 49 and 72 hr. The atypical pattern of protein expression from *dl*309/*rev* noted at the 16 hr time point of the HEK-293 cell experiment (i.e., the absence of a 35K band and presence of an additional 28K band, as described and discussed in 6.2) is again apparent.

If the Rev/RRE system is able to compensate for deletion of the E1B 55K gene, then levels of hexon protein expression might be expected to be elevated in the *dl*338/RRE*rev* infection with respect to levels in the *dl*338/*rev* infection, since hexon RNA from the former but not the latter should carry the RRE. However, comparison of the hexon band intensities 72 hr post-infection from the *dl*338/*rev* (lane J) and *dl*338/RRE*rev* (lane H) infections shows no significant elevation in the double recombinant (fig. 7.2a); in both lanes only very low levels of hexon expression are apparent. A gel artefact prevents assessment of hexon band intensities in the 48 hr lanes of fig. 7.2a, but comparison of the equivalent bands in the 48 hr lanes (A, B) of a repeat experiment (fig. 7.2b) also shows no significant

Fig. 7.2. Analysis of protein expression in infected KB8a cells at 49 hr and 72 hr post-infection

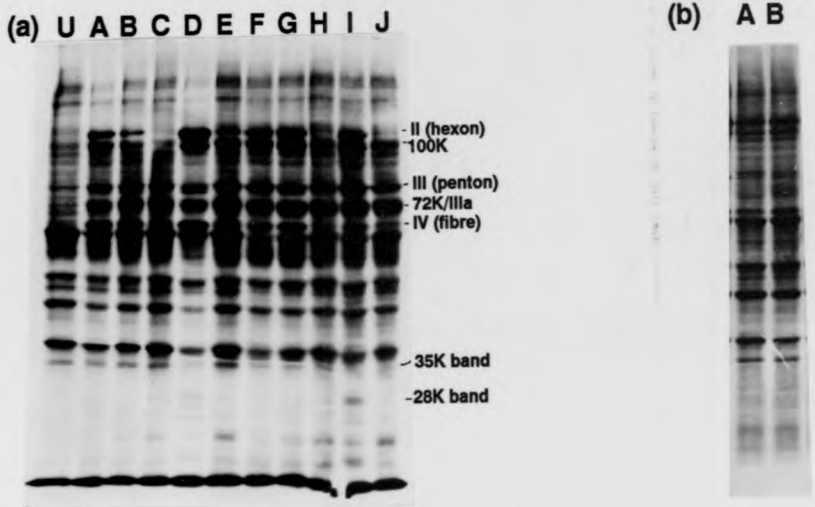


Fig. 7.2. Analysis of protein expression in infected KB8a cells at 49 hr and 72 hr post-infection. **(a)** KB8a cell monolayers were infected at a multiplicity of 10 pfu/cell with *dl*312 (lanes A, F), *dl*309/RRE*rev* (B, G), *dl*338/RRE*rev* (C, H), *dl*309/*rev* (D, I), *dl*338/*rev* (E, J), or mock-infected (U). Proteins were labelled *in vivo* with ^{35}S-Met, harvested at 49 hr (A-E, U) and 72 hr (F-J) post-infection, analysed by SDS-PAGE, and visualised by fluorography. **(b)** Infections with *dl*338/RRE*rev* (lane A), and *dl*338/*rev* (B) harvested at 49 hr. Conditions and experimental details as for (a), above.

elevation of hexon levels in *dl*338/RRE*rev*. Thus no efficient complementation of the 55K defect by the Rev/RRE system as assessed by hexon expression at this level of analysis is evident (low band intensities and high background levels prevent accurate quantitation by densitometry).

7.3. Analysis of RNA expression in infected KB8a cells

As in chapters 6 and 7, the analysis of gene expression was extended to the RNA level using RNase protection assays. To obtain a complete picture of RNA expression, cytoplasmic and nuclear RNAs were analysed using riboprobes to all regions of the major late transcription unit, and to E1A, E3 and IVa2 species (i.e., all the probes listed in table 2.1). Since no gross changes in the pattern of protein expression in the presence of the complete Rev/RRE system were detected at the protein level, it was expected that any effect at the RNA level was likely to be small and limited in scope. Moreover, such an effect is more likely to be detectable when comparing expression from *dl*338/RRE*rev* and *dl*338/*rev* than when examining expression from *dl*309/RRE*rev* and *dl*309/*rev*. The *rev*-recombinants derived from *dl*309 already express viral genes relatively efficiently; any small additional effect resulting from the action of the Rev/RRE system is likely to be swamped by normal levels of expression. Since Rev appears to be inactive in the absence of the RRE (chapter 6), *dl*338/*rev* serves as a negative control for any effects evident in the *dl*338/RRE*rev* infection. The interpretation of results in the following sections concentrates on comparisons between gene expression in these two recombinant viruses, although other viruses were included in the analyses.

7.3.1. Analysis of L3 RNA species from infected KB8a cells: 3' terminus

Overall levels of L3 mRNAs were assessed using the pGEM4.L3/E2A probe (table 2.1) from which a 212 nt fragment is protected by all L3 RNAs (fig. 5.2). As shown in fig. 7.3, the *dl*309/RRE*rev* and *dl*309/*rev* infections give rise to similar levels of L3 mRNAs in the cytoplasm; levels in the *dl*312 infection are slightly

Fig. 7.3. Analysis of L3 RNA species from infected KB8a cells: 3' terminus

Fig. 7.3. Analysis of L3 RNA species from infected KB8a cells: 3' terminus. RNA was extracted at 44 hr post infection from the cytoplasm (lanes A-E) and nucleus (F-J) of KB8a cells infected at a multiplicity of 10 pfu/cell with *dl*312 (A, F), *dl*309/RRE*rev* (B, G), *dl*338/RRE*rev* (C, H), *dl*309/*rev* (D, I), or *dl*338/*rev* (E, J). 5 μg of RNA were hybridised with excess antisense α^{32}P-UTP-labelled riboprobe transcribed from the Ad5 22188-22444 *Kpn* I-*Dra* I fragment in pGEM4.L3/E2A (table 2.1). The fragments protected from digestion by single-strand-specific ribonucleases were separated on a denaturing polyacrylamide gel and visualised by autoradiography.

lower. Predictably, the cytoplasmic levels of L3 mRNAs in the E1B-55K-deficient *dl*338/RRE*rev* and *dl*338/*rev* infections are very markedly reduced in comparison with the levels in the 55K-intact infections. Moreover, cytoplasmic levels in the *dl*338/RRE*rev* infection are significantly higher than in the *dl*338/*rev* infection, suggesting a response to the presence of the complete Rev/RRE system. However, the levels of L3 mRNAs in the cytoplasm appear to be mirrored by their levels in nuclei from each of the five infections - i.e., the cytoplasmic : nuclear (C/N) ratio for total L3 mRNA is apparently similar for each infection. This suggests that the Rev/RRE system has no gross effect on L3 mRNA expression, since functional Rev activity would be expected to increase the C/N ratio by promoting the cytoplasmic accumulation of responsive RRE-containing mRNAs However, the intensity of adjacent bands prevents accurate comparisons and quantitation of the *dl*338/RRE*rev* and *dl*338/*rev* bands, and could mask a small effect on overall L3 levels.

7.3.2. Analysis of specific L3 RNA species from infected KB8a cells

Although the Rev/RRE system did not appear to affect significantly the cytoplasmic level of total L3 mRNA, it was possible that smaller or compensating effects were being exerted on one or more of the individual L3 species. Probes specific for single L3 species were therefore employed. A 322 nt fragment of the pGEM4.L2-3'/L3-5' probe (table 2.1) is protected by pVI mRNA (fig. 5.2). As expected, the nuclear and cytoplasmic levels of pVI RNA (fig. 7.4a) were lower in cells infected with the E1B-55K-deficient viruses (*dl*338 derivatives) than in cells infected with the E1B-intact viruses. The cytoplasmic : nuclear ratio for pVI mRNA appeared slightly higher for the *dl*338/RRE*rev* infection than for the *dl*338/*rev* infection. However, the proximity of strong bands in the adjacent *dl*309/RRE*rev* and *dl*309/*rev* lanes again made precise quantitation difficult and prevented clear reproduction of the original autoradiograph in fig. 7.4a (it should be noted that all comparisons of relative band intensities were made directly from the

Fig. 7.4. Analysis of specific L3 and L2 RNA species from infected KB8a cells

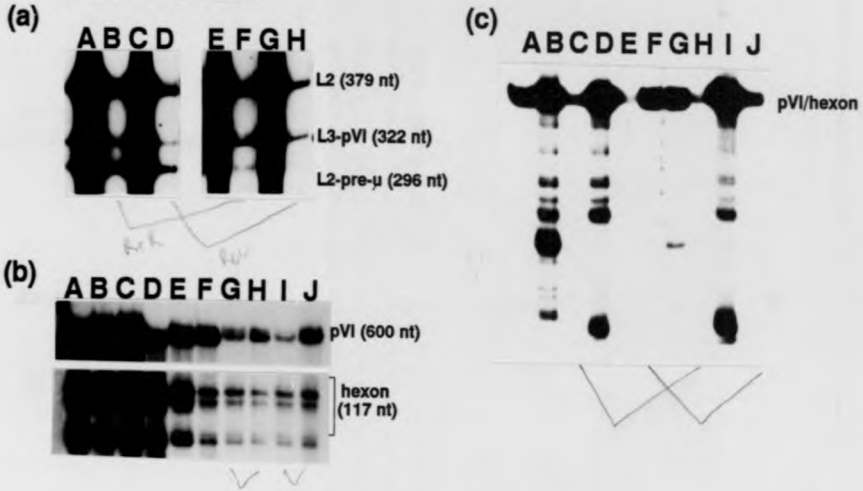


Fig. 7.4. Analysis of specific L3 and L2 RNA species from infected KB8a cells.
(a) pVI probe: RNA was extracted at 44 hr post infection from the cytoplasm (lanes A-D) and nucleus (E-H) of KB8a cells infected with *dl*309/RRE*rev* (A, E), *dl*338/RRE*rev* (B, F), *dl*309/*rev* (C, G), or *dl*338/*rev* (D, H). 5 μg of RNA were hybridised with excess antisense riboprobe transcribed from the Ad5 17068-18323 *Kpn* I-*Hin*d III fragment in pGEM4.L2-3'/L3-5' (table 2.1).

