HERPES SIMPLEX VIRUS RIBONUCLEOTIDE REDUCTASE DNA AND REQUIREMENTS FOR mRNA 3' END FORMATION

by

JOHN McLAUCHLAN

A Thesis presented for the degree of

Doctor of Philosophy

in

The Faculty of Science

at the University of Glasgow

Institute of Virology,

Church Street,

GLASGOW G11 5JR.

January 1986

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

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I thank Professor John H. Subak-Sharpe for provision of facilities within the Institute and for his overall supervision of this work.

I express my gratitude to my immediate supervisor Dr. J. Barklie Clements for his encouragement and helpful advice, and for critical reading of this manuscript.

I am grateful to Dr. Dairena Gaffney for her help in performing CAT assays and to Dr. Phil Taylor for calculating the Δ G values on Page 144.

Invaluable discussions with colleagues in the Institute, in particular Drs. Fraser Rixon, Lindsay Whitton and Bernadette Dutia were appreciated.

I wish to thank Drs. Duncan McGeoch and Phil Taylor for their assistance with computer programs.

I am grateful to Mr. Graham Hope for proof reading this manuscript.

My thanks to the office staff for their cheerful advice on the format for this thesis.

I thank my friends both within and outside the Institute for their help and encouragement.

To my family, especially my father and mother, I express my gratitude for their support and advice.

Finally, this work is dedicated to my long-suffering wife, Gladys, and my sons, Paul and Stephen, who were a constant source of encouragement.

During the period of this work, the author was employed by the Medical Research Council. Unless otherwise stated, all the results presented in this thesis were obtained by the author's own efforts.

SUMMARY

This study was undertaken to examine the organisation and structures of mRNAs mapping at the herpes simplex virus type 2 (HSV-2) ribonucleotide reductase locus. From comparisons between equivalent HSV-2 and HSV-1 nucleotide sequences, putative transcription control signals were identified, and functional analysis of certain control elements was performed. The positions of translated regions within the mRNAs encoding ribonucleotide reductase were identified as were conserved amino acid domains between viral and cellular reductases. During the course of these comparisons, a conserved DNA sequence, YGTGTTYY (Y = pyrimidine), located 30nuc downstream from the polyadenylation (AATAAA) signal of both viral and cellular genes, was identified. Functional tests showed that the YGTGTTYY signal was required for efficient formation of mRNA 3' termini.

Arrangement of mRNAs and DNA Sequence Comparisons at the HSV-2 Ribonucleotide Reductase Locus.

Four overlapping mRNAs have been identified at the ribonucleotide reductase locus. These mRNAs are unspliced and share common 5' and 3' termini. Two early mRNAs with sizes of 4.5kb and 1.2kb have a common 3' terminus and encode HSV-2 polypeptides Vmw138 and Vmw38 respectively which are almost certainly components of the viral ribonucleotide reductase. The other two mRNAs are late, 5' co-terminal species with sizes of 6.4kb and 1.7kb which

appear to encode an identical 54,000 mol. wt. protein; the 6.4kb transcript is 3' co-terminal with the early mRNAs. As a consequence of this mRNA arrangement, the 5' termini of the 4.5kb and 1.2kb mRNAs and the 3' terminus of the 1.7kb mRNA are located within larger, overlapping transcripts. The genome region adjacent to the 6.4kb and 1.7kb mRNAs specifies three late transcripts which are synthesised in the opposite orientation; the 5' terminal region of one of these species overlaps with the 6.4kb and 1.7kb mRNAs.

The 3' terminus of the late 1.7kb mRNA is located within the transcribed region of the 6.4kb species. Previously, it has been proposed that read-through of HSV poly A sites is due to inefficient processing at certain 3' termini. The ability of sequences flanking the 1.7kb mRNA 3' terminus to produce functional mRNA was tested using a plasmid containing the bacterial chloramphenicol acetyltransferase (CAT) protein coding sequences fused to an HSV-2 immediate early promoter. CAT activities produced in HeLa cells by a plasmid carrying the 1.7kb mRNA 3' terminal sequences were comparable to activities obtained with a plasmid containing the 3' processing signals from a non-internal poly A site. Read-through of HSV poly A sites may therefore reflect a general reduction in mRNA 3' processing efficiency rather than differences in the 3' processing signals of individual genes.

The HSV-1 transcripts equivalent to the 4.5kb and 1.2kb mRNAs have sizes of 5.0kb and 1.2kb and encode HSV-1 polypeptides Vmw136 and Vmw38 respectively. Nucleotide

sequence comparisons between the 5' flanking regions of these equivalent HSV-2 and HSV-1 mRNAs have revealed the following homologies:

1) upstream from the 4.5kb and 5.0kb mRNA 5' termini, blocks of conserved sequences are present which resemble transcription control signals at the promoter regions of other genes; TATA box homologues, C-rich tracts and A+C-rich elements have been identified. The C-rich tracts are similar to the consensus sequence for the binding site of Sp1, a transcription control factor, and, from other studies, the A+C-rich sequences could be a component of other HSV early promoter regions.

2) the 5' terminal region of the HSV-2 1.2kb mRNA is located within the polypeptide coding region of Vmw138; an identical situation exists for the equivalent HSV-1 transcript. The HSV-2 and HSV-1 5' flanking sequences show high homology, and both sets of nucleotide sequences contain inverted repeats which could form stem and loop structures. Possible effects of these structures on transcription of the HSV-2 and HSV-1 1.2kb mRNAs are discussed.

Amino Acid Conservation Between Herpesvirus and Cellular Ribonucleotide Reductases.

The amino acid sequences for the Vmw138 N- and C-terminal regions and all of the Vmw38 polypeptide have been predicted from the nucleotide sequences. Comparisons with coding regions of the corresponding HSV-1 proteins,

Vmw136 and Vmw38, reveal high homology apart from areas at the N-termini of equivalent proteins. Poorly conserved N-terminal regions have been identified in other equivalent HSV-2 and HSV-1 proteins; these may be non-essential for function.

In common with other ribonucleotide reductases with similar enzymatic properties, HSV ribonucleotide reductase appears to contain two non-identical polypeptides, and HSV-2 polypeptides Vmw138 and Vmw38 almost certainly represent components of the virus enzyme. The amino acid sequences of Vmw138 and Vmw38 share areas of homology with coding regions on the Epstein-Barr virus (EBV) and varicella-zoster virus genomes; this allows identification of putative proteins which appear to be ribonucleotide reductase components specified by these viruses. These blocks of homology extend to cellular ribonucleotide reductases of eukaryotic and prokaryotic origin and thus may represent domains involved in enzyme function.

Interestingly, Vmw138 and Vmw136 both contain large N-terminal domains which are absent in equivalent herpesvirus and in cellular ribonucleotide reductase proteins. These N-terminal domains could have arisen by gene fusion and may represent functions not essential for ribonucleotide reductase activity.

The EBV polypeptide which shares homology with HSV-2 Vmw138 has a size of 93,000 mol. wt. The C-terminal portions of both proteins contain the 5' terminal regions of the mRNAs specifying the proposed small subunit of

ribonucleotide reductases. However, these 5' termini are not located at equivalent positions in the Vmw138 and 93,000 mol. wt. polypeptides and separate evolutionary routes are proposed in the formation of these internal promoters.

HSV has the ability to transform cells and is potentially oncogenic; the transformating ability of HSV may be linked to its action as a mutagen. One possibility is that this mutagenic effect is due to the expression of HSV ribonucleotide reductase and this is discussed.

Sequences Downstream from the AATAAA Signal Required for mRNA 3' End Formation.

Comparisons of the regions flanking HSV mRNA 3' termini identified conserved sequences downstream from poly A sites. From more extensive comparisons, 67% of the mammalian and viral genes examined contained 3' flanking sequences similar to those conserved in HSV. A consensus sequence, YGTGTTYY (Y = pyrimidine), was derived for a conserved downstream element, preferentially located 24 to 35nuc beyond the AATAAA signal. Downstream from the YGTGTTYY sequence, other G-rich and T-rich segments also were conserved but to a lesser extent. To test the functional significance of YGTGTTYY in mRNA 3' end formation, deletions produced in the 3' flanking sequences of HSV-2 IE gene-5 were inserted into a CAT plasmid. Levels of correctly-terminated mRNA from plasmid constructions, both with and without deletions, were estimated by assaying

CAT activities and by nuclease S1 analysis. Deletion of downstream sequences which extended up to the YGTGTTYY sequence reduced CAT levels to values 35% of those obtained with undeleted plasmids. Removal of a further 14bp, which deleted YGTGTTYY but not the poly A site, reduced CAT activities to levels of 1% to 4%. The levels of CAT mRNA 3' termini reflected the reductions in CAT activities, but, in contrast, levels of mRNA 5' termini were unaffected by these deletions.

These results demonstrate a requirement of 3' flanking sequences, in particular the YGTGTTYY signal, for efficient 3' end formation of poly A mRNAs. An analogy is drawn between those sequences required for efficient 3' end formation of poly A mRNAs with those required for 3' end processing of the non-polyadenylated histone genes. Alternative functions for the YGTGTTYY sequence in either mRNA termination or mRNA processing are discussed. These proposals lead to the possibility that the presence of the trinucleotide, GTG, a component of the consensus which is associated with local alterations in DNA structure, could be a more general feature of the sequences involved in DNA-protein interactions. The possible interaction of precursor RNAs with small nuclear ribonucleoprotein particles and the limits of sequences involved in mRNA 3' end formation also are discussed.

ABBREVIATIONS

The following list gives the key to abbreviations used in the text. A number of abbreviations for buffers and media are given in the Materials and Methods section, and those describing the segments comprising herpesvirus genomes are on Pages 5 to 7 and Page 13. The key to abbreviations in figures is provided on accompanying figure legends.

A	adenine
acetyl CoA	acetyl coenzyme A
Ad	adenovirus
ADP	adenosine diphosphate
ALV	avian leukosis virus
amp^r	ampicillin resistance
ATP	adenosine triphosphate
b, bp	base(s), base pairs(s)
BHK	baby hamster kidney
bisacrylamide	N,N'-methylene bisacrylamide
BPV	bovine papilloma virus
BSA	bovine serum albumin
C	cytosine
^{14}C	radiolabelled carbon
CAT	chloramphenicol acetyltransferase
CCV	channel catfish virus
CDP	cytidine diphosphate
Ci	Curie(s)
cm	centimetre
CGRP	calcitonin gene-related peptide

dADP	deoxyadenosine diphosphate
dATP	deoxyadenosine triphosphate
dCDP	deoxycytidine diphosphate
dCTP	deoxycytidine triphosphate
dGDP	deoxyguanosine diphosphate
dGTP	deoxyguanosine triphosphate
DMSO	dimethyl sulphoxide
DNase	deoxyribonuclease
dNTP(s)	deoxyribonucleoside triphosphate(s)
DTT	dithiothreitol
dTTP	thymidine triphosphate
dUDP	deoxyuridine diphosphate
dUMP	deoxyuridine monophosphate
dUTP	deoxyuridine triphosphate
EBV	Epstein-Barr virus
EDTA	sodium ethylenediamine tetra-acetic acid
EHV	equine herpesvirus
EM	electron microscope
G	guanine
Δ G	free energy of formation
g	gram(s)
g(D)	glycoprotein (D)
G+C	moles per cent guanosine + cytidine moieties
GDP	guanosine diphosphate
h	hour(s)
HCMV	human cytomegalovirus
HEPES	N-2-hydroxyethyl piperazine-N'-2-ethane sulphonic acid

HSV	herpes simplex virus
HVA	herpesvirus ateles
HVP	herpesvirus papio
HVS	herpesvirus saimiri
IE	immediate early
Ig	immunoglobulin
ILB	isotonic lysis buffer
IS	insertion sequence
IVS	intervening sequence
kb, kbp	kilobase(s), kilobase pair(s)
l	litre
LC	light chain
LTR	long terminal repeat
m^7G	methyl-7-guanosine
mcs	multiple cloning site
min	minute
ml	millilitre
mm	millimetre
moi	multiplicity of infection
mol. wt.	molecular weight
mtr	morphological transforming region
N	unspecified nucleotide (A, G, C or T)
NaAc	sodium acetate
NDP(s)	ribonucleoside diphosphate(s)
ng	nanogram
nm	nanometre
NMR	nuclear magnetic resonance
NP40	nonidet P40

npt	non-permissive temperature
NTP(s)	ribonucleoside triphosphate(s)
nuc	nucleotide(s)
O.D.	optical density
^{32}P	radiolabelled phosphate
PAA	phosphonoacetic acid
PAGE	polyacrylamide gel electrophoresis
PEB	phenol extraction buffer
pfu	plaque forming units
p.i.	post infection
PIPES	piperazine-N,N'-bis(2-ethane sulphonic acid)
poly A	polyadenylated
PRV	pseudorabies virus
R	purine moiety
RNase	ribonuclease
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
sec	second(s)
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
SV40	simian virus 40
T	thymine
TEMED	N,N,N',N'-tetramethylethylene diamine
tet^r	tetracycline resistance
TK	thymidine kinase
ts	temperature-sensitive
U	uracil

UDP	uridine diphosphate
UV	ultra-violet
V	volt(s)
v/v	volume/volume
VA RNA	virus-associated RNA
Vmw	molecular weight of viral polypeptide in kilodaltons
vol	volume(s)
VZV	varicella-zoster virus
W	watts
w/v	weight/volume
wt	wild type
Xgal	5-bromo 4-chloro 3-indolyl β D galactopyranoside
Y	pyrimidine moiety
µCi	microcurie
µg	microgram
µl	microlitre

INTRODUCTION

The herpesviruses are a family of DNA viruses with large double-stranded genomes which infect a wide range of organisms from lower order to higher order eukaryotes. This thesis is concerned with herpes simplex virus (HSV), in particular HSV-2 which causes clinical disease in man. The studies presented deal with the organisation of the mRNAs and the DNA sequences which specify a virus enzyme, ribonucleotide reductase. In addition, these studies identify transcription control signals involved in gene expression, in particular sequences required for 3' end formation of polyadenylated mRNAs.

This Introduction provides an overall review of herpesvirus biology, and focusses on the properties of HSV. The HSV-2 genomic region analysed has been implicated in cellular transformation, thus a more complete description of this topic is given. The results are interpreted in relation to data in addition to that obtained for HSV, hence the Introduction contains general sections on the properties of ribonucleotide reductases, and on eukaryotic transcription with emphasis on the arrangements of mRNAs and mRNA 3' end formation.

SECTION A.

PROPERTIES OF THE HERPESVIRUSES.

1. Classification of Herpesviruses.

Members of the Herpetoviridae family (Fenner, 1976) are identified by morphology of the virus particle which comprises four distinct components, namely:

1) the core which consists of an electron-dense toroid structure (Epstein, 1962; Furlong et al., 1972). The torus contains DNA wound around a less dense proteinaceous cylinder (Chai, 1971; Furlong et al., 1972).

2) surrounding the core is the capsid, an icosahedral structure with a diameter of 100nm, consisting of 162 capsomeres (Wildy et al., 1960).

3) a fibrous layer termed the tegument (Roizman and Furlong, 1974) the width of which varies markedly within the Herpetoviridae family (Nayak, 1971).

4) the outer envelope (Wildy et al., 1960) which herpesvirus virions acquire by budding through the inner nuclear membrane. This component is a bilayered membrane, probably composed of lipid, which has periodic surface projections (Spring et al., 1968). In addition, the envelope contains most of the virus-specified glycoproteins (Spear and Roizman, 1972).

To date, at least eighty different herpesviruses with a wide variety of biological properties have been isolated (listed in Roizman, 1982), hence classification

of such a diverse family is difficult. However, viruses within the Herpetoviridae family have been categorised on the basis of either a) their biological properties or b) structural arrangements of their genomes (Roizman et al., 1978; Roizman, 1982).

a) Classification on the basis of biological properties.

From the biological properties of host range, duration of reproductive cycle, cytopathology and characteristics of latent infection, herpesviruses can be divided into three sub-families, the alpha-, beta- and gammaherpesvirinae (Roizman et al., 1978; Matthews, 1982; Roizman, 1982).

The alphaherpesvirinae are cytolytic herpesviruses which have a variable host range both in vivo and in vitro, a short reproductive cycle, spread rapidly during infection and frequently establish latent infections in ganglia. Members of this sub-family include HSV-1 and HSV-2, equid herpesvirus 1 (EHV-1; Randall et al., 1953), bovine mammilitis virus (BMV; Martin et al., 1966), pseudorabies virus (PRV; Gustafsohn, 1970), varicella-zoster virus (VZV; Weller, 1953) and channel catfish virus (CCV; Wolf and Darlington, 1971; Chousterman et al., 1979). The betaherpesvirinae are typified by a narrow host range, a relatively long reproductive cycle, slow progression in culture and may establish a latent state in secretory glands and lymphoreticular cells. Murine cytomegalovirus (MCMV; Smith, 1954) and human cytomegalovirus (HCMV;

Smith, 1956) are members of this group. The gammaherpesvirinae are specific for either T- or B-lymphocytes in vitro, have a variable reproductive cycle and cytopathology, and frequently give rise to a latent infection in lymphoid tissue. Viruses within this group are Epstein-Barr virus (EBV; Epstein et al., 1965), Marek's disease virus (MDV; Marek, 1907; Churchill and Biggs, 1967), herpesvirus saimiri (HVS; Melendez et al., 1968) and herpesvirus ateles (HVA; Melendez et al., 1972).

b) Classification by genome structure.

Herpesviruses contain DNA genomes of at least 130kbp in length which consist of unique and reiterated sequences. Furthermore, the G+C content of herpesvirus DNAs varies from 33% to 75% (reviewed in Honess and Watson, 1977). Variability in the above genomic features forms the basis of a further method for categorisation of the Herpetoviridae family, in particular the different arrangements of unique and reiterated sequences within genomes. A limited number of herpesviruses have been studied in sufficient detail to allow classification by this method, however, five distinct groups have been identified (Fig. 1).

1) Group A. CCV is the sole identified member of this group and CCV DNA consists of a unique segment of 90kbp flanked by direct repeats of 20kbp which are not highly reiterated (Chousterman et al., 1979).

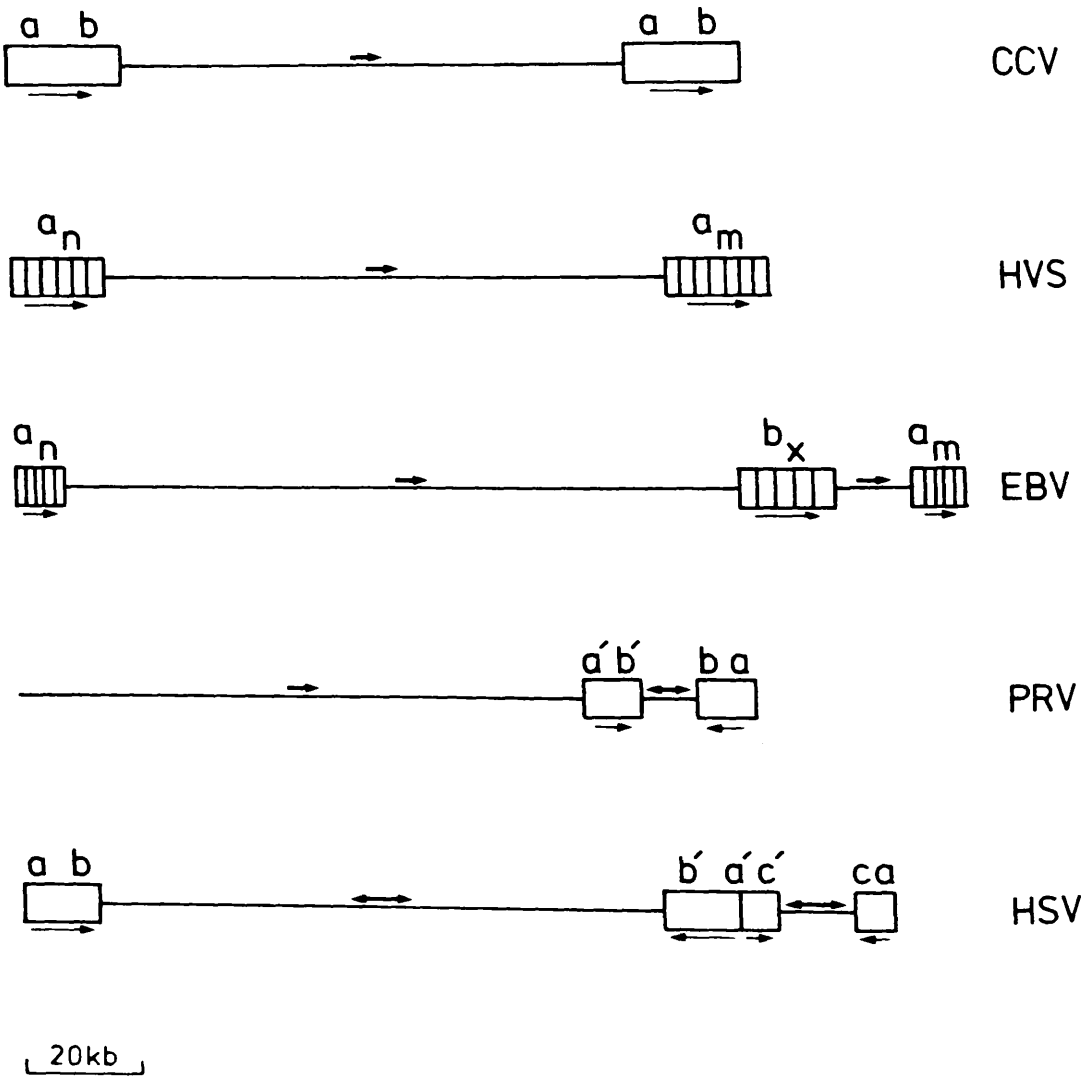

a b
a b
CCV
a_n
a_m
HVS
a_n
b_x
a_m
EBV
a´b´
b a
PRV
a b
b´ a´c´
ca
HSV
20kb

Figure 1. Genome structures representing five herpesvirus groups. Virus examples are given for each group and genomes are drawn to scale. Boxed areas represent repetitive sequences and single lines indicate unique segments. The letters, a, b and c show the arrangement of sequences within the repeats of each group; sequences arranged in inverse orientation relative to a, b and c are denoted by a', b' and c'. Arrows below boxed areas indicate the relative orientations of repeats. Tandemly reiterated sequences within repeats are shown by vertical lines within boxed regions. The number of copies of tandemly reiterated sequences is represented by subscripts (n, m and x) and different subscripts within a group indicate that copy numbers differ between repeats. Segments which invert are shown by double-headed arrows above unique sequences; unique sequences which do not invert are indicated by single-headed arrows. The key to abbreviations is: CCV, channel catfish virus; HVS, herpesvirus saimiri; EBV, Epstein-Barr virus; PRV, pseudorabies virus; HSV, herpes simplex virus.

2) Group B. In this group, the genome arrangement consists of a unique segment flanked by direct repeats which are composed of a series of reiterated sequences arranged as multiple tandem copies. These reiterated sequences are about 1.5kbp in length in HVS, a member of this group (Fleckenstein and Wolf, 1974; Fleckenstein and Bornkamm, 1975; Simonds et al., 1975; Bornkamm et al., 1976). The HVS genome is approximately 160kbp in length (Fleckenstein and Wolf, 1974) with most of the genetic information contained within the long unique (U_L) segment which is 112kbp (Knust et al., 1983). Thus, reiterated sequences account for approximately 30% of the unit-length genome, however, the repeat copy numbers vary at either terminus (Fleckenstein and Mulder, 1980). The G+C content of HVS DNA varies between 36% in the unique segment to 71% in the repetitive sequences (Fleckenstein and Wolf, 1974). The nucleotide sequence of the 1.5kbp repeat unit appears not to contain transcription control signals and transcripts generated from the repetitive sequences have not been detected in HVS-infected cells (Bankier et al., 1985). Two other herpesviruses of New World primates, herpesvirus aotus type 2 (Barahona et al., 1973; Fuchs et al., 1985) and HVA (Fleckenstein et al., 1978) and also bovine herpesvirus type 4 (Todd and Storz, 1983; Storz et al., 1984) have similar genome arrangements to HVS DNA.

3) Group C. This group includes EBV, whose DNA has a mean G+C content of 58% (Schulte-Holthausen and zur Hausen,

1970), and its simian counterpart, herpesvirus papio (HVP; Falk et al., 1976). The entire sequence of EBV strain B95-8 has been determined and is about 172kbp (Baer et al., 1984). The termini of the EBV genome consist of multiple tandem repeats (up to twelve copies) of a 500bp nucleotide sequence (Given and Kieff, 1978; Given et al., 1979; Kintner and Sugden, 1979) through which DNA circularisation is mediated in the infected cell nucleus (Lindahl et al., 1976). Another major set of up to twelve direct repeats, each of 3kbp, are located internally within the genome (Given and Kieff, 1979; Cheung and Kieff, 1981; Hayward et al., 1982). These internal repeats are not related to the terminal repeats (Cheung and Kieff, 1982; Jones and Griffin, 1983; Baer et al., 1984) and therefore, the unique tracts of EBV DNA do not invert (see Page 13). The terminal and major internal repeats separate the genome into large and small unique segments with lengths of 125kbp and 14kbp respectively (Baer et al., 1984). In addition, shorter repetitive regions are scattered throughout the genome (Baer et al., 1984). The HVP genome is structurally equivalent to that of EBV (Lee et al., 1980; Heller et al., 1981); both genomes appear to be colinear (Heller and Kieff, 1981; Lee et al., 1981) and exhibit approximately 40% homology (Falk et al., 1976).

4) Group D. Characteristically, members of this group have genomes comprising sets of long and short unique sequences, U_L and U_S, one of which is bracketed by terminal (TR) and

inverted internal (IR) repeat sequences. Included in this group are the genomes of PRV and EHV-1 which have sizes of approximately 130kbp (Darlington and Randall, 1963; Rubenstein and Kaplan, 1975) and G+C contents of 73% (Ben-Porat and Kaplan, 1962) and 56% (Darlington and Randall, 1963) respectively. The U_L segments of these genomes are not flanked by inverted repeats and are unable to invert (see Page 13). However, the U_S portions are bracketed by inverted repeats and do invert, generating two isomeric forms of both the PRV (Stevely, 1977; Ben-Porat et al., 1979) and EHV-1 genomes (Henry et al., 1981; Whalley et al., 1981). The termini of PRV DNA do not appear to be terminally redundant as sequence homology cannot be detected either by hybridisation or sequence analysis of cloned terminal DNA fragments (Ben-Porat and Kaplan, 1985).

The VZV genome closely resembles those of PRV and EHV-1, and consists of two unique segments, U_L and U_S, with the U_S region flanked by inverted repeats, TR_S and IR_S (Dumas et al., 1981; Ecker and Hyman, 1982; Straus et al., 1982). However, U_L is flanked by a short inverted repeat of 89bp which allows limited inversion of the L segment (Davison, 1984). Thus, a genomic population of VZV will contain two major forms of VZV DNA (95% of the genomic pool), resulting from free inversion of the S segment and two minor forms (5% of the genomic pool), resulting from limited inversion of the L segment (Davison, 1984).

5) Group E. HSV is a member of this group and its genome structure will be fully described on Page 13. The arrangement of unique and reiterated sequences in HCMV DNA resembles that of HSV (Kilpatrick and Huang, 1977; LaFemina and Hayward, 1980) and both genomes have other common structural features. HCMV DNA is 240kbp in length (Geelen et al., 1978; DeMarchi et al., 1978; Westrate et al., 1980), approximately 50% longer than the HSV genome, and has a mean G+C content of 59% (Crawford and Lee, 1964; Huang et al., 1973). In common with HSV DNA, there are four isomeric forms of HCMV DNA (LaFemina and Hayward, 1980; Westrate et al., 1980) and the genome circularises following brief digestion with a processive exonuclease (Geelen and Westrate, 1982), suggesting that the DNA ends are terminally redundant. This terminal redundancy is 600 to 750bp in length as suggested by the 'stepwise' heterogeneity of the repeat sequences (LaFemina and Hayward, 1980; Spector et al., 1982). DNA sequence analysis of HCMV AD169 (Tamashiro et al., 1984) and Towne strains (Spaete and Mocarski, 1985) has located the terminal redundancy to a fragment which contains dispersed repeat sequences, however, precise limits for the terminal redundancy have not been defined. These data indicate that the HCMV terminal redundancy has features similar to the HSV 'a' sequence (see Page 16). Multiple copies of a 200bp fragment adjacent to the HCMV terminal redundancy have also been identified (Tamashiro et al., 1984). The AD169, Towne and C87 strains of HCMV show approximately 80% DNA homology

(Huang et al., 1976), and there is inter-strain size heterogeneity of up to 8kbp which is associated primarily with the TR_L and IR_L repeats (Pritchett, 1980; LaFemina and Hayward, 1980).

2. Biology and Pathogenesis of Herpesviruses.

Herpesviruses are known to infect a wide range of eukaryotic organisms from fungi (Kazama and Schornstein, 1972) to man. A total of five herpesviruses have been isolated from humans, namely HSV-1, HSV-2, VZV, HCMV and EBV.

HSV-1 is commonly associated with facial lesions, primarily around the mouth (Fiddian et al., 1983) but the virus can infect many organs. Primary infection may be asymptomatic or lead to development of a range of clinical symptoms including fever, sore throat, ulcerative and vesicular lesions, oedema, localised lymphadenopathy and malaise (Whitley, 1985). Following primary infection, a latent state is normally established with the virus residing in the trigeminal ganglia (Bastian et al., 1972). Analysis of the banding patterns of viral DNAs cleaved with restriction endonucleases has indicated that viruses isolated from recurrent and primary lesions are identical (Lonsdale et al., 1979). Therefore, most recurrent outbreaks are due to reactivation of latent virus rather than to secondary infection.

The majority of genital HSV infections are caused by HSV-2 (Kessler, 1977), however a variable number are

attributable to HSV-1 (Kalinyak et al., 1977; Wolontis and Jeansson, 1977; Peutherer et al., 1981). In one study, HSV-1 was identified in half of the isolates from female genitalia (Chaney et al., 1983). HSV-2 genital infections can recur due to reactivation of latent virus from the sacral ganglia (Baringer, 1974); the number of HSV-2 genital infections attributed to reactivation as compared to primary infection varies between 43% to 75% (Rawls et al., 1971; Josey et al., 1972; Smith et al., 1981). HSV-2 has also been associated with cervical carcinoma, details of which are described in Section C (Page 45).

VZV is the causative agent of chickenpox (varicella) usually in childhood, and of shingles (herpes zoster; Weller, 1976) which is normally found in adults. In common with other alphaherpesvirinae, VZV can establish a latent state and reactivation of the virus is at least partly dependent upon the immunological status of the infected host (Rifkind, 1966). VZV is difficult to propagate in tissue culture (Weller et al., 1958; Gilden et al., 1978; Grose et al., 1979), however, it does have the ability to transform cells in vitro (Gelb et al., 1980).

Infection by HCMV induces a number of syndromes, in particular intrauterine death and congenital defects (Weller, 1971; Rapp, 1980) and is associated with Kaposi's sarcoma (Giraldo et al., 1975). Reactivation of the virus is usually associated with immuno-suppression due to a variety of situations such as malignant disease and chemotherapy (Weller, 1971; Plummer, 1973; Gold and Nankervis, 1976).

EBV infects specifically B-lymphocytes *in vitro* and is the causative agent of infectious mononucleosis (glandular fever; Evans *et al*., 1968; Henle *et al*., 1968; Evans and Niederman, 1976). It has also been associated with the incidence of two cancers, Burkitt's lymphoma (Epstein *et al*., 1964; de The *et al*., 1978) and nasopharyngeal carcinoma (zur Hausen *et al*., 1970; Nonayama and Pagano, 1973).

3. Structure of the HSV Genome.

The relationship between the HSV-1 and HSV-2 genomes has been studied by both intertypic recombination and intertypic complementation data (Timbury and Subak-Sharpe, 1973; Esparza *et al*., 1976; Marsden *et al*., 1978; Schaffer *et al*., 1978) along with DNA/DNA hybridisation studies (Davison and Wilkie, 1983a). Results indicate that the HSV-1 and HSV-2 genomes are essentially colinear with viral genes mapping at equivalent positions on both genomes. Most studies have focussed on the HSV-1 genome arrangement, however, the information available for HSV-2 DNA will be presented where appropriate.

a) Chemical composition.

HSV-1 DNA is a linear duplex molecule (Ben-Porat and Kaplan, 1962; Russell, 1962; Russell and Crawford, 1963 and 1964) which is approximately 155kbp in length (Becker *et al*., 1968; Kieff *et al*., 1971; Wilkie, 1973 and 1976; Clements *et al*., 1976) and has a mean G+C content of 67%

(Kieff et al., 1971). The G+C content of HSV-2 DNA is slightly higher at 69% (Goodheart et al., 1968; Halliburton, 1972) and the HSV-2 short genome segment is 3kbp longer than that of HSV-1 (Cortini and Wilkie, 1978). These additional HSV-2 nucleotide sequences have been located to an area of the short unique segment of HSV-2 DNA (for a description of the HSV genome arrangement, see Page 13); the equivalent region of the HSV-1 genome encodes U_S gene-4 (Fig. 3; Rixon and McGeoch, 1985) and most of the additional HSV-2 sequences probably contribute to the coding region of a g92K glycoprotein (see Page 40; D.J. McGeoch, personal communication).

b) Alkaline lability.

Both HSV-1 and HSV-2 DNAs are fragmented by treatment with alkali (Kieff et al., 1971; Wilkie, 1973), suggesting that the HSV genome may contain single-stranded nicks (Ecker and Hyman, 1981) or alternatively, that ribonucleotides are present in the genome (Gordin et al., 1973; Muller et al., 1979). The sedimentation profile in neutral sucrose solutions is similar for DNA denatured either by formamide or alkali, indicating that the presence of ribonucleotides alone could not account for fragmentation of the DNA (Spear and Roizman, 1980). It has been suggested that alkali-sensitive regions are positioned at specific sites (Frenkel and Roizman, 1972), however, other studies have shown them to be distributed randomly (Wilkie, 1973; Ecker and Hyman, 1981).

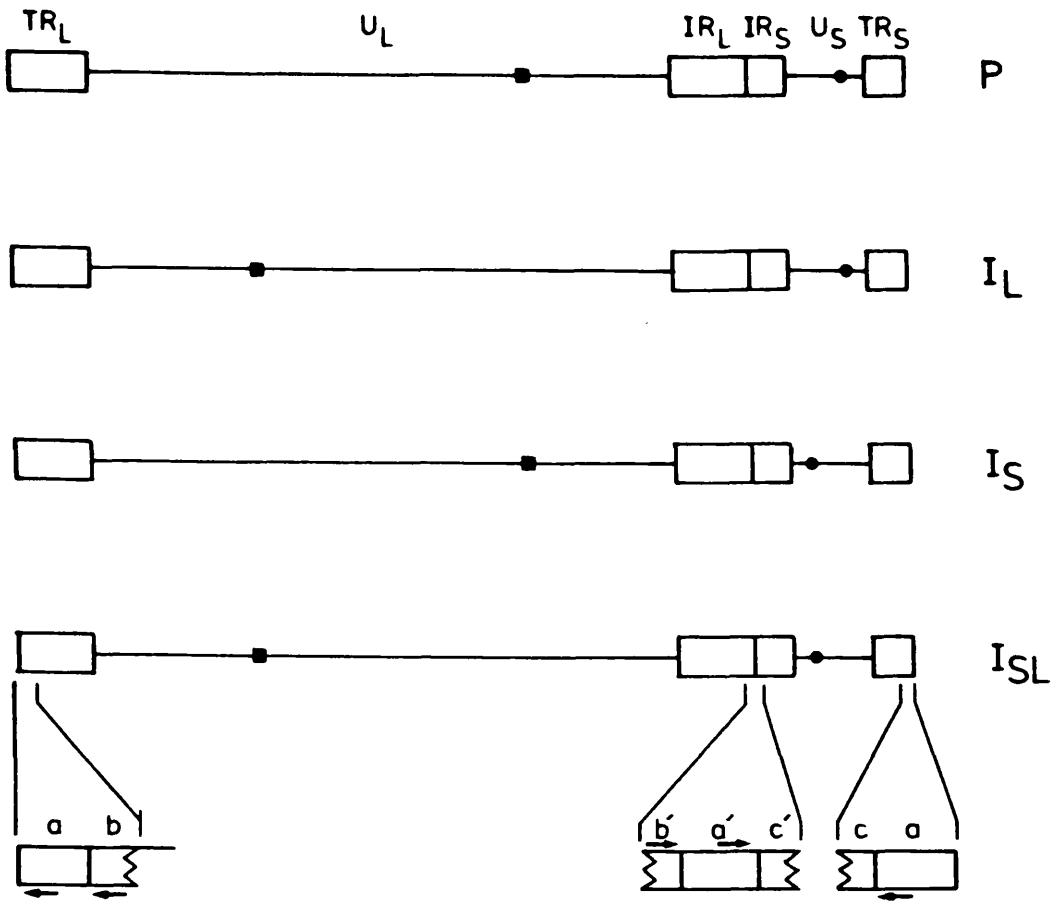


Figure 2. Four possible isomeric forms of the HSV genome generated by inversion of the long and short unique segments (U_L and U_S). The terminal and internal repeats of L and S segments are represented by TR_L/TR_S and by IR_L/IR_S. The distance between markers in U_L (■) and U_S (●) components is dependent upon the relative orientations of unique regions. Sequence arrangements at the termini and joint regions are shown below the various isomeric forms. The letter, a, represents the terminal redundancy and b and c denote sequences flanking the terminal redundancy in the long repeat (R_L) and short repeat (R_S) respectively. Primed letters at the joint indicate that sequences are inverted relative to those at the termini. Arrows show the location, orientation and number of copies of direct repeat, DRI.

c) Genome arrangement.

Electron microscopy studies of structures produced by intra-molecular hybridisation of single-stranded DNA (Sheldrick and Berthelot, 1974) have shown that the HSV genome consists of long and short segments termed L and S respectively; the region at which these segments are linked is termed the joint. The L and S segments are composed of unique regions (U_L and U_S) flanked by inverted repeat sequences which are located internally (IR_L and IR_S) and terminally (TR_L and TR_S). Sheldrick and Berthelot (1974) further suggested that either intermolecular or intramolecular recombination occurring within repetitive sequences flanking the U_L and U_S segments could result in the inversion of unique DNA regions between participating repetitive regions thus generating four possible isomeric forms of HSV DNA (Fig. 2). These isomeric forms are termed P (prototype), I_S, I_L and I_{SL} (Roizman et al., 1979), each of which is likely to produce viable progeny. Unless stated otherwise, diagrams in this thesis will depict only the P genome arrangement. Partial denaturation mapping confirmed that four isomeric forms do exist in populations of HSV DNA molecules and in approximately equal amounts (Hayward et al., 1975; Delius and Clements, 1976). Additional evidence for inversion of L and S segments was derived from restriction endonuclease analysis of HSV-1 and HSV-2 DNA which showed that the relative molarities of terminal and junction fragments were 0.5 and 0.25 respectively with respect to fragments derived from other regions of the

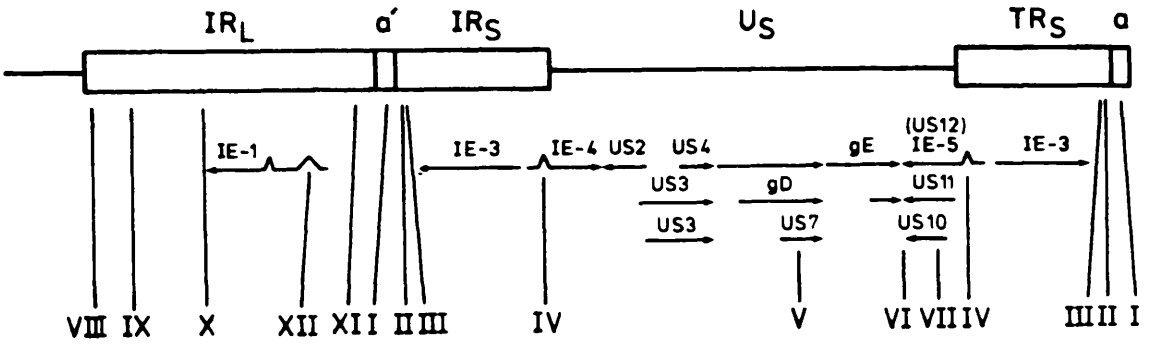


Figure 3. Locations of short tandemly reiterated sequences within the S segment and R_L regions of HSV-1 strain 17^+ DNA. Positions of reiterated sequences (numbered I to XII) are shown relative to the locations of mRNAs mapping in this region of the HSV-1 genome. Transcripts referred to in the text are labelled. The locations of spliced regions in immediate early (IE) genes-1, -4 and -5 are shown by (∧). gD and gE are the abbreviations for glycoproteins D and E. mRNA mapping data were taken from the following references: transcripts other than IE mRNAs, Rixon and McGeoch, 1984 and 1985; IE mRNA-1, F.J. Rixon, personal communication; IE mRNA-3, Rixon et al., 1982; IE mRNAs-4 and -5, Watson et al., 1981a; Rixon and Clements, 1982. Positions for the reiterated sequences are taken from Rixon et al., 1984 apart from reiterations VIII, IX and XI whose locations were determined by L. Perry (personal communication).

genome (Hayward et al., 1975; Clements et al., 1976; Wilkie and Cortini, 1976; Cortini and Wilkie, 1978).

d) Locations of short tandem reiterations.

In addition to large repeat units, the HSV-1 and HSV-2 genomes contain shorter segments of tandemly reiterated sequences which vary in length between 5bp (Whitton and Clements, 1984b) and 54bp (L. Perry, personal communication) and generally have high G+C contents (Rixon et al., 1984). The locations of these repeats are shown in Fig. 3. Tandem repeat sequences have been identified in the translated regions of U_S genes-7, -10 and -11 (Rixon and McGeoch, 1984; McGeoch et al., 1985), the introns of immediate early (IE) mRNAs-1, -4 and -5 (Watson et al., 1981a, b; Murchie and McGeoch, 1982; Rixon and Clements, 1982; L. Perry and F.J. Rixon, personal communication), the terminal redundancy (Davison and Wilkie, 1981; Mocarski and Roizman, 1981) and downstream from mRNA 3' termini (Davison and Wilkie, 1981; Rixon et al., 1984; McGeoch et al., 1985); additional reiterated sequences have been located within the long repeat (R_L) segments, however, their positions with respect to HSV genes have not been rigorously determined (Fig. 3; L. Perry, personal communication). The copy number of repeated sequences varies between virus isolates (Davison and Wilkie, 1981; Watson et al., 1981b; Murchie and McGeoch, 1982) and, comparing HSV-1 and HSV-2 genomes, the nucleotide sequences of reiterations at equivalent positions are

either poorly conserved (reiteration IV in Fig. 3; Whitton and Clements, 1984b) or absent in the HSV-2 genome (reiterations I and VI in Fig. 3; Davison and Wilkie, 1981; Whitton, 1984).