(b) pVI and hexon probe: RNA was extracted at 44 hr post infection from the cytoplasm (lanes A, C, E, G, I) and nucleus (B, D, F, H, J) of KB8a cells infected with *dl*309/RRE*rev* (A, B), *dl*309/*rev* (C, D), *dl*312 (E, F), *dl*338/RRE*rev* (G, H), or *dl*338/*rev* (I, J). 5 μg of RNA were hybridised with excess antisense riboprobe transcribed from the Ad5 18318-18922 *Hin*d III-*Sma* I fragment in pGEM3.L3-hexon (table 2.1)

(c) 23K probe: RNA was extracted at 44 hr post infection from the cytoplasm (lanes A-E) and nucleus (F-J) of KB8a cells infected with *dl*312 (A, F), *dl*309/RRE*rev* (B, G), *dl*338/RRE*rev* (C, H), *dl*309/*rev* (D, I), or *dl*338/*rev* (E, J). 5 μg of RNA were hybridised with excess antisense riboprobe transcribed from the Ad5 21562-21775 nt *Bam*H I-*Bgl* II fragment in pGEM3.L3-23K (table 2.1).

All other details were as for fig. 7.3.

autoradiographs). There was no apparent difference in cytoplasmic : nuclear ratio between pVI RNA levels in the *dl*309/RRE*rev* and *dl*309/*rev* infections.

pVI mRNA protects a 600 nt fragment and hexon mRNA several fragments of 110-130 nt derived from the pGEM3.L3-hexon probe (table 2.1; fig. 5.2). The cytoplasmic and nuclear levels of the pVI mRNA (600 nt) are, as in the cases of the other L3 species assayed for (above), highest in the infections with *dl*309 derivatives, slightly lower in the *dl*312 infection, and greatly reduced in the infections with *dl*338 derivatives (fig. 7.4b). Comparing *dl*338/RRE*rev* and *dl*338/*rev* assays, higher levels of pVI mRNA are apparent in the cytoplasmic assay for *dl*338/RRE*rev* than for *dl*338/*rev*, whereas the corresponding nuclear levels for these two infections are comparable. This suggests that the complete Rev/RRE system in this virus is capable of partially complementing the 55K$^-$ defect in cytoplasmic accumulation of pVI mRNA. Comparison of the levels of the hexon-specific species detected in the *dl*338/RRE*rev* and *dl*338/*rev* infections shows only a small increase in the cytoplasmic : nuclear ratio for the *dl*338/RRE*rev* infection over that observed for the *dl*338/*rev* infection.

All 214 nt of the probe transcribed from pGEM3.L3-23K (table 2.1) are protected by pVI and hexon mRNAs, while a 170 nt fragment is specifically protected by the 23K RNA. The difficulties in RNA quantitation caused by the partial cleavage of these fragments into 163 nt, 119 nt and 51 nt fragments due to the RRE have been described above (section 5.3.2). However, assuming that the extent of cleavage is similar for cytoplasmic and nuclear RNA assays of a particular infection, the C/N ratio will not be affected and these ratios can be compared between infections. Comparing *dl*338/RRE*rev* and *dl*338/*rev* infections (fig. 7.4c), the cytoplasmic : nuclear ratios for the 214 nt hexon and pVI species are greater for the *dl*338/RRE*rev* double recombinant. The levels of the 170 nt (L3-23K) fragment

from *dl*338/RRE*rev* and *dl*338/*rev*, and the levels of the 119 nt and 51 nt cleavage products from *dl*338/RRE*rev* were too low for quantitation.

Taken together, these data suggest that a small RRE-dependent, Rev-mediated enhancement of accumulation of both L3 pVI and L3 hexon mRNAs is detected, when the Ad5 E1B 55K function is inactivated and late mRNA accumulation is very inefficient. The individual results obtained with the L3-hexon probe suggested that a greater effect was exerted on pVI than on hexon mRNA.

7.3.3. Analysis of L1 and L2 RNA species from infected KB8a cells

As controls for the L3 analysis, the relative levels of L1 and L2 species in cells infected with the *rev*-recombinants were assessed (processed L1 and L2 mRNAs do not carry the RRE, so their expression would be expected to be independent of Rev/RRE-specific effects). The pGEM4.L2-3'/L3-5' probe (table 2.1) is protected by the 3' end of L2 mRNAs, (fig. 5.2) yielding protected fragments of 296 nt (pre-μ) and 379 nt (all other L2 mRNAs). As shown in fig. 7.4a, the general pattern noted previously - increased levels of viral RNA in the *dl*309-derived recombinant infections, compared with lower levels in the *dl*338-derived recombinant infections - was again obtained. As discussed in 7.3.2, the very large variation in band intensities made precise quantitation of *dl*338/RRE*rev* and *dl*338/*rev* bands difficult and prevented clear reproduction of the autoradiograph in fig. 7.4a (nuclear pre-μ levels are not visible for *dl*338/*rev* in the figure). However, within these limitations there was no evidence of an effect by the Rev/RRE system on L2 mRNA expression. The cytoplasmic : nuclear ratios of each L2 species as assessed from the original autoradiographs were comparable in the *dl*338/RRE*rev* and *dl*338/*rev* infections; similarly, uniform ratios were found in the *dl*309/RRE*rev* and *dl*309/*rev* infections.

Fig. 7.5. Analysis of L1 and L2 RNA species from infected KB8a cells.

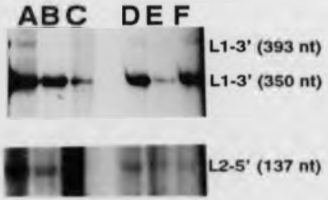


Fig. 7.5. Analysis of L1 and L2 RNA species from infected KB8a cells. RNA was extracted at 44 hr post infection from the cytoplasm (lanes A, C, E) and nucleus (B, D, F) of KB8a cells infected with *dl*312 (A, B) *dl*338/RRE*rev* (C, D), or *dl*338/*rev* (E, F). 5 μg of RNA were hybridised with excess antisense riboprobe transcribed from the Ad5 13721-14291 *Hin*d III-*Kpn* I fragment in pGEM3.L1-3'/L2-5' (table 2.1). All other details were as for fig. 7.3.

The pGEM3.L1-3'/L2-5' probe is protected by the common 3' ends of L1 mRNAs, and by the 5' end of the L2-III 3' exon (6.3.3 and fig. 6.4c(i)). Levels of the L2-III species are too low for accurate quantitation in the infections with *dl*338-derived recombinants, but no obvious difference in levels is apparent between the *dl*338/RRE*rev* and *dl*338/*rev* infections (fig. 7.5). Comparison of L1 levels detected as the ≈350 nt fragment shows that the cytoplasmic : nuclear RNA ratios for the *dl*338/RRE*rev* and *dl*338/*rev* infections are indistinguishable. This indicates that the Rev/RRE system has no effect on the cytoplasmic expression of L1 mRNA in this context, a result consistent with that obtained for L2 mRNA.

7.3.4 Analysis of L4 and E3 species from infected KB8a cells

As further controls for the L3 analysis, the expression of mRNA from late region L4 was examined (processed L4 mRNAs, like L1/L2 mRNAs, do not carry the RRE, and were therefore not expected to be affected by Rev/RRE-specific effects). A 312 nt fragment of the probe transcribed from pGEM3.L4-5' is protected by sequences at the 5' end of the L4-100K 3' exon (fig. 6.4c(ii) & 6.3.3). If the levels of L4-100K RNA are compared for the *dl*338/RRE*rev* and *dl*338/*rev* infections (fig. 7.6a), a dramatic difference in the cytoplasmic levels is apparent. The nuclear levels, however, are very similar. The cytoplasmic : nuclear ratio of RNA levels is therefore much greater for the *dl*338/RRE*rev* infection, suggesting that the Rev/RRE system is facilitating the cytoplasmic expression of L4-100K mRNA. This result was surprising, as the position of the RRE in the Ad5 genome precludes its incorporation into processed L4-100K mRNA, although it will be present in the unspliced primary transcript.

To extend the L4-5' analysis to other L4 species and mRNAs from the adjacent E3 transcription unit the pGEM4.E3/L4 probe (table 2.1) was employed (5.3.3 and fig. 5.5b). As shown in fig. 7.6b, the cytoplasmic : nuclear ratio for total L4 mRNA (802 nt fragment) appears slightly higher in the *dl*338/RRE*rev* infection than in the

Fig. 7.6. Analysis of L4 and E3 species from infected KB8a cells

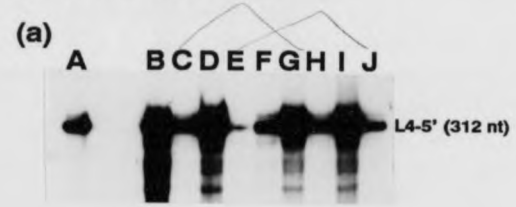


Fig. 7.6. Analysis of L4 and E3 species from infected KB8a cells.

(a) L4-5' probe: RNA was extracted at 44 hr post infection from the cytoplasm (lanes A-E) and nucleus (F-J) of KB8a cells infected with *dl*312 (A, F), *dl*309/RRE*rev* (B, G), *dl*338/RRE*rev* (C, H), *dl*309/*rev* (D, I), or *dl*338/*rev* (E, J). 5 μg of RNA were hybridised with excess antisense riboprobe transcribed from the Ad5 23286-24369 *Pst* I-*Pst* I fragment in pGEM3.L4-5'. (table 2.1).