IE mRNAs-1 and -3 are located entirely within the R_L and R_S segments respectively and thus are diploid genes (Fig. 4; Rixon and Clements, 1982 and 1984). Reiterations flanking these genes may serve to increase the level of genetic exchange between the repeats, thus mutations arising in one copy of the large repeat regions would spread to both copies of the repeats (Rixon et al., 1984). By this mechanism, the probability of accumulating recessive mutations within repeat regions would be reduced as deleterious mutations would generate a non-viable product. This model for the role of reiterations is based on two assumptions (Rixon et al., 1984):

1) reiterations are located in non-essential regions of the genome; recent evidence indicates that removal of sequences between the terminal redundancy and the 3' terminus of IE mRNA-3, a region which contains repetitive elements (reiterations II and III in Fig. 3), does not reduce the viability of virus progeny (Hubenthal-Voss and Roizman, 1985).

2) reiterations promote a high frequency of recombination; in support of this suggestion, one set of repeat sequences beyond the 3' terminus of IE mRNA-3 (reiteration II in Fig. 3) has homology with sequences at immunoglobulin (Ig) class-switch recombination sites (Gomez-Marquez et al., 1985).

e) The 'a' sequence.

Treatment of HSV-1 genomic termini with 5'- or 3'-exonucleases followed by renaturation generates circular DNA molecules of genomic length, suggesting that the genome is terminally redundant (Grafstrom et al., 1974 and 1975; Sheldrick and Berthelot, 1974). This terminal redundancy, termed the 'a' sequence, is present in the direct orientation at the termini and in the inverse orientation at the junction between IR_L and IR_S (Fig. 2; Grafstrom et al., 1974 and 1975; Wadsworth et al., 1976; Wagner and Summers, 1978). The regions which bound the 'a' sequence in TR_L/IR_L and TR_S/IR_S are termed the 'b' and 'c' sequences respectively (Fig. 2; Davison and Wilkie, 1981). Nucleotide sequences for the 'a' sequences of four HSV-1 strains (Glasgow strain 17^+ and USA-8, Davison and Wilkie, 1981; strain F, Mocarski and Roizman, 1981; strain KOS, Varmuza and Smiley, 1985) and HSV-2 strain HG52 (Davison and Wilkie, 1981) have been determined. These data indicate firstly, that a direct repeat (DRI) brackets the 'a' sequence at the joint and secondly that the HSV-1 'a' sequence varies in size from 250bp to 500bp due to a variable number of tandemly reiterated sequences (DRII; reiteration I in Fig. 3). Nucleotide sequences of the various 'a' sequences are not identical for different HSV-1 strains, with the region of lowest homology occurring in the DRII reiterations. The HSV-2 'a' sequence does not contain a reiteration and has a constant size of 251bp. Most HSV molecules have a single 'a' sequence at both termini and joint regions, however

some molecules possess two or more 'a' sequences present as tandem direct repeats at the joint and L terminus with a single copy at the S terminus (Wagner and Summers, 1978; Davison and Wilkie, 1981; Mocarski and Roizman, 1982a).

The 'a' sequence is thought to play a role in two processes: 1) inversion of the L and S segments and 2) DNA maturation and encapsidation.

1) Genome inversion. The 'a' sequence appears to contain a cis-acting signal which functions in inversion of the L and S genome segments (Mocarski et al., 1980; Mocarski and Roizman, 1981 and 1982a; Smiley et al., 1981). Insertion of HSV-1 DNA fragments from either the L or S termini into the HSV-1 thymidine kinase (TK) locus within U_L resulted in formation of twelve isomers, the predicted number from free inversion about the various 'a' sequences (Mocarski and Roizman, 1982a). Moreover, inversion requires DNA fragments to be flanked by inverted 'a' sequences; DNA fragments bounded by 'a' sequences in the same orientation do not invert (Mocarski et al., 1980; Mocarski and Roizman, 1981; Smiley et al., 1981). Studies with intertypic HSV-1/HSV-2 recombinants have further indicated that inversion requires a DNA segment to be flanked by homotypic 'a' sequences. Recombinant RE4, which possesses HSV-1 'a' sequences at the termini and an inverted HSV-2 'a' sequence at the joint region, fails to invert normally and RE4 DNA is present predominantly in the I_S arrangement (Davison and Wilkie, 1983b). Removal of most of DRI does not prevent inversion

(Mocarski and Roizman, 1982a) and further deletion studies indicate that tandem, direct repeats within the 'a' sequence may constitute the sites through which inversion is mediated (Chou and Roizman, 1985).

Inversion could occur either by a recombination mechanism or by circularisation of the genome (see Page 19) and subsequent recleavage. The presence of two complete directly repeated 'a' sequences at the joint region suggests that circular DNA would be formed by direct ligation of the termini rather than by processive exonuclease digestion and ligation (Davison and Wilkie, 1981).

2) DNA maturation and encapsidation. DNA replication is thought to occur by a rolling circle mechanism (see Page 23; Jacob et al., 1979), generating concatemers of DNA which are cleaved to produce unit-length molecules. Viral DNA termini are then generated by two separate cleavage events which occur at a distance from cleavage signals located within the 'a' sequence (Varmuza and Smiley, 1985). The cleavage sites are located within the direct repeat, DRI, such that essentially the L terminus contains two copies of DRI with one copy at the S terminus (Fig. 2; Davison and Wilkie, 1981; Mocarski and Roizman, 1982a).

DNA maturation appears to be tightly coupled to encapsidation as indicated from studies of PRV and HSV-1 mutants (Ladin et al., 1980; V.G. Preston and N.D. Stow, personal communication). Plasmids containing an HSV-1

origin of replication and fragments from either genome terminus are amplified and packaged as concatemers into capsids (Stow et al., 1983). The 'a' sequence is the only DNA common to both termini suggesting that the signals required for encapsidation are contained within this sequence. The nucleotide sequences which direct DNA maturation and encapsidation have not been separated by deletion analysis of the 'a' sequence (N.D. Stow, personal communication). Studies with defective HSV genomes have shown that only capsids containing DNA of approximately the same size as unit-length non-defective molecules are transported to the cytoplasm (Vlazny et al., 1982).

f) Interaction of proteins with the 'a' sequence.

The functions of the 'a' sequence presumably require the interaction of specific proteins. For example, genomic inversion requires at least one trans-acting viral component (Mocarski and Roizman, 1982b) and circularisation, which occurs soon after infection in the absence of de novo protein synthesis (Poffenberger and Roizman, 1985), may be mediated either by a host enzyme or by a viral protein in the inoculum. Moreover, a protein which may be bound close to one or both HSV DNA termini inhibits digestion from the 3' ends of duplex DNA (Kudler and Hyman, 1979). Electron microscopy studies have suggested that protein is bound close to the joint region and the termini (Wu et al., 1979) and therefore possibly to the 'a' sequence. Two virus-induced late polypeptides with

mol. wts. of 21,000 and 22,000, which may correspond to the 20,000 mol. wt. DNA binding protein identified by Hyman (1980), interact specifically with the HSV-1 'a' sequence (Dalziel and Marsden, 1984). One of these polypeptides appears to be encoded by U_S mRNA-11 (Fig. 3) which contains a reiterated amino acid triplet X-Pro-Arg (X = any amino acid) in the central portion of the protein coding region (Rixon and McGeoch, 1984).

SECTION B.

PROCESSES INVOLVED IN PRODUCTIVE HSV INFECTIONS.

4. Effects of HSV Infection on Host Cell Metabolism.

Productive infection with either HSV-2 or HSV-1 results in alterations in the synthesis and metabolism of cellular macromolecules. Following infection, cellular DNA synthesis ceases (Aurelian and Roizman, 1964; Roizman and Roane, 1964) and disaggregation of host polyribosomes results in decreased polypeptide synthesis (Sydiskis and Roizman, 1967 and 1968); the polyribosomes which reform contain predominantly virus-encoded mRNAs (Stringer et al., 1977). A decrease in the steady-state levels of host RNAs also occurs, however, the degree of inhibition is less marked than that observed with DNA and protein synthesis (Hay et al., 1966; Flanagan, 1967). In contrast to the general reduction in levels of RNA and protein synthesis, production of heat shock proteins is stimulated in HSV-1- (Notarianni and Preston, 1982) and HSV-2-infected cells (LaThangue et al., 1984). Moreover, certain host RNAs continue to be synthesised at late times post infection (p.i.; Stringer et al., 1977) and the promoters of genes integrated into the genomes of biochemically transformed cells are activated on virus infection (Everett, 1985).

Analysis of the steady-state levels of individual host cell mRNAs indicates that the abundance of polyadenylated species is greatly reduced by 4h p.i.

(Nishioka and Silverstein, 1977 and 1978; Inglis, 1982; Nakai et al., 1982; Mayman and Nishioka, 1985) and it has been suggested that a decrease in host mRNA stability is due to deadenylation (Nakai et al., 1982; Mayman and Nishioka, 1985). In support of this proposal, the steady-state level of histone H3 mRNA, a non-polyadenylated transcript, does not decrease in Friend erythroleukaemia cells at early times p.i. (Mayman and Nishioka, 1985), however, infection of Vero cells induced similar rates of reduction of H3 and H4 mRNA levels as that observed with actin poly A mRNA (Schek and Bachenheimer, 1985).

The effects of HSV infection on cellular macromolecular synthesis appear to be mediated by a virus-encoded 'shut-off' function (Halliburton and Timbury, 1973; Yamanishi et al., 1975) and 'shut-off' is more rapid following infection by HSV-2 than by HSV-1 (Powell and Courtney, 1975; Pereira et al., 1977; Fenwick et al., 1979; Schek and Bachenheimer, 1985). A 'shut-off' function has been mapped between 0.52 and 0.59 map units on the HSV-2 genome (Morse et al., 1978; Fenwick et al., 1979), however, it is possible that more than one 'shut-off' function is associated with HSV infection. Inhibition of host RNA and protein synthesis appears to be induced by a virion function for both HSV-2 (Fenwick and Walker, 1978) and HSV-1 (Nishioka and Silverstein, 1977 and 1978; Read and Frenkel, 1983; Schek and Bachenheimer, 1985) while, in certain cell lines, viral gene expression is required for host RNA degradation (Nishioka and Silverstein, 1978; Stenberg and Pizer, 1982; Nishioka et al., 1983).

5. Virus DNA Replication.

Under one-step growth conditions, HSV DNA replication initiates some 2h p.i. and increases in rate for the following 8h (Munk and Sauer, 1964; Wilkie, 1973). A small fraction of the input HSV DNA is utilized for replication (Jacob and Roizman, 1977) but this is not the general situation for all herpesviruses as the majority of input DNA is replicated following PRV infection (Jean and Ben-Porat, 1976).

The mode of DNA replication is unclear, however, it has been suggested that HSV DNA circularises after infection and thereafter replication proceeds by a 'rolling circle' mechanism (Jacob et al., 1979). Supportive evidence for this model has been the detection of circular unit-length HSV DNA molecules by electron microscopy (Shlomai et al., 1976; Friedmann et al., 1977; Hirsch et al., 1977; Jacob and Roizman, 1977). Furthermore, restriction endonuclease data has demonstrated that replicative DNA contains fewer terminal fragments than genomic DNA, indicative of concatemeric or circular DNA forms (Jacob et al., 1979; Jongeneel and Bachenheimer, 1981).

The approximate locations of HSV DNA origins of replication have been deduced by characterisation of two classes of defective genomes which accumulate in virus stocks following serial passage at high multiplicities of infection. Class I defectives contain a replication origin from TR_S and IR_S and class II defectives possess an origin

from U_L (Vlazny and Frenkel, 1981; Spaete and Frenkel, 1982), thus, the HSV genome contains three potential origins of replication.

The accurate location of the IR_S/TR_S origin has been mapped by measuring levels of plasmid DNA amplification following deletion of sequences from cloned HSV-1 fragments (Stow, 1982; Stow and McMonagle, 1983). From these data, nucleotide sequences essential for DNA replication have been located to a 90bp segment which has a 45bp palindromic sequence containing 18 centrally-located A or T residues flanked by G+C-rich sequences (Murchie and McGeoch, 1982; Watson and Vande Woude, 1982; Stow and McMonagle, 1983). This 90bp segment is located in a non-transcribed region between the 5' termini of HSV-1 IE mRNAs-3 and -4/-5 (Fig. 4; Rixon and Clements, 1982; Rixon et al., 1982). The equivalent HSV-2 region contains two copies of a DNA sequence which is almost identical to the 45bp palindrome in HSV-1 (Whitton and Clements, 1984a). Moreover, HSV-2 DNA fragments spanning U_S and TR_S/IR_S which contain these sequences are amplified by HSV-1 superinfection (J.L. Whitton and N.D. Stow, personal communication).

The U_L origin has been located between the 5' termini of mRNAs encoding the DNA polymerase and major DNA binding protein (see Fig. 5; Holland et al., 1984; Quinn, 1984) and the DNA sequences within this region have been determined for two HSV-1 wild-type (wt) strains (Quinn, 1984; Weller et al., 1985). A portion of the DNA sequence

contains a 72bp palindrome which has homology to sequences comprising the IR_S/TR_S origin of replication; a defective of HSV-1 Angelotti strain contains two copies of this palindromic sequence (Gray and Kaerner, 1984). Cloned DNA from this region contains deletions of variable size which arise spontaneously (Spaete and Frenkel, 1982; Weller et al., 1983; Quinn, 1984) and plasmid constructions with such deletions are unable to support DNA replication (Weller et al., 1985).

Following replication, HSV DNA is cleaved and packaged into capsids; the signals involved in these events are described on Page 18.

6. HSV mRNA Biosynthesis.

Following transport to the nucleus (Hummeler et al., 1969), mRNA is synthesised from viral DNA by RNA polymerase II as indicated by the sensitivity of transcription to α-amanitin (Alwine et al., 1974; Ben-Zeev et al., 1976); this sensitivity continues throughout all stages of virus infection (Costanzo et al., 1977). Failure to detect a novel RNA polymerase activity in infected cells further implies that HSV transcription is mediated by cellular RNA polymerase II (Lowe, 1978). Moreover, HSV mRNA synthesis is maintained in α-amanitin treated cells possessing an endogenous α-amanitin-resistant RNA polymerase (Ben-Zeev and Becker, 1977; Costanzo et al., 1977). However, it is possible that the cellular polymerase is modified during the course of virus infection.

HSV transcription is temporally regulated and can be sub-divided into three broad classes, IE, early and late (Swanstrom and Wagner, 1974; Clements et al., 1977; Jones and Roizman, 1979).

a) Immediate early transcription.

IE transcripts are operationally defined as those virus mRNAs which are transcribed by an unmodified cellular RNA polymerase II in the absence of de novo protein synthesis (Kozak and Roizman, 1974; Roizman et al., 1974). The locations of IE transcripts have been mapped using cycloheximide, a protein synthesis inhibitor, (Clements et al., 1977 and 1979; Anderson et al., 1980; Easton and Clements, 1980) and also temperature sensitive (ts) mutants (Watson and Clements, 1978). Hybridisation of in vivo-labelled RNA to Southern blots (Southern, 1975) containing HSV-1 DNA fragments generated by restriction enzymes has shown that IE transcription is limited to restricted regions of the genome (Clements et al., 1977; Anderson et al., 1980). The mapping and orientation of these transcripts on the HSV-1 genome has revealed that there are five IE mRNAs (Watson et al., 1979; Clements et al., 1979; Anderson et al., 1980), two of which, IE mRNAs-1 and -3, are contained entirely within the TR_L/IR_L and TR_S/IR_S segments respectively and are therefore diploid genes (Fig. 4; see Page 15). The 5' termini of IE mRNAs-4 and -5 are located within TR_S/IR_S, thus these mRNAs have common promoter and 5' untranslated leader sequences (Fig.

4; Watson et al., 1981a; Rixon and Clements, 1982). From structural analysis using nuclease digestion techniques (Berk and Sharp, 1978a, b), three IE transcripts have been shown to be spliced (Page 34; Watson et al., 1981a; Rixon and Clements, 1982; F.J. Rixon, personal communication). The polypeptides encoded by the HSV-1 IE mRNAs have been identified by in vitro translation of purified mRNA species (Watson et al., 1979). HSV-2 also specifies five IE mRNAs which map at equivalent genomic locations and encode similar-sized polypeptides to their HSV-1 counterparts (Easton and Clements, 1980). Following the switch to early mRNA synthesis, the abundance of IE polypeptides decreases (Honess and Roizman, 1974; Marsden et al., 1976; Preston, 1979; Dixon and Schaffer, 1980); IE (Dixon and Schaffer, 1980), early (Honess and Roizman, 1974) and late products (DeLuca et al., 1984) have been implicated in this negative regulation of IE polypeptide synthesis. However, the cytoplasmic levels of IE mRNAs do not decrease at late times p.i. (see Page 38; Whitton, 1984; Harris-Hamilton and Bachenheimer, 1985; F.J. Rixon, personal communication).

A great deal of attention has focussed on the sequences which activate IE genes and the factors mediating this activation. Studies with the HSV-1 TK gene show that this early viral gene is regulated as an IE product following fusion of IE promoter sequences to the TK structural region (Post et al., 1981; Mackem and Roizman, 1982a; Cordingley et al., 1983). Linkage of far upstream regions of IE promoters to a truncated TK promoter

similarly confers an IE pattern of regulation to TK gene expression (Mackem and Roizman, 1982a; Cordingley et al., 1983), demonstrating that IE promoter and regulatory sequences are separable. From deletion mutant studies (Cordingley et al., 1983; Preston et al., 1984a) and DNA sequence comparisons of regions upstream from IE mRNA 5' termini (Mackem and Roizman, 1982b; Whitton et al., 1983; Whitton and Clements, 1984a), an 'activator' sequence with the core consensus TAATGARAT (R = purine) has been derived which is present as one or more copies between -100 and -300 upstream from IE mRNA cap sites. A fragment upstream from IE mRNA-3 containing this sequence possesses enhancer-like properties (Lang et al., 1984), however, the role of TAATGARAT motifs in this enhancer activity is unclear. The equivalent HSV-2 IE genes also contain activator elements which have an identical consensus sequence to that of HSV-1, however, HSV-1 and HSV-2 early and late promoters lack the TAATGARAT sequence as does the 25kb nucleotide sequence of the HSV-1 U_S region (Whitton and Clements, 1984a; McGeoch et al., 1985). Further evidence for the similarity between the HSV-1 and HSV-2 activator sequences is demonstrated by the ability to stimulate HSV-2 IE gene expression by HSV-1 infection and the converse also is true (Cordingley et al., 1983; Gaffney et al., 1985). It has been suggested that the sequence TAATGARAT alone is not sufficient for IE gene activation and that G+C-rich motifs flanking this activator element are required for full induction (Kristie and Roizman, 1984;

Preston et al., 1984a). Stimulation of chloramphenicol acetyltransferase (CAT) enzyme activity under IE conditions has been demonstrated from plasmids containing a synthetic activator sequence but lacking G+C-rich flanking regions when linked to an IE promoter in which TAATGARAT homologues are absent (Gaffney et al., 1985).

The trans-activating component which induces IE gene expression is present in the virus inoculum (Post et al., 1981; Mackem and Roizman, 1982a, b; Cordingley et al., 1983; Preston et al., 1984a) and has been identified as a tegument protein with a mol. wt. of 65,000 (Campbell et al., 1984). This protein has been mapped to a sub-fragment of HSV-1 Bam HI f (Campbell et al., 1984) and is possibly specified by a 1.9kb mRNA (Fig. 4; Hall et al., 1982). The nucleotide sequence specifying the trans-activating component has been determined (Dalrymple et al., 1985; Pellett et al., 1985b) and the predicted amino acid sequence of this protein shares homology with a proposed VZV 40,000 mol. wt. polypeptide (Dalrymple et al., 1985).

b) Early transcription.

Early mRNAs are produced in the presence of IE gene products prior to the onset of DNA replication (Wagner, 1972; Swanstrom and Wagner, 1974; Swanstrom et al., 1975; Clements et al., 1977). Early transcripts are readily detectable at 2h p.i. and increase in abundance until 6h to 8h p.i., after which the levels of individual early mRNAs either decrease or remain unaltered (McLauchlan and

Clements, 1982; Sharp et al., 1983; Harris-Hamilton and Bachenheimer, 1985). Inhibition of DNA replication results in an accumulation of early species within the cytoplasm (Swanstrom and Wagner, 1974; Swanstrom et al., 1975, Stringer et al., 1977) and an over-representation of individual transcripts (Harris-Hamilton and Bachenheimer, 1985). Early mRNAs map across the entire length of the genome (Clements et al., 1977) and a number of their protein products are involved in priming the infected cell for DNA replication (reviewed in Wagner, 1985).

At least two IE polypeptides are essential for maximal synthesis of early messages (Everett, 1984b; O'Hare and Hayward, 1985; Quinlan and Knipe, 1985). The presence of IE polypeptide Vmw175 is required throughout infection to maintain early transcription (Watson and Clements, 1978; Preston, 1979; Watson and Clements, 1980) and cell lines which constitutively produce this protein stimulate expression of early genes (Davidson and Stow, 1985; Persson et al., 1985). Activation in trans of plasmid-borne HSV early promoters by IE gene products, synthesised from co-transfected plasmids, has indicated that full activity of the glycoprotein D (gD) promoter is mediated by a combination of the IE Vmw175 and IE Vmw110 polypeptides (Everett, 1984b). Individually, IE Vmw175 increases transcription from the gD promoter poorly and IE Vmw110 fails to activate (Everett, 1984b). However, the promoter for the HSV-2 1.2kb mRNA which encodes Vmw38 (see Page 59) is stimulated by either IE Vmw175 or IE Vmw110 and maximal

expression from this promoter requires IE Vmw12 in addition to IE Vmw175 and IE Vmw110 (O'Hare and Hayward, 1985). Vmw110 also stimulates transcription from plasmids carrying the promoters for the TK (Gelman and Silverstein, 1985; O'Hare and Hayward, 1985) and major DNA binding protein genes (Quinlan and Knipe, 1985). Thus, the various early mRNA promoters may not have identical IE polypeptide requirements for stimulation.

Deletion mutant studies on the HSV-1 gD promoter linked to rabbit β-globin coding sequences have revealed that the DNA sequences required for full activity of this promoter, following infection by either HSV-1 or HSV-2, lie within 83bp of the mRNA 5' terminus (Everett, 1983). Further deletions within this 83bp region failed to detect regulatory sequences specific for early mRNA synthesis (Everett, 1984a). Moreover, virus infection stimulates transcription from a plasmid-borne rabbit β-globin promoter to similar levels as the gD promoter (Everett, 1983). Additional evidence for the lack of HSV-specific early regulatory sequences is that the gD promoter is activated by a number of herpesviruses and by adenovirus type 2 (Ad2; Everett and Dunlop, 1984) which synthesises proteins functionally equivalent to HSV-1 IE Vmw175 (Feldman et al., 1982). Deletion analysis of the HSV-1 TK promoter has revealed three regions required for constitutive mRNA synthesis (McKnight et al., 1981; McKnight and Kingsbury, 1982; McKnight et al., 1984). Two of these regions, distal signals II and III, consist of

G+C-rich tracts which are recognised by the cellular factor, Sp1 (see Page 63; Jones et al., 1985). ElKareh et al. (1985) have suggested that sequences which respond to trans-activation by viral polypeptides lie between -39 and -109 from the TK mRNA 5' terminus, however, an alternative interpretation of similar data indicates that early-specific regulatory signals do not exist in the TK promoter (Eisenberg et al., 1985).

c) Late transcription.

HSV-1 and HSV-2 late mRNAs map across the entire length of the genome (Clements et al., 1977; Easton and Clements, 1980) and a number of their encoded proteins are structural virion polypeptides. Late transcripts have been operationally divided into two classes, the early-late and true lates. Early-late mRNAs are synthesised at low levels prior to DNA replication while true late transcripts are detected only after the onset of DNA replication (Holland et al., 1979 and 1980; Jones and Roizman, 1979). Both classes of mRNA increase in abundance following the onset of DNA replication and inhibitors of DNA replication such as phosphonoacetic acid (PAA) and cytosine arabinoside (ara C) prevent full expression of late genes (Clements et al., 1977; Jones and Roizman, 1979; Holland et al., 1980). Individual late transcripts are either barely detectable or undetectable at early times p.i., however, they increase in abundance until about 10h p.i. and persist in the cytoplasm at a high level for up to 14h p.i. (Harris-Hamilton and Bachenheimer, 1985).

It is possible that late transcription is not directly coupled to DNA replication. Viral genes, under the control of late promoters which have been integrated into the cellular genome, are stimulated by HSV infection (Sandri-Goldin et al., 1983; Dennis and Smiley, 1984; Silver and Roizman, 1985), thus amplification of DNA sequences does not appear to be necessary for promoter activity. Moreover, plasmid-borne HSV late promoters do not require a cis-acting origin of DNA replication in order to initiate mRNA synthesis (P. Johnson, personal communication). Two viral products have been suggested as having a direct role in late gene expression. Firstly, mutations within the IE Vmw63 protein do not inhibit DNA replication but late polypeptide synthesis is significantly reduced in cells infected with these mutants (Sacks et al., 1985). Secondly, prior to DNA replication, the major DNA binding protein appears to repress synthesis of the late gene encoding gC; the mechanism of this repression is unclear (Godowski and Knipe, 1985).

7. Post-transcriptional Modification of HSV mRNAs.

a) Capping and methylation.

In common with most other eukaryotic mRNAs, HSV transcripts are capped at their 5' termini (Bartkoski and Roizman, 1976; Moss et al., 1977). Cap structures contain a methyl-7-guanosine (m^7G) residue which is attached to the 5' terminal base through a 5'-5' linkage (see Page 60);

major 5' terminal structures detected in HSV mRNAs are m^7GpppG^mpNp, $m^7Gpppm^6A^mpNp$ and m^7GpppA^mpNp (Moss et al., 1977). In addition, HSV transcripts can be modified by methylation of internal A residues (Moss et al., 1977), however, viral mRNAs labelled at late times during infection are not methylated, possibly as a result of inhibition by a virus-specified product (Bartkoski and Roizman, 1976).

b) Splicing.

Three HSV IE mRNAs, some minor mRNA species of the HSV-1 gC gene and a late 2.7kb mRNA are spliced. HSV-1/HSV-2 IE mRNAs-4 and -5 have a single splice within their common transcribed regions located within IR_S/TR_S (Fig. 3; Watson et al., 1981a; Rixon and Clements, 1982; Whitton and Clements, 1984b); this splice is 5' to the polypeptide coding region (Murchie and McGeoch, 1982; Watson and Vande Woude, 1982; Whitton and Clements, 1984b), however, minor unspliced species have been detected in HSV-1 (Rixon and Clements, 1982). The HSV-1 intron DNA sequence is approximately 200bp in length and is composed of tandemly reiterated copies of a short 22bp sequence in HSV-1 strain 17^+ (Murchie and McGeoch, 1982) which, in strain Patton, varies in copy number and is one residue longer (Watson et al., 1981b). The HSV-2 IEmRNA-4/-5 intron is approximately 540bp in length and comprises multiple tandem copies of three short G+C-rich DNA sequences, the number and arrangement of which differ between

independently-cloned isolates (Whitton and Clements, 1984b). HSV-1 IE mRNA-1 contains two introns with sizes of 765nuc and 135nuc, both of which are located in the translated portion of the gene (F.J. Rixon and L. Perry, personal communication). The 5' proximal intron DNA sequence possesses three tandemly arranged copies of an imperfect repeat of 54bp (L. Perry, personal communication).

Minor spliced species of HSV-1 gC have introns located close to the 5' terminus which range in size from 100 to 300nuc (Frink et al., 1981 and 1983). Apart from one spliced mRNA, these minor species would not encode a full-length gC polypeptide. A late 2.7kb transcript which maps at approximately 0.2 map units contains a large intron of 4kb (Fig. 4; Costa et al., 1985); interestingly, the intron encodes a family of unspliced 3' co-terminal mRNAs which are transcribed from the opposite DNA strand. The predicted amino acid sequence specified by this 2.7kb transcript shares homology with proposed protein coding regions in EBV (Baer et al., 1984; Costa et al., 1985).

c) Polyadenylation.

HSV mRNAs are polyadenylated (Bachenheimer and Roizman, 1972; Stringer et al., 1977) with the majority of cytoplasmic poly A transcripts containing a poly A tail of approximately 155 A residues (Silverstein et al., 1976). To date, DNA sequences encoding the 3' portions of HSV mRNAs contain the polyadenylation signal, AATAAA (Proudfoot and

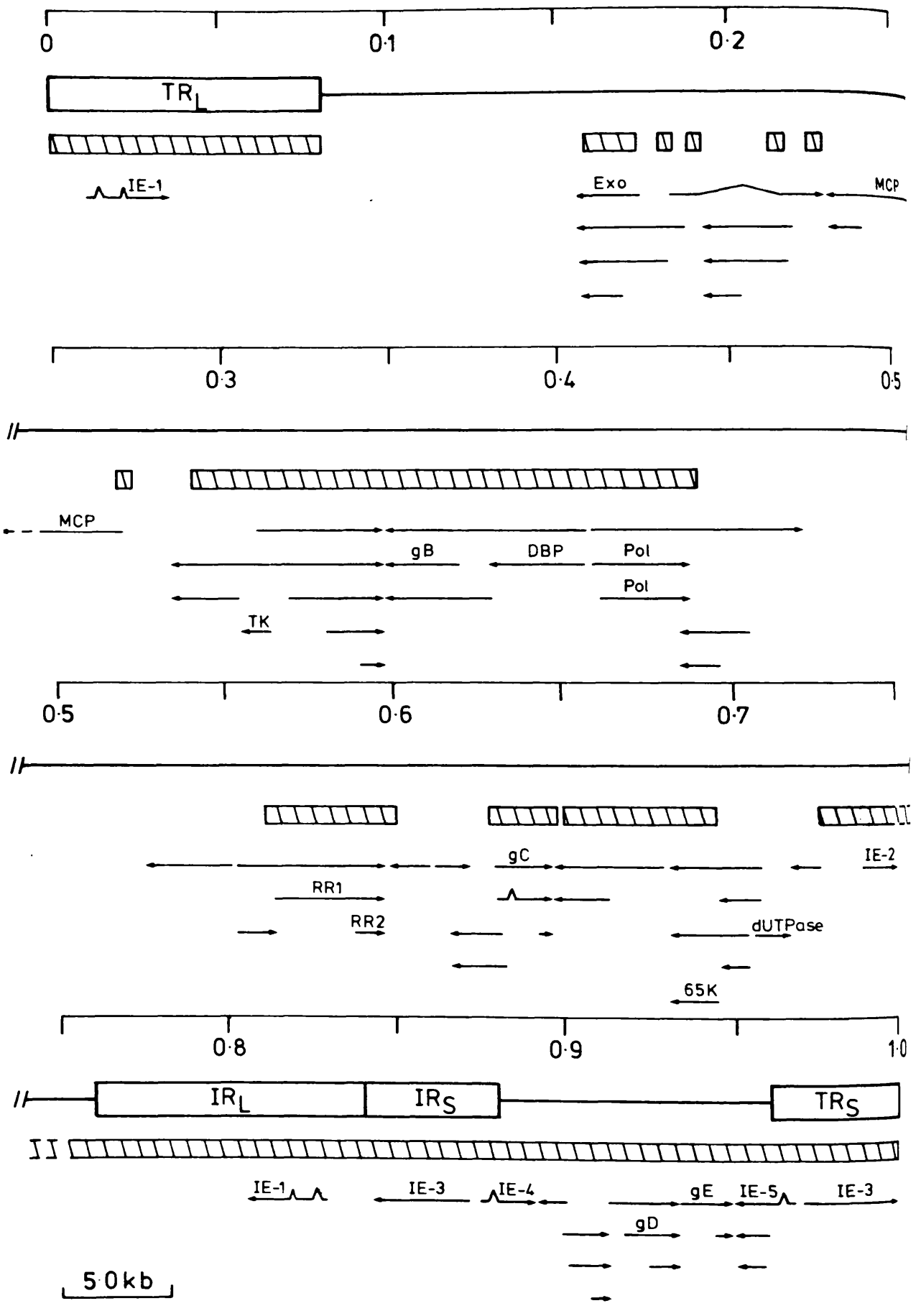

0
0·1
0·2
TRL
IE-1
Exo
MCP
0·3
0·4
0·5
MCP
gB
DBP
Pol
Pol
TK
0·5
0·6
0·7
gC
IE-2
RR1
RR2
dUTPase
65K
0·8
0·9
1·0
IRL
IRS
TRS
IE-1
IE-3
IE-4
gE
IE-5
IE-3
gD
5·0kb

Figure 4. Organisation of HSV-1 mRNAs and regions of the HSV-1 genome for which nucleotide sequences have been determined. Locations of transcripts are shown below the genome and spliced regions within mRNAs are shown by (∧). Hatched areas indicate genomic regions which have been sequenced. Transcripts which specify known virus-encoded proteins are shown and the key to abbreviations is as follows: IE, immediate early; Exo, alkaline exonuclease; MCP, major capsid protein; TK, thymidine kinase; gB to gE, glycoproteins B to E; DBP, major DNA binding protein; Pol, DNA polymerase; RR, ribonucleotide reductase; 65K, IE stimulatory polypeptide. The following references do not represent a complete list of the literature on HSV transcript mapping and nucleotide sequence data and, where appropriate, only those references which describe HSV-1 strain 17^+ data are given. References used to compile the mRNA mapping data were: the Exo 3' co-terminal family, Costa et al., 1983; mRNAs at 0.2 map units including the spliced species, Costa et al., 1985; the MCP 3' co-terminal family, Costa et al., 1981; mRNAs mapping between 0.29 to 0.46 map units were taken from Holland et al., 1984; mRNAs at the RR locus, Anderson et al., 1981, McLauchlan and Clements, 1982 and 1983a; mRNAs mapping between 0.60 to 0.65 map units, Frink et al., 1981; mRNAs mapping between 0.65 to 0.72, Hall et al., 1982; IE mRNA-2, Whitton et al., 1983; IE mRNA-1, F.J. Rixon, personal communication; IE mRNA-3, Rixon et al., 1982; IE mRNAs-4 and -5, Rixon and Clements, 1982; mRNAs located within U_S, Rixon and McGeoch,

1984 and 1985. References used to compile the regions which have been sequenced were: Exo locus, Costa et al., 1983, E.K. Wagner, personal communication; region encoding spliced species at 0.2 map units, Costa et al., 1985; 5' end of MCP mRNA, Costa et al., 1984; TK locus, McKnight, 1980, Wagner et al., 1981, Sharp et al., 1983; gB gene, Bzik et al., 1984, Pellett et al., 1985a; the remaining sequences between 0.29 to 0.35 have been determined by D.J. McGeoch, personal communication; DBP and Pol loci, Quinn, 1984; RR locus, Draper et al., 1982, McLauchlan and Clements, 1982 and 1983a, I. Nikas, personal communication; gC gene, Frink et al., 1983; region between 0.65 to 0.69 map units, including the 65K gene, Dalrymple et al., 1985, M.L. Dalrymple, personal communication; region between 0.73 to 0.76 map units, Whitton et al., 1983, Debroy et al., 1985, L. Perry, personal communication; the TR_L and IR_L repeats, L. Perry, personal communication; the S segment; Davison and Wilkie, 1981, Murchie and McGeoch, 1982, McGeoch et al., 1985.

Brownlee, 1976), with the exception of gE mRNA which possesses the hexanucleotide sequence, ATTAAA, upstream from the poly A site (McGeoch et al., 1985; Rixon and McGeoch, 1985).

8. Locations and Arrangements of HSV Transcription Units.

Individual HSV transcripts have been mapped using the techniques of Northern blot analysis (Alwine et al., 1977; Spandidos and Paul, 1982) and nuclease digestion of DNA/RNA hybrids (Berk and Sharp, 1978a, b; Weaver and Weissman, 1979). The structural analysis and map locations of HSV mRNAs is incomplete, however, the majority of transcripts are unspliced (reviewed by Wagner, 1985) and the transcripts in certain regions of the genome have been fully characterised (Rixon and McGeoch, 1985). In addition, nucleotide sequences for large areas of the HSV-1 genome have been determined, allowing precise localisation of mRNA termini and splice sites and the predicted translated regions of HSV genes. HSV-1 transcripts which have been mapped and genomic regions for which nucleotide sequence data are available are indicated in Fig. 4; the known functions of certain genes also are indicated.

A large number of HSV mRNAs are arranged as families of overlapping, unspliced transcripts with common 3' termini and unique 5' termini. In certain cases, 5' co-terminal mRNAs have been identified which have alternative 3' termini (Anderson et al., 1981; Hall et al., 1982; Holland et al., 1984). The transcripts within a mRNA

family do not invariably belong to the same temporal class. For example, the U_S10, -11 and -12 mRNAs share an identical 3' terminus and are early, late and IE species respectively (Rixon and McGeoch, 1984), thus the pattern of transcription from a mRNA family may alter during the course of infection. Overlaps between transcripts synthesised from opposite strands of HSV DNA have been identified (see Page 35). For example, U_S2 and U_S3 mRNAs are divergently transcribed and their 5' termini overlap (Fig. 3; Rixon and McGeoch, 1985). Similarly, the 3' terminal portion of the DNA polymerase transcript overlaps with that of a transcript which is synthesised in the opposite direction (Fig. 4; Quinn, 1984).

9. HSV Polypeptide Synthesis.

HSV-1 infection induces synthesis of about 50 novel polypeptides as identified by one-dimensional SDS-PAGE analysis (Honess and Roizman, 1973; Powell and Courtney, 1975; Marsden et al., 1976). Increased resolution by two-dimensional polyacrylamide gel techniques has revealed that the infected cell contains some 230 virus-induced polypeptide species, however a number of these species represent related forms of the same polypeptide (Haarr and Marsden, 1981).

HSV polypeptide synthesis is temporally regulated and there are three broad groups of coordinately-synthesised polypeptides, IE (or α), early (or β) and late (or γ; Honess and Roizman, 1974). Both the

β and γ proteins have been further sub-divided into subsets of β_1, β_2 (Pereira et al., 1977) and γ_1, γ_2 (Powell et al., 1975). β_1 proteins are produced in the presence of canavanine and azetidine, amino acid analogues which result in the production of non-functional α polypeptides; β_2 protein synthesis requires fully functional α products (Pereira et al., 1977). The γ_1 and γ_2 proteins are analogous to early-late and true late mRNAs respectively. HSV infection is regulated predominantly at the transcriptional level, although, there is evidence that translational regulation occurs also. At late times p.i., gD mRNA continues to accumulate while the level of gD polypeptide synthesis is reduced (Johnson and Spear, 1984). Similarly, the levels of IE Vmw175 are reduced at early and late times p.i. (see Page 27), possibly as a result of this polypeptide regulating IE mRNA synthesis (Dixon and Schaffer, 1980). However, IE mRNA-3, which encodes IE Vmw175, is abundant at later stages of the lytic cycle (Harris-Hamilton and Bachenheimer, 1985; F.J. Rixon, personal communication); this lack of correlation between the levels of IE mRNA-3 and IE polypeptide Vmw175 may be due to translational regulation although this possibility has not been rigorously examined.

a) Post-translational modification of proteins.

A number of HSV polypeptides produced in vitro do not co-migrate with their in vivo counterparts (Preston, 1977), indicating that HSV polypeptides undergo

modification (Morse et al., 1978; Marsden et al., 1982). Four types of modification have been identified, namely phosphorylation, glycosylation, sulphation and cleavage.

There are eighteen HSV-2 phosphoproteins and sixteen HSV-1 phosphoproteins (Pereira et al., 1977; Marsden et al., 1978), nine of which have been mapped by analysis of HSV-1/HSV-2 intertypic recombinants (Marsden et al., 1978). The phosphorylation pattern of a number of proteins differs at various stages of infection (Wilcox et al., 1980). Moreover, phosphorylation decreases the DNA binding ability of Vmw136, a protein involved in ribonucleotide reductase activity (see Page 59), however, a 54,000 mol. wt. protein has an increased affinity for DNA following phosphorylation (Wilcox et al., 1980).

HSV-1 specifies at least five major glycoproteins, gB, gC, gD (Spear, 1976), gE (Baucke and Spear, 1979), gH (Buckmaster et al., 1984) and possibly a sixth, gY (Palfreyman et al., 1983). Apart from gH and gY, precise map locations and the DNA sequences of the HSV-1 glycoproteins have been determined (Fig. 4; gB, Bzik et al., 1984; Pellett et al., 1985a; gC, Frink et al., 1983; Draper et al., 1984a; gD, Watson et al., 1982; McGeoch et al., 1985; gE, McGeoch et al., 1985). Immunoprecipitation data using type common antibodies have shown that HSV-2 infection induces equivalent glycosylated species to HSV-1 gB (Pereira et al., 1981), gC (Para et al., 1983; Zezulak and Spear, 1983; Zweig et al., 1983), gD (Balachandran et al., 1982) and gE (Para et al., 1982). In addition, HSV-2

specifies a g92k glycoprotein which maps in the U_S region (Marsden et al., 1978 and 1984); this protein may be identical to the gG protein identified by Roizman et al. (1984). To date, an HSV-1 glycoprotein equivalent to g92K has not been identified, however, this HSV-2 genome region contains additional nucleotide sequences compared with the equivalent locus on HSV-1 DNA (see Page 12; D.J. McGeoch, personal communication). Glycosylation is either partially or completely inhibited by tunicamycin (Pizer et al., 1980; Norrild and Pedersen, 1982; Hope and Marsden, 1983; Kousoulas et al., 1983; Roizman et al., 1984) which specifically blocks production of an intermediate required for transfer of oligosaccharides to asparagine residues (Struck and Lennarz, 1980). Hence, glycosylation in HSV probably occurs by the addition to asparagine residues of a core oligosaccharide containing N-acetyl glucosamine (NAG) and mannose; this oligosaccharide may be further modified by addition of sugar residues (Marshall and Neuberger, 1964; Marshall, 1974). Both N- and O-linked oligosaccharides have been identified (Pizer et al., 1980; Hope and Marsden, 1983; Johnson and Spear, 1983) and glycosylation occurs in discrete steps (Cohen et al., 1980; Haarr and Marsden, 1981; Palfreyman et al., 1983).

Most HSV glycoproteins are sulphated (Hope et al., 1982) with the majority of sulphate residues attached to N-linked oligosaccharides (Hope and Marsden, 1983). The major sulphated protein is gE (Hope et al., 1982).

b) HSV structural polypeptides.

The HSV-1 virion contains approximately 33 structural proteins, ranging in mol. wt. up to 290,000 (Spear and Roizman, 1972; Heine et al., 1974). The majority or all of the virus-specified glycoproteins are located in the envelope (Spear and Roizman, 1972). In addition, a 65,000 mol. wt. polypeptide, involved in IE gene activation, has been assigned to the tegument (Campbell et al., 1984). Seven major capsid proteins ranging from 12,000 to 155,000 mol. wt. have been identified (Cohen et al., 1980). The 155,000 mol. wt. major capsid protein, which maps between 0.23 and 0.27 map units (Marsden et al., 1978) and is encoded by a 6kb mRNA (Fig. 4; Costa et al., 1981 and 1984), is involved in disulphide bonding within the virion structure (Zweig et al., 1979). A 51,000 mol. wt. capsid protein has DNA binding properties and may function in the packaging or anchoring of DNA in the capsid (Braun et al., 1984). Finally, a third capsid protein with a mol. wt. of 40,000 is involved in encapsidation of viral DNA (Preston et al., 1983).

c) HSV-induced enzymes.

The HSV genome encodes a number of enzymic activities whose biochemical specificities differ from those of endogenous cellular enzymes. The HSV enzymes identified to date include DNA-dependent DNA polymerase (Keir and Gold, 1963), deoxynucleoside pyrimidine kinase (thymidine kinase; Kit and Dubbs, 1963; Jamieson et al.,

1974), alkaline deoxyribonuclease (Morrison and Keir, 1968), nucleoside phosphotransferase (Jamieson et al., 1974), deoxyuridine triphosphate nucleotidylhydrolase (dUTPase; Caradonna and Cheng, 1981; Williams, 1984), uracil-DNA glycosylase (Caradonna and Cheng, 1981), cyclic AMP independent protein kinase (Blue and Stobbs, 1981), a topoisomerase (Biswal et al., 1983; Leary and Francke, 1984) and ribonucleotide reductase (Cohen, 1972; Dutia, 1983).

1) DNA polymerase. HSV DNA polymerase differs from the cellular polymerase in its requirement for high salt (Keir et al., 1966) and sensitivity to PAA (Mao et al., 1975). The purified enzyme has a mol. wt. of 150,000 (Powell and Purifoy, 1977). Additional polypeptides co-purify with DNA polymerase, in particular a 54,000 mol. wt. species appears to be a consistent component of polymerase preparations (Powell and Purifoy, 1977; Knopf, 1979; Derse et al., 1982). The genomic region encoding the polymerase has been mapped to between 0.396 to 0.418 map units (Chartrand et al., 1979 and 1980), and two overlapping early mRNAs with sizes of 4.3kb and 4.2kb probably encode the enzyme (Fig. 4; Holland et al., 1984; Quinn, 1984). The DNA sequence of this portion of the genome has been determined and the presumed reading frame encoding the polymerase has been identified (Quinn, 1984).

2) Thymidine kinase. The viral TK enzyme maps between 0.30 and 0.31 map units (Halliburton et al., 1980) and has a

number of different properties from its cellular counterpart, including an insensitivity to inhibition by dTTP which is widely applied for assay purposes (Klemperer et al., 1967). The enzyme consists of a dimer with subunits of 42,000 mol. wt. (Honess and Watson, 1974; Jamieson and Subak-Sharpe, 1978). The mRNA encoding HSV-1 TK is 1.5kb in length (McKnight, 1980) and the DNA sequences encoding both HSV-1 (McKnight, 1980; Wagner et al., 1981; D.J. McGeoch, personal communication) and HSV-2 enzymes (Swain and Galloway, 1983) have been determined. The HSV-1 TK mRNA specifies three proteins with mol. wts. of 43,000, 39,000 and 38,000 (Preston and McGeoch, 1981; Marsden et al., 1983) which are translated from three in-phase initiation codons (Haarr et al., 1985). Comparison of DNA sequences encoding the reading frames of the HSV-1 and HSV-2 enzymes has revealed that the initiation codons for the 39,000 and 38,000 mol. wt. HSV-1 proteins are not present in HSV-2 DNA (Swain and Galloway, 1983).

3) Alkaline deoxyribonuclease. The viral enzyme is a DNA-specific exonuclease (Morrison and Keir, 1968) which increases in activity during the course of virus infection (Keir and Gold, 1963; Keir, 1968). The enzymatic activity has been mapped to a region between 0.12 and 0.21 map units (Moss et al., 1979) which specifies a 85,000 mol. wt. polypeptide (Preston and Cordingley, 1982) encoded by an early 2.3kb mRNA (Fig. 4; Costa et al., 1983). Monoclonal antibodies directed against the 85,000 mol. wt. protein neutralise exonuclease activity, implying that this protein is involved in enzymatic activity (Banks et al., 1983).

4) dUTPase. HSV dUTPase converts dUTP to dUMP, thereby preventing the incorporation of dUTP, a substrate for HSV DNA polymerase (Ostrander and Cheng, 1980), into HSV DNA. The enzyme has a mol. wt. of 53,000 (Williams, 1984) and is specified by an early 1.5kb mRNA which maps at 0.69 to 0.71 map units on the HSV-1 genome (Fig. 4; Hall et al., 1982; Preston and Fisher, 1984).

5) Topoisomerase. HSV-1 topoisomerase binds to single-stranded DNA through a 3' phosphoryl link and alters the linking number of a unique topoisomer in steps of one (Muller et al., 1985). In addition, the enzyme does not require either ATP or $MgCl_2$ for activity. These properties indicate that the HSV enzyme is a type I topoisomerase which probably catalyses the breaking and rejoining of single-stranded DNA. Enzymatic activity is associated with a 65,000 mol. wt. protein which binds to DNA (H.S. Marsden, personal communication). The suggestion that HSV-1 topoisomerase may be the virion component involved in IE gene activation (see Page 29; Muller et al., 1985) appears unlikely as these polypeptides map at different genomic locations (H.S. Marsden, personal communication). While the precise map position for the topoisomerase is not known, it has been tentatively located to between 0.60 to 0.63 map units (H.S. Marsden, personal communication).

6) Ribonucleotide reductase. Properties of the ribonucleotide reductase are described in Section 17 (Page 58).

SECTION C.

HSV CELLULAR TRANSFORMATION AND ONCOGENICITY.

10. The Association of HSV with Cervical Carcinoma.

HSV-2 is the primary cause of genital herpes and has also been implicated as an aetiological agent in development of cervical carcinoma. Preliminary evidence for the association of HSV with cervical neoplastic disease was obtained from epidemiological and serological studies (Naib et al., 1966; Rawls et al., 1969; Nahmias et al., 1970; Adam et al., 1972). There is also an increased incidence of HSV antibodies and specific HSV antigens in patients with cervical intraepithelial neoplasia (CIN) or cancer (Rawls et al., 1969; Royston and Aurelian, 1970; Skinner et al., 1977; Dreesman et al., 1980; Gilman et al., 1980). In situ hybridisation techniques demonstrated that HSV-specific RNA was present in tissue from cervical biopsies (McDougall et al., 1980; Eglin et al., 1981b) and that the most abundant transcripts were derived from limited regions of the genome. These regions were the IR_S/TR_S repeats and from between 0.07 to 0.4 map units (Eglin et al., 1981a; McDougall et al., 1982) and, in one study, from between 0.58 to 0.63 map units (McDougall et al., 1982). Viral DNA has not been detected in all tumours (zur Hausen et al., 1974; Pagano, 1975), however, DNA probes from the genomic region between 0.58 to 0.612 map units hybridised to DNA isolated from one out of eight carcinoma tissues (Park et

al., 1983); other DNA fragments spanning most of the genome failed to detect viral DNA in these samples. In a separate study, DNA from three out of nine cervical tumours had homology to specific HSV-2 DNA probes (McDougall et al., 1982), one of which, Bgl II n, maps between 0.58 and 0.63 map units.

11. Cellular Transformation In Vitro.

The transformation potential of HSV has been studied in rodent embryo cells transformed with U.V.-inactivated virus, ts mutants (Macnab, 1974) and sheared HSV DNA (Wilkie et al., 1974). The ability to rescue ts^+ virus from HSV-transformed cells infected with ts mutants has indicated that viral sequences are maintained in transformed cell lines (Park et al., 1980). However, these data contrast with the failure to detect viral DNA in all transformants (Minson et al., 1976; Cameron et al., 1985), possibly as a result of either the low level of DNA retained in transformed cell lines (<1 copy/cell) or insertion of short HSV fragments into the cell genome. Several HSV-specific proteins have been detected in cells transformed with inactivated virus (Duff and Rapp, 1973; Reed et al., 1975; Jamieson et al., 1976; Flannery et al., 1977; Macnab et al., 1980; Lewis et al., 1982) although no polypeptide was invariably present in all transformants.

The use of specific HSV DNA fragments has identified three genomic regions which are involved in

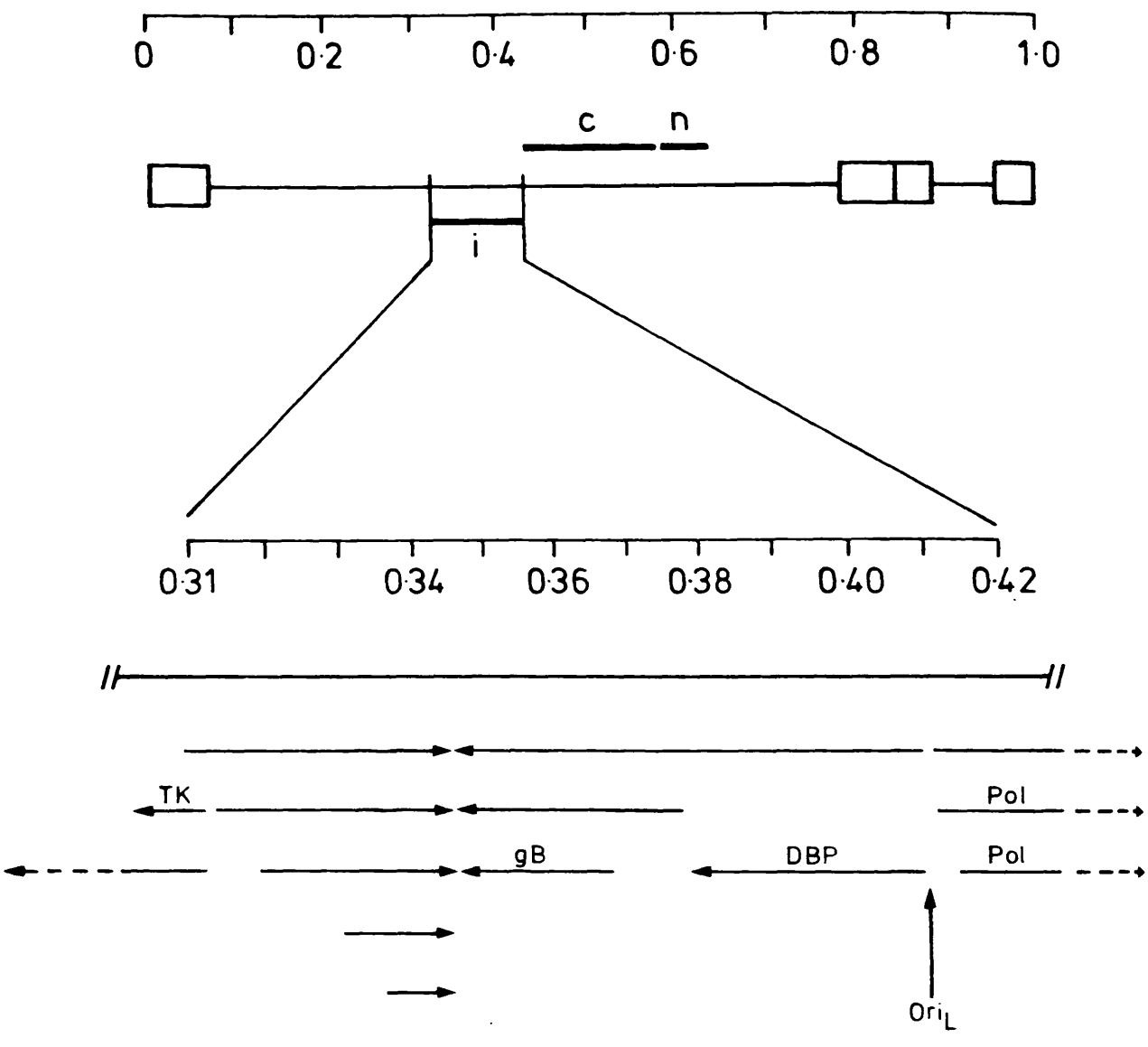


Figure 5. Locations and arrangements of mRNAs which map in HSV-1 transforming region, Bgl II i. The positions of HSV-2 transforming fragments, Bgl II c and n, are shown above the HSV genome. mRNAs which encode specific viral functions are labelled and the location of an origin of DNA replication within U_L (Ori_L) is shown. The key to abbreviations and the references which describe mRNAs mapping in this region are given in the legend to Fig. 4.

cellular transformation; two of these regions, morphological transforming region (mtr) II (Bgl II n, 0.58 to 0.63 map units; Reyes et al., 1979; Cameron and Macnab, 1980; Galloway and McDougall, 1981; Cameron et al., 1985) and Bgl II c (0.43 to 0.58 map units; Jariwalla et al., 1980 and 1983), are located on the HSV-2 genome and the third, mtr I (Bgl II i, 0.31 to 0.42 map units; Camacho and Spear, 1978; Reyes et al., 1979), is on the HSV-1 genome (Fig. 5). Interestingly, the mtr I and mtr II regions do not map at equivalent positions on HSV-1 and HSV-2 DNAs. Cells transformed with fragments encompassing mtr I and mtr II exhibit different properties from those transformed with Bgl II c. Mtr I- and mtr II-transformed cells are selected by focus formation of morphologically-altered cells in low serum (Camacho and Spear, 1978; Reyes et al., 1979; Galloway and McDougall, 1981; Cameron et al., 1985) whereas Bgl II c induces rapidly proliferating Syrian hamster cells which do not senesce and overgrow non-transformed cells (Jariwalla et al., 1979 and 1980).

The mtr-1 region encodes a number of viral proteins including gB (Fig. 5; Little et al., 1981; DeLuca et al., 1982) which has been detected in cells transformed by inactivated virus (Lewis et al., 1982) and by mtr-1 DNA (Camacho and Spear, 1978). Cells transformed by mtr-1 appear to be tumorigenic in rodents (J.C.M. Macnab, personal communication).

It has been proposed that immortalisation of cells by Bgl II c is an early transformation event, however,

tumorigenicity in hamsters is acquired only upon further *in vitro* passage (Jariwalla *et al.*, 1980). Two separate transforming functions have been identified within Bgl II *c* (Jariwalla *et al.*, 1983). The leftmost 64% of the fragment is sufficient to induce immortalisation of hamster cells which are, however, non-tumorigenic. Therefore, Jariwalla *et al.* (1983) have proposed that conversion to tumorigenicity requires the right-hand 36% of the Bgl II *c* fragment. In a small number of transformed cell lines tested, radiolabelled Bgl II *c* hybridised to restriction enzyme fragments with sizes identical to those of fragments derived from the DNA probe. These data suggested that viral sequences were present in transformants, although Bgl II *c* has a background of homology with normal Syrian hamster DNA which appears to be located in the left-hand 64% portion of the fragment (Jariwalla *et al.*, 1983). One viral polypeptide with a mol. wt. of 120,000 has been detected in Bgl II *c*-transformed cells (Jariwalla *et al.*, 1980).

Cells transformed by mtr II are oncogenic in rodents, the latent period for the appearance of tumours being similar to that for cells transformed with U.V.-inactivated virus (Macnab, 1979; Cameron *et al.*, 1985). A transforming region within mtr II has been located to a 2.1kb Bam HI/Pst I subfragment of Bam HI *t* (0.584 to 0.602 map units), however, the resultant transformed cell lines appear not to retain viral sequences (Galloway and McDougall, 1983). *In vitro* translation of hybrid-selected RNA has shown that three proteins with mol. wts. of

138,000, 38,000 and 61,000 are partially or wholly encoded within this fragment (Docherty et al., 1981; Galloway et al., 1982). A DNA segment of 737bp located within the coding sequences of the 61,000 mol. wt. protein induces transformation in NIH 3T3 cells; a portion of this sequence could form a hairpin structure with a loop of 129b and a stem of 8bp flanked by repeated sequences of 7b (Galloway et al., 1984). This structure has been compared to insertion sequence (IS) elements which may initiate transformation by insertion into the cellular genome (Rechavi et al., 1982).

Viral sequences, which are present in cells transformed with HSV-2 DNA fragments at early passages, are not detectable in later passages (Cameron et al., 1985). Therefore, a 'hit and run' mechanism is favoured whereby maintenance of the transformed phenotype does not appear to require the continuous presence of viral sequences which may be lost by the cell following induction of the transformation process (Galloway and McDougall, 1983). Alternatively, HSV may initiate transformation by introducing chromosomal aberrations into cellular DNA. HSV infection causes chromosomal damage (Hampar and Ellison, 1963; Stich et al., 1964; Schlehofer and zur Hausen, 1982) which is possibly mediated by early functions (Waubke et al., 1968; O'Neill and Rapp, 1971); thus, HSV enzymes involved in DNA biosynthesis may act as cellular mutagens and Huszar and Bacchetti (1983) have suggested that HSV-2 ribonucleotide reductase may be involved in the

transformation process. Infection with U.V.-inactivated HSV-1 and HSV-2 (Boyd et al., 1978) and transfection with isolated HSV-1 restriction enzyme fragments (Boyd et al., 1980) also activates synthesis of mouse xenotropic (MuX) retrovirus, an endogenous C-type virus present in mouse Balb/c-3T3 cells. While xenotropic viruses appear to be non-pathogenic, this activation raises the possibilty that other endogenous retroviruses associated with malignant disease (reviewed in Teich et al., 1982) are induced by HSV infection and thus lead to transformation.

SECTION D.

RIBONUCLEOTIDE REDUCTASES.

12. General Properties of Ribonucleotide Reductases.

Conversion of the four ribonucleotides to the corresponding deoxyribonucleotides by ribonucleotide reductase represents the first unique step in DNA biosynthesis (Larsson and Reichard, 1966a, b; reviewed by Thelander and Reichard, 1979) and the level of DNA synthesis correlates closely with but is not regulated solely by ribonucleotide reductase activity (Thelander and Reichard, 1979). The enzyme catalyses replacement of the hydroxyl group by a hydrogen atom at the 2' position of the ribose residue on the ribonucleotide. Certain ribonucleotide reductases may be associated with a multienzyme complex *in vivo* which channels high concentrations of dNTPs to sites of DNA replication (Flanegan and Greenberg, 1977; Forsdyke and Scott, 1979). For example, the bacteriophage T4-induced ribonucleotide reductase is bound along with most of the enzymes involved in deoxyribonucleotide synthesis to a multienzyme aggregate (Greenberg and Chiu, 1978; Allen *et al.*, 1980; Chiu *et al.*, 1980 and 1982), from which ribonucleotide reductase is readily dissociable (Chiu *et al.*, 1982).

Reductases have been isolated from a number of sources and two classes of enzyme have been characterised according to the components required for activity (Thelander and Reichard, 1979):

1) the first class is represented by the Lactobacillus leichmanii (L. leichmanii) enzyme which is a monomer composed of a single polypeptide with a mol. wt. of 76,000 (Panagou et al., 1972; Chen et al., 1974). The enzyme requires adenosylcobalamin (coenzyme B_{12}) as a dissociable co-factor (Vitols et al., 1967; Singh et al., 1977) and uses ribonucleoside triphosphates as substrates (NTPs; Chen et al., 1974; Singh et al., 1977). This class of enzyme is more commonly found among prokaryotes than eukaryotes (Gleason and Hogenkamp, 1972).

2) the second class is represented by the Escherichia coli (E. coli) enzyme which has a mol. wt. of 240,000 (Thelander, 1973) and consists of two non-identical subunits, B1 and B2 (Brown and Reichard, 1969b). Both subunits are required for enzyme activity (Brown and Reichard, 1969a; Thelander, 1973) and the substrates are ribonucleoside diphosphates (NDPs; Larsson and Reichard, 1966a, b). The E. coli enzyme has no requirement for coenzyme B_{12}, however, each B2 subunit contains two bound iron atoms (Brown et al., 1969a) and a free radical located on a tyrosine residue (Ehrenberg and Reichard, 1972; Sjoberg et al., 1977) which are involved in the enzyme reaction. Results by Reichard and Ehrenberg (1983) have suggested that generation of the free radical requires oxygen. Removal of iron leads to loss of the tyrosine radical (Atkin et al., 1973) while removal of the radical by hydroxyurea treatment leaves the iron centre intact (Atkin et al., 1973; Sjoberg et al., 1982). Eukaryotic enzymes which have similar properties to the E. coli

reductase have been isolated from green algae, yeast, higher plants and mammalian species (Gleason and Hogenkamp, 1972; Feller and Follman, 1976; Gleason and Wood, 1976).

Ribonucleotide reductases of cellular origin are tightly regulated by the reaction products, the deoxyribonucleoside triphosphates (dNTPs), which act as both positive and negative effectors, modifying substrate specificity as well as enzyme activity (see following section).

13. E. coli Ribonucleotide Reductase.

The B1 and B2 subunits of the E. coli enzyme (Brown and Reichard, 1969b; Brown et al., 1969b) are encoded by the nrdA and nrdB genes respectively (Fuchs and Karlstrom, 1976) which are located at adjacent sites on E. coli DNA (Bachman et al., 1976; Yamada et al., 1982). The B1 subunit contains three classes of binding site; one for the substrates (von Dobeln and Reichard, 1976) and two different sites for the allosteric effectors (Brown and Reichard, 1969b). There are two substrate binding sites on each B1 subunit which bind all four NDPs (von Dobeln and Reichard, 1976). The two classes of allosteric effector sites are termed h- and l-sites with each class consisting of two subsites (Brown and Reichard, 1969a, b). The h-sites are defined by their high affinity for dATP while the l-sites have a low affinity for this metabolite. Furthermore, the h-sites bind ATP, dTTP and dGTP whereas the l-sites bind only ATP in addition to dATP. The l-sites

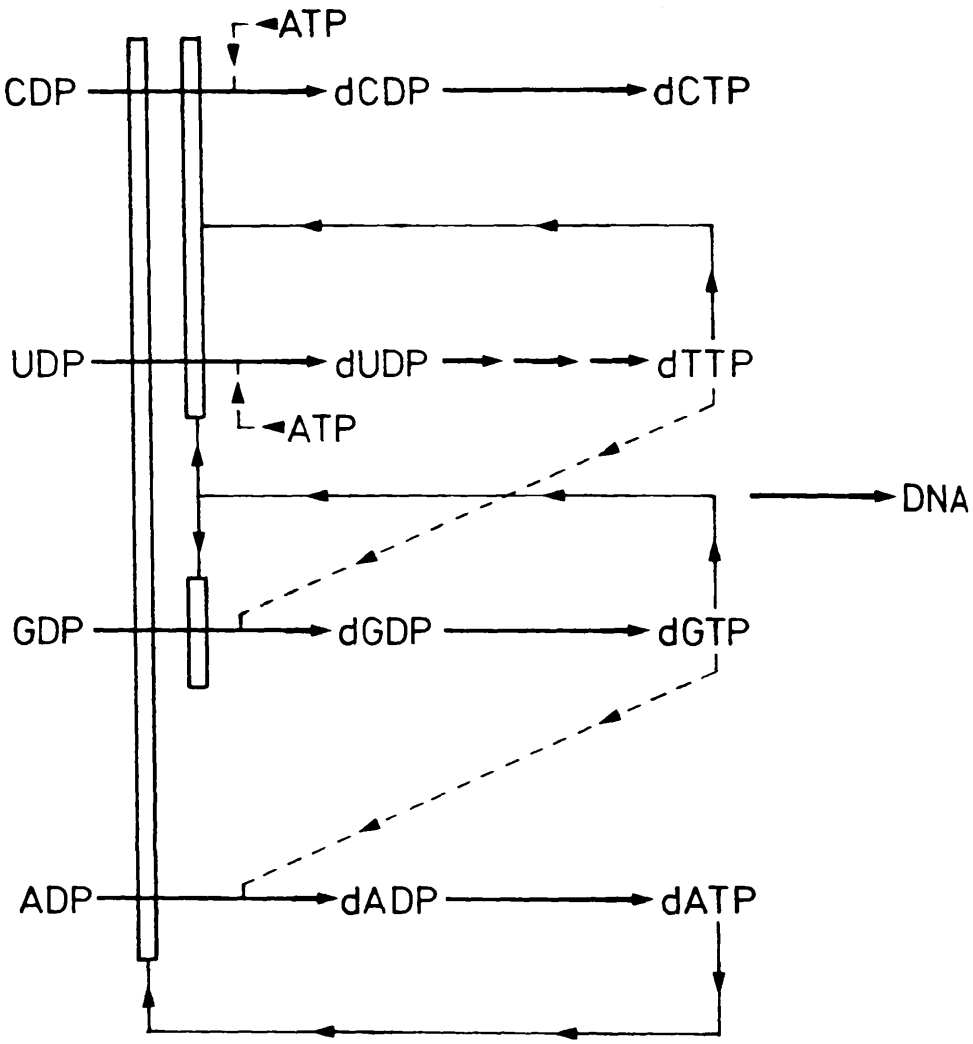


Figure 6. Allosteric control of E. coli ribonucleotide reductase. Substrates and reaction products are shown, with the direction of reaction pathways indicated by thick, arrowed lines. Reactions inhibited by negative effectors are shown by open bars; dashed lines indicate reactions stimulated by positive effectors.

regulate the general enzyme activity with ATP acting as a positive effector and dATP as a negative effector (Fig. 6). The presence of effector molecules at h-sites changes the active site conformation thereby altering substrate specificity (Brown and Reichard, 1969b). The active site of the enzyme comprises portions of both B1 and B2 subunits (von Dobeln and Reichard, 1976).

The B1 subunit consists of two polypeptides with mol. wts. of 80,000 which have identical C-termini but different N-termini (Thelander, 1973) and DNA sequence analysis of the nrd locus indicates that the different N-termini are produced by differential processing of a precursor polypeptide (Carlson et al., 1984). The B2 subunit is composed of two identical polypeptides with mol. wts. of 40,000 (Thelander, 1973) and requires Mg^{2+} for binding to the B1 subunit (Brown and Reichard, 1969a; Thelander, 1973).

The polypeptides which comprise the two subunits are both specified by a 3.2kb dicistronic mRNA (Hanke and Fuchs, 1983b) and their synthesis is coordinately regulated (Fuchs, 1977). Inhibition of DNA synthesis induces an increase in enzyme activity which is reflected in higher levels of mRNA synthesis (Hanke and Fuchs, 1983a); this increase in transcription requires continuous protein synthesis (Hanke and Fuchs, 1984).

14. Bacteriophage Ribonucleotide Reductases.

Bacteriophages T2, T4, T5 and T6 encode viral ribonucleotide reductases (Cohen and Barner, 1962; Biswas et al., 1965; Berglund et al., 1969; Eriksson and Berglund, 1974), however, bacteriophages T7 and λ do not synthesise a novel enzyme during infection. T4 ribonucleotide reductase has a mol. wt. of 225,000 and a structure similar to that of the E. coli enzyme (Berglund, 1975) represented as $\alpha 2 \beta 2$. The α and β polypeptides are encoded by the nrdA and nrdB genes which map at adjacent sites on the T4 genome (Yeh et al., 1969). In contrast to the E. coli enzyme, the T4 subunits do not require Mg^{2+} for binding and are not easily dissociable (Berglund, 1975). The α polypeptides synthesised by nrdB mutants which fail to produce an $\alpha 2 \beta 2$ structure are degraded to three fragments with sizes of 61,000, 57,000 and 24,500, indicating that the $\beta 2$ subunit protects the $\alpha 2$ moiety from proteolytic cleavage (Cook and Greenberg, 1983). Moreover, the C-terminal portion of the α polypeptide is required for binding to dATP (Cook and Greenberg, 1983). The T4 enzyme is inhibited neither by dATP, which acts only as a positive effector in reduction of CDP and UDP (Berglund, 1972), nor by the accumulation of dNTPs. Therefore, it is possible that the T4 enzyme may not contain an allosteric site analogous to the l-site of E. coli ribonucleotide reductase which regulates enzyme activity (see Page 53). T4 also encodes a thioredoxin (nrdC; Berglund et al., 1969) which does not map at the nrdA/B locus (Yeh and Tessman, 1972) but is the specific hydrogen donor for T4 ribonucleotide reductase (Berglund and Sjoberg, 1970).

15. Mammalian Ribonucleotide Reductases.

Ribonucleotide reductase activity has been detected in a range of cell types from different mammalian species including rat hepatoma (Moore, 1977), calf thymus (Engstrom et al., 1979; Eriksson et al., 1979; Thelander et al., 1980), rabbit bone marrow (Hopper, 1972 and 1978) and human tissue culture cell lines (Chang and Cheng, 1979). The mammalian enzyme consists of two non-identical subunits, M1 and M2 (Thelander et al., 1980), and has similar properties to those of the E. coli enzyme. For example, the M1 and M2 subunits are easily separable (Hopper, 1972), enzymatic activity does not require cobalamin coenzymes and the enzyme contains non-heme iron centres (Thelander et al., 1985) and a tyrosine free radical (Graslund et al., 1982). Homogenous preparations of both subunits have been isolated. The M1 subunit consists of two polypeptides, each with a mol. wt. of 90,000 which contain the nucleotide binding sites responsible for allosteric regulation of the enzyme (Thelander et al., 1980). Similarly, the M2 subunit is composed of two proteins with mol. wts. of 44,000 which contain the iron centres and free radical (Thelander et al., 1985). Allosteric regulation of the mammalian enzyme appears to be similar to that of E. coli ribonucleotide reductase; results suggest that the mammalian enzyme contains two classes of effector sites (Thelander et al., 1980) which are analogous to the h- and l-sites of the E. coli B1 subunit (see Page 53). Photolabelling techniques have shown that the M1 subunit is directly involved in

substrate binding and that a non-functional M2 subunit prevents the attachment of NDPs to this site (Caras et al., 1983).

While both subunits of the E. coli enzyme are coordinately expressed, the M1 and M2 subunits are differentially regulated. During cellular proliferation, the levels of the M1 subunit do not increase at the same rate as those of the M2 subunit (Eriksson and Martin, 1981) and the increased abundance of the M2 subunit is a result of de novo protein synthesis (Eriksson et al., 1984). The subcellular location of mammalian reductases is currently the subject of debate. Reddy and Pardee (1980 and 1982) have proposed that ribonucleotide reductase is associated with a multienzyme complex in the nucleus, termed the replitase, which channels NDPs to sites of DNA replication. More recently, results from immunofluorescent staining of cells using monoclonal antibodies directed against the M1 subunit have located this protein in the cytoplasm at all stages of the cell cycle (Engstrom et al., 1984).

The mutagenic effect of ribonucleotide reductase has been demonstrated in cell lines which possess an altered enzymatic activity. In such cell lines, the levels of dNTPs are abnormal and mutations accumulate in marker genes at a higher than normal frequency (Meuth et al., 1979; Weinberg et al., 1981). Results from one study have shown that two out of three cell lines contained M1 subunits with an altered allosteric control (Weinberg et al., 1981).

16. Animal Virus Ribonucleotide Reductases.

Virus-induced ribonucleotide reductase activity, which differs from the endogenous cellular enzyme, has been identified in cells infected with vaccinia virus (Slabaugh and Mathews, 1984; Slabaugh et al., 1984) and a number of herpesviruses, namely HSV-1 (Cohen, 1972), HSV-2 (Cohen et al., 1974), PRV (Lankinen et al., 1982), EBV (Henry et al., 1978), and EHV serotypes 1 and 3 (Cohen et al., 1977; Allen et al., 1978). Direct evidence that the novel enzymatic activities in infected cells are virus-encoded has not been determined for most of these viruses. Of the herpesvirus-induced enzymes only those of PRV and HSV have been well characterised. The PRV-induced reductase contains a tyrosine free radical and the lack of allosteric control of CDP-reductase activity (Lankinen et al., 1982) is very similar to that of the HSV enzyme (Ponce de Leon et al., 1977; Langelier et al., 1978; Huszar and Bacchetti, 1981).

17. HSV Ribonucleotide Reductase.

The enzymatic properties of the HSV-1- and HSV-2-induced enzymes appear to be identical (Ponce de Leon et al., 1977; Langelier et al., 1978; Huszar and Bacchetti, 1981; Averett et al., 1983 and 1984). In contrast to the cellular enzyme, the HSV reductase does not require Mg^{2+} for activity and is not inhibited by dNTPs (Ponce de Leon et al., 1977, Langelier et al., 1978; Huszar and Bacchetti, 1981; Langelier and Buttin, 1981; Averett et al., 1983), suggesting that there is no allosteric control. ATP, which

activates the cellular enzyme, inhibits the HSV enzyme in the absence of Mg^{2+}, however this inhibition is abolished in the presence of Mg^{2+}, possibly due to formation of a Mg.ATP complex (Ponce de Leon et al., 1977; Huszar and Bacchetti, 1981; Averett et al., 1983). Analysis by marker rescue techniques of HSV-1 ts mutant, 17tsVP1207, which induces a thermolabile enzymatic activity (Dutia, 1983), has shown that a virus-encoded early protein, Vmw136 (mol. wt. 136,000), is a component of the ribonucleotide reductase (Preston et al., 1984b). This protein is a phosphoprotein (Marsden et al., 1978) which has cleavage products with sizes of 100,000 and 90,000 mol. wt. in infected cells (H.S. Marsden, personal communication). Vmw136 is encoded by an unspliced 5.0kb mRNA (RR1 in Fig. 4) which is 3' co-terminal with a 1.2kb mRNA (RR2 in Fig. 4) specifying a 38,000 mol. wt. polypeptide (Vmw38; Anderson et al., 1981; McLauchlan and Clements, 1982 and 1983a); this protein is likely to be a second component of the HSV-1 ribonucleotide reductase (Frame et al., 1985).

SECTION E.

ARRANGEMENT AND REGULATION OF OVERLAPPING EUKARYOTIC mRNAS.

18. Transcription of RNA Polymerase II Genes.

RNA synthesis is mediated by a number of processes, the simplest pathway requiring an initiation event followed by elongation and finally termination. In eukaryotes, the following additional processes occur during the synthesis of mRNAs by RNA polymerase II (reviewed in Darnell, 1982 and Nevins, 1983):

1) following initiation, a m^7G cap is added to the 5' terminal nucleotide by a 5'-5' hydroxyl group linkage through a triphosphate bridge (reviewed in Shatkin, 1976). The presence of this cap structure may be required for translation (Both et al., 1975).

2) the mature mRNA 3' terminus is generated by cleavage of a primary transcript which extends beyond the cleavage site. In poly A^+ mRNAs, about 200 residues are added to the mature 3' terminus (Brawerman, 1981). The cleavage/polyadenylation process is described in detail in Section 23 (Page 84).

3) in mammalian mRNAs, a methyl group is acquired by about 1% of A residues at the N^6 position (reviewed in Banerjee, 1980).

4) finally, introns are removed by splicing mechanisms to generate mature mRNA (Gilbert, 1978).

The locations at which these events occur are largely determined by nucleotide sequences serving as recognition signals for processing complexes. Consensus sequences for a number of these signals have been derived from DNA sequence comparisons and their functions evaluated by both in vivo and in vitro analysis. The following sections describe those conserved sequences which are present in most genes and have a role in transcription.

a) Signals involved in the initiation of mRNA synthesis.

The control signals required to initiate transcription can be divided into two categories; firstly, sequences which are involved in constitutive mRNA synthesis and secondly, sequences which modulate the level of transcription.

1) Constitutive control sequences. A highly conserved sequence at mRNA 5' termini is the TATA box homology which has the consensus G$\frac{C}{T}$ATAA$\frac{T}{A}$AG (Busslinger et al., 1980; Proudfoot, 1978; Gannon et al., 1979) and is located 25 to 35nuc upstream from the mRNA start site. Both in vivo and in vitro studies have shown that point mutations within the TATA box homology significantly reduce the level of transcription (Wasylyk et al., 1980; Grosschedl et al., 1981). Furthermore, alteration of the TATA box produces heterogeneous mRNA 5' termini in many genes (Mathis and Chambon, 1981; Grosveld et al., 1982a; Lebowitz and Ghosh, 1982) but not in every case (Wasylyk et al., 1980).

Clearly, the TATA box homology has a crucial role in the initiation of mRNA synthesis, however, absence of a TATA box homology in a number of genes (Haegeman and Fiers, 1978; Baker and Ziff, 1981; Ghosh and Lebowitz, 1981) indicates that this sequence alone is not an obligatory requirement *in vivo*.

A second sequence homology, the CAAT box (consensus GGYCAATCT where Y = pyrimidine; Benoist *et al*., 1980; Efstratiadis *et al*., 1980) is present at -70 to -80 upstream from the 5' termini of a number of genes, in particular the α- and β-globin genes (Liebhaber *et al*., 1981; van Ooyen *et al*., 1979; Efstratiadis *et al*., 1980), HSV-1 TK gene (McKnight, 1980), simian virus 40 (SV40) early mRNAs (Benoist and Chambon, 1981) and the sea urchin histone H2A gene (Grosschedl and Birnstiel, 1980a, b). This signal is important *in vivo* for the expression of β-globin mRNA (Dierks *et al*., 1981; Grosveld *et al*., 1982a, b) and possibly α-globin transcription (Mellon *et al*., 1981). In contrast, removal or mutation of CAAT box homologues had no effect on the expression of SV40 early (Benoist and Chambon, 1981) or HSV-1 TK mRNAs (McKnight and Kingsbury, 1982). Indeed, transcription of the sea urchin histone H2A gene was increased by deleting the CAAT box (Grosschedl and Birnstiel, 1980a).

The 5' flanking sequences of some promoters contain tracts of either G-rich or C-rich residues which are of functional significance. Analysis of linker scanning mutants of the HSV-1 TK promoter identified two upstream

G-rich and C-rich segments as being important transcriptional elements (see Page 31); distal signal II, GGGGCGGCGCGG, at -47 to -61 and distal signal III which included the sequence CCCCGCCC at -80 to -105 (McKnight and Kingsbury, 1982). Similar G- and C-rich sequences contained within two perfect 21bp repeats and an imperfect 22bp copy upstream from the SV40 early promoter are essential for the efficient expression of early transcripts (Fromm and Berg, 1982; Everett et al., 1983). These repeat units contain a total of six CCGCCC segments, each of which is functional in increasing mRNA synthesis and can operate in either orientation (Everett et al., 1983; Baty et al., 1984). G-rich and C-rich sequences within the SV40 early promoter are binding sites for the cellular protein, Sp1, a factor required for synthesis of SV40 mRNA (Dynan and Tjian, 1983). Similar sequences present in the intergenic region between HSV-1 IE mRNAs-3 and -4/-5 and in the HSV-1 TK promoter also bind Sp1 (Jones and Tjian, 1985; Jones et al., 1985); from these studies, the consensus $\frac{G}{T}GGGCGG\frac{GGC}{AAT}$ has been proposed for Sp1 binding sites (Dynan and Tjian, 1985). Within the HSV-1 gD promoter (see Page 31), two tracts of a G-rich motif between -83 and -40 are required for constitutive mRNA synthesis and function in either orientation (Everett, 1983). Sequences similar to these G-rich motifs or their complement are also present upstream from other HSV mRNA 5' termini (Everett, 1983).

Efficient expression of rabbit β-globin mRNA requires a C-rich element with consensus CCNCACCC (N = any

base; Dierks et al., 1983) which is also present in mouse (van Ooyen et al., 1979) and human β-globin genes (Orkin et al., 1982). This element is located approximately 100bp from the mRNA start site and a C to G mutation within this sequence which alters CCACACCC to CCACACGC results in β-thalassaemia (Orkin et al., 1982).

2) Sequences involved in modulating gene expression. In addition to the sequences necessary for constitutive mRNA synthesis, regulatory signals which modulate the level of transcription in response to various stimuli have been identified. Examples of such control elements are enhancers which confer tissue- or cell-specific properties on some genes (de Villiers et al., 1982; Banerji et al., 1983; Gillies et al., 1983) and sequences that respond to induction by steroid hormones (Schrader et al., 1981; Compton et al., 1983; Scheidereit and Beato, 1984), heat shock (Ashburner and Bonner, 1979; Pelham, 1982; Amin et al., 1985), heavy metal ions (Mayo and Palmiter, 1981; Karin et al., 1984; von der Ahe et al., 1985), light (Morelli et al., 1985) and viral proteins (Nevins, 1982; Elkaim et al., 1983; Imperiale et al., 1983; Campbell et al., 1984). The metallothionein gene (Karin et al., 1984) contains more than one class of regulatory elements and therefore responds directly to induction by at least two different stimuli. Regulatory sequences are frequently located upstream from mRNA 5' termini, however, these signals have been identified within the intron of the mouse

heavy chain Ig gene (Banerji et al., 1983; Gillies et al., 1983), downstream from the 5' termini of human α- and β-globin genes (Charnay et al., 1984; Wright et al., 1984) and beyond the 3' terminus of bovine papillomavirus (BPV) early mRNA (Lusky et al., 1983). Synthetic oligomers containing sequences corresponding to consensus signals for heavy metal ion (Searle et al., 1985), heat shock (Pelham and Bienz, 1982) and HSV IE gene stimulation (Gaffney et al., 1985) can confer inducibility on linked promoters. In general, induction of modulator sequences will increase mRNA expression, however, in constructs containing viral enhancers linked to heterologous promoters, transcription was repressed by products of the Ad2 Ela transcription unit (Borrelli et al., 1984). Therefore, transcription can be both positively and negatively regulated through these sequences. In addition, regulatory sequences which have the ability to selectively inhibit transcription in the presence of cholesterol have been located to within 500bp upstream from the 5' terminus of 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMG CoA reductase) mRNA (Osborne et al., 1985).

b) Sequences involved in mRNA splicing.