(b) L4/E3 probe: RNA was extracted at 44 hr post infection from the cytoplasm (lanes A, C, E) and nucleus (B, D, F) of KB8a cells infected with *dl*312 (A, B), *dl*338/RRE*rev* (C, D), or *dl*338/*rev* (E, F). 5 μg of RNA were hybridised with excess antisense riboprobe transcribed from the Ad5 27331-28137 *Eco*R I-*Bgl* II fragment in pGEM4.E3/L4 (table 2.1).

All other details were as for fig. 7.3.

*dl*338/*rev* infection, confirming that the Rev/RRE system is able to facilitate the cytoplasmic expression of L4 mRNA species. However, the magnitude of the change in ratio is much smaller than was detected with the L4-100K probe. This may be due to an unequal effect of Rev/RRE on the cytoplasmic expression of different L4 species; a weak effect or no effect on L4 33K and/or pVIII expression could mask a relatively strong effect on 100K when the probe used hybridises with sequences common to all three.

The two E3 species detected by this probe appear to show different patterns of response to the Rev/RRE system (fig. 7.6b). The cytoplasmic : nuclear RNA level ratio of the E3 species that protects a 372 nt probe fragment (fig. 5.5b) is similar for the *dl*338/RRE*rev* and *dl*338/*rev* infections. However, the ratio for the E3 species that protects a 570 nt fragment is markedly greater for the *dl*338/RRE*rev* infection, indicating Rev/RRE mediated accumulation. This difference in response to Rev between E3 mRNAs may be a consequence of the different patterns of expression of the two species. The 372 nt fragment defines a 'conventional' E3 RNA expressed at early times in infection from the E3 promoter, whereas the 570 nt fragment defines an mRNA that is expressed predominantly at late times from the major late transcription unit (Tollefson *et al.*, 1992), and could therefore be expected to behave similarly to L4.

7.3.5 Analysis of L5 RNA species from infected KB8a cells

To determine if the Rev/RRE-mediated effect on cytoplasmic RNA expression detected for L4 extended to L5, a probe transcribed from pGEM2.L5 (table 2.1) was employed. Fragments of 800 nt and 1091 nt are protected by the 3' ends of the two major L5 mRNA species (fig. 7.7b). As shown in fig. 7.7a, the cytoplasmic : nuclear RNA level ratios of the 800 nt species are not increased for the *dl*338/RRE*rev* infection with respect to the *dl*338/*rev* infection (levels of the 1091 nt

Fig. 7.7. Analysis of L5 RNA species from infected KB8a cells

(a)

(b)

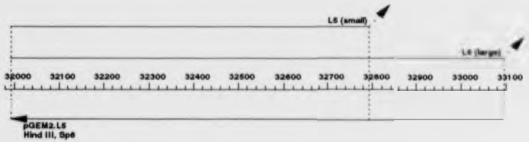


Fig. 7.7. Analysis of L5 RNA species from infected KB8a cells. (a) RNA was extracted at 44 hr post infection from the cytoplasm (lanes A-D) and nucleus (E-H) of KB8a cells infected with *dl*309/RRE*rev* (A, E), *dl*338/RRE*rev* (B, F), *dl*309/*rev* (C, G), or *dl*338/*rev* (D, H). 5 μg of RNA were hybridised with excess antisense riboprobe transcribed from the Ad5 31933-33096 *Hind* III-*Sma* I fragment in pGEM2.L5 (table 2.1). All other details were as for fig. 7.3. (b) Map of viral mRNA regions that protect L5 riboprobe.

species are too low to quantify). This result indicates that the Rev/RRE-mediated facilitation of cytoplasmic RNA expression does not extend to L5.

7.3.6 Analysis of IVa2 RNA species from infected KB8a cells

Although L1, L2, L4 and L5 mRNAs do not carry the RRE in their mature forms, the major late primary transcript from which they are derived does contain the response element. Therefore the L1/2/4/5 control experiments described above are not entirely independent of the RRE; indeed the L4 analysis provided evidence of a significant Rev response. As a completely independent control, IVa2 RNA expression was examined. IVa2 is transcribed from its own promoter (fig. 1.4) active during the late phase of infection. Since the RRE will not be incorporated into IVa2 RNA at any stage of gene expression, no Rev/RRE-mediated effect was expected. pGEM3.IVa2 probe fragments of 87 nt and 241 nt are protected by the first and second exons of the IVa2 mRNA respectively (fig. 7.8b). As shown in fig. 7.8a, the cytoplasmic : nuclear RNA level ratios of the 241 nt species in the *dl*338/RRE*rev* and *dl*338/*rev* infections were similar, indicating that the cytoplasmic expression of IVa2 mRNA is unaffected by the presence of Rev/RRE system acting in *trans*.

7.4. Discussion

The results presented in this chapter demonstrate that the HIV-1 Rev/RRE regulatory system has a functional effect on adenovirus gene expression in the test system employed. In the context of the *dl*338/RRE*rev* recombinant virus, the Rev/RRE system was shown to complement partially a defect in cytoplasmic late RNA expression caused by deletion of the E1B 55K gene. Unexpectedly, the strongest effect observed was on expression of L4-100K mRNA. The cytoplasmic : nuclear ratio of RNA levels was markedly increased for this species when expressed from *dl*338/RRE*rev*, in comparison with expression from *dl*338/*rev*. Similar but smaller effects were also noted on the expression of L3 pVI and hexon mRNAs, and

Fig. 7.8. Analysis of IVa2 RNA species from infected KB8a cells

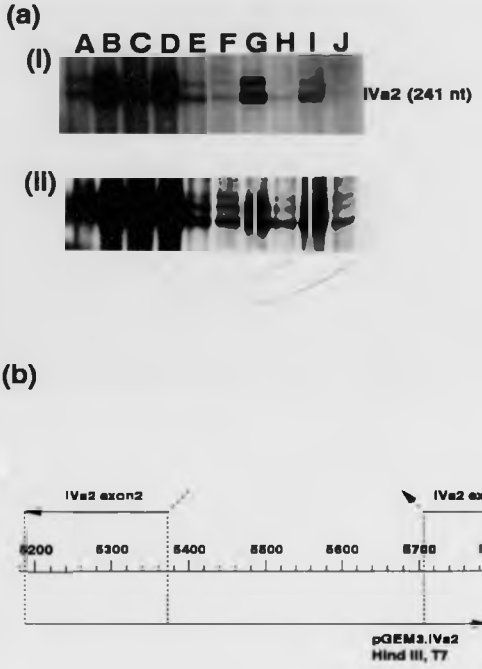


Fig. 7.8. Analysis of IVa2 RNA species from infected KB8a cells. (a) RNA was extracted at 44 hr post infection from the cytoplasm (lanes A-E) and nucleus (F-J) of KB8a cells infected with *dl*312 (A, F), *dl*309/RRE*rev* (B, G), *dl*338/RRE*rev* (C, H), *dl*309/*rev* (D, I), or *dl*338/*rev* (E, J). 5 μg of RNA were hybridised with excess antisense riboprobe transcribed from the Ad5 5186-5792 *Bst*E II-*Xho* I fragment in pGEM3.IVa2 (table 2.1). All other details were as for fig. 7.3. (i) Short autoradiograph exposure. (ii) Long exposure. (b) Map of viral mRNA regions that protect IVa2 riboprobe.

on an E3 species expressed from the major late promoter. The effect on L3 expression can be explained by the presence in L3 mRNAs of the RRE, which can bind the Rev protein in the nucleus to facilitate cytoplasmic mRNA expression. However, in the cases of L4 and E3, the RRE (although initially present) is in a region of the primary transcript removed from these mRNAs during processing. To explain the effects on L4 and E3, it is necessary to propose that Rev/RRE can determine the fate of an RNA before excision of the RRE (excision of sequences 5' to the L4-100K sequences was demonstrated by the normal size of the protected fragment from the pGEM3.L4-5' probe). For example, L4-100K RNA may be committed to a specific nuclear export pathway by Rev/RRE action before processing is complete. This could be mediated by the type of RRE-nucleated ordered assembly of Rev molecules along an RNA chain described by Heaphy *et al.* (1991) and Wingfield *et al.* (1991). Such a process could package the RNA for interaction with a specific nuclear export pathway (or block splice site usage and prevent further splicing of L4-100K RNA to smaller L4 species).

Despite the clear difference in the cytoplasmic : nuclear level ratio of L4-100K mRNA between the *dl*338/RRE*rev* and *dl*338/*rev* infections, no significant difference in 100K protein levels was observed in the SDS-PAGE analysis of corresponding infections at a similar time point (fig. 7.2). However, as discussed in chapter 1, Rev may be a multifunctional protein required at several different stages of the gene expression pathway. It is possible that a Rev/RRE-mediated commitment to nuclear export of RNA is linked to Rev-dependent cytoplasmic events that lead to efficient translation. D'Agostino *et al.* (1992) have suggested that Rev-dependent HIV mRNAs follow a transport pathway distinct from that of cellular mRNAs and interact comparatively slowly with the translation machinery ; Rev is envisaged as acting to improve the efficiency of this interaction. Indeed, Arrigo and Chen (1991) and D'Agostino *et al.* (1992) have shown that in the absence of Rev or the RRE, Rev-dependent mRNAs are not efficiently loaded onto

polysomes but are instead associated with smaller 40-80 S complexes and/or sequestered in a cytoplasmic subcompartment. Rev/RRE could therefore act by releasing these mRNAs for interaction with the translational machinery. If excision of the RRE occurs after Rev/RRE-mediated commitment to nuclear export, transport may occur, but the mRNA will be unable to bind Rev in the cytoplasm. Without Rev binding, the transported mRNA may not be efficiently translated. This interpretation of the data implies that one of several potential Rev activities has been isolated in the *dl*338/RRE*rev* system.