Recent studies have suggested that the pathway for splicing involves the excision of introns as 'lariat' or 'branched-circle' structures (Grabowski et al., 1984; Padgett et al., 1984; Ruskin et al., 1984). This structure appears to be formed by cleavage at the splice donor site

followed by joining a G residue from the 5' end of the intron to an internal A residue through a 2'-5' phosphodiester bond (Konarska et al., 1985). This branch point is nuclease-resistant (Grabowski et al., 1984; Padgett et al., 1984; Ruskin et al., 1984) and blocks reverse transcriptase (Ruskin et al., 1984). Cleavage at the splice acceptor site removes the intron and exons flanking the intervening sequence are ligated. It is not known whether these events occur through transesterification reactions or by a series of cleavage/ligation steps. Small nuclear ribonucleoproteins (snRNPs) which contain U1 small nuclear RNA (snRNA) are thought to be involved in the selection of splice sites by base pairing (Mount et al., 1983) as removal of 8nuc from the 5' end of U1 snRNA abolishes splicing activity in vitro (Kramer et al., 1984).

Consensus sequences for both splice donor ($^{A}_{C}$AGGT$^{A}_{G}$AGT) and splice acceptor sites (Y_6NCAGG$^{G}_{T}$; Y = pyrimidine, N = any base) have been deduced from DNA sequence comparisons of splice sites from a large number of genes (Breathnach and Chambon, 1981; Sharp, 1981; Mount, 1982). Both mutations generated in vitro and naturally-occurring alterations within these sequence motifs either abolish or reduce splicing activity (Solnick, 1981; Felber et al., 1982; Montell et al., 1982; Treisman et al., 1983; Wieringa et al., 1983). Inactivation of authentic splice sites both in vivo (Treisman et al., 1983) and in vitro (Krainer et al., 1984) may stimulate the

utilisation of cryptic splice sites and, therefore, formation of incorrectly-spliced transcripts. Moreover, the creation of novel splice sites by mutations may alter the mRNA splicing pattern (Busslinger et al., 1981; Felber et al., 1982; Fukumaki et al., 1982; Treisman et al., 1983). Specific nucleotide sequences within the large intron of rabbit β-globin mRNA do not appear to be essential for efficient splicing (Wieringa et al., 1984), however, Ruskin et al. (1984) have noted that the sequence YNYTRAY (Y = pyrimidine, R = purine, N = any base) is conserved in the vicinity of mammalian β-globin splice acceptor sites. This motif is similar to the consensus sequence TACTAAC which is present upstream from the splice acceptor sites of yeast genes (Langford and Gallwitz, 1983; Pikielny et al., 1983; Langford et al., 1984) and is involved in branch point formation in the lariat structure (Pikielny et al., 1983; Domdey et al., 1984; Rodriguez et al., 1984). The lariat branch point maps at the sequence TGCTAAC in rabbit β-globin precursor RNA (Zeitlin and Efstratiadis, 1984) which is 31nuc upstream from the splice acceptor site and fits the consensus YNYTRAY proposed by Ruskin et al. (1984). However, precursor RNAs are accurately spliced in vitro following the removal of sequences at the branch point within the human β-globin gene (Ruskin et al., 1985). The deletion of these sequences activates cryptic branch points which are located 22 to 37nuc upstream from the splice acceptor site.

c) Termination of transcription.

The possible mechanisms and DNA sequences which are required to form mature mRNA 3' termini are described in detail in Section F (Page 80).

19. Arrangements of Transcription Units.

As described in the previous section, the processes and DNA sequences required for production of a mature mRNA are numerous, however, an additional level of complexity is created by overlapping mRNAs which share common transcription control signals. Many different arrangements of overlapping mRNAs have been identified and these range from simple transcription units which specify a single message to more complex arrangements from which a number of mRNA species can be produced by differential selection of alternative termini and splice sites. The following sections describe possible arrangements of mRNAs and provide examples for each type.

a) Simple transcription units.

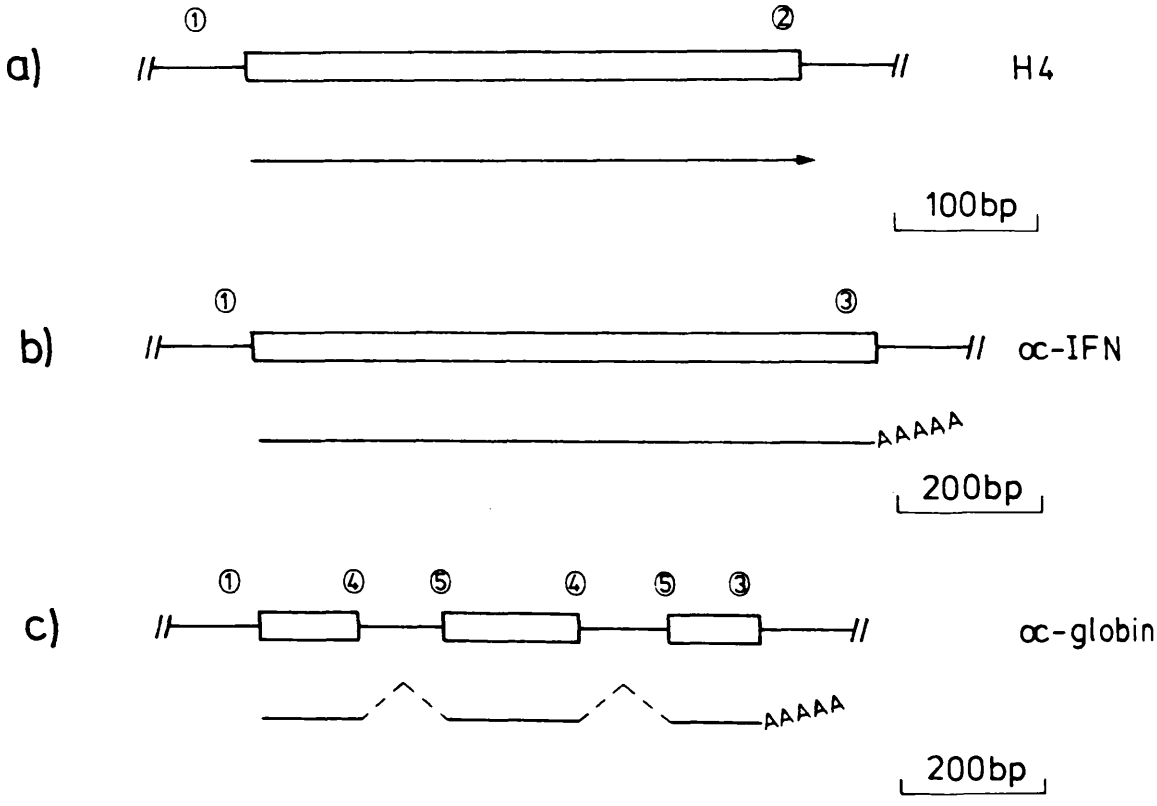
Simple transcription units do not contain alternative control sites which can be differentially selected and thus specify a single mRNA species. Previously, simple transcription units have been defined as encoding a single polypeptide (Darnell, 1982), however, this definition appears inadequate as the selection of alternative control sites outwith a protein coding region could generate a number of different mRNA species which

encode the same protein. For example, the Ad5 DNA binding protein is specified by two mRNAs which have different 5' termini at early and late times p.i. (Kruijer et al., 1981). Simple transcription units may be further categorised according to the processes required to form a mature mRNA (Fig. 7).

1) mRNAs which are neither polyadenylated nor spliced. Most histone mRNAs (reviewed in Hentschel and Birnstiel, 1981) are typical of this group. While the majority of histone mRNAs are not polyadenylated, their mature 3' termini are generated by post-transcriptional modification of a precursor mRNA (see Page 89).

2) Unspliced mRNAs which are polyadenylated. Examples of this type include HSV-1 and HSV-2 IE mRNA-2 (Whitton et al., 1983), HSV-1 IE mRNA-3 (Rixon et al., 1982), the human (Nagata et al., 1980) and mouse α-interferons (Shaw et al., 1983).

3) mRNAs which are both spliced and polyadenylated. A number of mammalian transcripts including the α - (Nishioka and Leder, 1979) and β-globins (Konkel et al., 1978) are produced in this way as is HSV-1 IE mRNA-1 (see Page 35; F.J. Rixon, personal communication).



1. TATAA box
2. palindrome / ACC(C)A
3. AATAAA signal
4. GT-splice donor site
5. AG-splice acceptor site

Figure 7. Arrangements of simple transcription units. Boxed regions indicate the regions of DNA present in mature mRNA. Poly A tails are shown by 'AAAA' and dashed lines indicate regions in precursor RNA which are removed by splicing. The locations of certain control signals, numbered 1 to 5, which are involved in mRNA synthesis and processing are indicated.

a) mRNAs which are neither polyadenylated nor spliced. The example shown is the human histone H4 gene (Heintz et al., 1981).

b) unspliced mRNAs which are polyadenylated. The example given is the human ∝-interferon gene (Nagata et al., 1980).

c) spliced mRNAs which are polyadenylated. The example shown is the mouse ∝-globin (Nishioka and Leder, 1979).

b) Complex transcription units.

A consequence of mRNAs sharing common signals is that fewer transcription control signals may be necessary for expression of a number of polypeptides. This form of gene organisation has been utilised by the adenovirus (reviewed in Shenk and Williams, 1984), papovavirus (Fiers et al., 1978) and retrovirus groups (Coffin, 1982) whose genomes contain tightly-packed translated regions. Some functionally related cellular proteins are also produced from mRNAs generated by differential selection of control sites. There are a total of seven possible arrangements of overlapping mRNAs, each of which is described below and shown in Fig. 8.

1) mRNAs with alternative 3' termini. Overlapping transcripts which have common 5' termini and identical splicing patterns but different poly A sites invariably encode the same protein. Alternative poly A sites are located in the untranslated regions of genes specifying chick pro α^2(1) collagen (Aho et al., 1983), β 2-microglobulin (Parnes et al., 1983) and mouse dihydrofolate reductase (Setzer et al., 1980). In most examples, the multiple 3' termini are selected with different efficiencies. The distribution of mRNAs with different 3' termini for the β 2-microglobulin gene, which contains three possible poly A sites, is identical in a number of tissues (Parnes et al., 1983). However, the sequence content of the different 3' untranslated regions

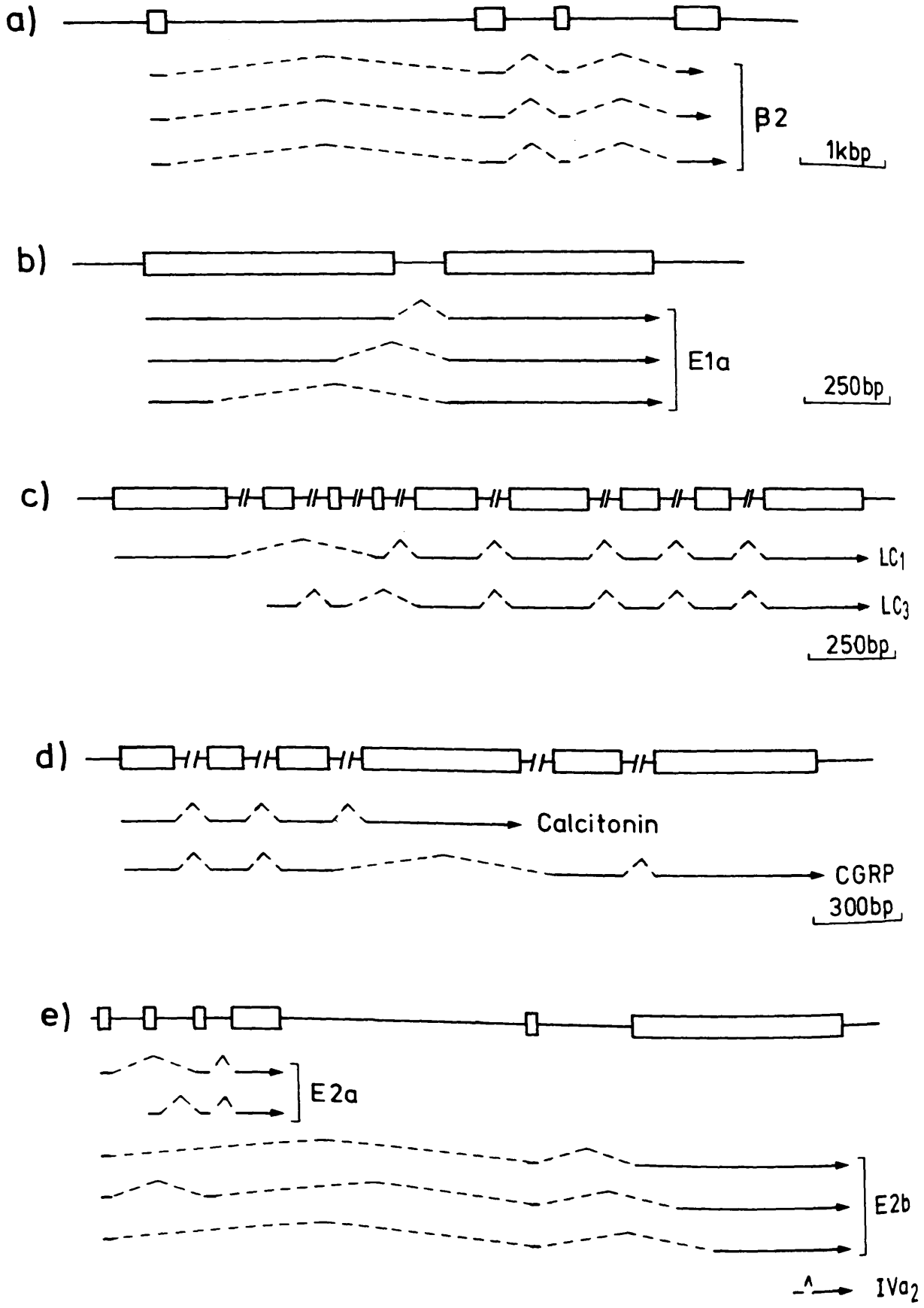

a)
β2
1kbp
b)
E1a
250bp
c)
LC1
LC3
250bp
d)
Calcitonin
CGRP
300bp
e)
E2a
E2b
IVa2

Figure 8. Arrangements of mRNAs in complex transcription units. The organisation of exons (boxed regions) and introns (solid lines) on genomic DNA is shown above the arrangements of mRNAs. Dashed lines indicate sequences removed by splicing and solid lines interrupted by (//) in (c) and (d) denote introns of greater than 500nuc. Examples of mRNA families with alternative 5' termini, and alternative 5' and 3' termini are not given here; overlapping transcripts with these arrangements are shown in Fig. 4. The genes described below are drawn to scale apart from the Ad2 E2 transcription unit in (e). References for the mapping of these genes are given in the text.

a) shows the alternative 3' termini of the β2-microglobulin (β2) transcripts.

b) indicates the alternative splicing within Ad2 E1a transcription unit.

c) shows the alternative 5' termini and splice sites used in the myosin light chain (LC) gene, generating transcripts LC1 and LC3.

d) shows the alternative poly A sites and splice sites for the rat calcitonin gene. CGRP is the abbreviation for calcitonin gene-related peptide.

e) indicates the arrangement of the mRNAs in the Ad2 E2 transcription unit, comprising early families E2a and E2b, and the IVa_2 gene.

may affect either the subcellular location or stability of β2-microglobulin transcripts; to date, these possibilities remain untested.

2) mRNAs with alternative splice sites. In most cases, differential splicing allows expression of more than one protein from a region of DNA which specifies transcripts with common 5' and 3' termini. The use of alternative splice sites within a protein coding region may generate transcripts which encode polypeptides with common epitopes. For example, the mouse αA^{ins}-crystallin gene has an identical amino acid sequence to the αA_2-crystallin protein except for an additional 22 amino acids between residues 63 and 64 (Cohen et al., 1978; de Jong et al., 1980). The exon containing this small peptide is inserted into the αA_2-crystallin protein by the selection of splice sites within the first intron of the αA_2-crystallin gene (King and Piatigorsky, 1983). A further example is the Ad Ela region which specifies three mRNAs; two transcripts (13S and 12S) are early species while the third (9S) is synthesised only at late times (Berk and Sharp, 1978b; Chow et al., 1979; Perricaudet et al., 1979). Ela transcripts share common 5' and 3' termini, however, the different mRNAs are produced by splicing between three potential splice donor sites and one common splice acceptor site. The splice donor and acceptor sites are located in the translated regions of all three mRNAs, thus the different Ela polypeptides share common amino acid residues

(Perricaudet et al., 1979). The 13S transcript encodes a 51,000 mol. wt. protein which stimulates transcription of Ad genes during virus infection (Ricciardi et al., 1981; Montell et al., 1982). The 12S mRNA specifies a 48,000 mol. wt. polypeptide which is 46 amino acids shorter than the 51,000 mol. wt. product. To date, the splice donor site for the 9S mRNA has not been accurately positioned and a protein specified by this transcript has not been detected in infected cells.

3) mRNAs with alternative 5' termini. The majority of overlapping HSV mRNAs which are 3' co-terminal are unspliced and the individual members of each family have unique 5' termini (reviewed in Wagner, 1985). A description of the transcripts encoded by HSV DNA is given in Section 8 (Page 36).

4) mRNAs with alternative 5' and 3' termini. Again, the HSV genome specifies mRNA families with this type of arrangement. In addition, the transcripts of vaccinia virus are unspliced and form families of nested mRNAs (Cooper et al., 1981; Mahr and Roberts, 1984a, b; Morgan and Roberts, 1984).

5) mRNAs with alternative 5' termini and splice sites. In chicken muscle cells, myosin light chains are encoded by a single gene of size 18kbp (Nabeshima et al., 1984). From this gene, two distinct light chain proteins, LC1 and LC3,

are produced by differential splicing of two precursor RNAs which have different initiation sites (Nabeshima et al., 1984). LC1 mRNA differs from the LC3 transcript only at the 5' terminal region where the two 5' proximal exons of each mRNA species are encoded by unique DNA segments. The remaining exons are common to both transcripts and consequently, the predicted amino acid sequences for LC1 and LC3 proteins are identical except for the N-terminal region.

6) mRNAs with alternative 3' termini and splice sites. The Ad2 E3 transcription unit contains three possible poly A sites, two splice donor sites and five splice acceptor sites (Berk and Sharp, 1978b; Chow et al., 1979). In total, eight different E3 mRNAs have been identified, each with an identical 5' terminus. In eukaryotes, similar mRNA arrangements have been described for the Ig heavy chain genes (Alt et al., 1980; Early et al., 1980), the glycinamide ribotide transformylase gene of Drosophila melanogaster (D. melanogaster; Henikoff et al., 1983b) and the calcitonin gene (Amara et al., 1982 and 1984; Edbrooke et al., 1985).

7) mRNAs with alternative termini and splice sites. Such complex organisation occurs in the Ad E2 transcription unit which comprises two mRNA families designated E2a and E2b that utilise different poly A sites (Stillman et al., 1981); the poly A site for E2a mRNA is located within the

transcribed region of the E2b family. The E2a region specifies two mRNA species with different 5' termini and 5' terminal leader sequences which are selected at different times during the virus growth cycle (Kruijer et al., 1981). The E2b region produces three mRNA species which have identical leader sequences but differ in the location of their 3' proximal splice acceptor sites. A third promoter, within the coding sequences of the E2b transcripts, specifies the 5' terminus of IVa_2 mRNA.

These seven arrangements for overlapping trancripts relate only to those mRNAs synthesised from the same DNA strand. Overlaps occur between transcripts produced from opposite strands of DNA, for example transcription of the Ad E2 and late transcription units (reviewed in Shenk and Williams, 1984). Divergently transcribed mRNAs with overlapping 5' termini have been identified in the HSV-1 U_S region (Rixon and McGeoch, 1985) and the 3' termini of SV40 early and late mRNAs also overlap (Fiers et al., 1978).

20. Regulation of Overlapping mRNAs.

a) Poly A site selection in immunoglobulins.

The heavy chains of antibodies are composed of variable (V_H) and constant (C_H) regions; the C_H regions encode the 3' portion of Ig mRNA and contain at least one poly A site. The effector or class of antibody is determined by the selection of one out of eight possible C_H

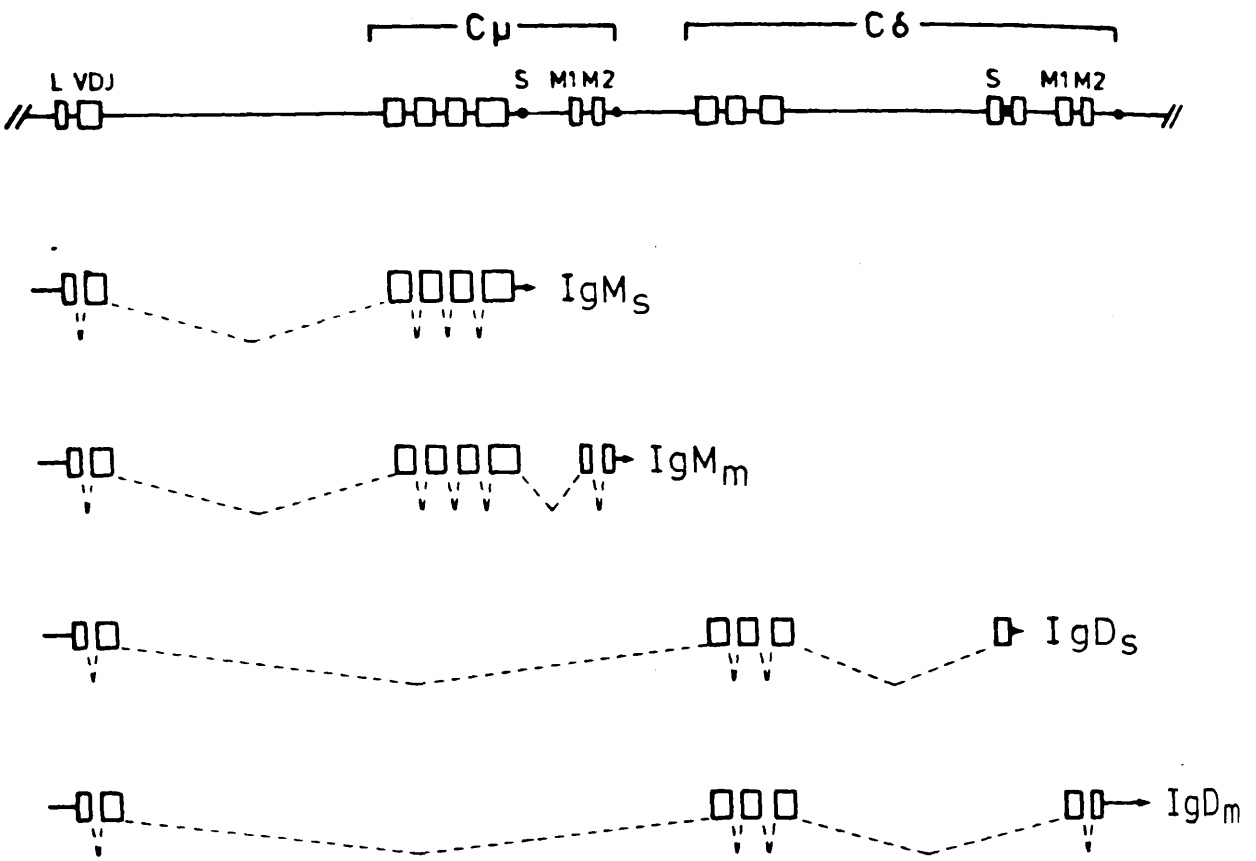


Figure 9. Differential selection of splice sites and 3' termini for IgM and IgD mRNAs. Locations of the polypeptide coding regions for variable (represented by L, V, D and J) and constant (Cμ and Cδ) regions on genomic DNA are shown by boxed regions on the top line. Positions for the poly A sites of membrane-bound (M) and secreted forms (S) of IgM and IgD are indicated by closed circles. Dashed lines denote the regions removed from precursor RNAs by splicing.

regions (reviewed in Marcu, 1982 and Shimizu et al., 1982). In most cases, a change in the class of antibody is due to DNA rearrangements mediated through switch recombination sites which juxtapose the V_H region to an alternative C_H region. Thus, removal of potential poly A sites by DNA rearrangements influences the selection of Ig mRNA 3' termini. In contrast, production of IgD, which is co-expressed with IgM in mature resting B cells (Cooper et al., 1976; Pernis et al., 1976), does not require the deletion of DNA sequences; rather, IgD expression is determined by differential processing of precursor mRNAs (Maki et al., 1981; Knapp et al., 1982). Both IgM and IgD exist as secretory and membrane-bound species and the species expressed is governed by poly A site selection and differential splicing processes. Therefore, there are at least four potential poly A sites involved in IgM and IgD expression; the 5' proximal poly A sites produce mRNAs encoding IgM_s and IgM_m respectively (Fig. 9; Alt et al., 1980; Early et al., 1980) while the 5' distal poly A sites generate transcripts specifying IgD_s and IgD_m respectively (Cheng et al., 1982).

It has been suggested that poly A site selection in IgM and IgD transcription is regulated by alterations in the level of an 'endase' enzyme which cleaves precursor RNAs at poly A sites (Blattner and Tucker, 1984). In resting B cells, this enzyme would be present at low levels and the choice of poly A site is determined by an affinity for the 'endase'. B cell differentiation increases the

level of 'endase' and subsequently the number of cleavage events at poly A sites on the precursor RNA. Thus, the poly A site proximal to the mRNA 5' terminus is preferentially selected and cleavage at this site defines the mRNA 3' terminus irrespective of further downstream cleavage events. Such a model would account for a number of observations including the disappearance of IgD following activation of resting B cells and the concomitant rise in IgM_S (Parker et al., 1979; Isakson et al., 1981) whose poly A site is the most proximal to the mRNA 5' terminus. Furthermore, putative RNA processing intermediates have been detected which correspond to the distance between two poly A sites; such products would be predicted from more than one cleavage event on a precursor RNA molecule (Kemp et al., 1983).

b) Formation of differentially spliced mRNAs.

Two possible mechanisms for the selection of alternative splice sites have been proposed (Sharp, 1981; Kuhne et al., 1983). Both models suggest that nucleotide sequences flanking the splice donor and acceptor sites influence the efficiency of splicing at a particular site, however, they differ on the additional factors which determine splice site selection.

The first model (Sharp, 1981) proposes a 'scanning-type' mechanism whereby a splicing complex attaches to a splice donor or acceptor site and moves along the RNA molecule until it encounters a splice acceptor or

donor site at which point the intervening sequence between the splice sites is removed. Further scanning of the RNA by the complex and additional splicing reactions may give rise to a number of mRNAs with different splicing patterns. Evidence for this proposal is derived from the probable stepwise excision of introns from procollagen RNA (Avvedimento et al., 1980).

To test this model, Kuhne et al. (1983) inserted a duplication of the rabbit β-globin splice donor site upstream from the normal donor site and, in a separate construction, the splice acceptor site was positioned downstream from the normal acceptor site. The 'scanning-type' model predicts that donor and acceptor sites proximal to the 5' and 3' termini would be utilised. However, analysis of the resultant β-globin mRNAs indicated that splicing occurred at splice sites distal to the 5' and 3' termini, thus arguing against a 'scanning-type' mechanism. From these data, a separate model has been proposed suggesting that the relative positions of splice sites determines the probability of a splicing reaction occurring between them. Therefore, stability of both the enzyme-splicing region complexes and the secondary or tertiary structure of the RNA would have a role in splice site selection. This model would account for the different splicing patterns of mRNAs which have alternative 5' and 3' termini such as the membrane and secretory forms of IgD and IgM, the calcitonin gene, the glycinamide ribotide transformylase gene of D. melanogaster

and the myosin light chain gene (see Pages 72 and 73 for references). However, neither model explains the shift in splicing pattern of Ad L1 mRNAs during the course of viral infection (Akusjarvi and Persson, 1981a).

c) Tissue-specific selection of termini and splice sites.

Expression of some eukaryotic genes which specify more than one transcript is controlled by tissue specificity. The chicken myosin light chain gene comprises LC1 and LC3 mRNAs which have different 5' termini and splicing patterns (Page 72; Nabeshima et al., 1984). Both transcripts are produced in skeletal muscle, however, cardiac muscle expresses a light chain homologous to LC1 while a light chain homologous to LC3 is synthesised in gizzard muscle. Similarly, the hormone calcitonin is produced in thyroid 'C' cells while the calcitonin gene-related peptide (CGRP) is expressed in brain cells (Rosenfeld et al., 1983). The mRNAs which encode these polypeptides have common 5' termini and the N-terminal protein sequences are identical (Fig. 8; Amara et al., 1982 and 1984; Edbrooke et al., 1985). Differential selection of two poly A sites and alternative splicing pathways enables expression of the unique C-terminal portions of both proteins. Selection of the 3' distal poly A site produces a CGRP-specific transcript, however, mRNAs which terminate at the 5' proximal poly A site encode calcitonin. Calcitonin-specific coding sequences and the 5' proximal

poly A site are removed from the CGRP-specific mRNA by splicing. The mouse ∝-amylase also is encoded by two tissue-specific transcripts (Young et al., 1981). In this case, the mRNAs utilise alternative 5' termini in the liver and parotid gland and have different splicing patterns. These differences in the mRNAs occur in the 5' untranslated region and therefore both transcripts contain identical protein coding sequences. The promoters which specify the different ∝-amylase 5' termini are differentially activated during parotid gland development (Shaw et al., 1985).

SECTION F.

RNA TERMINATION AND 3' END PROCESSING IN EUKARYOTES.

The processes and DNA sequences involved in termination of transcription and formation of mature RNA 3' termini are not yet fully understood. There are two possible mechanisms by which RNA 3' termini can be generated:

1) termination at the RNA 3' terminus and no further cleavage of the transcript. Yeast RNA may be further processed by the addition of a poly A tail to the 3' terminus (Zaret and Sherman, 1982).

2) synthesis of a precursor RNA which is cleaved to produce the mature 3' terminus. This 3' terminus may be further processed by addition of either a poly A tail to poly A mRNAs and some histone messages or a terminal CCA sequence to tRNAs.

The DNA sequences which control these events have been partially characterised, however their functions remain unclear. The following sections review the processes and DNA sequences involved in the 3' end formation of RNAs transcribed by RNA polymerases I, II and III.

21. Termination and Processing of RNA Polymerase I Transcripts.

RNA polymerase I transcribes a precursor RNA which is subsequently cleaved by a number of pathways to generate the mature ribosomal RNAs (Bowman et al., 1981; Fetherston

et al., 1984). The only rRNA not produced by RNA polymerase I is the 5S species which is synthesised by RNA polymerase III (McReynolds and Penman, 1974). Xenopus laevis (X. laevis) precursor RNA sediments at 40S (Wellauer and Dawid, 1974) and is cleaved to produce the mature 28S, 18S and 5.8S rRNAs (Landesman and Gross, 1968; Speirs and Birnstiel, 1974; Wellauer and Dawid, 1974). By nuclease S1 analysis, the 3' termini of 40S and 28S RNAs are identical and therefore, following termination of 40S RNA synthesis, there are no further cleavage events at the 3' terminus (Sollner-Webb and Reeder, 1979).

DNA sequences at the termination site consist of a cluster of T residues in a G+C-rich environment (Sollner-Webb and Reeder, 1979). Measurement of the length of transcripts produced by plasmids with deleted sequences immediately downstream from the 3' terminus has shown that the cluster of T residues is not only the termination site but also delimits the 3' boundary of the termination signal (Bakken et al., 1982). Plasmids in which the distal T residue of a four T residue cluster was deleted continued to produce correctly-terminated RNA. However, removal of an additional T residue abolished termination at this site and extended RNAs were detected. The presence of stretches of T residues within rRNA coding sequences suggests that T residues alone are not sufficient for termination (Sollner-Webb and Reeder, 1979; Gourse and Gerbi, 1980); at present, no additional signals have been identified.

22. Termination and Processing of RNA Polymerase III Transcripts.

RNA polymerase III is responsible for transcription of a number of RNA species including tRNAs, 5S RNAs (McReynolds and Penman, 1974), Alu RNAs (Haynes and Jelinek, 1981; Pan et al., 1981), virus-associated RNAs such as Ad VAI and VAII (Soderland et al., 1976) and other snRNAs (reviewed in Busch et al., 1982).

In almost all cases, the termination sites of 5S RNA transcripts are located within a cluster of T residues (Brown and Brown, 1976; Korn and Brown, 1978). Analysis of X. laevis 5S RNA in vitro has indicated that the DNA sequences involved in termination are very similar to those signals required for termination by RNA polymerase I, namely a cluster of T residues flanked by G or C nucleotides; stretches of A residues in the vicinity of the T cluster will reduce termination efficiency (Bogenhagen and Brown, 1981). Furthermore, no extensive sequences flanking the T cluster are necessary for efficient termination. The minimum sequence length required is 2 to 5nuc on the 5' side of the T cluster and 2nuc on the 3' side. Plasmids with deletions extending into the T cluster fail to terminate at the correct site. A common feature of Xenopus 5S genes is the presence of secondary termination sites downstream from the preferred site; these secondary termination sites are thought to prevent read-through transcription beyond the region encoding 5S RNA (Korn and Brown, 1978).

The 3' termini of precursor and mature 5S transcripts are identical in X. laevis (Birkenmeier et al., 1978), however, post-transcriptional processing does occur at the 3' termini of 5S RNAs in other organisms. Isolated nuclei from rat liver and HeLa cells synthesise a precursor RNA species which is 8nuc longer at the 3' terminus than the mature 5S form (Hamada et al., 1979) and, in D. melanogaster, 15nuc are removed from a precursor transcript to generate the mature 5S species (Jacq et al., 1977). In yeast, up to 10nuc are removed from the precursor 5S RNA (Tekamp et al., 1980).

Mature tRNA 3' termini are generated by endonuclease cleavage of a precursor form and by addition of the trinucleotide sequence CCA (Melton et al., 1980; Frendewey et al., 1985). Presumed terminator sequences in the precursor RNA are T-rich and therefore similar to the sequences at 5S RNA 3' termini (Hosbach et al., 1980). Efficient synthesis in vitro of a D. melanogaster $tRNA^{ARG}$ gene requires approximately 35bp of flanking sequences and additional factors other than RNA polymerase III are essential for formation of a stable transcription complex (Schaak et al., 1983 and 1984). This requirement for 3' flanking sequences in a tRNA gene contrasts with that of X. laevis 5S RNA in which removal of flanking sequences does not affect termination efficiency (Bogenhagen and Brown, 1981). Moreover, X. laevis RNA polymerase III, purified to greater than 90% homogeneity is able to accurately and efficiently terminate 5S RNA transcription (Cozzarelli et

al., 1983). A possible explanation for these differences is that the 3' flanking sequences and binding factors are required for processing of precursor $tRNA^{ARG}$ 3' terminus; the 3' terminus of the X. laevis 5S primary transcript is not modified and therefore may not require flanking sequences and additional factors other than RNA polymerase III.

23. Termination and Processing of RNA Polymerase II Transcripts.

RNA polymerase II is responsible for the synthesis of poly A mRNAs, histone messages (see Page 60) and snRNAs which have a high uridylic acid content (U-snRNAs; reviewed in Busch et al., 1982). Current models for 3' end formation in higher eukaryotes and animal DNA viruses suggests that mature 3' termini are generated by post-transcriptional cleavage of a longer precursor RNA, however, these processes are poorly understood. For poly A mRNAs, a stretch of up to 200 A residues is then added to the mature 3' terminus. The following sections describe those processes and DNA sequences which are involved in 3' end formation of poly A and histone mRNAs.

a) Termination of transcription.

Studies on a number of genes have shown that transcription continues beyond the poly A site (reviewed in Darnell, 1982 and Nevins, 1983). Examination of nuclear RNA produced at late times in Ad2-infected HeLa cells indicated

that the mature 3' termini did not correspond to the RNA polymerase termination site. Primary transcripts, which are initiated at the major late promoter, extend across most of the Ad2 genome from map coordinate 16.4 to between map coordinates 91.9 and 99.0 (Nevins and Darnell, 1978; Fraser et al., 1979); mature mRNA 3' termini are generated by selection of one of five possible poly A sites from within the primary transcripts. Again in Ad2, the primary transcripts for the E2 and E4 early transcription units extend for some distance beyond their respective poly A sites (Nevins et al., 1980). Analysis of late mRNAs in SV40-infected cells by short-term pulse-labelling methods has demonstrated that transcription continues for at least 1kb beyond the late RNA poly A site (Ford and Hsu, 1978). Termination has been studied extensively in the mouse β^{maj}-globin gene. RNase fingerprint analysis of in vitro-labelled RNA from isolated nuclei and hybridisation of in vivo-labelled nuclear RNA to sequences spanning the β^{maj}-globin gene have indicated that nascent RNA contains sequences downstream from the poly A site (Hofer and Darnell, 1981). Further analysis of this gene has focussed on identification of a downstream termination site and recent evidence suggests that termination occurs within a region 0.7 to 2kb downstream from the poly A site but not at particularly favoured sites (Citron et al., 1984). This downstream region contains A-rich and T-rich segments which are conserved in other globin genes, however their function is unknown. A portion of the DNA sequence which is not

transcribed and is approximately 2kb downstream from the poly A site contains a stem and loop structure followed by a stretch of fifteen consecutive T residues (Citron et al., 1984). Insertion of a mouse β^{maj}-globin fragment, which spans the poly A signal and proposed termination region, into the Ad5 Ela transcription unit generates mRNA which is processed exclusively at the β-globin poly A site (Falck-Pedersen et al., 1985). In contrast, transcripts are processed at the normal Ela poly A site in recombinants which contain the proposed β-globin termination region but lack the β-globin poly A signal. Thus, discrete termination events appear not to occur in the absence of a functional poly A site.

Mature histone mRNA 3' termini also are generated by processing of a primary transcript. RNA precursors, synthesised in vitro and subsequently injected into Xenopus oocytes, are processed to form mature 3' termini and it has been suggested that termination occurs heterogeneously within a 100 to 200nuc stretch of spacer DNA (Birchmeier et al., 1984; Krieg and Melton, 1984).

b) 3' end formation of poly A mRNAs.

The DNA sequences required for 3' end formation of poly A mRNAs can be divided into two categories: 1) the polyadenylation signal and 2) DNA sequences downstream from the polyadenylation signal.

1) The polyadenylation signal. The polyadenylation signal, AATAAA, is located 10 to 30nuc upstream from the poly A site of poly A mRNAs and is well conserved in mammalian and viral genes (Proudfoot and Brownlee, 1976). The functional significance of this sequence has been demonstrated in a number of studies. SV40 mutants which lack an AATAAA signal in the late transcription unit fail to produce late poly A mRNA (Fitzgerald and Shenk, 1981). In other mutants which retain an AATAAA signal, removal of sequences between the AATAAA sequence and the poly A site does not dramatically affect SV40 late mRNA production. The mRNA 3' termini from these constructions are located at a similar distance from the AATAAA signal as in wild-type SV40 late mRNA, thus sequences between the AATAAA signal and the poly A site do not appear to be necessary for positioning the 3' terminus. Recent studies with these SV40 mutants indicate that deletion of sequences downstream from the poly A signal decreases the efficiency with which the normal poly A site is utilised (Sadofsky and Alwine, 1984; Conway and Wickens, 1985).

Point mutations generated in **vitro** within the poly A signals of SV40 late and Ad2 Ela transcription units dramatically reduce the level of detectable poly A mRNA (Montell et al., 1983; Wickens and Stephenson, 1984). The altered AATAAA sequences examined in SV40 were AACAAA, AATTAA, AATACA and AATGAA (Wickens and Stephenson, 1984) and in Ad2 the mutated sequence AAGAAA was analysed (Montell et al., 1983). These point mutations did not

prevent polyadenylation; the low level of detectable mRNA was polyadenylated, indicating that the poly A signal is possibly involved in cleavage of the primary transcript. The failure to detect cleaved RNA products which were not polyadenylated further suggests that polyadenylation rapidly follows cleavage and the two reactions may be coupled (Manley et al., 1982; Wickens and Stephenson, 1984). Evidence from in vivo studies reveal that a mutated signal with the DNA sequence AATAAG results in α-thalassaemia (Higgs et al., 1983). α-globin RNA produced in vivo extends beyond the normal poly A site and is unstable giving rise to deficient α-globin production.

Certain alterations within the AATAAA sequence can be tolerated. For example, a high proportion of RNA molecules are processed at the major poly A site of the X. laevis β1 globin gene following mutation of the AATAAA sequence to AATACA (Mason et al., 1985) and a number of genes utilise the sequence AATTAA as a poly A signal (Wickens and Stephenson, 1985).

2) DNA sequences downstream from the poly A signal. The hexanucleotide sequence, AATAAA, occurs in regions of DNA which do not contain a poly A site (Fiers et al., 1978; Perricaudet et al., 1980; Tosi et al., 1981), therefore, sequences other than AATAAA must be involved in determining the location of 3' termini. Benoist et al., (1980) have described a model sequence TTTTCACTGC which was present at the 3' termini of five out of nine mRNAs. In another

report, Taya et al., (1982) noted that the DNA sequence GTGTTG was present at approximately 10nuc downstream from the 3' termini of six mRNAs.

Analysis of transcripts produced by plasmids with deletions at 3' termini has indicated that DNA sequences downstream from the AATAAA signal are required for efficient cleavage and polyadenylation. Removal of 3' flanking sequences from the Ad5 E2a transcription unit has demonstrated that 35bp distal to the poly A site are necessary for the efficient formation of mature mRNA 3' termini (McDevitt et al., 1984). Deletion of an additional 15bp abolishes the ability to produce functional E2a mRNA and the level of correctly-terminated mRNA is reduced but remains detectable. Plasmids with a further deletion of 22bp, removing the poly A site, fail to produce correctly-terminated transcripts. Efficient production of rabbit β-globin 3' termini requires 31bp of 3' flanking DNA sequences (Gil and Proudfoot, 1984) and an 18bp region immediately downstream from the HSV-1 TK mRNA poly A site is critical for 3' end formation (Cole and Stacy, 1985). Sequences at a similar distance from the poly A site of SV40 late transcripts have been implicated in efficient 3' end processing (Sadofsky and Alwine, 1984; Conway and Wickens, 1985).

c) 3' end formation in histone mRNAs.

The majority of histone mRNAs are not polyadenylated (Adesnik and Darnell, 1972; Greenberg and

perry, 1972), however, examples of polyadenylated histone transcripts include X. laevis H4 mRNA, chicken H5 mRNA (Krieg et al., 1983) and yeast histone messages (Fahrner et al., 1980). DNA sequence comparisons have revealed that many but not all histone mRNAs have a conserved sequence ACC(C)A at the 3' terminus and a short palindromic sequence upstream from the 3' end which may form a stem and loop structure (Busslinger et al., 1979; Hentschel et al., 1980; Childs et al., 1982; Turner and Woodland, 1982; Krieg et al., 1983). Insertion of point mutations within the stem portion of this structure abolishes the ability to generate correctly-terminated sea urchin H2A mRNA following micro-injection of the mutated DNA into Xenopus oocytes (Birchmeier et al., 1983). Pseudorevertants with complementary point mutations within the palindrome produce authentic 3' termini, suggesting that the presence of a stem and loop structure within precursor RNA is essential for 3' end formation rather than a requirement for specific DNA sequences (Birchmeier et al., 1983).

In addition to the dyad symmetry and the 3' terminal ACC(C)A motif, sequences up to 78nuc downstream from the 3' terminus are essential for efficient 3' end production (Birchmeier et al., 1982 and 1983). Within this 78nuc segment, correctly-processed mRNA is produced at a sub-optimal level by the 24nuc proximal to the 3' end (Birchmeier et al., 1983). This 24nuc sequence contains a CAAGAAAGA motif which is highly conserved in sea urchin histone mRNA species and is located approximately 8nuc

downstream from 3' termini (Busslinger et al., 1979; Birnstiel et al., 1985). This sequence motif is essential for RNA processing and increasing the distance between the CAAGAAAGA sequence and the palindrome abolishes 3' end formation (Georgiev and Birnstiel, 1985).

d) Termination and polyadenylation in lower eukaryotes.

Little is known of the DNA sequences and processing events which are essential for the generation of yeast mRNA 3' termini. The AATAAA sequence which is essential for processing of higher eukaryote poly A mRNAs is not present at the 3' termini of most yeast mRNAs (Zaret and Sherman, 1982). A spontaneous mutant of Saccharomyces cerevisae (S. cerevisae) has a 38bp deletion downstream from the translated region of the CYC1 locus. Absence of these DNA sequences results in the production of extended transcripts which are polyadenylated, leading Zaret and Sherman (1982 and 1984) to propose that polyadenylation of yeast mRNAs is directly coupled to termination and that primary transcripts are not cleaved. From comparisons of this 38bp sequence to regions upstream from other yeast poly A sites, alternative consensus sequences have been proposed for termination (Zaret and Sherman, 1982; Bennetzen and Hall, 1982). A third possible consensus sequence has been suggested from a separate study which examined the termination and polyadenylation of a Drosophila DNA segment in yeast using deletion mutant analysis (Henikoff et al., 1983a). The three proposed consensus sequences are;

1) (T-rich)..TAG.....TA(T)GT..(A/T-rich)..TTT (Zaret and Sherman, 1982).

2) TTTTTATA (Henikoff et al., 1983a; Henikoff and Cohen, 1984).

3) TAAATAA$\begin{matrix}A\\G\end{matrix}$ (Bennetzen and Hall, 1982).

24. In Vitro Formation of mRNA 3' Termini.

Synthesis in vitro of mRNA species in whole-cell extracts has demonstrated that efficient cleavage and polyadenylation does not require primary transcripts which extend to specific termination sites. Precursor RNAs synthesised in vitro from DNA templates which are cleaved either at or just downstream from the mature mRNA 3' terminus are accurately and efficiently polyadenylated by HeLa cell extracts (Manley, 1983; Manley et al., 1985). In contrast, longer precursor RNAs are processed inefficiently and the poly A tail is added to the 3' terminus of uncleaved primary transcripts (Moore and Sharp, 1984; Manley et al., 1985). However, cell extracts, possessing both transcription and processing activities, efficiently produce correctly-processed mRNAs from precursor RNAs synthesised from DNA templates which extend at least 50bp beyond the mature 3' end; these results suggest that processing rapidly follows transcription beyond the poly A site (Moore and Sharp, 1984).

Processing of precursor RNAs probably requires cleavage of the primary transcript followed by polyadenylation. Purified poly A polymerase does not

possess either of these activities (Brawerman, 1981). The inhibition of cleavage/polyadenylation by mouse monoclonal antibodies which immunoprecipitate ribonucleoproteins containing U1, U2, U4, U5 and U6 RNAs suggests a role for snRNAs in 3' end processing (Moore and Sharp, 1984) and it has been proposed that U4 snRNA may hybridise to complementary sequences at poly A sites (Berget, 1984). Processing of a histone H3 precursor RNA synthesised *in vitro* requires the prior injection into Xenopus oocytes of a 60nuc RNA (termed U7 RNA; Strub *et al.*, 1984) which appears to be a component of a snRNP in sea urchin embryos (Galli *et al.*, 1983; Birchmeier *et al.*, 1984). Similar experiments with purified Drosophila H3 precursor RNA demonstrated that authentic 3' termini were produced by a nuclear extract from Drosophila cultured cells (Price and Parker, 1984). This processing activity was isolated and two RNA species were present in highly purified preparations.

MATERIALS AND METHODS

SECTION A.

MATERIALS

1. Viruses.

HSV-2 strain HG52 (Timbury, 1971) and HSV-1 Glasgow strain 17^+ (Brown et al., 1973) were used during the course of this work.

2. Cells.

Baby hamster kidney cells (BHK-C13; MacPherson and Stoker, 1962) were used for growth and assay of virus and in the preparation of viral RNA. HeLa cells (a gift from R. Everett) were used for short-term transfection assays of plasmid constructions.

3. Tissue Culture Media.

BHK cells were propagated using modified Eagle's medium (Busby et al., 1964) containing 100µg/ml streptomycin, 100units/ml penicillin, 0.2µg/ml n-butyl p-hydroxybenzoate and 0.002% (w/v) phenol red. The growth medium, ETC10, comprised 80% Eagle's, 10% tryptose phosphate and 10% calf serum. To titrate virus, tryptose phosphate and calf serum were replaced with human serum which was added to a final concentration of 5% to give EHu5.

HeLa cells were grown in Dulbecco's medium (Flow Laboratories), a modified version of Eagle's medium,

supplemented with 20mM glutamine, 100units/ml of streptomycin and penicillin and 0.55ml of 0.025% (w/v) amphotericin. In addition to the above ingredients, growth medium contained 2.5% calf serum and 2.5% foetal calf serum (Gibco Ltd).

PBS contained 170mM NaCl, 3.4mM KCl, 1mM Na_2HPO_4, 2mM KH_2PO_4 (pH 7.2) in distilled water. Versene consisted of PBS containing 6mM EDTA with 0.0015% (w/v) phenol red. Trypsin comprised 25% (w/v) Difco trypsin dissolved in Tris-saline. The constituents of Tris-saline were 140mM NaCl, 30mM KCl, 0.28mM Na_2HPO_4, 1mg/ml dextrose, 25mM Tris HCl (pH 7.4), 0.005% (w/v) phenol red supplemented with 100units/ml penicillin and 100µg/ml streptomycin.

4. Bacterial Culture Media.

Bacteria were grown in L-broth which consisted of 177mM NaCl, 10g/l bactopeptone and 5g/l yeast extract (pH 7.5, presterilisation). For agar plates, L-broth contained 1.5% (w/v) agar. Where appropriate, ampicillin was added to both the culture medium and L-broth/agar at a final concentration of 100µg/ml.

5. Buffer Solutions.

TE	1mM Tris HCl (pH 7.5), 0.05mM EDTA
HBS	140mM NaCl, 50mM HEPES (pH 7.12), 1.5mM Na_2HPO_4
TNE	40mM Tris HCl (pH 7.8), 120mM NaCl, 10mM EDTA
ILB	150mM NaCl, 1.5mM $MgCl_2$, 10mM Tris HCl (pH 7.8), 0.65% (v/v) NP40

PEB 7.0M urea, 350mM NaCl, 10mM EDTA, 10mM Tris HCl (pH 7.9), 1% (w/v) SDS

TBE 120mM Tris (pH 9.3), 40.4mM boric acid, 2.7mM EDTA

TAE 0.04M Tris HAc (pH 8.0), 0.002M EDTA

STET 50mM EDTA, 50mM Tris HCl (pH 8.0), 8% (w/v) sucrose, 5% (v/v) Triton X100

SSC 0.15M NaCl, 0.015M sodium citrate (pH 7.5)

MOPS 20mM sodium morpholino-propane sulphonate, 5mM NaAc (pH 7.0), 1mM EDTA

STES 200mM NaCl, 25mM Tris HCl (pH 7.9), 5mM EDTA, 0.1% (v/v) sarkosyl,

Denhardt's solution 0.02% (w/v) Ficoll, 0.02% (w/v) polyvinylpyrrolidone, 0.02% (w/v) bovine serum albumin (BSA; Denhardt, 1966)

6. Enzymes.

All enzymes were obtained from Bethesda Research Laboratories (BRL) or New England Biolabs with the exception of nuclease S1, calf intestinal alkaline phosphatase and Hinc II which were supplied by Boehringer Corporation Limited. Unless otherwise stated, enzyme digests were performed in 1x core buffer (50mM Tris HCl, pH 8.0, 10mM $MgCl_2$, 50mM NaCl), supplied by BRL.

7. Radiochemicals.

Radioisotopes were supplied by Amersham International PLC with the exception of ^{14}C chloramphenicol

which was obtained from NEN Research Products, Boston, Mass. Specific activities for the various radiochemicals were:

5' [α-^{32}P] deoxynucleoside triphosphates, 3000Ci/mmol

5' [γ-^{32}P] adenosine triphosphate, 5000Ci/mmol

^{14}C chloramphenicol, 45.5mCi/mmol

8. Chemicals and Miscellaneous Reagents.

Sigma Chemical Company supplied the following materials: agarose, SDS, Tris base, DTT, lysozyme, acetyl co-enzyme A (acetyl CoA), yeast RNA, BSA, dimethyl formamide, 5-bromo 4-chloro 3-indolyl β D galactopyranoside (Xgal), polyvinylpyrrolidone, piperazine-N,N'-bis[2-ethane sulphonic acid] (PIPES) and all ribo- and deoxyribonucleoside triphosphates.

Koch-Light laboratories supplied acrylamide, boric acid, sodium hydroxide and caesium chloride.

Methylenebisacrylamide, ammonium persulphate, TEMED (N,N,N',N'-tetramethylethylene diamine) and Bradford reagent were obtained from BioRad laboratories.

Low melting point agarose was supplied by BRL.

Analar alcohol was supplied by James Burroughs Ltd.

0.25mm silica gel thin layer chromatography plates were obtained from Camlab, Cambridge.

Tissue culture materials were supplied by Sterilin Ltd. and Nunc.

'Eppendorf' reaction tubes were obtained from Sarstedt Ltd.

Nitrocellulose paper (BA85) was obtained from Schleicher and Schuell.

Intensifying screens were supplied by DuPont.

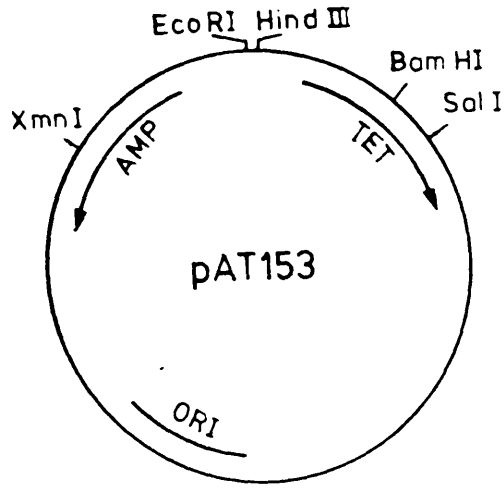
All other reagents were of analytical grade and were supplied by BDH Chemicals Ltd.

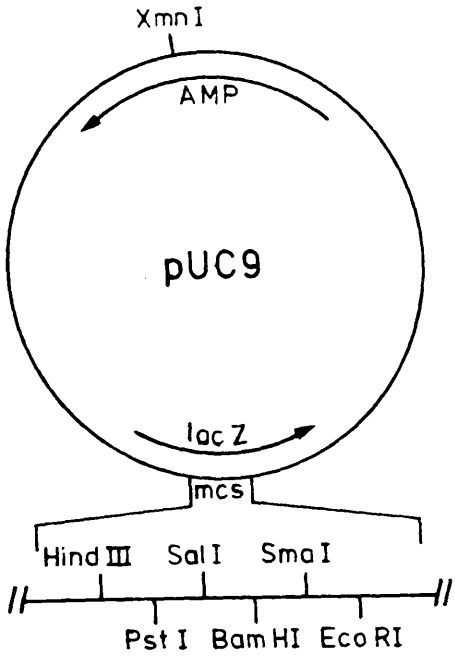
9. Cloning Vectors.

The vectors used for cloning purposes were pAT153 and constructs based on the pUC plasmid series. Fig. 10 shows the various vectors and a number of useful cloning sites.

Plasmid pAT153 is a derivative of pBR322 (Bolivar et al., 1977), an E. coli plasmid carrying genes which confer ampicillin (amp^r) and tetracycline (tet^r) resistance. Plasmid pAT153 was generated by removal of two adjacent Hae II fragments between residues 1646 to 2351 (Sutcliffe, 1978) from pBR322 DNA. The resultant plasmid retains both amp^r and tet^r genes. Useful cloning sites on pAT153 are unique Bam HI and Sal I sites within the tet^r gene and a unique Xmn I site in the amp^r gene (Fig. 10).

A more versatile plasmid series are the pUC vectors (Vieira and Messing, 1982). These plasmids comprise fragments from both pBR322 and M13mp bacteriophage vectors (Messing et al., 1981). The pBR322 segment consists of a Pvu II/Eco RI fragment (positions 2067 to 4362; Sutcliffe, 1978) which carries the amp^r gene and origin of replication; unique Hinc II, Pst I and Acc I sites were removed from this fragment by in vivo mutagenesis and Bal31


EcoRI Hind III
Bam HI
Sal I
Xmn I
AMP
TET
pAT153
ORI


Xmn I
AMP
pUC9
lac Z
mcs
Hind III
Sal I
Sma I
Pst I
Bam HI
Eco RI

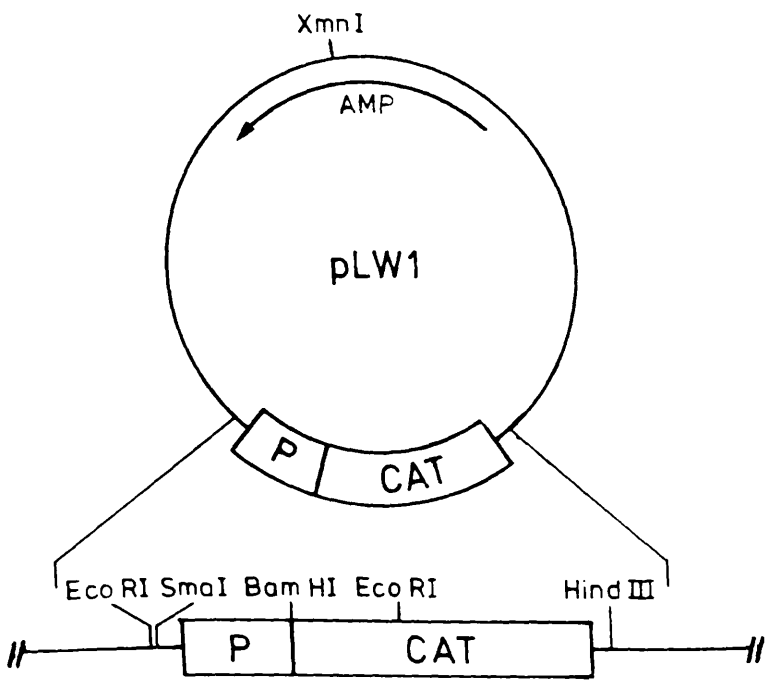

Xmn I
AMP
pLW1
P
CAT
Eco RI
Sma I
Bam HI
Eco RI
Hind III
P
CAT

Figure 10. Details of plasmids used for cloning purposes. Locations of useful restriction enzyme sites, antiobiotic resistant genes (AMP, ampicillin and TET, tetracycline) and the N-terminal portion of β-galactosidase (lac Z) are shown for each vector. The location of the multiple cloning site in pUC9 is shown by mcs. The recognition sequence for Sal I in pUC9 is also cleaved by Hinc II and Acc I restriction enzymes. Fragments carrying promoter and chloramphenicol acetyltransferase sequences in pLW1 are denoted by P and CAT respectively.

digestion. From M13mp DNA, a Hae II fragment which carries the promoter and N-terminal portion of the β-galactosidase gene was inserted into the Hae II site at position 2352 on a mutagenised Pvu II/Eco RI fragment. This N-terminal portion is able to complement a defective β-galactosidase encoded by the host bacterium thus generating β-galactosidase activity (Langley et al., 1975). Within the polypeptide coding region carried by the plasmid is a multiple cloning site (mcs) which comprises a series of unique restriction enzyme cleavage sites. Cloning sites within pUC9 are shown in Fig. 10; pUC8 contains the same restriction enzyme sites but in reverse order to those of pUC9. β-galactosidase activity is detected by hydrolysis of an indicator dye, Xgal (Miller, 1972), which produces bacterial colonies that are blue in colour. Insertion of a fragment into the mcs abolishes the ability to form the N-terminal portion of β-galactosidase resulting in a loss of enzyme activity, thus generating a colony which is white.

The final vector used for cloning, pLW1 (a gift from J.L. Whitton), is based on the pUC plasmid series and contains a promoter fragment from HSV-2 IE gene-4/-5 linked to the polypeptide coding region of the bacterial CAT gene (Gaffney et al., 1985). The CAT gene fragment was cloned into the Acc I site of pUC8 to yield plasmid pUC_8CAT and the promoter sequences inserted upstream from the CAT gene between the Bam HI and Sma I sites in pUC_8CAT to yield pLW1; unique restriction enzyme sites in pLW1 are shown in Fig. 10.

10. Additional Plasmids.

A number of other plasmids were used in this study for either stimulating the level of gene expresion or as experimental controls. Important features of these constructs are shown in Fig. 11.

Plasmid pPRVKpnh contains Kpn I h (constructed by A.J. Davison and supplied by R. Everett), a fragment which maps in the inverted repeat segment of PRV DNA (Davison and Wilkie, 1983a). This region of the PRV genome encodes at least one viral product, an IE protein of 180,000 mol. wt. (Ihara et al., 1983). Co-transfection with this plasmid stimulates transcription of plasmid-borne HSV genes and other viral and cellular promoters (Green et al., 1983; Imperiale et al., 1983; Everett et al., 1984b), presumably as a result of expression of the PRV polypeptide(s).

Plasmid p $\beta(244^{+})\beta$ contains two copies of the rabbit β-globin gene, inserted in the same orientation into pBR322 DNA; these two copies are separated by a fragment carrying the polyoma virus enhancer (de Villiers and Schaffner, 1981). A Bst NI fragment from plasmid pRED4, which contains the HSV-1 gD promoter region fused to the rabbit β-globin structural gene, was used as a DNA probe to detect mRNA produced from plasmid p $\beta(244+)\beta$ (see Fig. 11). These plasmids were gifts from R. Everett.

Plasmid pLW2 contains promoter and CAT fragments as described for pLW1 in the previous section and in addition, downstream from the CAT gene, a 100bp Sma I/Xba I fragment which includes the sequences flanking the poly A site of

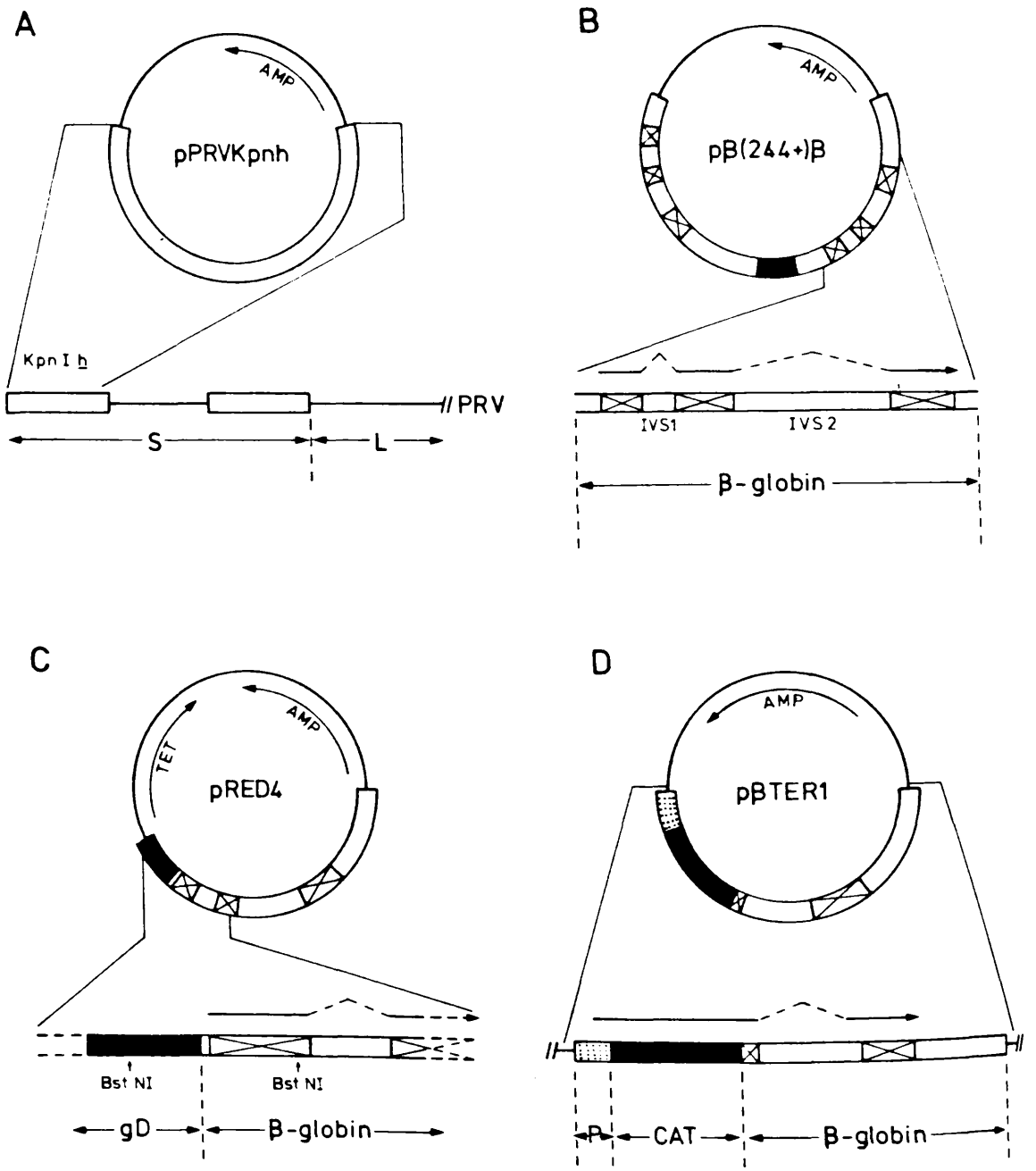

A
AMP
pPRVKpnh
Kpn I h
// PRV
S
L
B
AMP
pβ(244+)β
IVS1
IVS2
β-globin
C
TET
AMP
pRED4
Bst NI
Bst NI
gD
β-globin
D
AMP
pβTER1
P
CAT
β-globin

Figure 11. Details of plasmids used in transfection experiments. In all plasmids, cloning vector sequences are shown by a solid line and fragments inserted into vectors by boxed regions; boxed regions which contain a cross represent exons of the rabbit β -globin gene. Locations of antiobiotic resistance genes (AMP, ampicillin and TET, tetracycline) are indicated for each plasmid.

A) The location of Kpn I h, a terminal fragment, is shown within the short (S) segment of the PRV genome.

B) The fragment carrying the polyoma virus enhancer is shown by the shaded region. The structure of β -globin mRNA is indicated with the positions of intervening sequences (IVS 1 and IVS 2) within β -globin mRNA shown by dashed lines.

C) The dark area shows the position of the HSV-1 gD mRNA promoter region. Also indicated is the Bst NI fragment used as a DNA probe in hybridisations with pβ(244+)β mRNA.

D) The dark area indicates the fragment carrying CAT coding sequences and the stippled region represents the promoter region from HSV-2 IE mRNAs-4/-5. Plasmid pβTER1 contains only a portion of the second exon and a complete copy of the third exon of the rabbit β -globin gene. The proposed structure of CAT mRNA from pβTER1 is indicated, however, it should be noted that structural analysis of this transcript has not been performed.

HSV-2 IE mRNA-5. The arrangement of DNA fragments in pLW2 is identical to that in plasmid pTER5 (see Fig. 66). Plasmid pLW2 was a gift from J.L. Whitton.

Plasmid pβTER1 contains a 1250bp Bam HI/Xho I fragment from the 3' portion of the rabbit β-globin gene (Banerji et al., 1981), inserted downstream from the CAT gene in pLW1. These β-globin sequences include all of the IVS 2 sequences and 600bp of flanking DNA. This plasmid was a gift from D. Gaffney.

11. Bacterial Hosts for Plasmids.

The host bacterium for pAT153 was E. coli K12 strain HB101 (ramC1F^{-}Pro^{-}Gal^{-}Str^{r}Rec^{-}; Boyer and Roulland-Dussoix, 1969).

The host bacterium for the pUC plasmids and their derivatives was E. coli K12 strain JM83 (ara, lac-pro^{del}, strA, thi, ϕ80dlacM15^{del}; Messing, 1979) a β-galactosidase deficient strain which contains a phage encoding the β-peptide of the enzyme integrated into the chromosome.

SECTION B.

METHODS

12. Virus Growth and Assay.

BHK-C13 cells in 850cm^2 roller bottles were infected with HSV-2 strain HG52 at a multiplicity of infection (moi) of 1 plaque forming unit (pfu) per 350 cells. Infected cells were maintained in 20ml of ETC10 for 3 to 4days at 31^{o}C. Procedures for the harvesting of virus from infected cells were performed at 4^{o}C. Cells were removed by gentle shaking and pelletted in the cold by centrifugation at 1500rpm for 10min. The supernatant was removed and centrifuged at 12,000rpm for 2h. The resultant pellet was resuspended in approximately 5ml of ETC10 and stored at -70^{o}C as cell released virus (CRV). The pelletted material from the initial low speed centrifugation was resuspended in an equal vol of ETC10 and the suspension sonicated. Following removal of cell debris by centrifugation at 3000rpm for 15min, the supernatant was recovered and stored at -70^{o}C as cell associated virus (CAV).

Both CRV and CAV fractions were titrated for infectious virus by determining the number of pfu in serial dilutions of virus preparations (Brown et al., 1973).

13. Standard Procedures.

A number of procedures were used routinely which are referred to in various sections; the following is a description of these procedures.

a) Restriction enzyme digestion of DNA.

DNA was cleaved at 37°C or, in certain cases, at 65°C (for example, digestion with Bst EII) for 1h to 4h. Digests were performed in 1x core buffer (see Page 96), except for digestions with Sma I which were carried out in 15mM Tris HCl (pH 8.0), 6mM $MgCl_2$ and 15mM KCl.

b) Phenol/chloroform extraction.

Proteins and ethidium bromide were removed by addition of an equal vol of phenol/chloroform (1vol: 1vol) which had been equilibrated three times with either TE or water. The mixture was shaken and the phases separated by centrifugation at either 3000rpm for 5min or 12,000rpm for 1min. The top (aqueous) layer was then removed and either re-extracted or ethanol precipitated. Frequently, a final extraction with an equal vol of chloroform was performed to remove phenol from the aqueous layer prior to ethanol precipitation.

c) Precipitation of nucleic acid.

DNA and RNA were precipitated by the addition of 3vol of Analar ethanol and 1/20thvol of 5M NaCl followed by incubation at -20°C for 16h or in dry ice for a minimum of

5min. Alternatively, an equal vol of isopropanol was added to precipitate DNA and the mixture incubated at room temperature for 15min to 25min. Nucleic acid was recovered by centrifugation at either 12,000rpm for 5min or 7,000rpm for 15min and the pellet was washed with 70% ethanol prior to drying.

d) Visualisation of non-radioactive nucleic acid.

Ethidium bromide, an intercalating agent, was used for staining non-radioactive DNA and RNA. Agarose and acrylamide gels were soaked in a solution of 0.5µg/ml ethidium bromide and bands were visualised by transmitted UV light. For preparative samples, long wave UV light was used which could detect 10ng of DNA. For analytical samples, short wave UV light was utilised, allowing detection down to 2ng of DNA. Both sucrose and CsCl gradients contained ethidium bromide and bands were visualised under long wave UV light.

e) Visualisation of radiolabelled DNA and chloramphenicol.

Gels containing radiolabelled DNA and TLC plates containing ^{14}C chloramphenicol were covered with Clingfilm and the surface of the Clingfilm was washed with 70% ethanol. Kodak XS film was pre-flashed, laid on top of the Clingfilm and covered with an intensifying screen. The film was exposed at -70°C (for long exposures) or at room temperature (for short exposures).

f) Recovery of DNA fragments from polyacrylamide gel slices.

Gel slices were extruded through a 0.7ml reaction tube, punctured by a 23 **gauge** hypodermic needle, into a 1.5ml reaction tube by centrifugation at 12,000rpm for 10sec. 2vol of H_2O were added, the mixture vortexed and incubated at 42°C for 16h. DNA was recovered by filtration through glass wool to remove polyacrylamide followed by the addition of ethanol and salt to the eluate.

g) Removal of DNA from gradients.

Bands were recovered from gradients by puncturing the side of the centrifuge tube with a 23 **gauge** hypodermic needle and DNA was removed by a syringe attached to the needle.

14. Gel Electrophoresis.

a) Non-denaturing agarose gels.

Profiles of DNA cleaved with restriction enzymes were obtained with these gels from which maps for DNA fragments were constructed. The concentration of gels varied from 0.6% to 1.2% (w/v) agarose dissolved in 1x TBE buffer and gels also contained 0.5µg/ml of ethidium bromide. Samples were applied in 1x TBE, 10% (w/v) sucrose, 0.02% (w/v) bromophenol blue and xylene cyanol. Gels (260mm x 160mm) were electrophoresed in horizontal kits up to a maximum voltage of 12V/cm.

b) Denaturing agarose/formaldehyde gels.

RNA species separated by these gels were subsequently transferred to nitrocellulose paper. The concentration of gels was 1.5% and agarose was dissolved in 1x MOPS buffer containing 2.2M formaldehyde. Samples were prepared as described on Page 116. Gels (190mm x 160mm) were electrophoresed in vertical kits at 4V/cm for about 6h and thereafter stained with ethidium bromide prior to transfer to nitrocellulose paper.

c) Low melting point agarose gels.

These gels were used to obtain fragments of at least 500bp for cloning purposes. The concentration of gels was 1% and agarose was dissolved in 1x TAE buffer. Gels (260mm x 160mm) were electrophoresed in horizontal kits at a maximum voltage of 7.5V/cm.

d) Non-denaturing polyacrylamide gels.

DNA fragments of less than 1kbp were analysed on polyacrylamide gels which were prepared as described by Maniatis et al. (1975).

A stock of 30% (w/v) polyacrylamide (acrylamide: bisacrylamide, 29:1) was diluted to 6% in 75ml of 1x TBE buffer. Polymerisation was effected by addition of 0.5ml of 10% (w/v) ammonium persulphate along with 50µl of TEMED. Samples were prepared as described on Page 105 and gels (1mm x 160mm x 260mm) were electrophoresed in vertical kits up to a maximum voltage of 16V/cm.

e) Strand-separation polyacrylamide gels.

The strands of end-labelled DNA fragments were separated using gels which were prepared as described in the previous section with the following modifications. The ratio of acrylamide: bisacrylamide was adjusted to 59:1 and the gels (1mm x 230mm x 450mm) were electrophoresed in the cold at 10V/cm. Samples were prepared by dissolving pelletted DNA in 30% (v/v) DMSO and 1mM EDTA followed by heating to 90°C for 3min. Prior to loading onto the gel, samples were chilled in ice.

f) Denaturing polyacrylamide gels.

These gels were used to determine the nucleotide sequences of DNA fragments and the lengths of DNA/RNA hybrids following nuclease digestion. Gels were prepared as described by Maxam and Gilbert (1980).

A stock of 30% (w/v) polyacrylamide (acrylamide: bisacrylamide, 29:1) was diluted to a range of concentrations from 6% to 16% in 1x TBE buffer containing 9M urea. Following polymerisation, gels (0.35mm x 230mm x 450mm) were pre-run at 40W for 1h before the samples were loaded. Samples, dissolved in 90% (v/v) formamide, 0.1% (w/v) bromophenol blue and xylene cyanol, were heated to 90°C for 3min and chilled rapidly in ice prior to loading onto the gel. Gels were electrophoresed in vertical kits at 40W.

15. Preparation of Plasmid DNA.

10ml of L-Broth supplemented with 100µg/ml ampicillin was inoculated with a loop of bacteria containing the appropriate plasmid and the culture was incubated overnight at 37°C. This culture was used to seed 800ml to 1600ml of L-Broth containing ampicillin which was shaken vigorously for 6h to 8h at 37°C. Plasmids were amplified (Clewell, 1972) by the addition of chloramphenicol to 50µg/ml and the culture was maintained overnight at 37°C.

Plasmid DNA was purified essentially by the 'boiling' method (Holmes and Quigley, 1981). Bacteria were recovered by centrifugation at 8000rpm for 10min and resuspended in 30ml of STET followed by the addition of 4ml of freshly prepared lysozyme (10mg/ml). The suspension was boiled for approximately 1min and cell debris removed by centrifugation at 15,000rpm for 1h. Plasmid DNA in the supernatant was precipitated with isopropanol and pelletted by centrifugation at 3000rpm for 15min. The pellet was resuspended in 10.5ml of H_2O followed by the addition of 11g of caesium chloride and 0.5ml of ethidium bromide (10mg/ml). DNA was banded in a Ti50 angle rotor for 3days at 40,000rpm after which plasmid DNA was visualised under long wave UV light and recovered as described on Page 105. Ethidium bromide was removed by extraction with an equal vol of n-butanol which had been previously equilibrated with TE and DNA was dialysed against H_2O for at least 7h. Samples were treated with 50µg/ml RNase A at 60°C for 1h

and thereafter with 50µg/ml proteinase K at 37°C in the presence of 0.1% (w/v) SDS, again for 1h. These enzymes were removed by phenol/chloroform extraction and the DNA was ethanol precipitated twice. Finally, the DNA was pelletted, dried and dissolved in H_2O; the concentration of DNA was estimated by spectrophotometry.

16. ^{32}P-labelling of DNA Fragments.

a) Nick-translation.

Cleaved DNA was radiolabelled (Rigby et al., 1977) by addition of 1µCi of all four [α-^{32}P] dNTPS and 2units of E. coli DNA polymerase I. The mixture was left at room temperature for 30min and the sample applied to a polyacrylamide gel. The appropriate radiolabelled band was recovered following visualisation by autoradiography.

b) 5' end-labelling.

DNA was digested with the appropriate restriction enzymes in 200µl and then loaded onto a pre-formed 5% to 20% (w/v) sucrose gradient containing STES and 0.5µg/ml ethidium bromide. Samples were centrifuged for 16h at 30,000rpm in an AH650 swing-out rotor. The banded DNA was recovered and ethidium bromide removed by extraction with phenol/chloroform. DNA was precipitated by the addition of 3vol of ethanol and the mixture incubated at -20°C for at least 16h. The DNA was pelletted, dried and dissolved in 40µl of H_2O. Core buffer was added along with 20units of

calf intestinal alkaline phosphatase and the sample incubated at 37°C for 1h. Phosphatase was removed by three extractions with phenol/chloroform and the DNA was ethanol precipitated. Following recovery of DNA, the pellet was dissolved in a total vol of 30µl containing 50mM Tris HCl (pH 7.6), 10mM $MgCl_2$, 5mM DTT, 50 to 100µCi of [γ-^{32}P] ATP and 5units of T4 polynucleotide kinase. After incubation at 37°C for 1h, DNA was either loaded directly onto a polyacrylamide gel or precipitated by the addition of 200µl of 2.5M ammonium acetate and 3vol of ethanol. Precipitated DNA was denatured and applied to a strand-separation gel (see Page 107).

c) 3' end-labelling.

Digested DNA was 3' end-labelled by the addition of 10µCi of the appropriate [α-^{32}P] dNTP(s) along with 4units of large fragment DNA polymerase I. The mixture was incubated at room temperature for 30min after which the sample was either loaded onto a polyacrylamide gel or cleaved with a second restriction enzyme. Prior to recleavage, enzymes in the radiolabelled sample were inactivated by heating to 65°C for 10min followed by rapid cooling in ice. The second enzyme was added and the DNA digested before application to a polyacrylamide gel.

17. Production of Deletions in Plasmids.

Digested DNA was purified by phenol/chloroform extraction and ethanol precipitation. DNA was resuspended

in a small vol (11μl) containing 10mM Tris HCl (pH 8.0), 50mM NaCl, 3mM $MgCl_2$ and 1mM β-mercaptoethanol along with 32units of exonuclease III. The mixture was incubated at 22.5°C and 4μl aliquots were removed after 1min, 4min and 8min. 1μl of 5x Mung bean nuclease buffer (1M NaCl, 5mM $ZnCl_2$, 150mM NaAc, pH 5.0) along with 10units of Mung bean nuclease were added to each aliquot and samples were digested at 37°C for 10min. DNA was purified by phenol/chloroform extraction and by ethanol precipitation. The extent of deletions was analysed by cleaving a portion of each sample with a second restriction enzyme followed by electrophoresis on a polyacrylamide gel alongside standard size markers. Samples which appeared to contain deletions of the required length were pooled, ligated and transfected into competent bacteria.

18. DNA Sequencing.

The nucleotide sequences of DNA fragments, uniquely end-labelled at a restriction enzyme site, were determined by the chemical degradation procedures of Maxam and Gilbert (1977 and 1980).

Chemical degradation of DNA fragments was performed in two steps. Firstly, DNA bases were specifically modified or removed by chemical reagents. Specifically, guanine residues were methylated using dimethyl sulphate and a depurination mix containing diphenylamine and formic acid removed purine bases. For pyrimidines, hydrazine was used to open thymine and cytosine rings; C-specific reactions

were performed in the presence of salt as, under these conditions, hydrazine reacts appreciably with only cytosine residues. Reactions were performed at 25°C for periods which varied between 4min to 12min. Under these conditions, the chemical reactions do not reach completion, generating a series of molecules which have bases modified at random sites on the DNA fragments. The second step required reaction with 1M piperidine at 90°C which removes both altered bases and phosphate residues resulting in cleavage of the DNA backbone. Thus, a population of cleaved DNA molecules was generated which varied in length but were labelled at an identical site. Fragments resulting from each base-specific reaction were run in separate lanes on a denaturing polyacrylamide gel, producing a ladder of bands from which the DNA sequence was determined.

19. Ligation of DNA Fragments.

DNA fragments were produced for ligations by cleavage with restriction enzymes or digestion with Mung bean nuclease. Restriction enzyme digestion generated termini with either a single-stranded 5' overhang or a blunt-end. Alternatively, blunt-ended DNA fragments were produced by 'filling-in' the 5' overhang using all four dNTPs along with 4units of large fragment DNA polymerase I and incubating the mixture at room temperature for 15min. Mung bean nuclease treatment of DNA also generated blunt-ended fragments.

DNA fragments were ligated to a cleaved plasmid vector which contained termini compatible with those of the fragments. The fragment and vector were mixed in a 10:1 ratio and incubated overnight at 15°C in ligation buffer consisting of 66mM Tris HCl (pH 7.6), 6.6mM $MgCl_2$, 10mM DTT and 0.4mM ATP along with 2units of DNA ligase.

20. Ligation in Low Melting Point Agarose.

DNA fragments excised from low melting point agarose gels did not require removal from gel slices for ligation purposes. Gel slices, containing DNA fragments to be ligated, were mixed and an equal vol of H_2O was added. The mixture was heated to 65°C until the agarose melted, at which point the sample was cooled to 37°C. Ligation buffer and 2units of DNA ligase were added and the sample was incubated overnight at 15°C.

21. Preparation of Competent Bacteria.

100ml of L-Broth was seeded with 2ml of an overnight culture of bacteria and the bacteria were grown at 37°C with vigorous aeration until the $O.D._{650}$ reached 0.2. Following incubation in ice for 10min, bacteria were pelletted by centrifugation at 7,000rpm for 10min. The pellet was resuspended in 40ml of ice-cold 50mM $CaCl_2$ and left on ice for a further 20min. Bacteria were pelletted again and taken up in 4ml of 50mM $CaCl_2$ and this mixture left on ice overnight. Bacteria treated in this way were stored for up to 5days in ice and used for transformation by plasmid DNA.

22. Transformation of Bacteria by Plasmid DNA.

Plasmid DNA was added to 100µl to 200µl of competent bacteria, the mixture was vortexed and left on ice for 40min. After heating the sample to 37°C for 3min, 70µl aliquots were spread onto L-Broth agar plates containing ampicillin and incubated overnight at 37°C. For ligations performed in low melting point agarose, $CaCl_2$ was added to a final concentration of 50mM to the ligation mixture prior to addition to competent bacteria.

Selection of pUC plasmids containing inserted fragments was facilitated by spreading 100µl of 2.5% (w/v in dimethylformamide) Xgal onto plates before spreading the transformation mixture.

23. Preparation of DNA by 'Minilysis'.

2ml of L-Broth containing ampicillin was inoculated with a colony of bacteria and the culture grown at 37°C for a minimum of 6h. Bacteria from half of the culture were pelletted by centrifugation at 12,000rpm for 30sec. The pellet was resuspended in 200µl of STET and 20µl of lysozyme (10mg/ml) added. The mixture was heated in a boiling water bath for 40sec and then centrifuged at 12,000rpm for 10min. An equal vol of isopropanol was added to the supernatant to precipitate plasmid DNA. Following recovery of the DNA by centrifugation, the pellet was dissolved in 300µl of H_2O and phenol/chloroform extracted. The DNA was ethanol precipitated, restricted with the appropriate restriction enzyme(s) and run on either agarose

or polyacrylamide gels. Large scale preparations were produced from colonies containing the desired plasmids by seeding L-Broth with the remainder of the original bacterial culture.