It is unclear why the most pronounced effect of Rev/RRE was on L4-100K mRNA. It was originally expected that any responses to Rev/RRE would be confined to L3, since the position of RRE insertion places the response element in the pVI, hexon, and 23K exons. However, the magnitudes of the effects observed on L3 were smaller than the L4-100K response, whereas no effects were observed for L1, L2 or L5. The lack of any detectable effect on L1 and L2 may be a result of the pattern of adenovirus RNA processing. Since cleavage, polyadenylation and splicing, or at least commitment to these events, occur co-transcriptionally, transcripts that give rise to L1 and L2 mRNAs may be cleaved and polyadenylated 5' to L3 and therefore not contain the RRE at any stage of expression. The differences between the comparatively large effect observed for L4-100K and the smaller effect found for overall L4 expression may be explained by differences in the dependence of individual RNA species on mechanisms that facilitate nuclear export. Leppard (1993) has shown that major late L1, L2 and L3 RNA species that contain intronic sequences or unused splice sites are more dependent on E1B 55K for efficient cytoplasmic expression than related, fully processed RNAs. Similarly, L4-100K mRNA is more likely than other L4 species to be dependent on Rev/RRE for efficient cytoplasmic expression in the absence of E1B 55K.

Chapter 8

General discussion

8.1. Summary and interpretation of results

In the preceding chapters I have described experiments designed to test the function of the HIV-1 Rev/RRE system in an Ad5 context. More specifically, these experiments were directed towards examining the potential functional analogy between the HIV-1 Rev and Ad5 E1B 55K proteins. Six Ad5 recombinant viruses containing an HIV-1 *rev* cassette, the HIV-1 RRE, and the Ad5 E1B 55K gene in various combinations were constructed and isolated to provide the reagents needed to address this question. The four recombinants containing a *rev* cassette were shown to express functional Rev protein by a CAT reporter gene assay. Very recently, it has been shown in this laboratory that infection with the *dl*309/*rev* recombinant induces Env protein synthesis from a simian immunodeficiency virus *env* expression vector transfected into HEK-293 cells (C. Caravokyri, personal communication). This result independently confirms the functional activity of the Rev protein expressed from this system. This experiment also illustrates the potential usefulness of the *rev*-recombinants as general purpose Rev expression systems for experiments in E1A-complementing cells.

Using the six recombinant viruses, the effects of the Rev/RRE system on adenovirus gene expression were analysed at the protein and RNA levels. These analyses focused on three specific areas of investigation. Firstly, the effects of inserting the RRE into region L3 of the adenovirus genome were assessed. The patterns of viral late protein expression in cells infected with *dl*309 and *dl*309/RRE were similar and levels of individual viral late proteins did not differ significantly between parallel infections with the two viruses. Furthermore, the cytoplasmic levels of L3, L4 and E3 RNA species in these infections were comparable. From these results, it was concluded that insertion of the RRE does not detectably affect adenovirus gene expression. Secondly, the effects of inserting a *rev* expression cassette into early region E1A of the genome were evaluated. No significant Rev-specific effects on viral gene expression were detected by either protein or RNA analysis. Recombinant

virus *dl*309/*rev* displayed an atypical pattern of protein expression characterised by the synthesis of a novel protein and the absence of another. These changes were clearly not Rev-specific, since both *dl*309/RRE*rev* and *dl*338/*rev* infections produced the normal array of viral proteins. These results could reflect uncharacterised changes introduced into the genome during virus construction (any such changes have no effect on virus growth), or might indicate selective virion protein processing by the viral protease (an explanation for which it is difficult to envisage a mechanism). Thus it was concluded that insertion of the *rev* construct and the expression of active Rev protein in the absence of the RRE do not significantly perturb adenovirus gene expression (except, obviously, E1A expression).

Having established that neither Rev nor the RRE alone had any effect on Ad5 gene expression, the effects of the complete Rev/RRE system on adenovirus gene expression in the presence or absence of E1B 55K were investigated. No Rev/RRE-specific differences in the pattern of viral protein expression were observed. However, RNA analyses showed that the cytoplasmic : nuclear RNA ratios for L3, L4 and E3-late mRNAs (but not L1, L2, L5 or E3-early species) were elevated in *dl*338/RRE*rev* infections, with respect to the ratios in parallel *dl*338/*rev* infections. This apparent Rev-mediated increase in cytoplasmic levels of responsive adenovirus mRNAs did not restore expression to wild type levels. However, full complementation could not be expected in these experiments due to the pleiotropic effects of E1B 55K deletion observed at this time point, which resulted in reduced nuclear as well as cytoplasmic viral RNA levels.

Although the response of L3 mRNA to the Rev/RRE system can be explained by a conventional mechanism in which Rev-RRE binding mediates a process such as facilitated nuclear mRNA export, the effects on L4 and E3-late species are more difficult to interpret since the RRE is removed during processing from the primary

transcripts that give rise to these mRNAs. A mechanism was suggested (7.4, above) in which a Rev/RRE-mediated process results in commitment of *L4* and E3-late species to, for example, a facilitated nuclear export pathway before processing is complete. It was further envisaged that several mechanisms might operate to determine the relative responses of different species to the Rev/RRE system. The presence of intron sequences and unused splice sites (potential nuclear retention signals) in certain mRNAs could make them more responsive to Rev/RRE in the same way that such species show a greater dependence on E1B 55K function (Leppard, 1993). This could explain the unusually large effect observed for L4-100K mRNA. The lack of effect on L1 and L2 RNA species transcribed 5' to L3 may be explained in terms of the pattern of adenovirus RNA processing. Co-transcriptional commitment to splicing and polyadenylation may lead to production of mature L1 and L2 mRNAs before transcription of the RRE is completed. It is also possible that Rev/RRE-mediated facilitation of cytoplasmic mRNA expression is limited in range to species close to the RRE even if the RRE is removed from the mature message. Heaphy *et al.* (1991) and Wingfield *et al.* (1991) have suggested that Rev function is dependent on polymerisation of Rev molecules along RNA from a nucleation site in the RRE. RNA regions such as L5 may be too far away from the RRE in the primary transcript to be reached by the polymeric 'coating' of Rev molecules. Indeed, the distance in the Ad5 genome from the RRE insertion site in L3 to the end of the L5 region is more than 11.3 kbp (greater than the entire length of the HIV-1 genome).

One apparent difference between the mechanism suggested here for Rev/RRE function in adenovirus and the normal operation of the system in HIV is the way mRNAs that do not contain the RRE are handled. For example, cytoplasmic expression of the HIV *tat* and *rev* mRNAs (from which the RRE has been removed by splicing) is not elevated by Rev. However, the multiply-spliced HIV mRNAs, as well as lacking the RRE, do not contain the *cis*-acting repression sequences (unused

splice sites, intronic sequences, or more specific sequences - see 1.6, above) found in the unspliced and singly-spliced mRNAs. Conversely, adenovirus L4-100K mRNA, which showed the most marked response to the Rev/RRE system, contains unused splice sites that could mediate retention of mRNA in the nucleus. Thus adenovirus mRNAs such as L4-100K are in this respect similar to the Rev-responsive incompletely-spliced HIV mRNAs and would be expected to behave similarly if commitment to cytoplasmic expression can indeed occur before RRE excision.

As also noted earlier (7.4), the observed difference in the cytoplasmic : nuclear ratio for L4-100K mRNA between *dl*338/RRE*rev* and *dl*338/*rev* infections was not reflected in a difference in protein levels. However Rev/RRE action may be required at several stages of a facilitated gene expression pathway, including nuclear RNA export and polysome loading of responsive mRNAs. If Rev/RRE action leads to a commitment to nuclear export before RRE excision, mRNAs transported to the cytoplasm via a Rev-mediated pathway may not be efficiently translated in the absence of Rev binding (precluded by removal of the RRE). Thus levels of L4-100K protein in the *dl*338/RRE*rev* infection will be similar to the levels of protein translated from the background levels of cytoplasmic mRNA exported from the nucleus in the *dl*338/*rev* infection by more conventional mechanisms. If this interpretation is correct, then a single facet of Rev function (facilitation of mRNA transport) has been isolated in these experiments.

8.2. Implications for the Rev / E1B 55K analogy

It is possible from the results presented in this thesis to draw some general conclusions about the potential functional analogy between Rev/RRE and E1B 55K. Rev/RRE can function in the context of an adenovirus system to augment the cytoplasmic accumulation of certain mRNAs. However, not all of the Ad mRNAs known to be dependent on E1B 55K are affected (presumably reflecting the chosen

site of insertion of the RRE) and the degree of augmentation varies between different mRNA species, suggesting an unequal dependence of these species on mechanisms that facilitate cytoplasmic RNA accumulation (as has been observed for 55K). It therefore appears that the Rev/RRE and E1B 55K systems act to overcome difficulties in viral gene expression in similar ways and that the HIV-1 system can, to an extent, compensate for the lack of the E1B 55K system in an Ad infection.

It was not possible to draw definite conclusions on any changes in levels of protein translated from L3 mRNAs whose cytoplasmic accumulation was enhanced by the Rev/RRE system. The effect on cytoplasmic accumulation of hexon mRNA was small and a corresponding effect on protein synthesis might reasonably be expected to be difficult to detect. pVI protein, whose mRNA levels were also enhanced by the Rev/RRE system, could not be unambiguously identified in these experiments. As already discussed, the non-correspondence between RNA levels and protein synthesis for the L4-100K mRNA that apparently accumulates in the cytoplasm under the influence of Rev/RRE can be explained if Rev has multiple functions in the gene expression pathway that can be separated. An implication of this explanation is that the multiple activities of Rev are interdependent. If an mRNA transported to the cytoplasm by a Rev/RRE-facilitated mechanism also requires Rev for efficient polysome loading, then these processes must be linked in some way. One possibility is that inefficient polysome loading is intrinsic to Rev-dependent HIV mRNAs. However, the results discussed above suggest that Ad L4-100K mRNA that accumulates in the cytoplasm by a Rev/RRE-mediated process is not efficiently translated, a result suggested to be due to RRE excision after commitment to nuclear export. If this interpretation is correct, then poor translation in the absence of Rev/RRE binding may be a consequence of the transport mechanism, since it occurs with a non-HIV RNA species. Possibly, Rev acts to transport dependent RNAs directly into a nuclear subcompartment in which they are retained without further Rev activity. Such a mechanism of action appears more

complex than that of E1B 55K, whose effects appear to be confined to nuclear export (probably by induction of RNA release from the nuclear matrix) and possibly cytoplasmic stabilisation of late mRNAs (see 1.9.6-1.9.7).