24. DNA-mediated Transfer into Tissue Culture Cells by $CaPO_4$ Precipitation.

DNA transfections into HeLa cells were performed essentially as described by Wigler et al. (1978) and Corsalo and Pearson (1981).

15µg of plasmid DNA was dissolved in 420µl of TE and 60µl of 2M $CaCl_2$ was added dropwise with constant agitation. For transfections which required the use of additional plasmids either as controls or for trans-activation, 10µg of the appropriate construction was added to the TE. This mixture was then added dropwise with constant agitation to 480µl of 2x HBS and the precipitate allowed to form at room temperature for 30min. This solution was pipetted dropwise onto 90mm Petri dishes containing HeLa cells at approximately 30% confluency which were then incubated at 37°C. The medium was replaced 24h after addition of the $CaPO_4$ precipitate with fresh medium and cells were incubated at 37°C for a further 24h. For virus infection of transfected cells, the medium was removed after 24h and virus added in 5ml of fresh medium. Following adsorption at 37°C for 1h, a further 10ml of fresh medium was added and cells were harvested at the appropriate time p.i.

25. Preparation of Cytoplasmic RNA.

Cytoplasmic RNA was prepared from 90mm Petri dishes by a modification of the Kumar and Lindberg (1972) method.

Following removal of medium, cells were washed with 5ml of cold PBS and then harvested by scraping into 3.5ml of PBS. Cells were recovered by centrifugation in the cold at 2000rpm for 2min. The pellet was washed with a further 4ml of PBS and centrifugation was repeated. Cells were carefully resuspended in 0.3ml of cold isotonic lysis buffer (ILB) and left on ice for 4min. Nuclei were separated from components of the cytoplasmic fraction by centrifugation in the cold at 3000rpm for 5min. Following addition of an equal vol of phenol extraction buffer (PEB), the mixture was extracted three times with phenol/chloroform and once with chloroform. Cytoplasmic RNA was precipitated by the addition of 3vol of ethanol and incubation overnight at -20°C. RNA was recovered by centrifugation, dissolved in 200µl of H_2O and re-ethanol precipitated. Finally, RNA was pelleted, dried and dissolved in 30µl of H_2O. The concentration of RNA was estimated by spectrophotometry.

26. Northern Blot Analysis.

RNA was transferred to nitrocellulose paper and hybridised with radiolabelled DNA essentially as described by Spandidos and Paul (1982).

Cytoplasmic RNA was dissolved in 50µl of 1x MOPS buffer which contained 50% (v/v) formamide and 2.2M

formaldehyde and the mixture heated to 65°C for 2min. Sucrose was added to a final concentration of 10% (w/v) and the sample applied to a 1.5% denaturing agarose/formaldehyde gel (see Page 106). Following electrophoresis, the gel was shaken gently in 1 to 2vol of 20x SSC for 40min at room temperature and RNA was blotted onto nitrocellulose paper for 16h using 10x SSC. The nitrocellulose filter was washed with 6x SSC, dried in air and baked at 80°C for 2h under vacuum.

Prior to hybridisation with a radiolabelled DNA probe, the nitrocellulose filter was soaked in 6x SSC, 5x Denhardt's solution, 0.1% (w/v) SDS and 20μg/ml of denatured calf thymus DNA at 68°C for 6h. The DNA probe was denatured in 90% (v/v) formamide by heating to 90°C for 3min followed by rapid cooling in ice. Hybridisation of the DNA probe to the nitrocellulose filter was performed overnight at 68°C in 6x SSC, 10x Denhardt's solution, 20mM Tris HCl (pH 7.5), 1mM EDTA, 0.5% (w/v) SDS and 50μg/ml of denatured calf thymus DNA.

Following hybridisation, the nitrocellulose filter was washed four times in 2x SSC, 0.2% (w/v) SDS at 68°C for 4h. The filter was then dried and autoradiographed. mRNA sizes obtained with this technique are inaccurate as the poly A^+ mRNAs transferred to the nitrocellulose paper contain tracts of adenosine residues at their 3' termini.

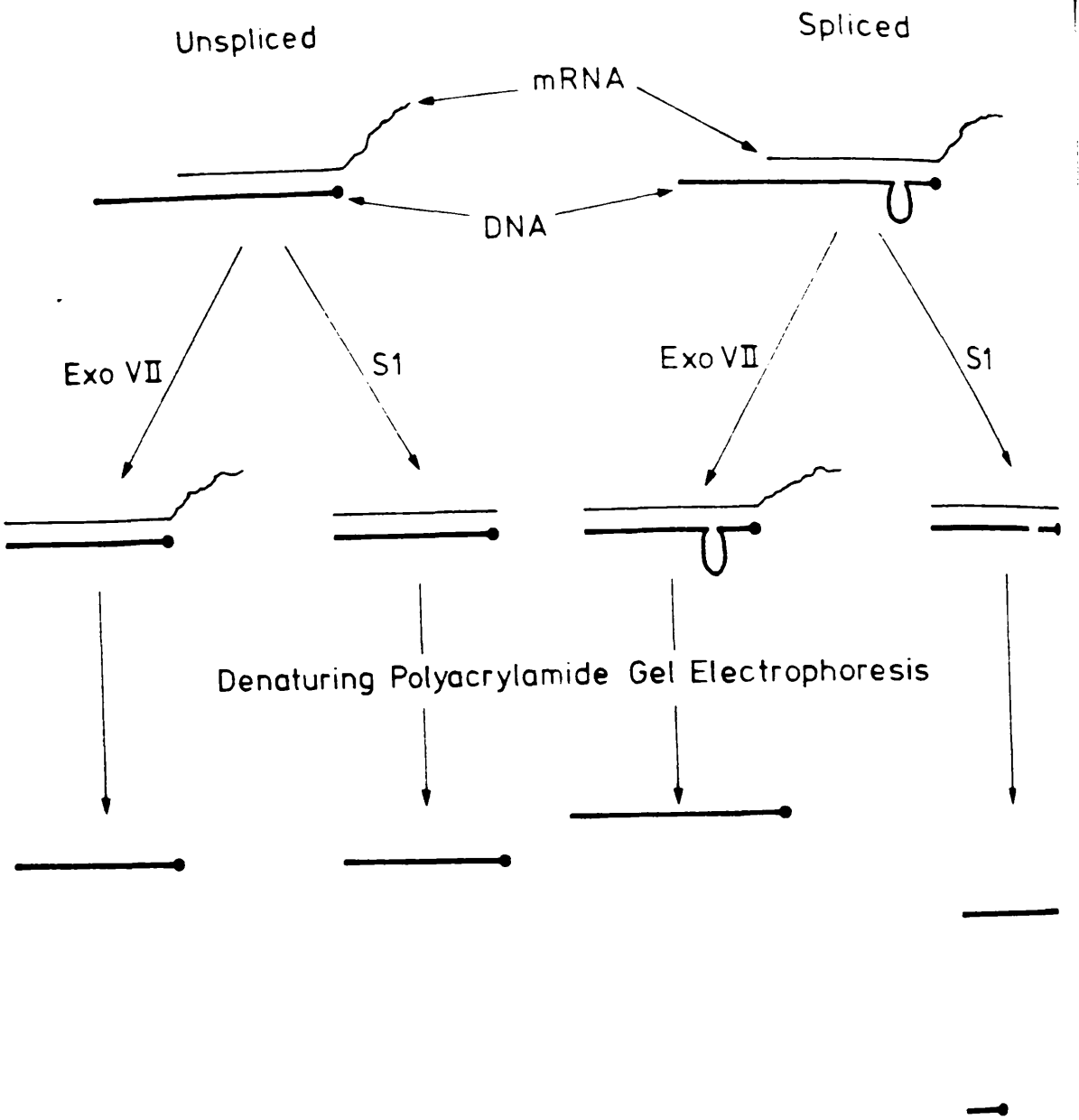


Figure 12. Strategy for performing structural analysis of spliced and unspliced transcripts. Thin lines indicate mRNA and thicker lines represent DNA. Regions of DNA, which do not hybridise to mature mRNA due to RNA splicing, are represented by loops. Closed circles indicate the DNA termini (either 5' or 3' termini) which are radiolabelled.

27. Structural Analysis of mRNAs.

Structural analysis of mRNAs was performed using the nuclease S1 procedure of Berk and Sharp (1978a, b), modified by using either 5'- or 3'-labelled DNA probes instead of uniformly-labelled DNA (Weaver and Weissmann, 1979). The rationale of this technique is that RNA hybridises to stretches of DNA containing complementary sequences in the presence of salt. Regions of nucleic acid which do not form a double-stranded hybrid can be removed by nuclease digestion and the length of the stable hybrid determined by electrophoresis alongside standard size markers. Nucleases used in this study were nuclease S1 and exonuclease VII. Nuclease S1 degrades single-stranded nucleic acid to yield 5' phosphoryl mono- or oligonucleotides (Vogt, 1973). The enzyme also removes looped regions generated by hybridisation of genomic DNA to spliced mRNA (see Fig. 12). Exonuclease VII digests the 5' and 3' termini of single-stranded DNA (Chase and Richardson, 1974), therefore looped regions of DNA are not substrates for this nuclease (see Fig. 12). The enzyme releases short oligonucleotides and has a preference for digestion at T residues. Subsequently, digestion products frequently do not contain flush-ended termini and as a result, exonuclease VII-treated hyrids are frequently larger than nuclease S1-digested hybrids. By end-labelling DNA fragments at unique restriction enzyme sites, precise locations for mRNA termini and splice sites within regions of DNA can be determined. To position mRNA 5' termini and

splice acceptor sites, 5' end-labelled DNA probes are utilised. Similarly, 3' end-labelled DNA probes locate mRNA 3' termini and splice donor sites.

End-labelled DNA fragments were co-precipitated with 20μg to 30μg of cytoplasmic RNA and the pellet resuspended in 20μl of 90% (v/v) formamide (deionised with Amberlite monobed resin MB-1), 400mM NaCl, 40mM PIPES (pH 6.8) and 1mM EDTA for double-stranded DNA probes. For single-stranded DNA probes, the total vol was increased to 30μl and the formamide concentration adjusted to 50%. The mixture was heated to 90°C for 3min and then incubated for 16h at 56°C for double-stranded DNA probes and at 42°C for single-stranded DNA probes. Prior to nuclease treatment, hybridisation mixtures were quenched in ice.

Nuclease S1 digestion was performed at 30°C for 2h to 3h in 300μl of 400mM NaCl, 40mM sodium acetate pH 4.5, 1mM $ZnSO_4$ with 5000units of nuclease S1. Exonuclease VII treatment was performed at 37°C for 1h in 200μl of 6.7mM potassium phosphate (pH 7.9), 8.3mM EDTA, 10mM β-mercaptoethanol along with 0.5units of exonuclease VII. Following phenol/chloroform extraction and ethanol precipitation, nuclease-resistant hybrids were analysed by denaturing polyacrylamide gel electrophoresis.

28. CAT Assays.

Cell extracts were prepared and CAT assays performed essentially as described by Gorman et al. (1982).

Firstly, media was removed and the cell sheet

washed with 5ml of cold PBS. Cells were scraped into 3.5ml of TEN and pelletted by centrifugation at 2,000rpm for 2min. The pellet was resuspended in 200µl of 250mM Tris HCl (pH 7.8) and the cells were sonicated. Following removal of cell debris by centrifugation at 12,000rpm for 5min, the supernatant was frozen at -20°C for periods of up to several months without an appreciable loss in CAT activity.

Levels of CAT activity were measured by incubating aliquots from cell extracts in a mixture containing 1mM acetyl CoA, 250mM Tris HCl (pH 7.8) and 0.125µCi ^{14}C chloramphenicol (68nmoles/µCi) at 37°C. Aliquots were removed at various times, extracted with 200µl of ethyl acetate and the organic phase was dried down. ^{14}C chloramphenicol was resuspended in 30µl ethyl acetate and spotted onto 0.25mm silica gel thin layer chromatography plates. Ascending chromatography was performed in 95% chloroform: 5% methanol (v/v), after which chromatograms were autoradiographed to visualise the acetylated and non-acetylated forms of chloramphenicol. The percentage conversion to acetylated ^{14}C chloramphenicol was calculated by removing spots from chromatograms and counting in a scintillation counter. CAT activities were normalised by measuring the protein content of each extract which allowed calculation of specific activities for the CAT enzyme.

29. Measurement of Protein Content.

The protein content of cell extracts was determined as described by Bradford (1976). The assay is based on the ability of protein to bind to Coomassie Brilliant Blue G-250 under acidic conditions and thereby shift the absorbance maximum of Coomassie Blue from 465nm to 595nm.

Following the addition of 5ml of dye reagent to 100µl aliquots of cell extracts, the mixture was incubated at room temperature for 5min and the absorbance of each sample was read at $O.D._{595}$. To estimate the protein content in cell extracts, various concentrations of BSA ranging from 0 to 700µg/ml were prepared in 250mM Tris HCl (pH 7.8) and dye reagent was added. A standard curve was derived from $O.D._{595}$ values of the protein standards allowing the protein content of extracts to be determined.

SECTION C.

COMPUTER ANALYSIS OF SEQUENCES

31. Computer Programs.

During the course of this work, all nucleotide sequence data was stored in a Digital Research PDP-11/44 computer using the RSX-11M operating system. Interpretation of this data was made possible by the availability of the following programs designed and/or modified by Dr. P. Taylor.

1) **BASES**. The frequency of occurrence of a particular base could be counted either for an entire nucleotide sequence or at specified positions within the sequence.

2) **PROFIL.** A plot of the hydropathicity of protein sequences was produced. The range within which the hydropathicity of amino acid residues was determined, could be defined by the operator.

3) **SEQLIST.** Sequences were listed and numbered in blocks of 10 residues.

4) **PTRANS.** Amino acid sequences were predicted from a nucleotide sequence. Positions at which translation started and ended were determined by the operator. In addition, tables presenting the amino acid content and codon usage within a translated region were produced.

5) CINTHOM. A program for comparing sequences with the homology displayed on a graph. The similarity between sequences was plotted as an uppercase or lowercase letter of the alphabet with A representing the highest homology (100%) and z denoting the lowest homology. Areas with conserved sequences formed a diagonal on the graph and insertions/deletions between two sequences were identified by diagonals which, when extrapolated, crossed the x- and y-axes at different points. The range of sequence to be compared and minimum level of homology were set by the operator.

6) HOMOL. Sequences could be aligned and gaps inserted in order to maximise homology. By varying the parameters, the number and size of gaps allowed could be altered.

7) SEARCH. The locations and numbers of cleavage sites for a wide range of restriction enzymes within a nucleotide sequence could be analysed.

8) HAIRPN. Nucleotide sequences were analysed for the presence of inverted repeats which could form a stem and loop structure. Parameters for the minimum and maximum sizes of stems and loops were set by the operator.

RESULTS AND DISCUSSION

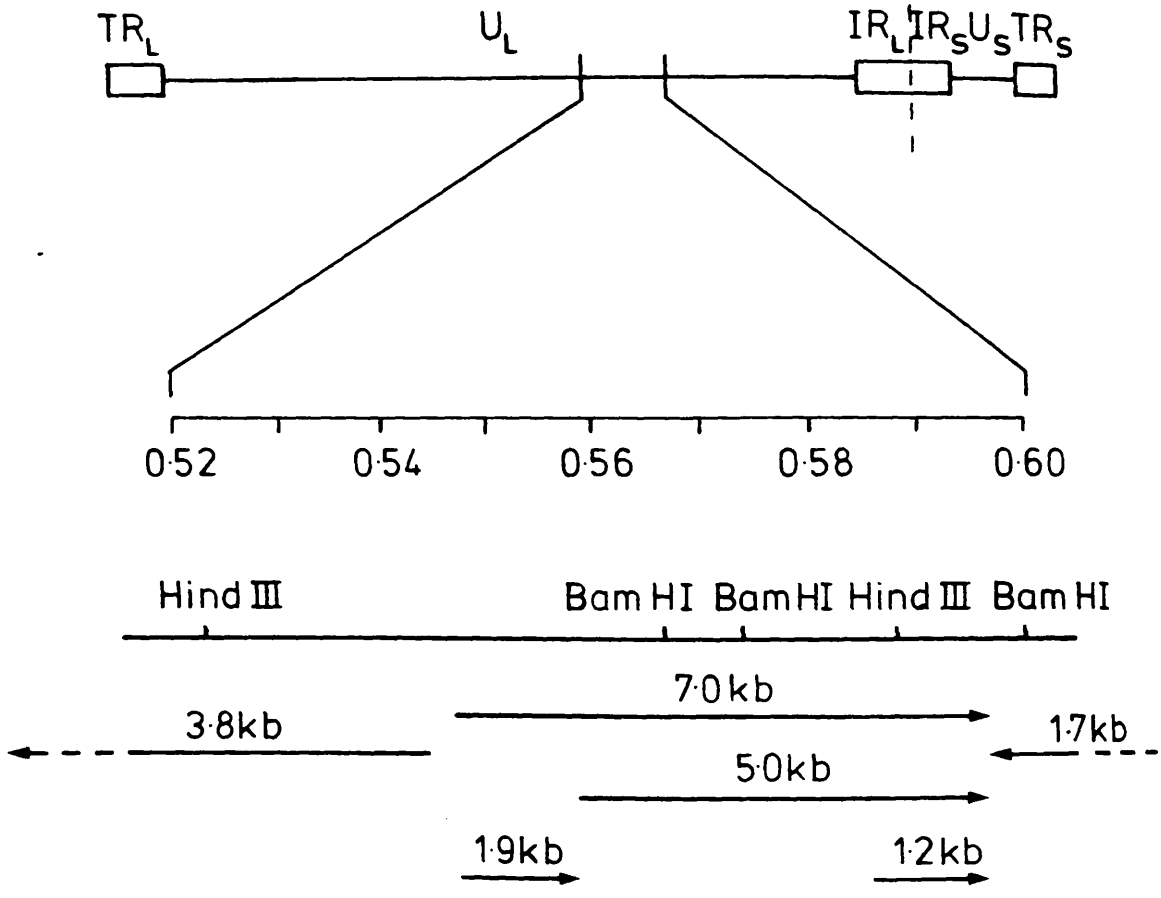


Figure 13. Arrangement and orientation of mRNAs on the HSV-1 genome between 0.52 to 0.60 map units. Also indicated are leftwards-transcribed species which do not map entirely within this region and the locations of Hind III and Bam HI restriction enzyme sites for this portion of HSV-1 DNA. The 5.0kb and 1.2kb mRNAs correspond to transcripts labelled RR1 and RR2 in Fig. 4. Restriction enzyme fragments shown on the diagram are: Hind III k (0.520 to 0.587 map units); Bam HI g' (0.566 to 0.572 map units); Bam HI o (0.572 to 0.599 map units).

SECTION A.

1. Identification of the mRNAs Spanning the Bgl II c and Bgl II n Fragments.

The HSV-1 locus between 0.54 and 0.60 map units specifies a family of four nested mRNAs (Fig. 13) which comprises two early and two late transcripts (Anderson et al., 1981; McLauchlan and Clements, 1982 and 1983a). The early mRNAs have sizes of 5.0kb and 1.2kb which encode polypeptides of 136,000 mol. wt. (Vmw136) and 38,000 mol. wt. (Vmw38) respectively; these mRNAs are unspliced and share a common 3' terminus. The late mRNAs, which have sizes of 7.0kb and 1.9kb, are presumed to be 5' co-terminal and encode the same 54,000 mol. wt. protein. The 7.0kb mRNA also appears to be 3' co-terminal with both early mRNAs.

The functions encoded by the HSV-1 and HSV-2 genomes are essentially colinear (see Introduction, Page 11). Therefore, to identify transcripts in the equivalent HSV-2 genomic region, a nick-translated Bgl II/Bam HI fragment from 0.580 to 0.583 map units, which is located at the left-hand end of the Bgl II n transforming region (Fig. 14), was hybridised to a Northern blot containing HSV-2 early mRNA. Two bands with approximate sizes of 5.4kb and 1.5kb were detected and the 1.5kb mRNA species was approximately five times more abundant; an additional band of 6.6kb was observed on prolonged exposure (Fig. 14). Jenkins et al. (1982) have detected mRNA species with similar sizes using radiolabelled Bgl II n. Since mRNA

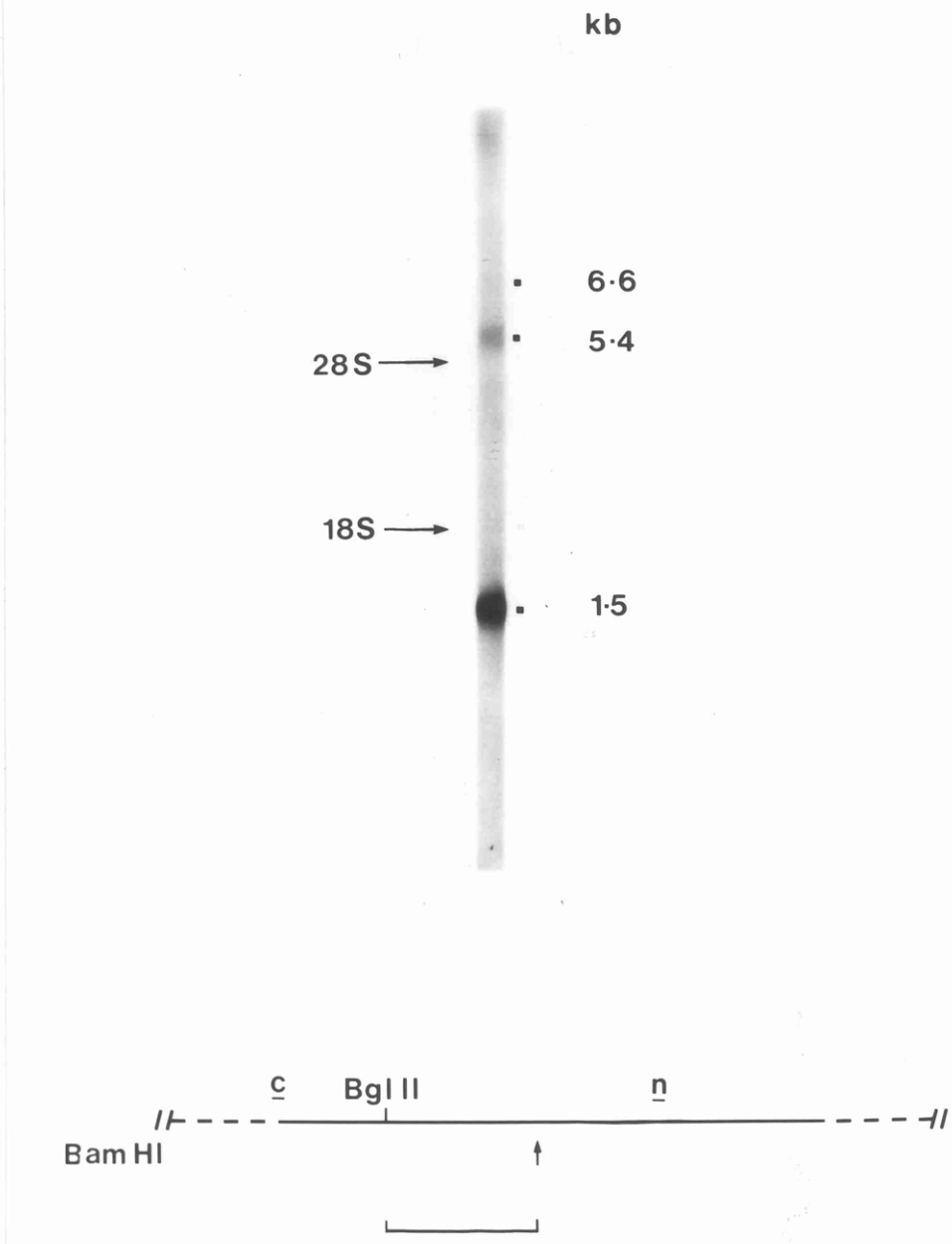

kb
6·6
5·4
28S
18S
1·5
c
Bgl II
n
Bam HI

Figure 14. Identification of mRNAs located within the left-hand portion of HSV-2 Bgl II n. The Northern blot contains 20µg of HSV-2 early cytoplasmic RNA and the location of the DNA probe, a nick-translated Bgl II/Bam HI fragment (0.580 to 0.583 map units), within Bgl II n is indicated. The sizes of mRNAs are shown with respect to 28S and 18S ribosomal RNA markers.

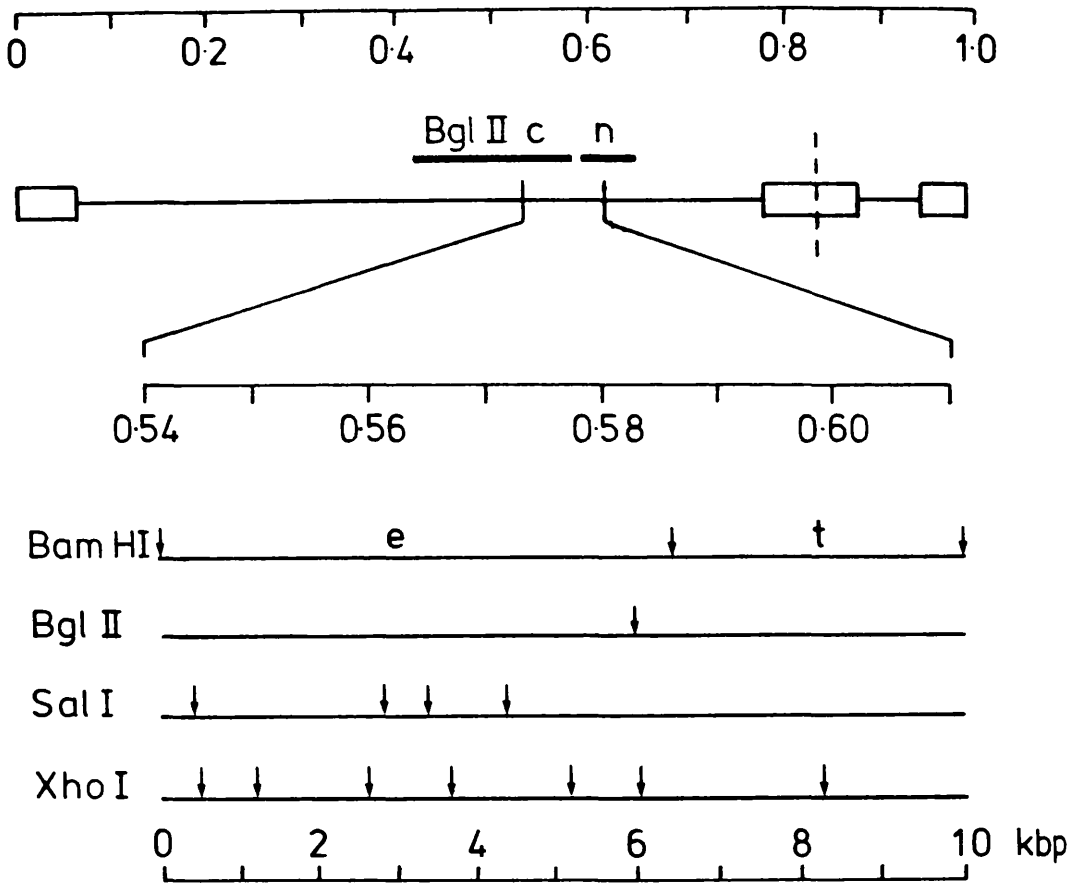


Figure 15. Locations of the cleavage sites for a number of restriction enzymes within HSV-2 Bam HI fragments e and t (0.54 to 0.61 map units). The locations of these Bam HI fragments on the HSV-2 genome with respect to HSV-2 transforming regions, Bgl II c and n, are shown.

sizes obtained with Northern blot analysis are inaccurate (see Methods, Page 117), nuclease digestion techniques were utilised to obtain more accurate size estimates and to study mRNA structures.

2. Structural Analysis of mRNAs.

For nuclease digestion analysis, two Bam HI fragments, Bam HI e and t, which span the region from 0.54 to 0.60 on HSV-2 DNA, were separately cloned into the Bam HI site of pAT153. Restriction enzyme cleavage sites in Bam HI e and t used to produce DNA probes are shown in Fig. 15.

a) An unspliced mRNA of 1.2kb corresponds to the 1.5kb species.

The Bgl II/Bam HI fragment, used in Northern blot analysis, was uniquely 5'-labelled at the Bam HI site and hybridised to infected and mock-infected cell RNAs. The resultant hybrids were digested with nuclease S1 and electrophoresed on a 6% denaturing polyacrylamide gel. A band of 143nuc was detected with infected cell RNA only (Fig. 16A, track 1), thus the 5' portion of a rightwards-transcribed mRNA is located within this fragment. The 5' portion of this mRNA was analysed further using Bam HI e, 5'-labelled at both ends. A nuclease S1 product of 143nuc and an exonuclease VII-resistant product of 148nuc were detected (Fig. 16B, tracks 2 and 4); the small size difference between these protected fragments is due to the processive nature of exonuclease VII activity

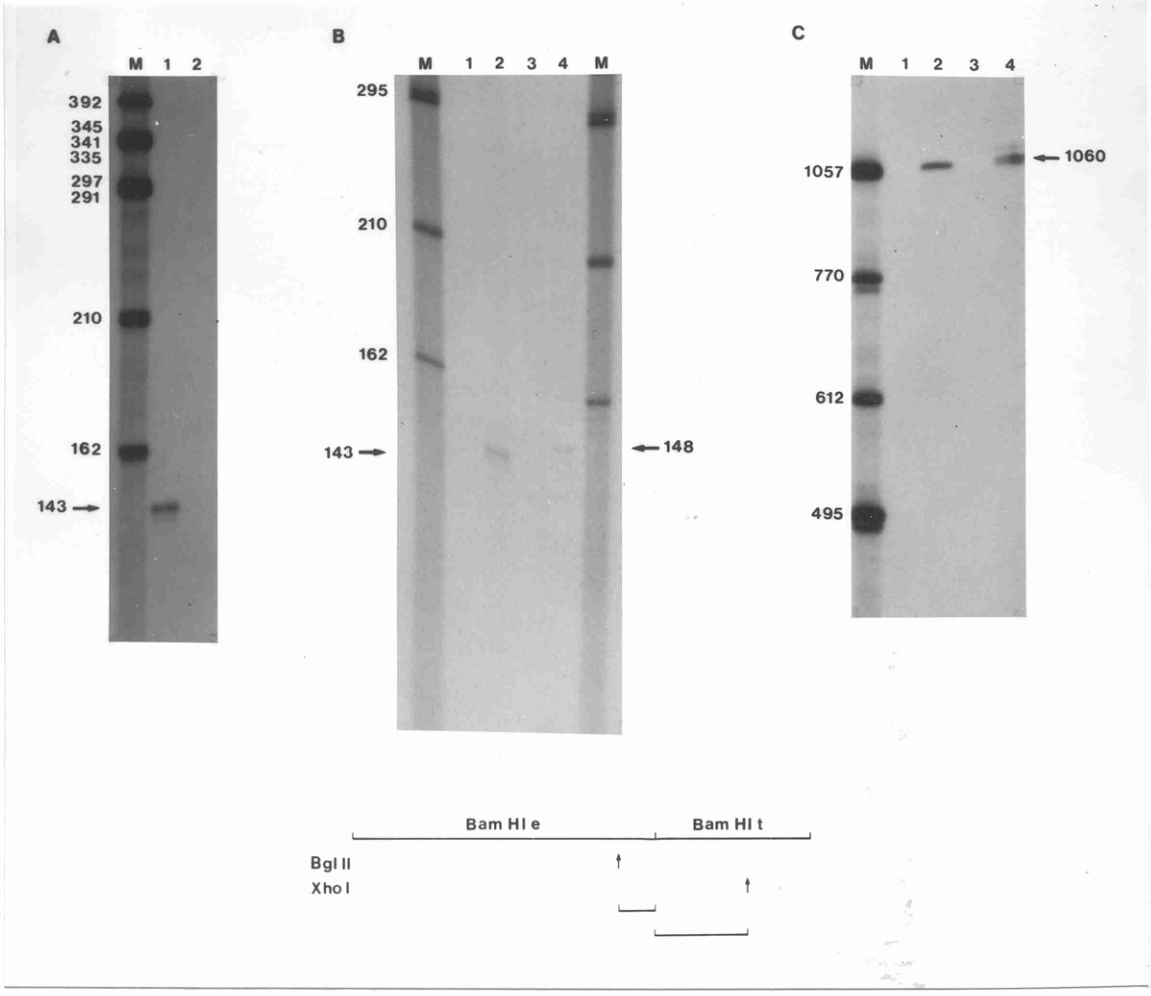

A
M 1 2
392
345
341
335
297
291
210
162
143
B
M 1 2 3 4 M
295
210
162
143
148
C
M 1 2 3 4
1060
1057
770
612
495
Bam HI e
Bam HI t
Bgl II
Xho I

Figure 16. Structural analysis of the 1.2kb mRNA.

A) Location of the 5' terminus of the 1.2kb transcript within the left-hand portion of Bgl II n. The DNA probe was a Bgl II/Bam HI fragment, as indicated on the diagram (0.580 to 0.583 map units), which was 5'-labelled at the Bam HI site. RNA samples used were: track 1, HSV-2 early infected cell cytoplasmic RNA; track 2, mock-infected cytoplasmic RNA.

B) Analysis of the 5' portion of the 1.2kb mRNA. The DNA probe was Bam HI e (see Fig. 15), 5'-labelled at both Bam HI sites. RNA samples used were: tracks 2 and 4, HSV-2 early infected cell cytoplasmic RNA; tracks 1 and 3, mock-infected cytoplasmic RNA.

C) Analysis of the 3' portion of the 1.2kb mRNA. The DNA probe was a Bam HI/Xho I fragment derived from Bam HI t, 3'-labelled at the Bam HI site. The location of the probe within Bam HI t is indicated on the diagram. RNA samples used were: tracks 2 and 4, HSV-2 early infected cell cytoplasmic RNA; tracks 1 and 3, mock-infected cytoplasmic RNA.

After hybridisation at 56°C, samples 1 and 2 in A, B and C were digested with nuclease S1 while samples 3 and 4 in B and C were treated with exonuclease VII. Nuclease-resistant material was electrophoresed on 6% denaturing polyacrylamide gels. Size standards (M) were ϕX174 DNA digested with Hinc II.

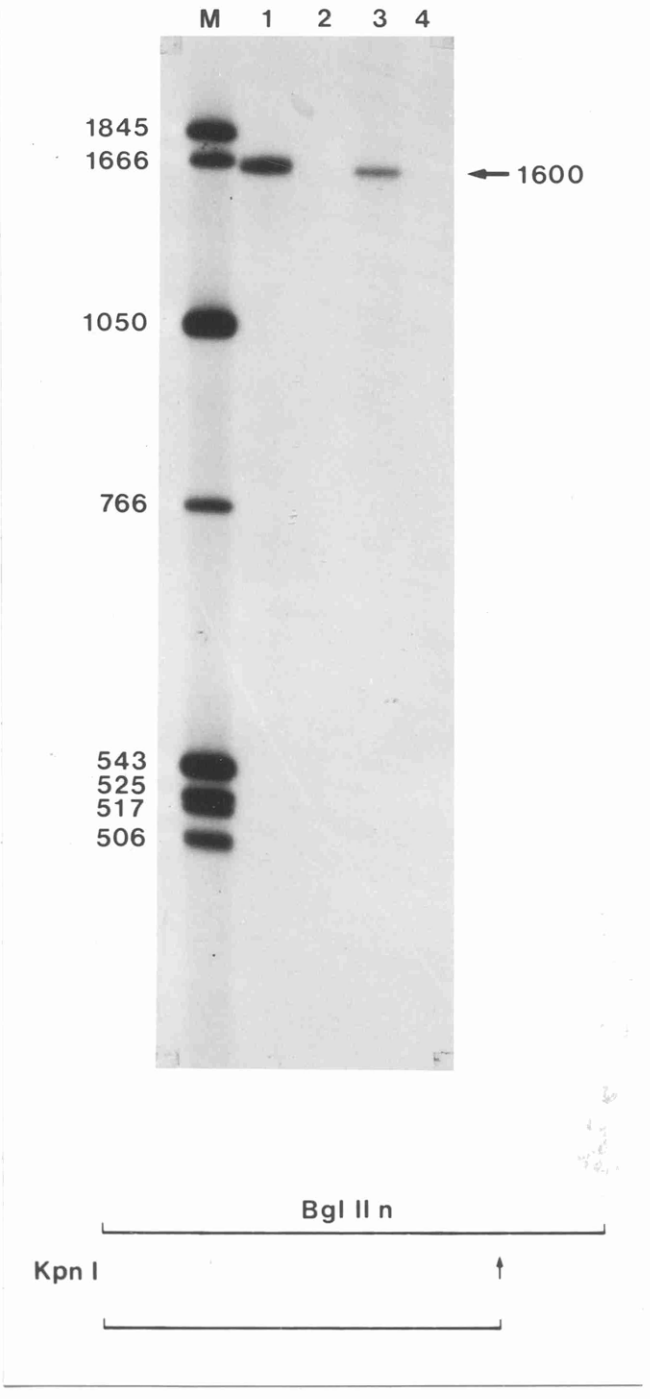
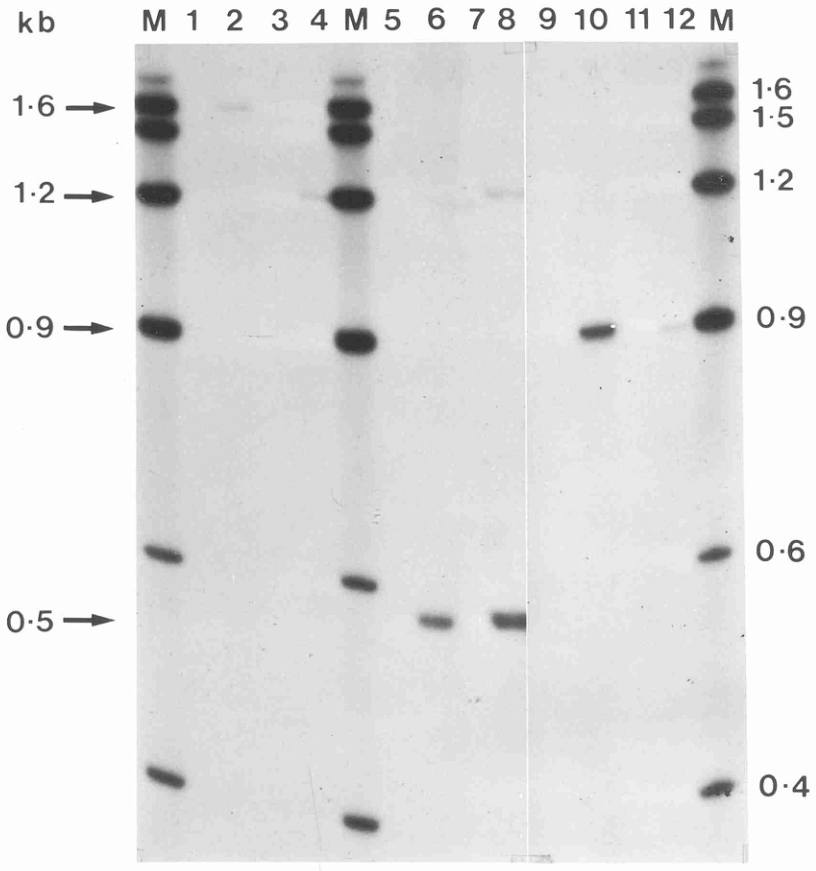

M
1
2
3
4
1845
1666
1600
1050
766
543
525
517
506
Bgl II n
Kpn I

Figure 17. Analysis of the 3' portion of the 4.5kb mRNA. The DNA probe, as indicated on the diagram, was a Bgl II/Kpn I fragment of Bgl II n, uniquely 3'-labelled at the Bgl II site. RNA samples used were: tracks 1 and 3, HSV-2 early infected cell cytoplasmic RNA; tracks 2 and 4, mock-infected cytoplasmic RNA. After hybridisation at 56°C, samples 1 and 2 were digested with nuclease S1 and samples 3 and 4 treated with exonuclease VII. Nuclease-resistant material was electrophoresed on a 6% denaturing polyacrylamide gel. Size standards (M) were pSV40 DNA digested with Hinf I.


kb
M 1 2 3 4 M 5 6 7 8 9 10 11 12 M
1·6
1·2
0·9
0·5
1·6
1·5
1·2
0·9
0·6
0·4

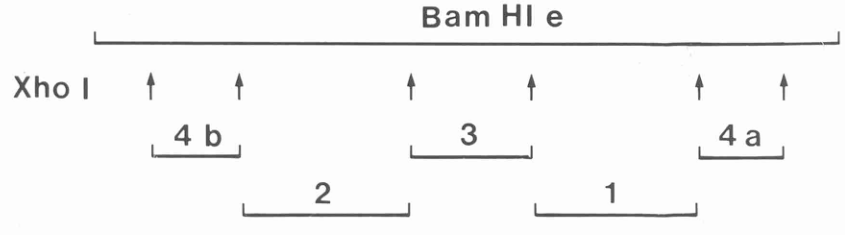

Bam HI e
Xho I
4 b
3
4 a
2
1

Figure 18. Analysis of mRNAs located upstream from the 1.2kb transcript. The DNA probes were a series of Xho I fragments derived from Bam HI e, numbered from 1 to 4 on the diagram, which were 5'-labelled at both termini. DNA probes for individual hybridisations were: tracks 1 and 2, Xho I fragment 1; tracks 3 and 4, Xho I fragment 2; tracks 5 to 8, Xho I fragment 3; tracks 9 and 10, Xho I fragment 4a; tracks 11 and 12, Xho I fragment 4b. RNA samples were HSV-2 early infected cell cytoplasmic RNA for the even-numbered tracks and mock-infected cytoplasmic RNA for odd-numbered tracks. After hybridisation at 56°C, samples were digested with nuclease S1 apart from samples 7 and 8 which were treated with exonuclease VII. Nuclease-resistant material was electrophoresed on a 6% denaturing polyacrylamide gel. Size markers (M) were the following Xho I fragments of Bam HI e:

the 1.6kb marker corresponds to Xho I fragment 1;

the 1.5kb marker corresponds to Xho I fragment 2;

the 1.2kb marker corresponds to Xho I fragment 3;

the 0.9kb marker corresponds to Xho I fragments 4a and 4b;

the 0.6kb and 0.4kb markers correspond to Xho I/Bam HI fragments.

(see Methods, Page 118). The result indicates that the 5' portion of a mRNA, transcribed across the Bam HI site, is unspliced.

The 3' portion of this mRNA was analysed using a Bam HI/Xho I fragment from Bam HI t, uniquely 3'-labelled at the Bam HI site. A band of 1.06kb was detected in both nuclease S1- and exonuclease VII-digested samples (Fig. 16C, tracks 2 and 4); it follows that the 3' portion is unspliced. Therefore, the total size of this mRNA is approximately 1.2kb and it corresponds to the 1.5kb species detected by Northern blot analysis. Additional minor products produced with this DNA probe (Fig. 16C, track 4) were detected also with other DNA probes (see Page 139 and Fig. 34, track 2).

b) An unspliced mRNA of 4.5kb corresponds to the 5.4kb species.

A mRNA was shown to be 3' co-terminal with and overlapping the 1.2kb mRNA by using a Bgl II/Kpn I fragment (0.580 to 0.612 map units) of Bgl II n, uniquely 3'-labelled at the Bgl II site. This Bgl II site is approximately 550bp upstream from the Bam HI site and therefore is located beyond the 1.2kb mRNA 5' terminus (Fig. 15). Nuclease S1- and exonuclease VII-resistant products of 1.6kb (Fig. 17, tracks 1 and 3) were detected which places the 3' terminus of this mRNA and that of the 1.2kb mRNA at the same location.

The 5' portion of this mRNA, which maps within the adjacent Bgl II c fragment, was analysed using Xho I fragments spanning Bam HI e. The Xho I fragments were 5'-labelled at both ends and are numbered from 1 to 4 in Fig. 18. Nuclease S1 analysis revealed that Xho I fragments 1 and 4a were fully protected by infected cell RNA (Fig. 18, tracks 2 and 10), indicating that mRNA is neither initiated nor spliced in these fragments. However, hybridisations with Xho I fragment 3 generated a 500nuc band (Fig. 18, track 6); exonuclease VII analysis with this fragment also produced a band of similar size (Fig. 18, track 8), indicating that the 5' terminus of a mRNA was located within Xho I fragment 3. The total size of this mRNA is 4.5kb and it corresponds to the 5.4kb species detected by Northern blot analysis.

c) An unspliced 6.4kb mRNA corresponds to the 6.6kb species.

Northern blot analysis indicated that a third mRNA of 6.6kb spans the Bgl II site at 0.58 map units; the equivalent HSV-1 region specifies a 7.0kb transcript which is 3' co-terminal with the early 5.0kb and 1.2kb species (Fig. 13). Nuclease S1 analysis with 5'-labelled Xho I fragment 2, which is adjacent to Xho I fragment 3 (Fig. 18), generated a band of approximately 1.2kb with infected cell RNA only (Fig. 18, track 4). Exonuclease VII digestions produced a protected DNA fragment of similar size (Fig. 19A, compare tracks 2 and 3), indicating that

the 1.2kb segment did not contain splice sites and that the 5' end of a HSV-2 mRNA, equivalent to the HSV-1 7.0kb species, is located within Xho I fragment 2. The predicted size of this HSV-2 transcript is 6.4kb and it corresponds to the 6.6kb species detected by Northern blot analysis.

d) Mapping of additional mRNA species within Bam HI e.

Nuclease S1 analysis with Xho I fragment 2 routinely generated three additional bands with sizes of 500nuc, 250nuc and 195nuc (Fig. 19A, track 3) which had similar sizes (500nuc, 270nuc and 198nuc) following exonuclease VII digestion (Fig. 19A, track 2). These nuclease-resistant fragments are thought to represent the 5' termini of additional transcripts mapping within Xho I fragment 2. To establish the orientation of these additional mRNAs, nuclease S1 digestions were performed using a 5'-labelled Xho I/Sst I DNA probe from the left-hand end of Xho I fragment 2 (Fig. 19). A 235nuc protected fragment (Fig. 19B, track 1) was not detected by nuclease S1 analysis with Xho I fragment 2 (Fig. 19A, track 3); this therefore corresponds to the 5' end of the 6.4kb transcript. The additional nuclease-resistant products in Fig. 19B, track 2 represent species detected by Xho I fragment 2 other than the rightwards-transcribed 6.4kb mRNA. Thus, these products locate the 5' termini of transcripts which are leftwards-transcribed, in the opposite orientation to the 6.4kb mRNA. The sizes of protected fragments suggest that the 5' terminus of the

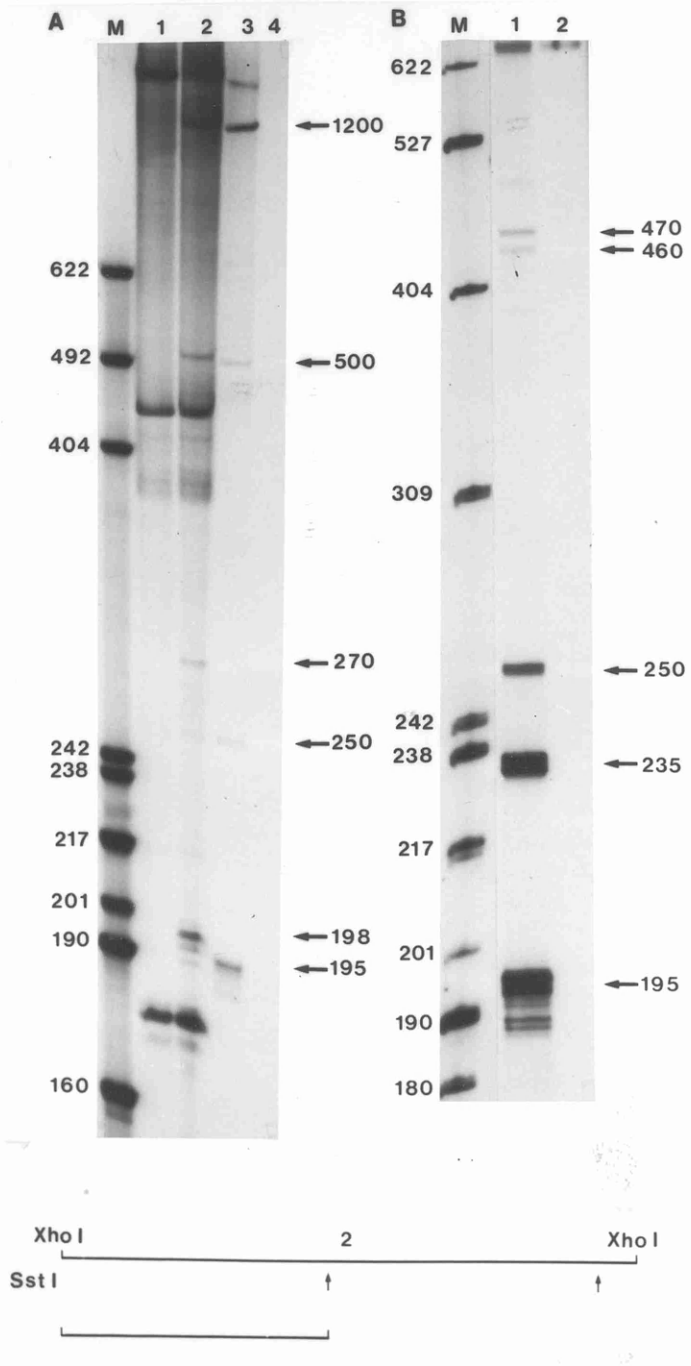

A
M
1
2
3
4
1200
500
270
250
198
195
622
492
404
242
238
217
201
190
160
B
M
1
2
622
527
404
309
242
238
217
201
190
180
470
460
250
235
195
Xho I
2
Xho I
Sst I

Figure 19. Analysis of the 5' portions of the 6.4kb mRNA and a series of leftwards-transcribed species.

A) Structural analysis of the mRNAs. The DNA probe was Xho I fragment 2, derived from Bam HI e (see Fig. 18), which was 5'-labelled at both sites. RNA samples used were: tracks 2 and 3, 16h infected cell cytoplasmic RNA; tracks 1 and 4, mock-infected cytoplasmic RNA.

B) Orientation of mRNAs within Xho I fragment 2. The DNA probe, shown on the diagram, was a Xho I/Sst I subfragment of Xho I fragment 2 which was 5'-labelled at both termini. RNA samples were 16h infected cell cytoplasmic RNA in track 1 and mock-infected cytoplasmic RNA in track 2.

After hybridisation at 56°C, samples 1 and 2 in A were digested with exonuclease VII while samples 3 and 4 in A and samples 1 and 2 in B were treated with nuclease S1. Digestion products were electrophoresed on 6% denaturing polyacrylamide gels. Size markers (M) were pAT153 DNA cleaved with Hpa II.

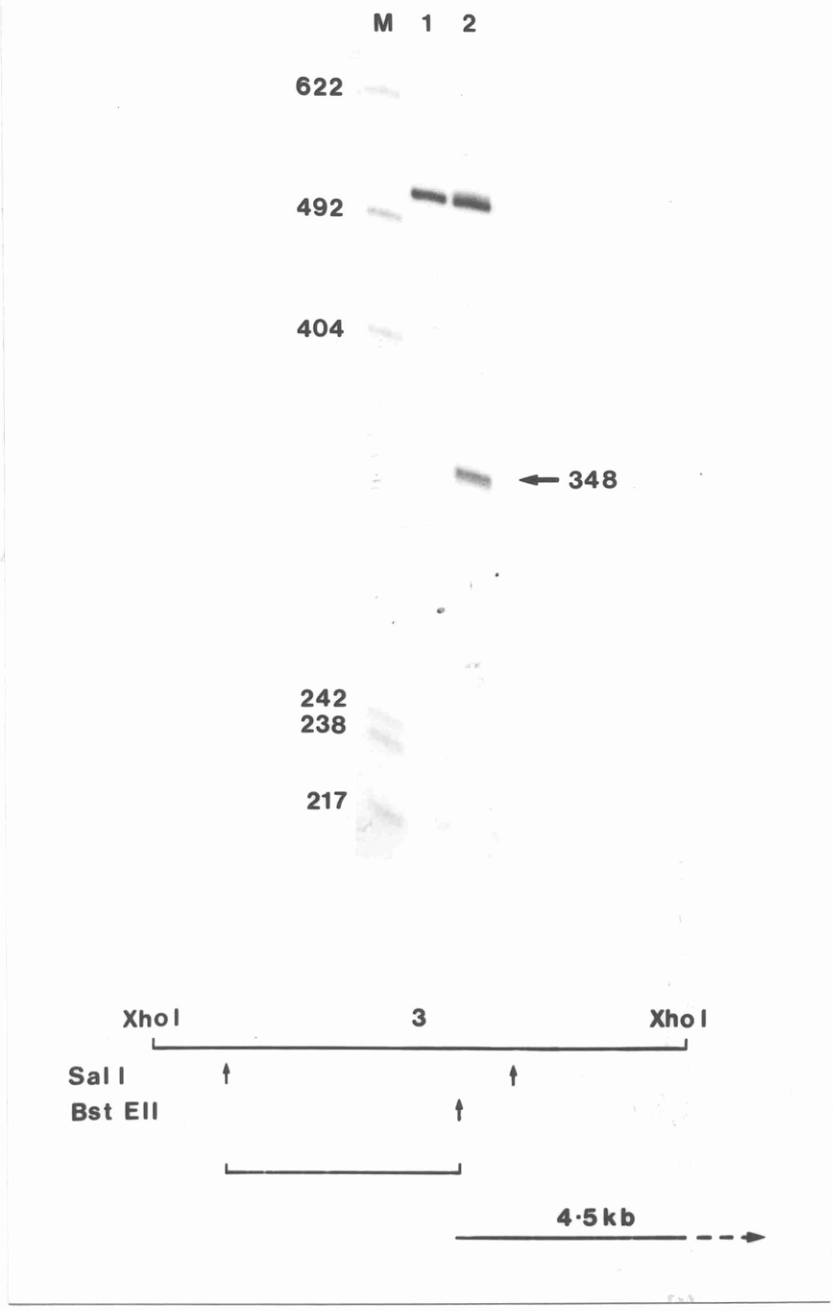

M 1 2
622
492
404
348
242
238
217
Xho I
3
Xho I
Sal I
Bst EII
4·5 kb

Figure 20. Location of the 3' end of the 1.7kb mRNA. The DNA probe, shown on the diagram, was a Sal I/Bst EII fragment derived from Xho I fragment 3 (see Fig. 18) which was uniquely 3'-labelled at the Sal I site. RNA preparations used were mock-infected cytoplasmic RNA in track 1 and 16h infected cell cytoplasmic RNA in track 2. After hybridisation at 56°C, samples were treated with nuclease S1 and digestion products were electrophoresed on a 6% denaturing polyacrylamide gel. Size markers (M) were pAT153 DNA digested with Hpa II.

6.4kb transcript overlaps with the 5' end of the mRNA species generating the 500nuc fragment by approximately 30 to 50 residues. The slight size difference between the 500nuc band in Fig. 19A (track 3), and the 470nuc and 460nuc nuclease S1-products in Fig. 19B (track 1) probably results from this 5' overlapping region where more efficient hybridisation of the abundant 6.4kb mRNA to the Xho I/Sst I subfragment prevents hybridisation of the less abundant leftwards-transcribed species, thus generating a truncated nuclease S1 product. Nuclease S1 analysis with Xho I fragment 4b produced a fully protected DNA probe (Fig. 18, track 12) which indicates that the three overlapping leftwards-transcribed mRNAs extend beyond this region. In the corresponding area of the HSV-1 genome, Northern blot analysis has identified a 3.8kb species (Fig. 13; Anderson et al., 1981).

e) Location of a poly A site upstream from the 4.5kb mRNA within the 6.4kb mRNA.

The HSV-1 locus equivalent to that analysed here, specifies a late 1.9kb mRNA which is presumed to be 5' co-terminal with the 7.0kb mRNA and terminates close to the 5.0kb mRNA 5' terminus (Anderson et al., 1981). To identify and map any equivalent HSV-2 transcript, a Sal I/Bst EII sub-fragment of Xho I fragment 3, which lies upstream from the 4.5kb mRNA, was 3'-labelled at the Sal I site and used for nuclease S1 analysis (Fig. 20). A protected DNA fragment of 348nuc in the infected cell RNA sample (Fig.

20, track 2) located a mRNA 3' terminus at approximately 200nuc upstream from the 4.5kb mRNA 5' terminus. As analysis of the 5' portion of the 6.4kb mRNA revealed a single mRNA initiation site (Fig. 19A, tracks 2 and 3), the 6.4kb mRNA and this overlapping transcript which terminates within the 6.4kb mRNA have identical 5' termini; this transcript has a size of 1.7kb.

3. Temporal Regulation of the Overlapping mRNAs.

The levels within the cytoplasm of the four nested mRNAs were examined by hybridising various DNA probes to infected cell RNA isolated at different times p.i. DNA probes used were:

1) 5'-labelled Bam HI/Bgl II fragment to detect the 1.2kb mRNA 5' terminus (see Fig. 16).

2) 5'-labelled Xho I fragment 3 to detect the 4.5kb mRNA 5' terminus (see Fig. 18).

3) 5'-labelled Xho I fragment 2 for the 5' portions of both 6.4kb and 1.7kb mRNAs (see Fig. 19).

4) the Sal I/Bst EII fragment, uniquely 3'-labelled at the Sal I site, to detect the 1.7kb mRNA 3' terminus (see Fig. 20).

The sizes of the various nuclease S1-resistant products generated from these hybridisations have been described in previous sections. The 4.5kb and 1.2kb mRNAs (indicated by ● and O respectively in Figs. 21A and B) are detectable at 2h p.i. and increase in abundance until 6h to 8h p.i., after which they no longer accumulate (Figs.

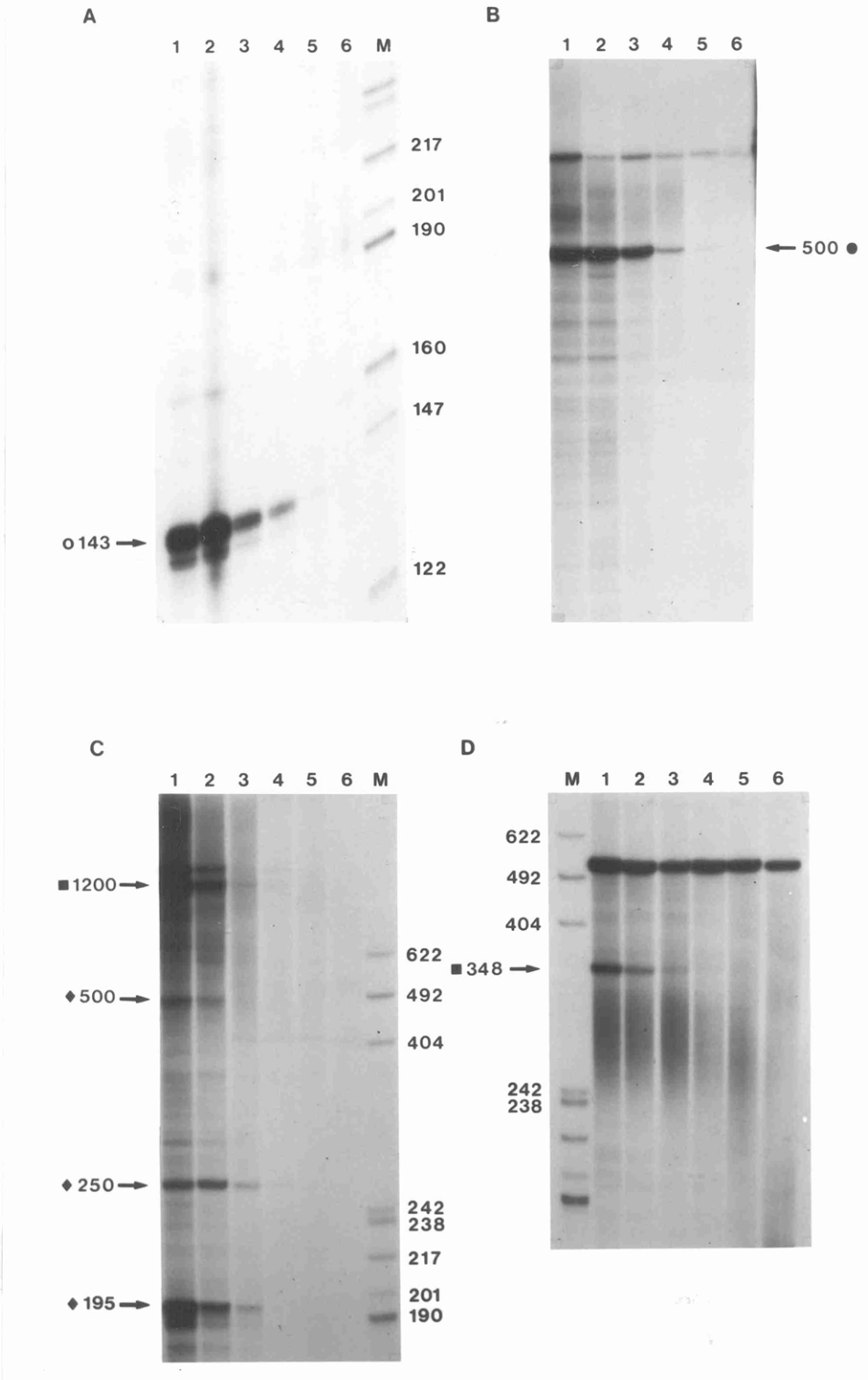

A
1 2 3 4 5 6 M
217
201
190
160
147
122
o 143
B
1 2 3 4 5 6
500 ●
C
1 2 3 4 5 6 M
■ 1200
♦ 500
♦ 250
♦ 195
622
492
404
242
238
217
201
190
D
M 1 2 3 4 5 6
622
492
404
■ 348
242
238

Figure 21. Temporal appearance of mRNAs located between 0.54 to 0.60 map units on the HSV-2 genome. The DNA probes used were:

A) a Bgl II/Bam HI fragment of Bgl II n (see legend to Fig. 16), which was 5'-labelled at the Bam HI site, to identify the 5' portion of the 1.2kb mRNA (represented by O);

B) 5'-labelled Xho I fragment 3 (see legend to Fig. 18) which identified the 5' end of the 4.5kb mRNA (represented by ●);

C) 5'-labelled Xho I fragment 2 (see legend to Fig. 19) which identified the 5' co-terminus of the 6.4kb and 1.7kb mRNAs (represented by ■) as well as the 5' ends of the leftwards-transcribed species (represented by ◆);

D) the Sal I/Bst EII fragment of Bam HI e, uniquely 3'-labelled at the Sal I site (see legend to Fig. 20), which identified the 3' terminus of the 1.7kb transcript (represented by ■).

RNA samples used were: track 1, 16h infected cell cytoplasmic RNA; track 2, 8h infected cell cytoplasmic RNA; track 3, 6h infected cell cytoplasmic RNA; track 4, 4h infected cell cytoplasmic RNA; track 5, 2h infected cell cytoplasmic RNA; track 6, mock-infected cytoplasmic RNA. Following hybridisation at 56°C, samples were digested with nuclease S1 and the nuclease-resistant products electrophoresed on 6% denaturing polyacrylamide gels. Size standards (M) were pAT153 DNA digested with Hpa II.

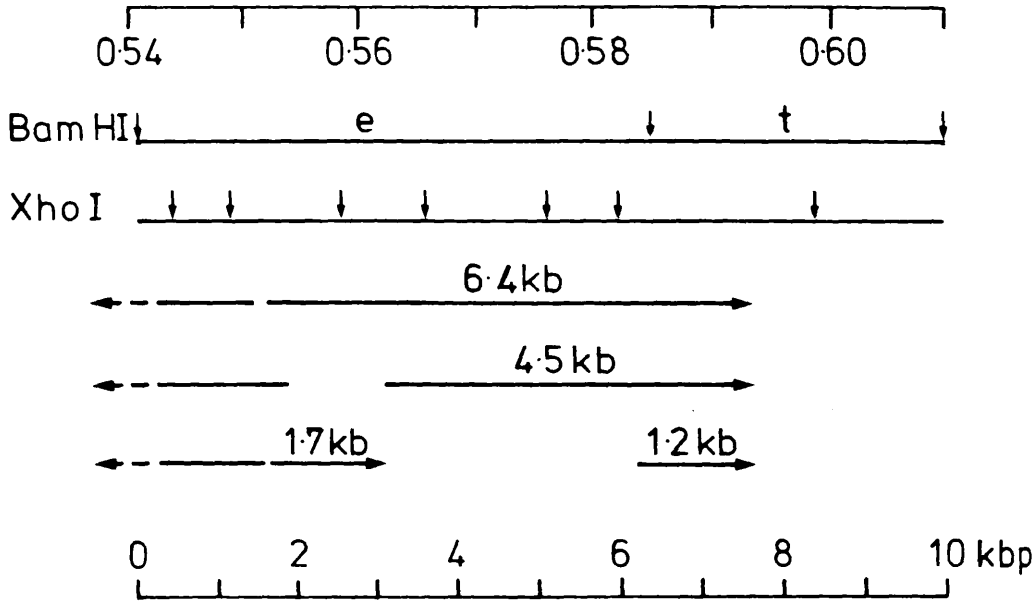


Figure 22. Arrangement and sizes of mRNAs mapping between 0.54 and 0.60 map units on the HSV-2 genome. Locations of the transcripts are shown with respect to Bam HI fragments e and t and the Xho I sites within these fragments.

21A and B, tracks 1 to 6). In contrast, the 6.4kb and 1.7kb mRNAs (indicated by ■ in Figs. 21C and D) do not appear in the cytoplasm until 6h p.i. and increase in abundance throughout the remainder of the infectious cycle (Figs. 21C and D, tracks 1 to 6). A similar pattern of temporal appearance was detected for the leftwards-transcribed species which map in Xho I fragment 2 (Fig. 21C, tracks 1 to 6, mRNAs indicated by ◆)∧ Thus, the 4.5kb and 1.2kb mRNAs are early species while the 6.4kb, 1.7kb and leftwards-transcribed transcripts are late products.

500nuc band does appear later than the other two transcripts. although the mRNA generating the

Fig. 22 details the organisation of this group of nested mRNAs between 0.54 to 0.60 map units; these HSV-2 mRNAs are arranged identically and have similar sizes to those of HSV-1. Furthermore, equivalent HSV-2 and HSV-1 transcripts share similar patterns of temporal appearance within the cytoplasm. HSV-2 polypeptides which map in this region have sizes of 138,000 mol. wt. (Vmw138) and 38,000 mol. wt. (Vmw38; Docherty et al., 1981; Galloway et al., 1982). Similarities in the sizes and arrangement of the HSV-1 and HSV-2 mRNAs suggest that the HSV-2 4.5kb mRNA encodes Vmw138 while the HSV-2 1.2kb mRNA specifies Vmw38. The 6.4kb and 1.7kb mRNAs are assumed to encode a polypeptide equivalent in size to the HSV-1 54,000 mol. wt. protein. Supportive evidence for this assumption is presented on Page 163.

4. DNA Sequencing Studies within Bam HI e and t.

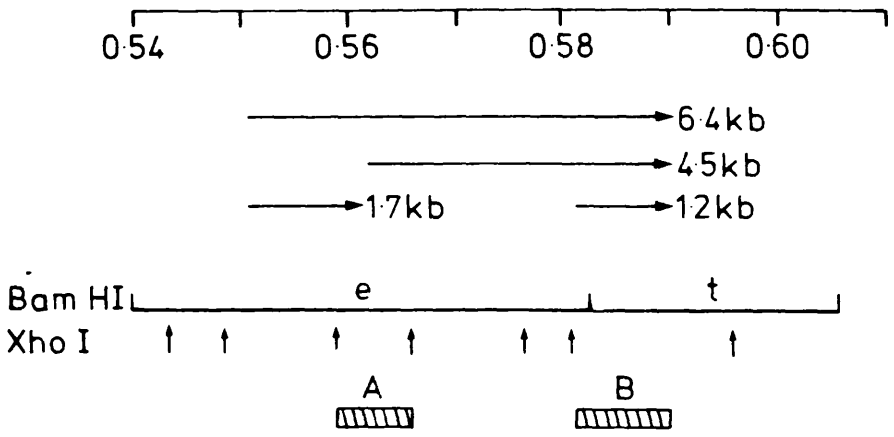
The nucleotide sequences of two HSV-1 DNA segments at this genome region have been published. These are:

1) DNA sequences flanking the 5' terminus of the early 5.0kb message (McLauchlan and Clements, 1983a) and
2) DNA sequences encoding the entire 1.2kb mRNA, including regions flanking the 5' and 3' termini (Draper et al., 1982; McLauchlan and Clements, 1982 and 1983a).

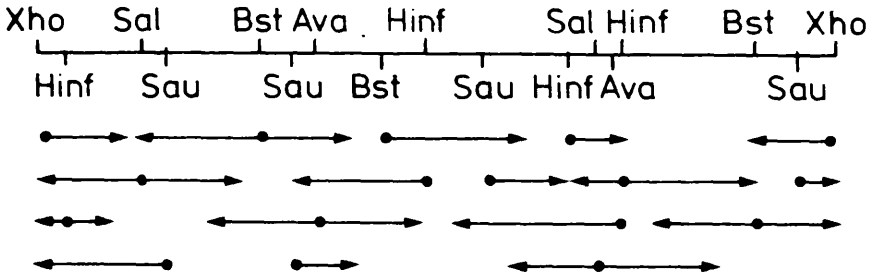
Nucleotide sequences were determined for the following corresponding regions encoding the equivalent HSV-2 mRNAs (Fig. 23):
1) a 1.12kbp DNA segment encompassing the 5' portion of the 4.5kb mRNA and the 3' terminal region of the 1.7kb mRNA.
2) a 1.6kbp segment which contains the transcribed region of the 1.2kb mRNA and the mRNA 5' and 3' terminal flanking sequences.

Restriction enzyme sites used to determine nucleotide sequences are shown in Figs. 23A and 23B and the complete nucleotide sequences are shown in Figs. 25 and 31.

The following section shows precise positions for the mRNA termini and identifies potential transcription control elements revealed by comparisons with the corresponding HSV-1 DNA sequences. Polypeptide coding regions within mRNAs have been predicted by computer studies and from comparisons with the corresponding HSV-1 DNA sequences. A complete analysis of polypeptide coding regions is presented in Section B (Page 148), but here, the positions of N- and C-termini are indicated as important reference points.


0·54
0·56
0·58
0·60
6·4kb
4·5kb
1·7kb
1·2kb
Bam HI
Xho I
e
t
A
B

A


Xho
Sal
Bst
Ava
Hinf
Sal
Hinf
Bst
Xho
Hinf
Sau
Sau
Bst
Sau
Hinf
Ava
Sau

B

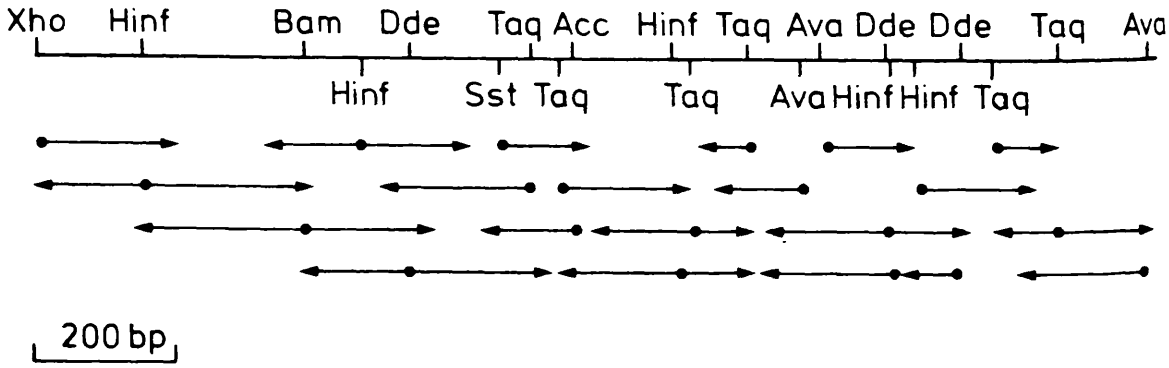

Xho
Hinf
Bam
Dde
Taq
Acc
Hinf
Taq
Ava
Dde
Dde
Taq
Ava
Hinf
Sst
Taq
Taq
Ava
Hinf
Hinf
Taq
200 bp

Figure 23. Regions within Bam HI e and t for which nucleotide sequences were determined. Locations of the two regions, (A) and (B), which were sequenced are shown below the restriction enzyme map and the positions of mRNAs are indicated with respect to Bam HI and Xho I restriction enzyme sites. Region A represents Xho I fragment 3 (see Fig. 18) which includes 3' and 5' terminal portions respectively of the 1.7kb and 4.5kb mRNAs. Region B includes the 1.2kb transcript and sequences flanking this transcript. (A) and (B) show the restriction enzyme sites used to determine nucleotide sequences for the respective regions and the extent of sequence data obtained from each site.

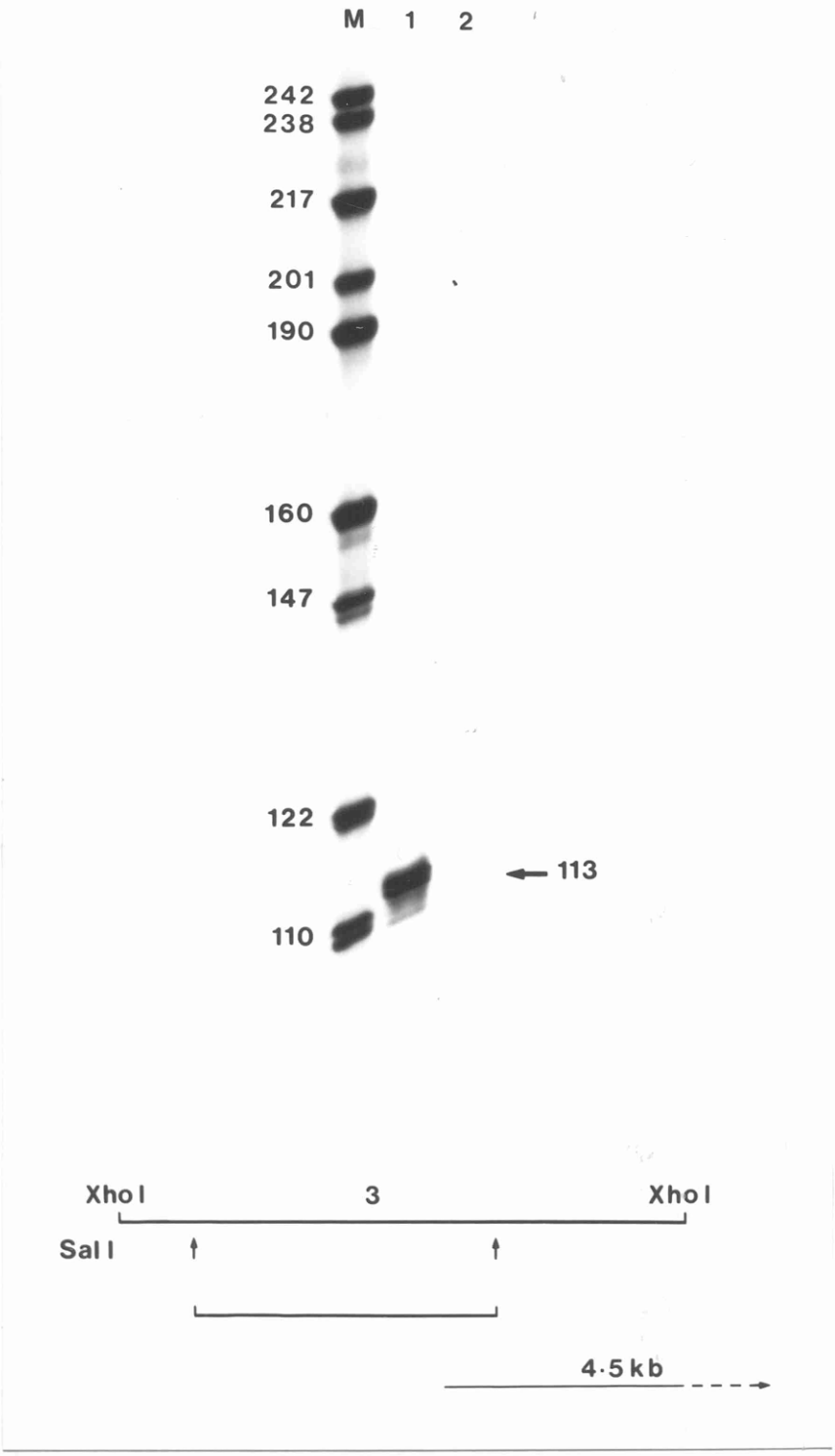

M
1
2
242
238
217
201
190
160
147
122
113
110
Xho I
3
Xho I
Sal I
4·5 kb

Figure 24. Precise location of the 5' terminus of the 4.5kb mRNA. The DNA probe was a 5'-labelled Sal I fragment whose position within Xho I fragment 3 (see Fig. 18) is indicated on the diagram. RNA samples were HSV-2 early infected cell cytoplasmic RNA for track 1 and mock-infected cytoplasmic RNA for track 2. After hybridisation at 56°C, samples were treated with nuclease S1 and the digestion products electrophoresed on a 6% denaturing polyacrylamide gel. The size markers (M) were pAT153 DNA digested with Hpa II.

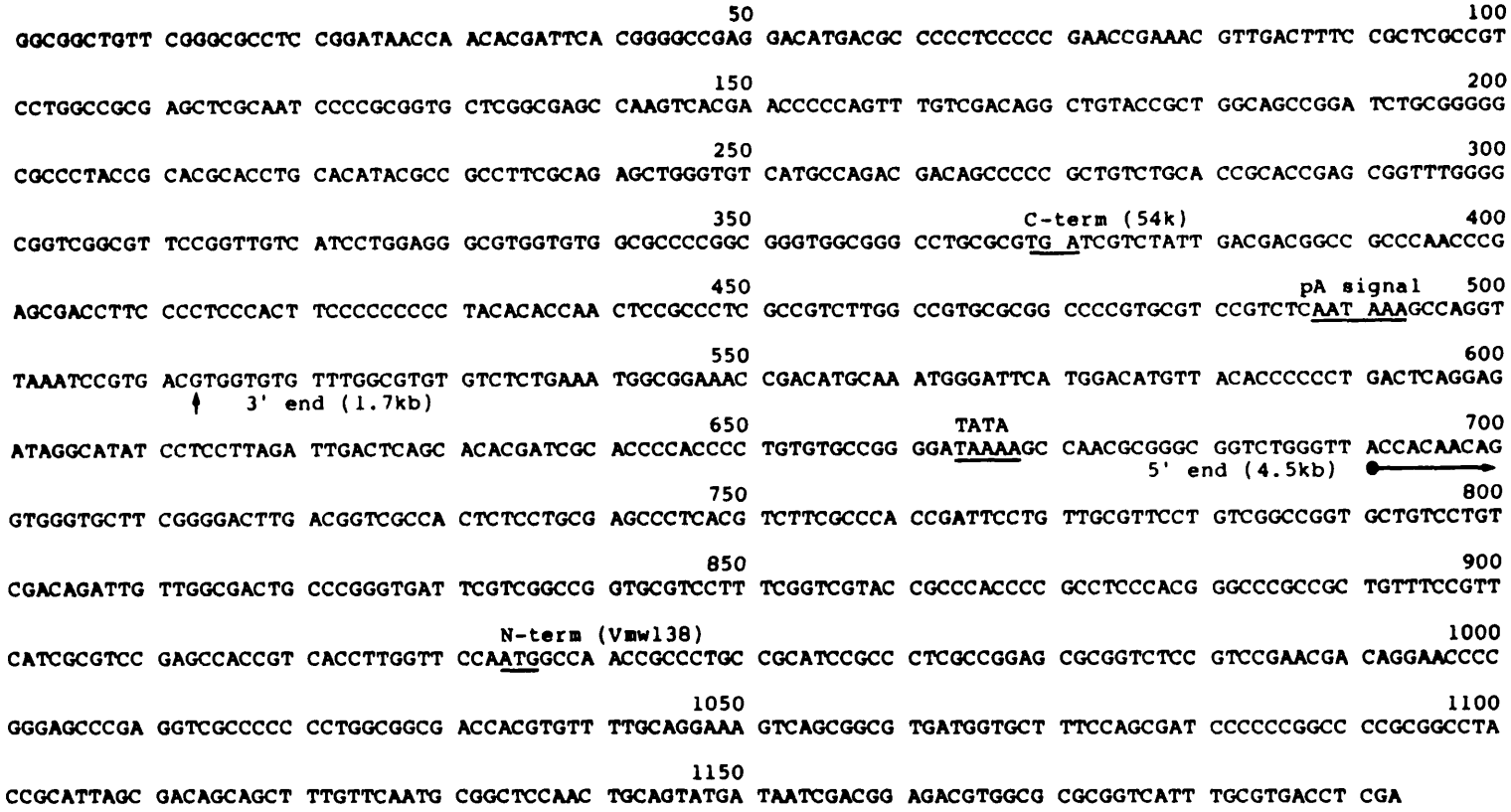

50
100
GGCGGCTGTT CGGGCGCCTC CGGATAACCA ACACGATTCA CGGGGCCGAG GACATGACGC CCCCTCCCCC GAACCGAAAC GTTGACTTTC CGCTCGCCGT
150
200
CCTGGCCGCG AGCTCGCAAT CCCCGCGGTG CTCGGCGAGC CAAGTCACGA ACCCCCAGTT TGTCGACAGG CTGTACCGCT GGCAGCCGGA TCTGCGGGGG
250
300
CGCCCTACCG CACGCACCTG CACATACGCC GCCTTCGCAG AGCTGGGTGT CATGCCAGAC GACAGCCCCC GCTGTCTGCA CCGCACCGAG CGGTTTGGGG
350
C-term (54k)
400
CGGTCGGCGT TCCGGTTGTC ATCCTGGAGG GCGTGGTGTG GCGCCCCGGC GGGTGGCGGG CCTGCGCGTG ATCGTCTATT GACGACGGCC GCCCAACCCG
450
pA signal
500
AGCGACCTTC CCCTCCCACT TCCCCCCCCC TACACACCAA CTCCGCCCTC GCCGTCTTGG CCGTGCGCGG CCCCGTGCGT CCGTCTCAAT AAAGCCAGGT
550
600
TAAATCCGTG ACGTGGTGTG TTTGGCGTGT GTCTCTGAAA TGGCGGAAAC CGACATGCAA ATGGGATTCA TGGACATGTT ACACCCCCCT GACTCAGGAG
3' end (1.7kb)
650
TATA
700
ATAGGCATAT CCTCCTTAGA TTGACTCAGC ACACGATCGC ACCCCACCCC TGTGTGCCGG GGATAAAAGC CAACGCGGGC GGTCTGGGTT ACCACAACAG
5' end (4.5kb)
750
800
GTGGGTGCTT CGGGGACTTG ACGGTCGCCA CTCTCCTGCG AGCCCTCACG TCTTCGCCCA CCGATTCCTG TTGCGTTCCT GTCGGCCGGT GCTGTCCTGT
850
900
CGACAGATTG TTGGCGACTG CCCGGGTGAT TCGTCGGCCG GTGCGTCCTT TCGGTCGTAC CGCCCACCCC GCCTCCCACG GGCCCGCCGC TGTTTCCGTT
N-term (Vmw138)
1000
CATCGCGTCC GAGCCACCGT CACCTTGGTT CCAATGGCCA ACCGCCCTGC CGCATCCGCC CTCGCCGGAG CGCGGTCTCC GTCCGAACGA CAGGAACCCC
1050
1100
GGGAGCCCGA GGTCGCCCCC CCTGGCGGCG ACCACGTGTT TTGCAGGAAA GTCAGCGGCG TGATGGTGCT TTCCAGCGAT CCCCCCGGCC CCGCGGCCTA
1150
CCGCATTAGC GACAGCAGCT TTGTTCAATG CGGCTCCAAC TGCAGTATGA TAATCGACGG AGACGTGGCG CGCGGTCATT TGCGTGACCT CGA

Figure 25. Nucleotide sequence of region (A) in Fig. 23. Indicated are the positions of the 3' and 5' termini of the 1.7kb and 4.5kb mRNAs respectively. The AATAAA signal and TATA box homologue are underlined as are proposed locations for the C-terminus of the 54,000 mol. wt. polypeptide and the N-terminus of Vmw138.