8.3. Further experiments

Several further experiments could be conducted in order to test and extend the hypotheses and conclusions presented in this thesis. One rather unlikely possibility not entirely ruled out by existing experiments is that the RRE has a small Rev-independent *cis*-acting effect on cytoplasmic RNA accumulation that becomes apparent in a *dl*338 (55K$^-$) background (the experiments using the *dl*309 background described in chapter 5 were designed to test for a negative, rather than a positive effect of the RRE on gene expression). Ideally, a *dl*338-derived E1A-deficient virus would be used as a parallel control for the RNase protection assays in KB8a cells (see below). It would, however, be possible to repeat the experiments described in chapter 5 using *dl*338 and *dl*338/RRE to test the effects of the RRE in a *dl*338 background.

Insertion of the RRE at other positions in the major late transcription unit would allow the effects to be compared with those of insertion in L3. However, due to the complexity of the pattern of gene expression from both strands of the adenovirus genome, the number of insertion sites that are theoretically suitable is strictly limited; even at such sites insertion could prove to disrupt gene expression in practice. It would also be useful to construct a recombinant based on *dl*338 that, like the *rev*-recombinants employed in this study, is deficient in E1A function and therefore has comparable expression levels in KB8a cells (an unsuccessful attempt to construct such a virus was in fact made in the course of the project but could not be repeated for lack of time). This would serve as a better negative control for the *dl*338/RRE*rev* experiments than *dl*338/*rev*. An alternative and much more laborious approach would be to construct helper-independent *rev*-recombinants with the *rev*

expression cassette inserted in a nonessential region of the genome. Such viruses could be used in experiments in noncomplementing cells (e.g. HeLa), where direct comparison could be made with E1A-intact *dl*309, *dl*338, *dl*309/RRE and *dl*338/RRE. However, Rev-expressing viruses competent for efficient replication in normal cells would be subject to additional safety considerations and perhaps to a higher level of containment. It was for these reasons that E1A-deficient rather than E1A-competent recombinants were originally designed and constructed.

In general, the analysis of Rev/RRE function in the recombinant viral systems already available could be extended by the use of additional riboprobes to detect a wider range of adenovirus RNA species. It would be especially useful to analyse directly the levels of L4-33K and pVIII mRNAs. Moreover, if antibodies could be obtained to proteins such as L3-pVI, which were difficult to detect in total protein extracts, the RNase protection assays could be correlated directly with protein expression. This would enable the translational fate of mRNAs that retain the RRE (rather than having it excised), and whose cytoplasmic accumulation is augmented by the Rev/RRE system, to be assessed.

References

Adachi, A., Gendelman, H.E., Koenig, S., Folks, T., Willey, R., Rabson, A., & Martin, M.A. (1986) Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J. Virol.* **59** 284-291

Adami, G., & Babiss, L.E. (1991) DNA template effect on RNA splicing: two copies of the same gene in the same nucleus are processed differently. *EMBO J.* **10** 3457-3465

Akusjärvi, G., & Persson, H. (1981) Controls of RNA splicing and termination in the major late adenovirus transcription unit. *Nature* **292** 420-426

Akusjarvi, G. (1993) Proteins with transcription regulatory properties encoded by human adenoviruses. *Trends Microbiol.* **1** 163-170

Anderson, C.W., Schmitt, R.C., Smart, J.E., & Lewis, J.B. (1984) Early region 1B of adenovirus 2 encodes two coterminal proteins of 495 and 155 amino acid residues. *J. Virol.* **50** 387-396

Arrigo, S.J., Weitsman, S., Rosenblatt, J.D., & Chen, I.S.Y. (1989) Analysis of *rev* gene function on human immunodeficieny virus type 1 replication in lymphoid cells by using a quantitative polymerase chain reaction method. *J. Virol.* **63** 4875-4881

Arrigo, S.J., & Chen, I.S.Y. (1991) Rev is necessary for translation but not cytoplasmic accumulation of HIV-1 *vif*, *vpr*, and *env/vpu* 2 RNAs. *Genes Dev.* **5** 808-819

Arrigo, S.J., Heaphy, S., & Haines, J.K. (1992) In vivo binding of wild-type and mutant human immunodeficiency virus type 1 Rev proteins: implications for function. *J. Virol.* **66** 5569-5575

Babiss, L.E., Young, C.S.H., Fisher, P.B., & Ginsberg, H.S. (1983) Expression of adenovirus E1a and E1b gene products and the *Escherichia coli* XGPRT gene in KB cells. *J. Virol.* **46** 454-465

Babiss, L.E., & Ginsberg, H.S. (1984) Adenovirus type 5 early region 1b gene product is required for efficient shutoff of host protein synthesis. *J. Virol.* **50** 202-212

Babiss, L.E., Ginsberg, H.S., & Darnell, J.E. (1985) Adenovirus E1B proteins are required for accumulation of late viral mRNA and for effects on cellular mRNA translation and transport. *Mol. Cell. Biol.* **5** 2552-2558

Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W., & Montagnier, L. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **230** 868-871

Bartel, D.P., Zapp, M.L., Green, M.R., & Szostak, J.W. (1991) HIV-1 Rev regulation involves recognition of non-Watson-Crick base pairs in viral RNA. *Cell* **67** 529-536

Berger, J., Aepinus, C., Dobrovnik, M., Fleckenstein, B., Hauber, J., & Bohnlein, E. (1991) Mutational analysis of functional domains in the HIV-1 Rev *trans*-regulatory protein. *Virology* **183** 630-635

Bevec, D., Dobrovnik, M., Hauber, J., & Böhnlein, E. (1992) Inhibition of human immunodeficiency virus type 1 replication in human T cells by retroviral-mediated gene transfer of a dominant-negative Rev trans-activator. *Proc. Natl. Acad. Sci. USA* **89** 9870-9874

Birnboim, H.C., & Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7** 1513-1525

Blobel, G. (1985) Gene gating: a hypothesis. *Proc. Natl. Acad. Sci. USA* **82** 8527-8529

Bogerd, H., & Greene, W.C. (1993) Dominant negative mutants of human T-cell leukemia virus type I Rex and human immunodeficiency virus type 1 Rev fail to multimerize in vivo. *J. Virol.* **67** 2496-2502

Bos, J., Polder, L.J., Bernards, R., Schrier, P.I., van den Elsen, P.J., van der Eb, A.J., & van Ormondt, H. (1981) The 2.2 kb E1b mRNA of human Ad12 and Ad5 codes for two tumor antigens starting at different AUG triplets. *Cell* **27** 121-131

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* **72** 248-254

Bridge, E., & Ketner, G. (1989) Redundant control of adenovirus late gene expression by early region 4. *J. Virol.* **63** 631-638

Bridge, E., & Ketner, G. (1990) Interaction of adenoviral E4 and E1b products in late gene expression. *Virology* **174** 345-353

Bridge, E., Hemstrom, C., & Pettersson, U. (1991) Differential regulation of adenovirus late transcriptional units by the products of early region 4. *Virology* **183** 260-266

Carter, K.C., Bowman, D., Carrington, W., Fogarty, K., McNeil, J.A., Fay, F.S., & Lawrence, J.B. (1993) A three-dimensional view of precursor messenger RNA metabolism within the mammalian nucleus. *Science* **259** 1330-1335

Chanda, P.K., Natuk, R.J., Mason, B.B., Bhat, B.M., Greenberg, L., Dheer, S.K., Molnar-Kimber, K.L., Mizutani, S., Lubeck, M.D., Davis, A.R. & Hung, P.P. (1990) High level expression of the envelope glycoproteins of the human immunodeficiency virus type 1 in presence of rev gene using helper-independent adenovirus type 7 recombinants. *Virology* **175** 535-547

Chang, D.D., & Sharp, P.A. (1989) Regulation by HIV Rev depends upon recognition of splice sites. *Cell* **59** 789-795

Chang, D.D., & Sharp, P.A. (1990) Messenger RNA transport and HIV *rev* regulation. *Science* **249** 614-615

Chinnadurai, G., Chinnadurai, S., & Brusca, J. (1979) Physical mapping of a large-plaque mutation of adenovirus type 2. *J. Virol.* **32** 623-628

Chroboczek, J., Bieber, F., & Jacrot, B. (1992) The sequence of the genome of adenovirus type 5 and its comparison with the genome of adenovirus type 2. *Virology* **186** 280-285

Ciejek, E.M., Nordstrom, J.L., Tsai, M-J., & O'Malley, B.W. (1982) Ribonucleic acid precursors are associated with the chick oviduct nuclear matrix. *Biochemistry* **21** 4945-4953

Cochrane, A., Kramer, R., Ruben, S., Levine, J., & Rosen, C.A. (1989a) The human immunodeficiency virus *rev* protein is a nuclear phosphoprotein. *Virology* **171** 264-266

Cochrane, A.W., Golub, E., Volsky, D., Ruben, S., & Rosen, C.A. (1989b) Functional significance of phosphorylation to the human immunodeficiency virus Rev protein. *J. Virol.* **63** 4438-4440

Cochrane, A.W., Perkins, A., & Rosen, C.A. (1990a) Identification of sequences important in the nucleolar localization of human immunodeficiency virus Rev: relevance of nucleolar localization to function. *J. Virol.* **64** 881-885

Cochrane, A.W., Chen, C-H., & Rosen, C.A. (1990b) Specific interaction of the human immunodeficiency virus Rev protein with a structured region in the env mRNA. *Proc. Natl. Acad. Sci. USA* **87** 1198-1202

Cochrane, A.W., Jones, K.S., Beidas, S., Dillon, P.J., Skalka, A.M., & Rosen, C.A. (1991) Identification and characterization of intragenic sequences which repress human immunodeficiency virus structural gene expression. *J. Virol* **65** 5305-5313

Constantoulakis, P., Campbell, M., Felber, B.K., Nasioulas, G., Afonina, E., & Pavlakis, G.N. (1993) Inhibition of Rev-mediated HIV-1 expression by an RNA binding protein encoded by the interferon-inducible 9-27 gene. *Science* **259** 1314-1318

Cook. K.S., Fisk, G.J., Hauber, J., Usman, N., Daly, T.J., & Rusche, J.R. (1991) Characterization of HIV-1 REV protein: binding stoichiometry and minimal RNA substrate. *Nucleic Acids Res.* **19** 1577-1583

Cullen, B.R., Hauber, J., Campbell, K., Sodroski, J.G., Haseltine, W.A., & Rosen, C.A. (1988) Subcellular localization of the human immunodeficiency virus *trans*-acting *art* gene product. *J. Virol.* **62** 2498-2501

Cullen, B.R. (1991) Regulation of human immunodeficiency virus replication. *Annu. Rev. Microbiol.* **45** 219-50

Cullen, B.R., & Malim, M.H. (1991) The HIV-1 Rev protein: prototype of a novel class of eukaryotic post-transcriptional regulators. *TIBS* **16** 346-350

Cutt, J.R., Shenk, T., & Hearing, P. (1987) Analysis of adenovirus early region 4-encoded polypeptides synthesized in productively infected cells. *J. Virol* **61** 543-552

Daefler, S., Klotman, M.E., & Wong-Staal, F. (1990) Trans-activating rev protein of the human immunodefiency virus 1 interacts directly and specifically with its target RNA. *Proc. Natl. Acad. Sci. USA* **87** 4571-4575

D'Agostino, D.M., Felber, B.K., Harrison, J.E., & Pavlakis, G.N. (1992) The Rev protein of human immunodeficieny virus type 1 promotes polysomal association and translation of *gag/pol* and *vpu/env* mRNAs. *Mol. Cell. Biol.* **12** 1375-1386

Daly, T.J., Cook, K.S., Gray, G.S., Maione, T.E., & Rusche, J.R. (1989) Specific binding of HIV-1 recombinant Rev protein to the Rev-responsive element *in vitro*. *Nature* **342** 816-819

Daly, T.J., Rusche, J.R., Maione, T.E., & Frankel, A.D. (1990) Circular dichroism studies of the HIV-1 Rev protein and its specific RNA binding site. *Biochemistry* **29** 9791-9795

Dayton, A.I., Terwilliger, E.F., Potz, J., Kowalski, M., Sodroski, J.G., & Haseltine, W.A. (1988) Cis-acting sequences responsive to the *rev* gene product of the human immunodeficieny virus. *J. AIDS* **1** 441-452

Dayton, E.T., Powell, D.M., & Dayton, A.I. (1989) Functional analysis of CAR, the target sequence for the Rev protein of HIV-1. *Science* **246** 1625-1629

Dayton, E.H., Konings, D.A.M., Powell, D.M., Shapiro, B.A., Butini, L., Maizel, J.V., & Dayton, A.I. (1992) Extensive sequence-specific information throughout the CAR/RRE, the target sequence of the human immunodeficiency virus type I Rev protein. *J. Virol.* **66** 1139-1151

Dix, I., & Leppard, K.N. (1992) Open reading frames 1 and 2 of adenovirus region E4 are conserved between human serotypes 2 and 5. *J. Gen. Virol.* **73** 2975-2976

Dix, I., & Leppard, K.N. (1993) Regulated splicing of adenovirus type 5 E4 transcripts and regulated cytoplasmic accumulation of E4 mRNA. *J. Virol.* **67** 3226-3231

Dyson, N., & Harlow, E. (1992) Adenovirus E1A targets key regulators of cell proliferation. In: Cancer surveys volume 12: Tumour Suppressor Genes, the Cell

Cycle and Cancer. A.J. Levine, ed. (Cold Spring Harbor Laboratory Press) pp 161-195

Eastman, A. (1987) An improvement to the novel rapid assay for chloramphenicol acetyltransferase gene expression. *BioTechniques* **5** 730-732

Emerman, M., Vazeux, R., & Peden, K. (1989) The *rev* gene product of the human immunodefiency virus affects envelope-specific RNA localization. *Cell* **57** 1155-1165

Fankhauser, C., Izaurralde, E., Adachi, Y., Wingfield, P., & Laemmli, U.K. (1991) Specific complex of human immunodeficiency virus type 1 Rev and nucleolar B23 proteins: dissociation by the Rev response element. *Mol. Cell. Biol.* **11** 2567-2575

Feinberg, M.B., Jarret, R.F., Aldovini, A., Gallo, R.C., & Wong-Staal, F. (1986) HTLV-III expression and production involve complex regulation at the levels of splicing and translation of viral RNA. *Cell* **46** 807-817

Feinberg, M.B., & Trono, D. (1992) Intracellular immunization: *trans*-dominant mutants of HIV gene products as tools for the study and interruption of viral replication. *AIDS Res. Human Retroviruses* **6** 1013-1022

Felber, B.K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T., & Pavlakis, G.N. (1989) rev protein of human immunodeficiency virus type 1 affects the stability and transport of the viral mRNA. *Proc. Natl. Acad. Sci. USA* **86** 1495-1499

Flint, S.J. (1986) Regulation of adenovirus mRNA formation. *Adv. Virus. Res.* **31** 169-228

Gallo, R., Wong-Staal, F., Montagnier, L., Haseltine, W.A., & Yoshida, M. (1988) HIV/HTLV gene nomenclature. *Nature* **333** 504

Gibson (1984) Studies on the Epstein-Barr virus genome. PhD thesis, Cambridge University, England.

Goh, W.C., Sodroski, J.G., Rosen, C.A., & Haseltine, W.A. (1987) Expression of the *art* gene product of human T-lymphotropic virus type III (HTLV-III/LAV) in Bacteria. *J. virol.* **61** 633-637

Gorman, C.M., Moffat, L.F., & Howard, B.H. (1982) Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2** 1044-1051

Graham, F.L., Smiley, J., Russell, W.C., & Nairn, R. (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36** 59-72

Gutman, D., & Goldenberg, C.J. (1988) Virus-specific splicing inhibitor in extracts from cells infected with HIV-1. *Science* **241** 1492-1495

Hadzopoulou-Cladaras, M., Felber, B.K., Cladaras, C., Athanassopoulos, A., Tse, A., & Pavlakis, G.N. (1989) The *rev* (*trs/art*) protein of human immunodefiency virus type 1 affects viral mRNA and protein expression via a *cis*-acting sequence in the *env* region. *J. virol.* **63** 1265-1274

Halbert, D.N., Cutt, J.R., & Shenks, T. (1985) Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. *J. Virol.* **56** 250-257

Hammarskjold, M-L., Heimer, J., Hammarskjold, B., Sangwan, I., Albert, L., & Rekosh, D. (1989) Regulation of human immunodeficiency virus *env* expression by the *rev* gene product. *J. Virol.* **63** 1959-1966

Hanahan, D. (1983) Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* **166** 557-580

Handa, H., Kingston, R.E., & Sharp, P.A. (1983) Inhibition of adenovirus early region IV transcription *in vitro* by a purified viral DNA binding protein. *Nature* **302** 545-547

Hauber, J., Bouvier, M., Malim, M.H., & Cullen, B.R. (1988) Phosphorylation of the *rev* gene product of human immunodeficiency virus type 1. *J. Virol.* **62** 4801-4804

Heaphy, S., Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Karn, J., Lowe, A.D., Singh, M., & Skinner, M.A. (1990) HIV-1 regulator of virion expression (Rev) protein binds to an RNA stem-loop structure located within the Rev response element region. *Cell* **60** 685-693

Heaphy, S., Finch, J.T., Gait, M.J., Karn, J., & Singh, M. (1991) Human immunodeficiency virus type 1 regulator of virion expression, rev, forms nucleoprotein filaments after binding to a purine-rich "bubble" located within the rev-responsive region of viral mRNAs. *Proc. Natl. Acad. Sci. USA* **88** 7366-7370

Holland, S.M., Ahmad, N., Maitra, R.K., Wingfield, P., & Venkatesan, S. (1990) Human immunodeficiency virus Rev protein recognizes a target sequence in Rev-responsive element RNA within the context of RNA secondary structure. *J. Virol.* **64** 5966-5975

Holland, S.M., Chavez, M., Gerstberger, S., & Venkatesan, S. (1992) A specific sequence with a bulged guanosine residue(s) in a stem-bulge-stem structure of Rev-responsive element RNA is required for *trans* activation by human immunodeficiency virus type 1 Rev. *J. Virol.* **66** 3699-3706

Hope, T.J., Huang, X., McDonald, D., & Parslow, T.G. (1990a) Steroid-receptor fusion of the human immunodeficiency virus type 1 Rev transactivator: Mapping cryptic functions of the arginine-rich motif. *Proc. Natl. Acad. Sci. USA* **87** 7787-7791

Hope, T.J., McDonald, D., Huang, X., Low, J., & Parslow, T.G. (1990b) Mutational analysis of the human immunodeficiency virus type 1 Rev transactivator: essential residues near the amino terminus. *J. Virol.* **64** 5360-5366

Hope, T.J., Bond, B.L., McDonald, D., Klein, N.P., & Parslow, T.G. (1991) Effector domains of human immunodeficiency virus type 1 Rev and human T-cell leukaemia virus type 1 Rex are functionally interchangeable and share an essential peptide motif. *J. Virol.* **65** 6001-6007

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. *J. Virol.* **66** 1849-1855

Horwitz, M.S. (1990a) Adenoviridae and their replication. In: Virology, second edition. Fields, B.N., Knipe, D.M., Chanock, R.M., Hirsch, M.S., Melnick, J.L., & Monath, T.P., eds. (Raven Press Ltd., New York) pp. 1679-1721

Horwitz, M.S. (1990b) Adenoviruses. In: Virology, second edition. Fields, B.N., Knipe, D.M., Chanock, R.M., Hirsch, M.S., Melnick, J.L., & Monath, T.P., eds. (Raven Press Ltd., New York) pp. 1723-1740

Huang, M-M., & Hearing, P. (1989) Adenovirus early region 4 encodes two gene products with redundant effects in lytic infection. *J. Virol.* **63** 2605-2615

Huang, X., Hope, T.J., Bond, B.L., McDonald, D., Grahl, K., & Parslow, T.G. (1991) Minimal Rev-response element for type 1 human immunodefiency virus. *J. Virol.* **65** 2131-2134

Itoh, M., Inoue, J-i, Toyoshima, H., Akizawa, T., Higashi, M., & Yoshida, M. (1989) HTLV-1 *rex* and HIV-1 *rev* act through similar mechanisms to relieve suppression of unspliced RNA expression. *Oncogene* **4** 1275-1279

Ivey-Hoyle, M., & Rosenberg, M. (1990) Rev-dependent expression of human immunodefiency virus type 1 gp160 in *Drosophila melanogaster* cells. *Mol. Cell. Biol.* **10** 6152-6159

Iwai, S., Pritchard, C., Mann, D.A., Karn, J., & Gait, M.J. (1992) Recognition of the high affinity binding site in rev-responsive element RNA by the human immunodeficiency virus type-1 rev protein. *Nucleic Acids Res.* **20** 6465-6472

Jones, N., & Shenk, T. (1979) Isolation of adenovirus type 5 host range deletion mutants defective for transformation of rat embryo cells. *Cell* **17** 683-689

Kjems, J., Brown, M., Chang, D.D., & Sharp, P.A. (1991a) Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element. *Proc. Natl. Acad. Sci. USA* **88** 683-687

Kjems, J., Frankel, A.D., & Sharp, P.A. (1991b) Specific regulation of mRNA splicing in vitro by a peptide from HIV-1 Rev. *Cell* **67** 169-178

Kjems, J., Calnan, B.J., Frankel, A.D., & Sharp, P.A. (1992) Specific binding of a basic peptide from HIV-1 Rev. *EMBO J.* **11** 1119-1129

Knight, D.M., Flomerfelt, F.A., & Ghrayeb, J. (1987) Expression of the art/trs protein of HIV and study of its role in viral envelope synthesis. *Science* **236** 837-840

Krausslich, H-G. (1992) Genetic analysis and gene expression of human immunodeficiency virus. *Curr. Opinion Genetics Dev.* **2** 82-89

Kubota, S., Tetsuya, N., Cullen, B.R., Maki, M., & Hatanaka, M. (1991) Effects of chimeric mutants of human immunodeficiency virus type 1 Rev and human T-cell leukemia virus type I Rex on nucleolar targeting signals. *J. Virol.* **65** 2452-2456

Kubota, S., Furuta, R., Maki, M., & Hatanaka, M (1992) Inhibition of human immunodeficiency virus type 1 Rev function by a Rev mutant which interferes with nuclear/nucleolar localization of Rev. *J. Virol.* **66** 2510-2513

Lazinski, D., Grzadzielska, E., & Das, A. (1989) Sequence-specific recognition of RNA hairpins by bacteriophage antiterminators requires a conserved arginine-rich motif. *Cell* **59** 207-218

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227** 680-685

Lawrence, J.B., Singer, R.H., & Marselle, L.M. (1989) Highly localized tracks of specific transcripts within interphase nuclei visualised by in situ hybridization. *Cell* **493-502**

Le, S-Y., Malim, M.H., Cullen, B.R., & Maizel, J.V. (1990) A highly conserved RNA folding region coincident with the Rev response element of primate immunodeficiency viruses. *Nucleic Acids Res.* **18** 1613-1623

Lechner, R.L., & Kelly, T.J. (1977) The structure of replicating adenovirus 2 DNA molecules. *Cell* **12** 1007-1020

Leppard, K.N. (1993) Selective effects on adenovirus late gene expression of deleting the E1b 55K protein. *J. Gen. Virol.* **74** 575-582

Leppard, K.N., & Shenk, T. (1989) The adenovirus E1B 55kd protein influences mRNA transport via an intranuclear effect on RNA metabolism. *EMBO J.* **8** 2329-2336

Logan, J., & Shenk, T. (1984) Adenovirus tripartite leader sequence enhances translation of mRNAs late after infection. *Proc. Natl. Acad. Sci. USA* **81** 3655-3659

Lu, X., Heimer, J., Rekosh, D., & Hammarskjold, M-L. (1990) U1 small nuclear RNA plays a direct role in the formation of a rev-regulated human immunodefiency virus *env* mRNA that remains unspliced. *Proc. Natl. Acad. Sci. USA* **87** 7598-7602

Maddon, P.J., Dalgleish, A.G., McDougal, J.S., Clapham, P.R., Weiss, R.A., & Axel, R. (1986) The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. *Cell* **47** 333-348

Maldarelli, F., Martin, M.A., & Strebel, K. (1991) Identification of posttranscriptionally active inhibitory sequences in human immunodeficiency virus type 1 RNA: novel level of gene regulation. *J. Virol.* **65** 5732-5743

Malette, P., Yee, S-P., & Branton, P.E. (1983) Studies on the phosphorylation of the 58 000 dalton early region 1B protein of human adenovirus type 5. *J. Gen. Virol.* **64** 1069-1078

Malim, M.H., Hauber, J., Le, S-Y., Maizel, J.V., & Cullen, B.R. (1989a) The HIV-1 *rev trans*-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. *Nature* **338** 254-257

Malim, M.H., Bohnlein, S., Hauber, J., & Cullen, B.R. (1989b) Functional dissection of the HIV-1 Rev *trans*-activator - derivation of a *trans*-dominant repressor of Rev function. *Cell* **58** 205-214

Malim, M.H., Tiley, L.S., McCarn, D.F., Rusche, J.R., Hauber, J., & Cullen, B.R. (1990) HIV-1 structural gene expression requires binding of the Rev *trans*-activator to its RNA target sequence. *Cell* **60** 675-683

Malim, M.H., & Cullen, B.R. (1991) HIV-1 structural gene expression requires the binding of multiple Rev monomers to the viral RRE: implications for HIV-1 latency. *Cell* **65** 241-248

Malim, M.H., McCarn, D.F., Tiley, L.S., & Cullen, B.R. (1991) Mutational definition of the human immunodeficiency virus type 1 Rev activation domain. *J. Virol.* **65** 4248-4254

Mariman, E.C.M, van Eekelen, C.A.G., Reinders, R.J., Berns, A.J.M., & van Venrooij (1982) Adenoviral heterogenous nuclear RNA is associated with the host nuclear matrix during splicing. *J. Mol. Biol.* **154** 103-119

McDonald, D., Hope, T.J., & Parslow, T.G. (1992) Posttranscriptional regulation by the human immunodeficiency virus type 1 Rev and human T-cell leukemia virus type I Rex proteins through a heterologous RNA binding site. *J. Virol.* **66** 7232-7238

McDougal, J., Kennedy, M., Sligh, J., Cort, S., Mawie, A., & Nicholson, J. (1986) Binding of the HTLV-III/LAV to $T4^+$ T cells by a complex of the 110K viral protein and the T4 molecule. *Science* **231** 382-385

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K., & Green, M.R. (1984) Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res* **12** 7035-7056

Mermer, B., Felber, B.K., Campbell, M., & Pavlakis, G.N. (1990) Identification of *trans*-dominant HIV-1 rev protein mutants by direct transfer of bacterially produced proteins into human cells. *Nucleic Acids Res.* **18** 2037-2044

Messing, J., Gronenborn, B., Müller-Hill, B., & Hofschneider, P. H. (1977) Filamentous coliphage M13 as a cloning vehicle: Insertion of a *Hin*dII fragment of

the *lac* regulatory region in M13 replicative form *in vitro*. *Proc. Natl. Acad. Sci. USA* **74** 3642-3646

Montell, C., Fisher, E.F., Caruthers, M.H., & Berk, A.J. (1984) Control of adenovirus E1B mRNA synthesis by a shift in the activities of RNA splice sites. *Mol. Cell. Biol.* **4** 966-972

Moore, M., Schaack, J., Baim, S.B., Morimoto, R.I., & Shenk, T. (1987) Induced heat shock mRNAs escape the nucleocytoplasmic transport block in adenovirus-infected HeLa cells. *Mol. Cell. Biol.* **7** 4502-4512

Moran, E., & Mathews, M.B. (1987) Multiple functional domains in the adenovirus E1A gene. *Cell* **48** 177-178

Nalin, C.M., Purcell, R.D., Antelman, D., Mueller, D., Tomchak, L., Wegrzynski, B., McCarney, E., Toome, V., Kramer, R., & Hsu, M-C. (1990) Purification and characterization of recombinant Rev protein of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. USA* **87** 7593-7597

Neumann, J.R., Morency, C.A., & Russian, K.O. (1987) A novel rapid assay for chloramphenicol acetyltransferase gene expression. *BioTechniques* **5** 444-447

Öhman, K., Nordqvist, K., & Akusjärvi, G. (1993) Two adenovirus proteins with redundant activities in virus growth facilitates tripartite leader mRNA accumulation. *Virology* **194** 50-58

Olsen, H.S., Nelbock, P., Cochrane, A.W., & Rosen, C.A. (1990a) Secondary structure is the major determinant for interaction of HIV *rev* protein with RNA. *Science* **247** 845-848

Olsen, H.S., Cochrane, A.W., Dillon, P.J., Nalin, C.M., & Rosen, C.A. (1990b) Interaction of the human immunodeficiency virus type 1 Rev protein with a structured region in *env* mRNA is dependent on multimer formation mediated through a basic stretch of amino acids. *Genes Dev.* **4** 1357-1364

Olsen, H.S., Cochrane, A.W., & Rosen, C. (1992) Interaction of cellular factors with intragenic *cis*-acting repressive sequences within the HIV genome. *Virology* **191** 709-715

Ornelles, D.A., & Shenk, T. (1991) Localization of the adenovirus early region 1B 55-kilodalton protein during lytic infection: association with nuclear viral inclusions requires the early region 4 34-kilodalton protein. *J. Virol.* **65** 424-439

Perkins, A., Cochrane, A.W., Ruben, S.M., & Rosen, C.A. (1989) Structural and functional characterization of the human immunodeficiency virus *rev* protein. *J. AIDS* **2** 256-263

Pilder, S., Moore, M., Logan, J., & Shenk, T. (1986) The adenovirus E1B-55K transforming polypeptide modulates transport or cytoplasmic stabilisation of viral and host cell mRNAs. *Mol. Cell. Biol.* **6** 470-476

Rosen, C.A., & Haseltine, W.A. (1988) Selective regulation of HIV gene expression by the *art* gene product. In: The Control of Human Retrovirus Gene Expression, R. Franza, B.R. Cullen, F. Wong-Staal, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 23-26

Rosen, C.A., Terwilliger, E., Dayton, A., Sodroski, J.G., & Haseltine, W.A. (1988) Intragenic cis-acting *art* gene-responsive sequences of the human immunodeficiency virus. *Biochemistry* **85** 2071-2075

Rosenfeld, P.J., O'Neill, E.A., Wides, R.J., & Kelly, T.J. (1987) Sequence-specific interactions between cellular DNA-binding proteins and the adenovirus origin of DNA replication. *Mol. Cell. Biol.* **7** 875-886

Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989) Molecular cloning, a laboratory manual, second edition. (Cold Spring Harbor Laboratory Press)

Samulski, R.J., & Shenk, T. (1988) Adenovirus E1B 55-M_r polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. *J. Virol.* **62** 206-210

Sarnow, P., Hearing, P., Anderson, C.W., Halbert, D.N., Shenk, T., & Levine, A.J. (1984) Adenovirus early region 1B 58,000-dalton tumor antigen is physically associated with an early region 4 25,000-dalton protein in productively-infected cells. *J. Virol.* **49** 692-700

Schröder, H.C., Bachmann, M., Diehl-Seifert, B., & Müller, W.G. (1987) Transport of mRNA from nucleus to cytoplasm. *Prog. Nucleic Acid Res.* **34** 89-142

Schwartz, S., Felber, B.K., & Pavlakis, G.N. (1992a) Distinct RNA sequences in the *gag* region of human immunodeficiency virus type 1 decrease RNA stability and inhibit expression in the absence of Rev protein. *J. Virol.* **66** 150-159

Schwartz, S., Campbell, M., Nasioulas, G., Harrison, J., Felber, B.K., & Pavlakis, G.N. (1992b) Mutational inactivation of an inhibitory sequence in human

immunodeficiency virus type 1 results in Rev-independent *gag* expression. *J. Virol.* **66** 7176-7182

Shaw, G., & Kamen, R. (1986) A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46** 659-667

Smiley, J.K., Young, M.A., & Flint, S.J. (1990) Intranuclear location of the adenovirus type 5 E1B 55-kilodalton protein. *J. Virol.* **64** 4558-4564

Smith, D.H., Kegler, D.M., & Ziff, E.B. (1985) Vector expression of adenovirus type 5 E1a proteins: evidence for E1a autoregulation. *Mol. Cell. Biol.* **5** 2684-2696

Sodroski, J., Goh, W.C., Rosen, C., Dayton, A., Terwilliger, E., & Haseltine, W. (1986) A second post-transcriptional *trans*-activator gene required for HTLV-III replication. *Nature* **321** 412-417

Stillman, B. (1989) Initiation of eukaryotic DNA replication in vitro. *Annu. Rev. Cell Biol.* **5** 197-245

Tan, R., Chen, L., Buettner, J.A., Hudson, D., & Frankel, A.D. (1993) RNA recognition by an isolated α-helix. *Cell* **73** 1031-1040

Tiley, L.S., Malim, M.H., & Cullen, B.R. (1991) Conserved functional organization of the human immunodeficiency virus type 1 and visna virus Rev proteins. *J. Virol.* **65** 3877-3881

Tiley, L.S., Malim, M.H., Tewary, H.K., Stockley, P.G., & Cullen, B.R. (1992) Identification of a high-affinity RNA-binding site for the human immunodeficiency virus type 1 Rev protein. *Proc. Natl. Acad. Sci. USA* **89** 758-762

Tollefson, A.E., Scaria, A., Saha, S., & Wold, W.S.M. (1992) The 11,600-M_W protein encoded by region E3 of adenovirus is expressed early but is greatly amplified at late stages of infection. *J. Virol.* **66** 3633-3642

Trono, D., & Baltimore, D. (1990) A human cell factor is essential for HIV-1 Rev action. *EMBO J.* **9** 4155-4160

Vaishnav, Y.N., & Wong-Staal, F. (1991) The biochemistry of AIDS. *Annu. Rev. Biochem.* **60** 577-630

Venkatesan, S., Gerstberger, S.M., Park, H., Holland, S.M., & Nam, Y-S. (1992) Human immunodeficiency virus type 1 Rev activation can be achieved without Rev-responsive element RNA if Rev is directed to the target as a Rev/MS2 fusion protein which tethers the MS2 operator RNA. *J. Virol.* **66** 7469-7480

Venkatesh, L.K., & Chinnadurai, G. (1990) Mutants in a conserved region near the carboxy-terminus of HIV-1 Rev identify functionally important residues and exhibit a dominant negative phenotype. *Virology* **178** 327-330

Venkatesh, L.K., Mohammed, S., & Chinnadurai, G. (1990) Functional domains of the HIV-1 *rev* gene required for *trans*-regulation and subcellular localization. *Virology* **176** 39-47

Virtanen, A., Gilardi, P., Näslund, A., LeMoullec, J.M., Pettersson, U., & Perricaudet, M. (1984) mRNAs from human adenovirus 2 early region 4. *J. Virol.* **51** 822-831

Virtanen, A. & Pettersson, U. (1985) Organization of early region 1B of human adenovirus type 2: identification of four differentially spliced mRNAs. *J. Virol.* **54** 383-391

Weichselbraun, I., Farrington, G.K., Rusche, J.R., Böhnlein, E., & Hauber, J. (1992) Definition of the human immunodeficiency virus type 1 Rev and human T-cell leukemia virus type I Rex protein activation domain by functional exchange. *J. Virol.* **66** 2583-2587

Weinberg, D.H., & Ketner, G. (1986) Adenoviral early region 4 is required for efficient viral DNA replication and for late gene expression. *J. Virol.* **57** 833-838

White, E., Cipriani, R., Sabbatini, P., & Denton, A. (1991) Adenovirus E1B 19-kilodalton protein overcomes the cytotoxicity of E1A proteins. *J. Virol.* **65** 2968-2978

White, E., Sabbatini, P., Debbas, M., Wold, W.S.M., Kusher, D.I., & Gooding, L.R. (1992) The 19-kilodalton adenovirus E1B transforming protein inhibits programmed cell death and prevents cytolysis by tumor necrosis factor α. *Mol. Cell. Biol.* **12** 2570-2580

Williams, J., Karger, B.D., Ho, Y.S., Castiglia, C.L., Mann, T., & Flint, S.J. (1986) The adenovirus E1B 495R protein plays a role in regulating the transport and stability of the viral late messages. *Cancer Cells* **4** 275-284

Winberg, G., & Shenk, T. (1984) Dissection of overlapping functions within the adenovirus type 5 E1A gene. *EMBO J.* **3** 1907-1912

Wingfield, P.T., Stahl, S.J., Payton, M.A., Venkatesan, S., Misra, M., & Steven, A.C. (1991) HIV-1 Rev expressed in recombinant *Escherichia coli*: purification, polymerisation, and conformational properties. *Biochemistry* **30** 7527-7534

Wold, W.S.M., & Gooding, L.R. (1991) Region E3 of adenovirus: a cassette of genes involved in host immunosurveillance and virus-cell interactions. *Virology* **184** 1-8

Wong-Staal, F. (1990) Human immunodeficiency viruses and their replication. In: Virology, second edition. Fields, B.N., Knipe, D.M., Chanock, R.M., Hirsch, M.S., Melnick, J.L., & Monath, T.P., eds. (Raven Press Ltd., New York) pp. 1529-1543

Xing, Y., Johnson, C.V., Dobner, P.R., & Lawrence, J.B. (1993) Higher level organization of individual gene transcription and RNA splicing. *Science* **259** 1326-1330

Yew, P.R., Kao, C.C., & Berk, A.J. (1990) Dissection of functional domains in the adenovirus-2 early 1B-55K-polypeptide by suppression linker insertional mutagenesis. *Virology* **179** 795-805

Yew, P.R., & Berk, A.J. (1992) Inhibition of p53 transactivation required for transformation by adenovirus early 1B protein. *Nature* **357** 82-85

Yoshida, K., Venkatesh, L., Kuppuswamy, M., & Chinnadurai, G. (1987) Adenovirus transforming 19-Kd T antigen has an enhancer-dependent *trans*-activation function and relieves enhancer repression mediated by viral and cellular genes. *Genes Dev.* **1** 645-658

Zapp, M.L., & Green, M.R. (1989) Sequence-specific RNA binding by the HIV-1 Rev protein. *Nature* **342** 714-716

Zapp, M.L., Hope, T.J., Parslow, T.G., & Green, M.R. (1991) Oligomerization and RNA binding domains of the type 1 human immunodeficiency virus Rev protein: a dual function for an arginine-rich binding motif. *Proc. Natl. Acad. Sci, USA* **88** 7734-7738

TITLE CONSTRUCTION AND ANALYSIS OF ADENOVIRUS/HIV-1 REV RECOMBINANTS

AUTHOR Richard Dafydd WILLIAMS

DEGREE Ph D

AWARDING BODY Warwick University

DATE 1993

THESIS NUMBER DX182473