5. Analysis of mRNA 5' and 3' Termini.

a) 5' end of the 4.5kb mRNA.

A precise location for the 5' terminus of this mRNA was determined by nuclease S1 analysis using a 5'-labelled Sal I fragment, derived from Xho I fragment 3 (Fig. 25, positions 162 to 799). A nuclease-resistant fragment of 113nuc (Fig. 24, track 1) positioned the 5' terminus on the A residue (Fig. 25, position 691) within the sequence GGTTACC, a Bst EII restriction endonuclease cleavage site. Approximately 23nuc upstream from the 5' end, a TATA box sequence (see Introduction, Page 61) is located between positions 664 to 668 (Fig. 25).

HSV-2/HSV-1 nucleotide sequence comparisons in this region indicate that both TATA box and cap site sequences are well conserved (Fig. 26). The nucleotide sequences are essentially colinear to the limit of the HSV-1 sequence except for an additional A residue at position 668 in the HSV-1 sequence and an extra G residue at position 629 of the HSV-2 sequence (Fig. 26). Upstream from the TATA box homologue, there are two regions of highly conserved DNA sequences which are separated by an area of lower homology. The 5' proximal area of high homology is located between positions -39 and -49 and has the sequence CCCCACCCCTG (Fig. 26, positions 642 to 652). The 5' distal conserved sequence is between positions -92 and -110 (Fig. 26, positions 581 to 599); this area of homology can be divided into two separate blocks. The first block has the sequence

ACACCCCC which resembles A+C-rich elements located approximately 90 to 120nuc upstream from the 5' termini of a number of early HSV mRNAs (Wagner, 1985). The second block has the sequence TGACTCAGGA (shown by R1 in Fig. 26), most of which is repeated (TGACTCAG) further downstream in HSV-2 DNA between positions -62 and -69, however, this second copy is not conserved in HSV-1 DNA (Fig. 26, positions 622 to 629).

The distance between the cap site of the HSV-2 4.5kb mRNA and proposed initiation codon for Vmw138 is 243nuc while the equivalent 5' untranslated leader sequence in the HSV-1 5.0kb mRNA is 17nuc shorter (Fig. 26). A feature of both HSV-2 and HSV-1 sequences is the occurrence of direct repeats which are not conserved in both sequences. An 18nuc segment, located between positions 781 to 798 (shown by R2 in Fig. 26) in the HSV-2 sequence is repeated between positions 833 to 849 apart from an additional T residue in the 5' proximal R2 segment; these repeats are not conserved in the corresponding region of the HSV-1 sequence. The HSV-1 sequence contains three copies of a CCCGCCG motif (shown by R3 in Fig. 26) upstream from the initiation codon of Vmw136. Of these three copies, only the 5' proximal repeat unit is present in the HSV-2 sequence.

b) 1.7kb mRNA 3' terminus.

Nuclease S1 analysis located the 3' terminus of this mRNA some 348nuc downstream from the Sal I site (Fig.

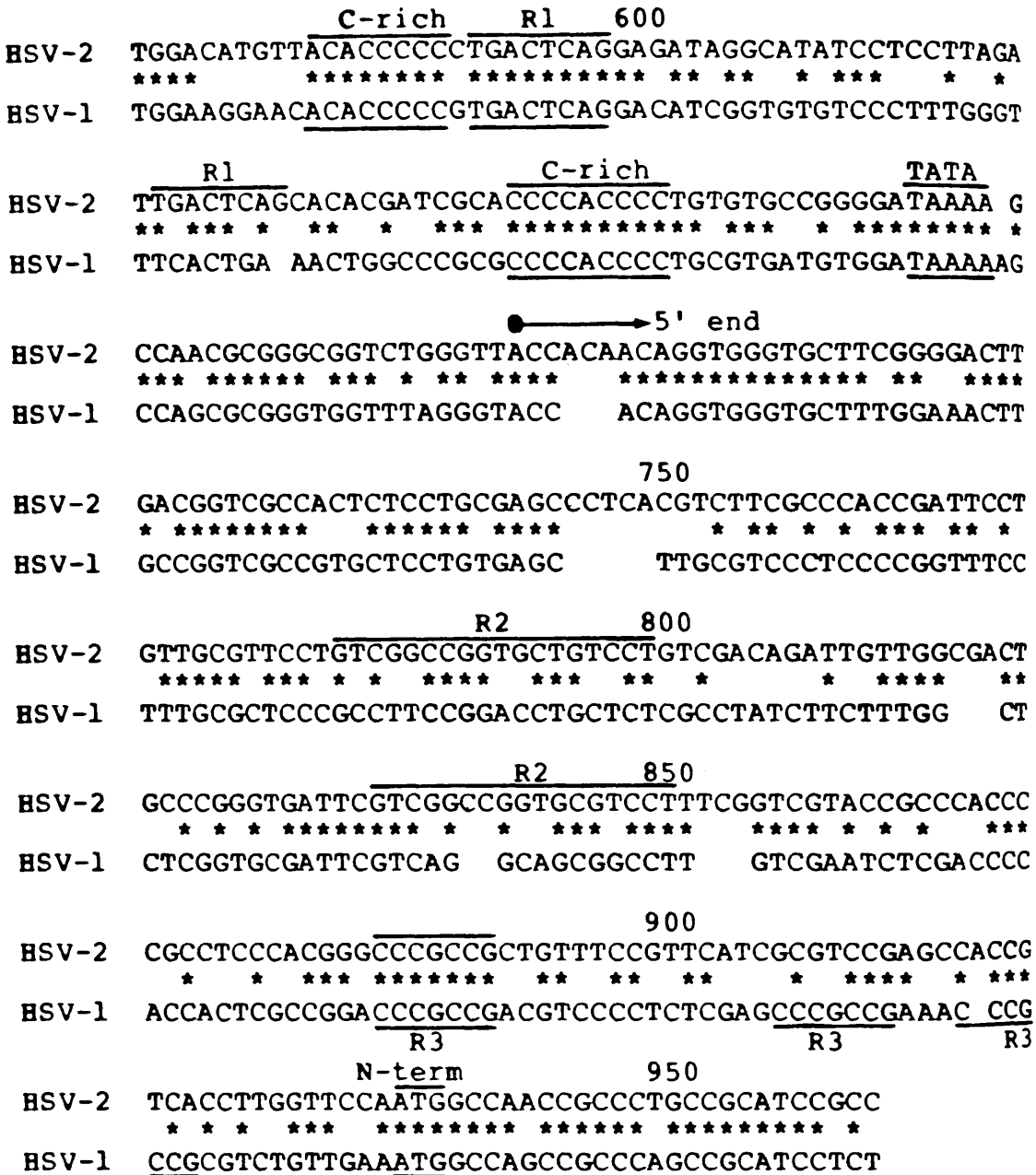

C-rich
R1
600
HSV-2 TGGACATGTTACACCCCCCTGACTCAGGAGATAGGCATATCCTCCTTAGA
HSV-1 TGGAAGGAACACACCCCCGTGACTCAGGACATCGGTGTGTCCCTTTGGGT
R1
C-rich
TATA
HSV-2 TTGACTCAGCACACGATCGCACCCCACCCCTGTGTGCCGGGGATAAAA G
HSV-1 TTCACTGA AACTGGCCCGCGCCCCACCCCTGCGTGATGTGGATAAAAAG
5' end
HSV-2 CCAACGCGGGCGGTCTGGGTTACCACAACAGGTGGGTGCTTCGGGGACTT
HSV-1 CCAGCGCGGGTGGTTTAGGGTACC ACAGGTGGGTGCTTTGGAAACTT
750
HSV-2 GACGGTCGCCACTCTCCTGCGAGCCCTCACGTCTTCGCCCACCGATTCCT
HSV-1 GCCGGTCGCCGTGCTCCTGTGAGC TTGCGTCCCTCCCCGGTTTCC
R2
800
HSV-2 GTTGCGTTCCTGTCGGCCGGTGCTGTCCTGTCGACAGATTGTTGGCGACT
HSV-1 TTTGCGCTCCCGCCTTCCGGACCTGCTCTCGCCTATCTTCTTTGG CT
R2
850
HSV-2 GCCCGGGTGATTCGTCGGCCGGTGCGTCCTTTCGGTCGTACCGCCCACCC
HSV-1 CTCGGTGCGATTCGTCAG GCAGCGGCCTT GTCGAATCTCGACCCC
900
HSV-2 CGCCTCCCACGGGCCCGCCGCTGTTTCCGTTCATCGCGTCCGAGCCACCG
HSV-1 ACCACTCGCCGGACCCGCCGACGTCCCCTCTCGAGCCCGCCGAAAC CCG
R3
R3
R3
N-term
950
HSV-2 TCACCTTGGTTCCAATGGCCAACCGCCCTGCCGCATCCGCC
HSV-1 CCGCGTCTGTTGAAATGGCCAGCCGCCCAGCCGCATCCTCT

Figure 26. Nucleotide sequence comparison between the 5' terminal portions of the HSV-2 4.5kb and HSV-1 5.0kb transcripts. The HSV-2 sequence is numbered as in Fig. 25 and asterisks denote conserved bases. The position of the 5' end of the 4.5kb mRNA is indicated; the 5' terminus of the 5.0kb transcript is at the same position in the HSV-1 sequence. TATA box homologues and conserved C-rich elements are overlined on the HSV-2 sequence and underlined on the HSV-1 sequence. The locations of repetitive sequences, R1, R2 and R3, are indicated in a similar manner in both sequences. Positions for the N-termini of HSV-2 Vmw138 and HSV-1 Vmw136 are also shown.

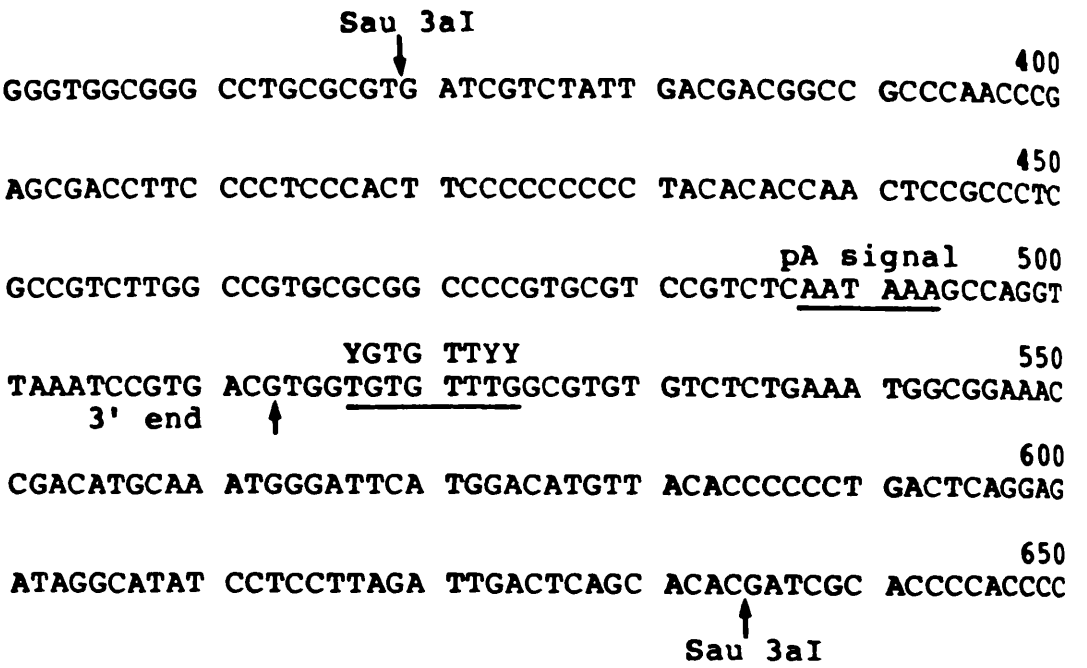


Figure 27. Nucleotide sequence of the Sau 3aI fragment which contains the 3' terminal portion and flanking sequences of the 1.7kb mRNA. The sequence is numbered as in Fig. 25 and the position of the mRNA 3' terminus is arrowed. Also shown are the locations of the AATAAA and YGTGTTYY signals.

20, Page 129) at position 513 (Fig. 25). This location for the mRNA 3' terminus is downstream from a polyadenylation signal, AATAAA (Fig. 25, positions 488 to 493), a sequence which is present upstream from the poly A site of most eukaryotic polyadenylated transcripts (see Introduction, Page 87). Moreover, 29nuc beyond the AATAAA signal is the DNA sequence TGTGTTTG (Fig. 25, positions 517 to 524) which fits the 3' terminal consensus sequence YGTGTTYY and is located at the appropriate distance from the AATAAA sequence (see Section E, Page 174). Thus, the 1.7kb mRNA 3' terminus, which lies within the transcribed region of the 6.4kb mRNA, possesses signals similar to those present at other eukaryotic polyadenylation sites. The distance between the 1.7kb mRNA 3' terminus and the 5' terminus of the 4.5kb transcript is 178nuc.

c) <u>DNA sequences flanking the 1.7kb mRNA 3' terminus produce functional transcripts</u>.

To verify that DNA sequences at the 1.7kb mRNA 3' terminus contain signals necessary to terminate transcription and to produce correctly-processed transcripts, a 265bp Sau 3aI fragment (Fig. 27) which comprises DNA sequences flanking the 1.7kb mRNA 3' terminus, including the AATAAA and YGTGTTYY motifs, was inserted into the CAT plasmid, pLW1. Plasmid pLW1 contains the HSV-2 IE gene-4/-5 promoter linked to the bacterial CAT gene but lacks signals necessary to produce correctly-terminated CAT mRNA (see Methods, Page 99 and

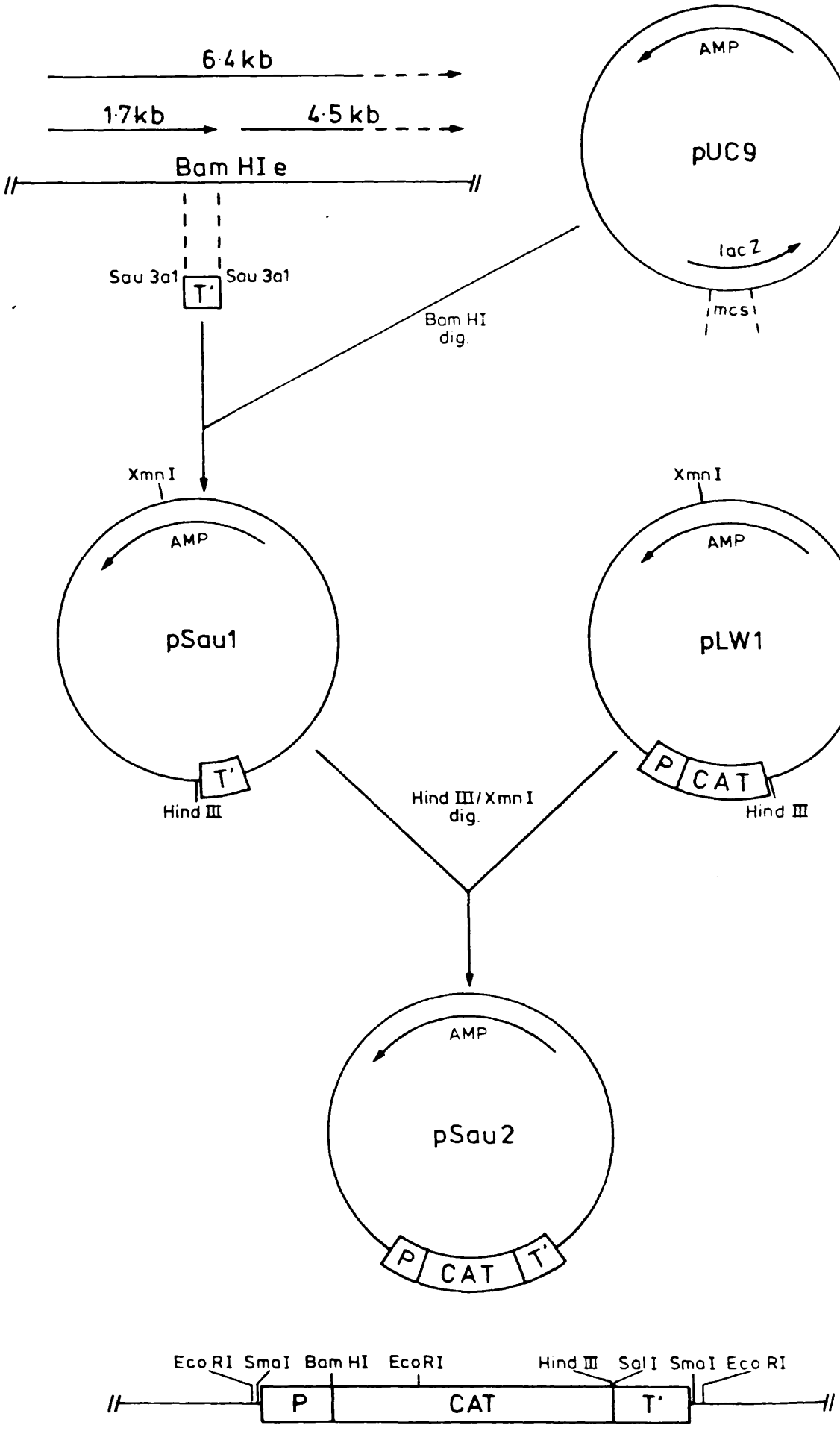

6·4kb
1·7kb
4·5kb
Bam HI e
Sau 3a1
T'
Sau 3a1
AMP
pUC9
lacZ
mcs
Bam HI
dig.
Xmn I
AMP
pSau1
T'
Hind III
Xmn I
AMP
pLW1
P
CAT
Hind III
Hind III/Xmn I
dig.
AMP
pSau2
P
CAT
T'
EcoRI SmaI Bam HI EcoRI
Hind III SalI SmaI EcoRI
P
CAT
T'

Figure 28. Construction of plasmid pSau2. The Sau 3a1 fragment (T') shown in Fig. 27 was inserted into the Bam HI site of pUC9 to yield pSau1. This plasmid was cleaved with Hind III and Xmn I and fused to the Hind III/Xmn I fragment of pLW1, which carries the promoter (P) and CAT modules, to give plasmid pSau2. A number of restriction enzyme sites within pSau2 are shown and details of plasmids pUC9 and pLW1 are given in Fig. 10.

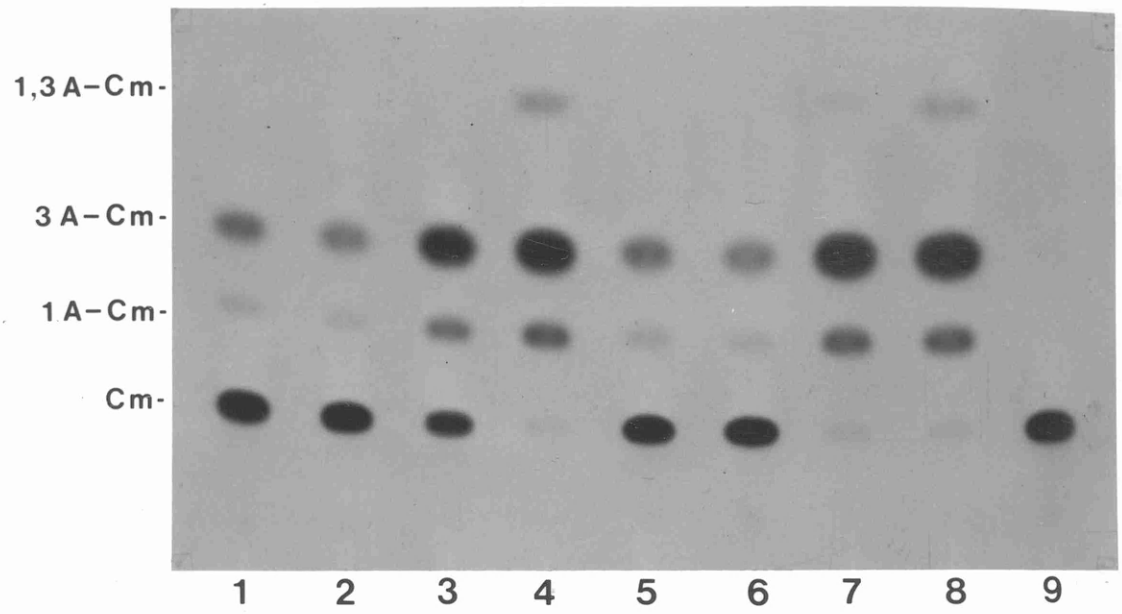


Figure 29. CAT activities produced by plasmids during the course of virus infection as determined by thin layer chromatography. Extracts were incubated at 37°C for 30min. The positions of acetylated (A-Cm) and non-acetylated forms (Cm) of ^{14}C chloramphenicol are indicated. HeLa cells were transfected with the following plasmids: tracks 1 to 4, pLW2; tracks 5 to 8, pSau2; track 9, pLW1. 24h after transfection, HSV-1 virus was added and cells were harvested at the following times p.i.: tracks 1, 5 and 9, mock-infected; tracks 2 and 6, 4h p.i.; tracks 3 and 7, 8h p.i.; tracks 4 and 8, 12h p.i.

Fig. 10). Firstly, the 1.7kb mRNA Sau 3aI fragment (T') was cloned into the Bam HI site of pUC9 to yield plasmid pSau1 (Fig. 28). A Hind III/Xmn I fragment containing the T' fragment was then fused to a Hind III/Xmn I fragment of pLW1 which contained the IE promoter and CAT gene; the resultant construct was termed pSau2 (Fig. 28).

The ability of pSau2 to produce functional CAT mRNA in HeLa cells was compared with CAT levels detected in cells transfected with plasmids pLW2 and pLW1. Plasmid pLW2 possesses a 'terminator' fragment derived from the 3' terminal region of HSV-2 IE gene-5 and is essentially identical to plasmid pTER5 (see Section E, Page 176). Results indicated that plasmids pLW2 and pSau2 produced similar levels of CAT activity which were significantly higher than those produced by pLW1 (Fig. 29, tracks 1, 5 and 9; Table 1). Thus, the 3' terminal region of the 1.7kb mRNA does not require additional DNA sequences to terminate mRNA synthesis and process primary transcripts. Cells transformed with either pLW2 or pSau2 were superinfected with HSV-1 strain 17^{+} virus and CAT activities in pLW2- and pSau2-transformed cells were analysed at various times p.i. The increase in CAT activity during virus infection was similar for both pLW2 and pSau2 (Fig. 29, tracks 1 to 8; Table 1), further indicating that the 1.7kb mRNA 'terminator' does not differ from other HSV 3' terminal sequences which are not located within overlapping transcripts.

	Plasmid		
Time p.i.	pLW2	pSau2	pLW1
MI	0.29	0.42	0.02
4h	0.33	0.53	ND
8h	1.87	2.43	ND
12h	2.8	2.87	ND

Table 1. CAT activities obtained with various plasmids during the course of virus infection. Activities are expressed as nmoles converted chloramphenicol/µg protein/h. ND indicates CAT activity not determined.

d) 5' terminus of the 1.2kb mRNA.

Nuclease S1 analysis using a Bam HI/Bgl II fragment, 5'-labelled at the Bam HI site generated a 143nuc band which represented the 1.2kb mRNA 5' terminus (see Figs. 16A and B). This nuclease S1-resistant band was electrophoresed alongside the sequence reaction products of the hybridisation probe (Fig. 30, track 2), thus locating the major cap site on the G residue within the sequence ATGTAC (Fig. 31, positions 245 to 250). Upstream from the 5' end, a TATA box sequence is located at positions -22 to -28 (Fig. 31, positions 221 to 225).

HSV-2/HSV-1 nucleotide sequence comparisons in this region reveal that the cap site and TATA box sequences are completely conserved and are located at identical positions relative to the 5' ends of the HSV-2 and HSV-1 1.2kb mRNAs (Fig. 32). The proposed stop codon for Vmw138, the protein encoded by the 4.5kb mRNA, is located at position 329 (Fig. 32), some 82nuc downstream from the 1.2kb mRNA 5' terminus. Therefore, transcription control elements for the 1.2kb mRNA lie within the translated region of the 4.5kb mRNA; this arrangement is identical to that for the equivalent HSV-1 mRNAs.

As the HSV-2 and HSV-1 nucleotide sequences share extremely high homology in this translated region, identification of transcription regulatory signals is difficult. However, conserved sequences upstream from the 5' termini of the HSV-1 and HSV-2 1.2kb mRNAs contain inverted repeat elements (R1 and R1', R2 and R2' in Figs.

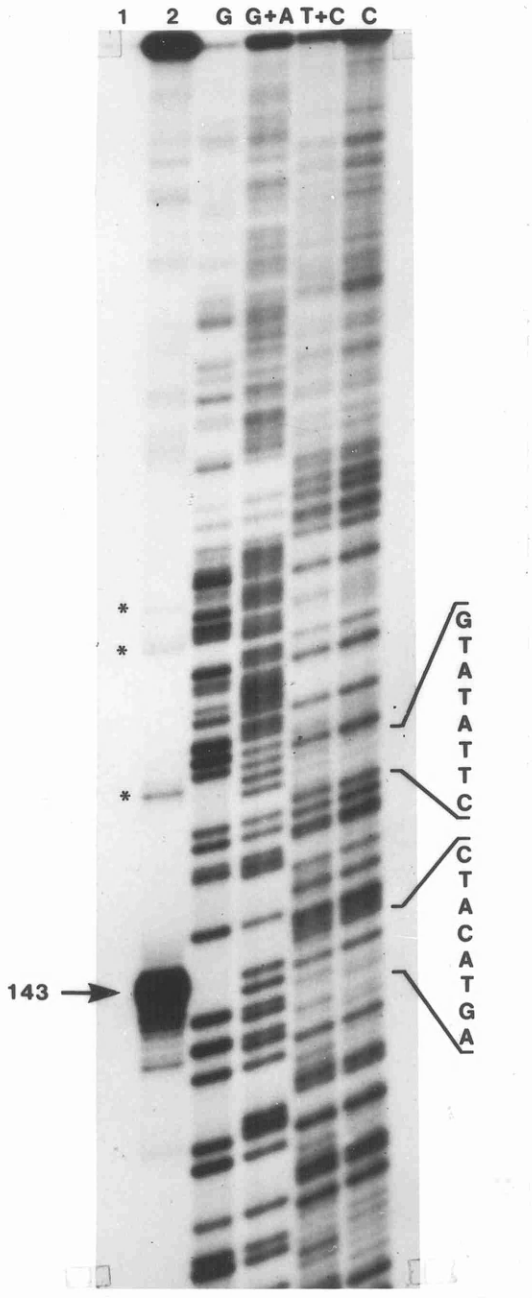

1
2
G
G+A
T+C
C
*
*
*
143
GTATATTC
CTACATGA

Figure 30. Precise location of the 5' terminus of the HSV-2 1.2kb mRNA. The DNA probe was the Bam HI/Bgl II fragment, uniquely 5'-labelled at the Bam HI site, which is described in Fig. 16. RNA samples were early infected cell cytoplasmic RNA in track 2 and mock-infected cytoplasmic RNA in track 1. Following hybridisation, samples were digested with nuclease S1 and the digestion products were electrophoresed on an 8% denaturing polyacrylamide gel alongside the sequence reaction products of the DNA probe. The position of the 5' terminus is arrowed and asterisks locate minor nuclease S1 products which are described in the text.

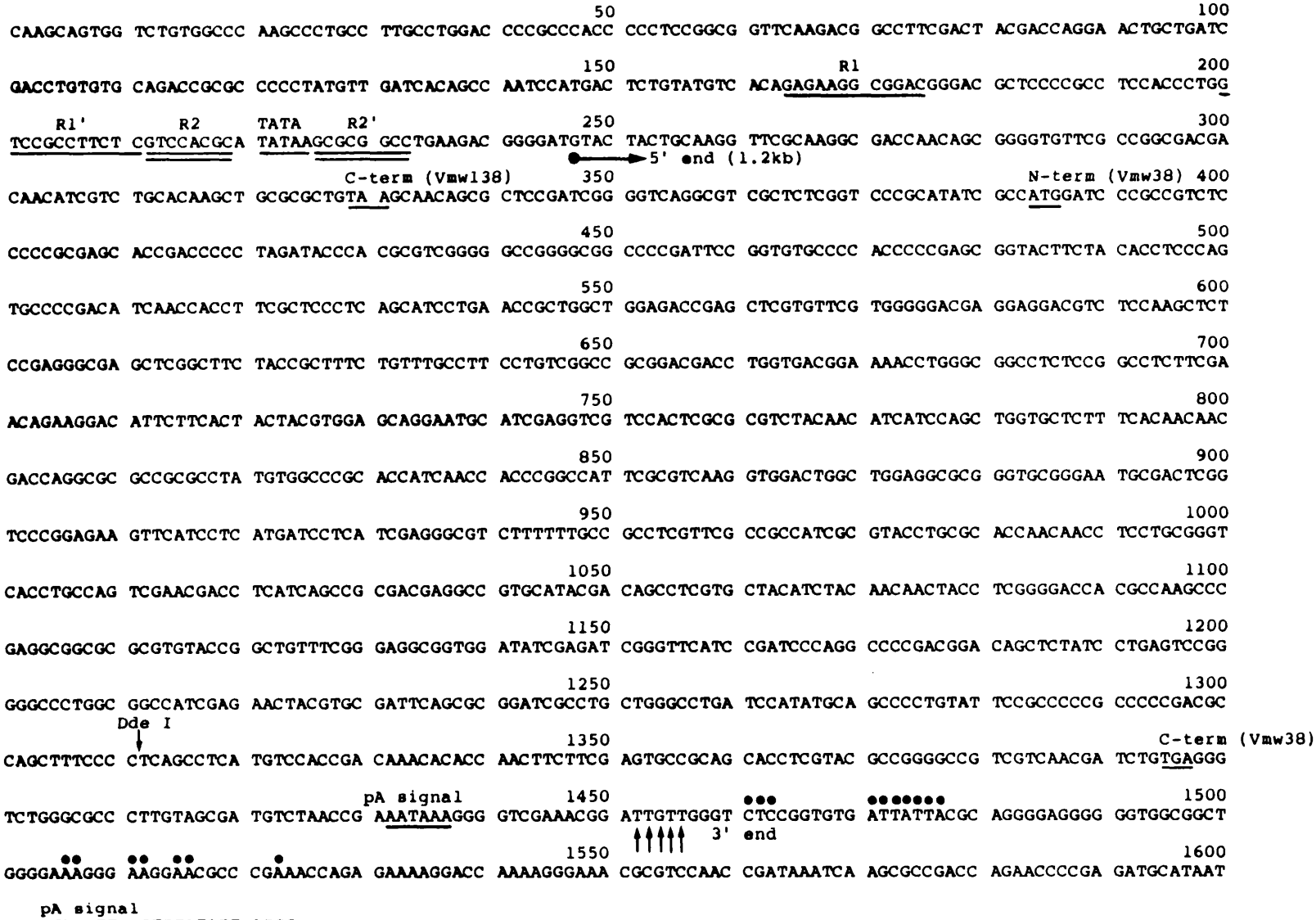

50
100
CAAGCAGTGG TCTGTGGCCC AAGCCCTGCC TTGCCTGGAC CCCGCCCACC CCCTCCGGCG GTTCAAGACG GCCTTCGACT ACGACCAGGA ACTGCTGATC
150
R1
200
GACCTGTGTG CAGACCGCGC CCCCTATGTT GATCACAGCC AATCCATGAC TCTGTATGTC ACAGAGAAGG CGGACGGGAC GCTCCCCGCC TCCACCCTGG
R1'
R2
TATA
R2'
250
300
TCCGCCTTCT CGTCCACGCA TATAAGCGCG GCCTGAAGAC GGGGATGTAC TACTGCAAGG TTCGCAAGGC GACCAACAGC GGGGTGTTCG CCGGCGACGA
5' end (1.2kb)
C-term (Vmw138)
350
N-term (Vmw38)
400
CAACATCGTC TGCACAAGCT GCGCGCTGTA AGCAACAGCG CTCCGATCGG GGTCAGGCGT CGCTCTCGGT CCCGCATATC GCCATGGATC CCGCCGTCTC
450
500
CCCCGCGAGC ACCGACCCCC TAGATACCCA CGCGTCGGGG GCCGGGGCGG CCCCGATTCC GGTGTGCCCC ACCCCCGAGC GGTACTTCTA CACCTCCCAG
550
600
TGCCCCGACA TCAACCACCT TCGCTCCCTC AGCATCCTGA ACCGCTGGCT GGAGACCGAG CTCGTGTTCG TGGGGGACGA GGAGGACGTC TCCAAGCTCT
650
700
CCGAGGGCGA GCTCGGCTTC TACCGCTTTC TGTTTGCCTT CCTGTCGGCC GCGGACGACC TGGTGACGGA AAACCTGGGC GGCCTCTCCG GCCTCTTCGA
750
800
ACAGAAGGAC ATTCTTCACT ACTACGTGGA GCAGGAATGC ATCGAGGTCG TCCACTCGCG CGTCTACAAC ATCATCCAGC TGGTGCTCTT TCACAACAAC
850
900
GACCAGGCGC GCCGCGCCTA TGTGGCCCGC ACCATCAACC ACCCGGCCAT TCGCGTCAAG GTGGACTGGC TGGAGGCGCG GGTGCGGGAA TGCGACTCGG
950
1000
TCCCGGAGAA GTTCATCCTC ATGATCCTCA TCGAGGGCGT CTTTTTTGCC GCCTCGTTCG CCGCCATCGC GTACCTGCGC ACCAACAACC TCCTGCGGGT
1050
1100
CACCTGCCAG TCGAACGACC TCATCAGCCG CGACGAGGCC GTGCATACGA CAGCCTCGTG CTACATCTAC AACAACTACC TCGGGGACCA CGCCAAGCCC
1150
1200
GAGGCGGCGC GCGTGTACCG GCTGTTTCGG GAGGCGGTGG ATATCGAGAT CGGGTTCATC CGATCCCAGG CCCCGACGGA CAGCTCTATC CTGAGTCCGG
1250
1300
GGGCCCTGGC GGCCATCGAG AACTACGTGC GATTCAGCGC GGATCGCCTG CTGGGCCTGA TCCATATGCA GCCCCTGTAT TCCGCCCCCG CCCCCGACGC
Dde I
1350
C-term (Vmw38)
CAGCTTTCCC CTCAGCCTCA TGTCCACCGA CAAACACACC AACTTCTTCG AGTGCCGCAG CACCTCGTAC GCCGGGGCCG TCGTCAACGA TCTGTGAGGG
pA signal
1450
1500
TCTGGGCGCC CTTGTAGCGA TGTCTAACCG AAATAAAGGG GTCGAAACGG ATTGTTGGGT CTCCGGTGTG ATTATTACGC AGGGGAGGGG GGTGGCGGCT
3' end
1550
1600
GGGGAAAGGG AAGGAACGCC CGAAACCAGA GAAAAGGACC AAAAGGGAAA CGCGTCCAAC CGATAAATCA AGCGCCGACC AGAACCCCGA GATGCATAAT
pA signal
AACGTTTATT ACTCTATATT ACGG

Figure 31. Nucleotide sequence of region (B) in Fig. 23. Positions of the 5' end of the 1.2kb mRNA and the 3' co-terminus are shown and the TATA box homologue and AATAAA signal are indicated; the AATAAA signal for a leftwards-transcribed mRNA also is underlined. Locations of additional 3' termini, as detected in Fig. 34, are indicated by closed circles. The codons which represent the N- and C-termini of Vmw38 and the C-terminus of Vmw138 are underlined as are the inverted repeats (R1, R1', R2 and R2') which may form the stem and loop structures shown in Fig. 33. The position of the Dde I site used to locate the mRNA 3' co-terminus in Fig. 34 is arrowed.

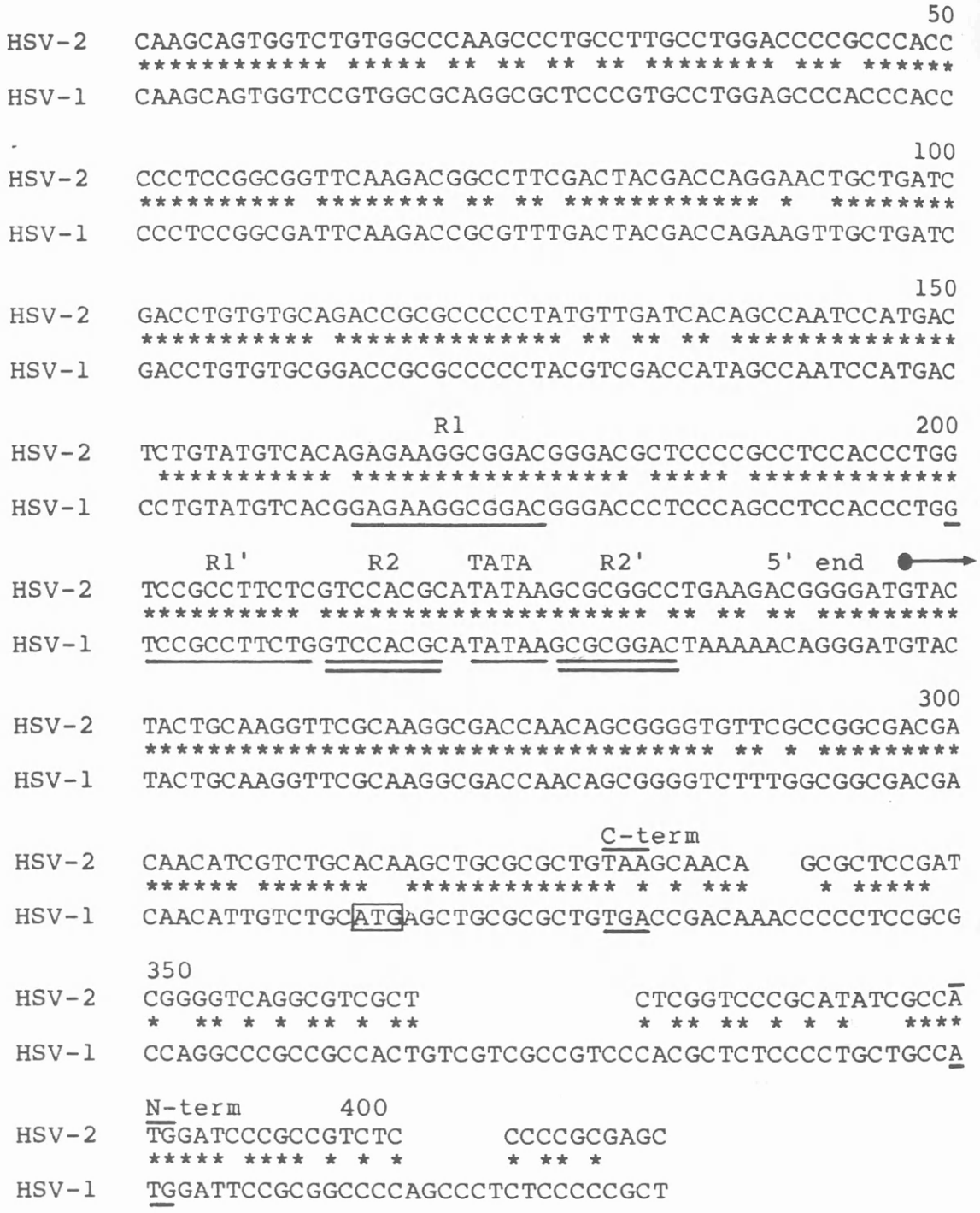
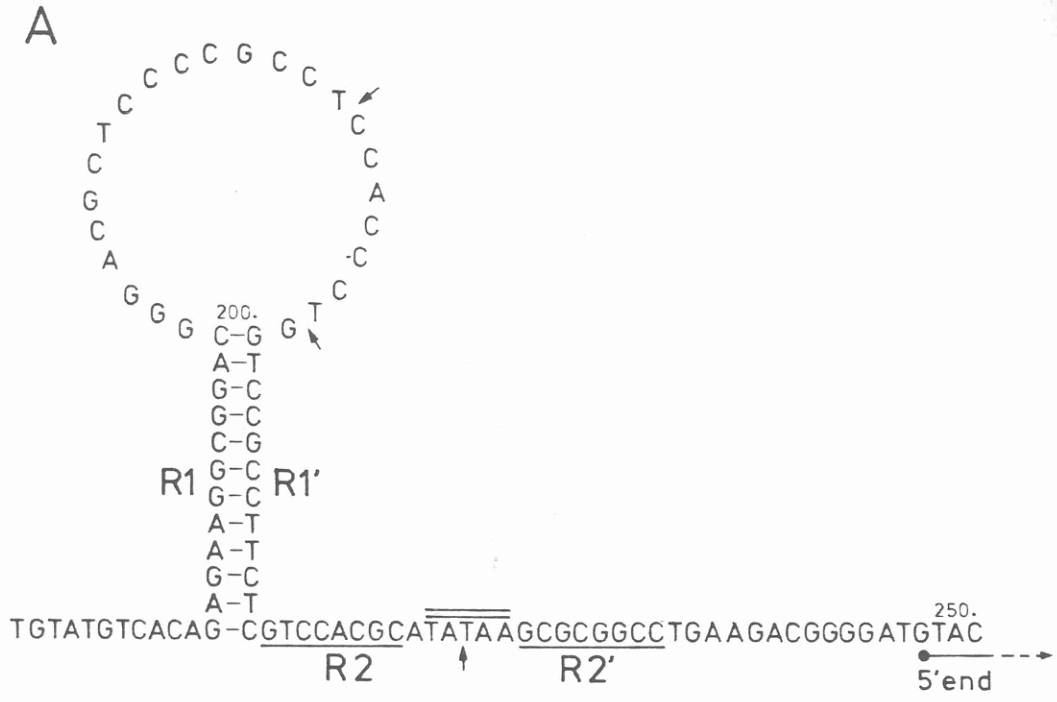

50
HSV-2 CAAGCAGTGGTCTGTGGCCCAAGCCCTGCCTTGCCTGGACCCCGCCCACC
HSV-1 CAAGCAGTGGTCCGTGGCGCAGGCGCTCCCGTGCCTGGAGCCCACCCACC
100
HSV-2 CCCTCCGGCGGTTCAAGACGGCCTTCGACTACGACCAGGAACTGCTGATC
HSV-1 CCCTCCGGCGATTCAAGACCGCGTTTGACTACGACCAGAAGTTGCTGATC
150
HSV-2 GACCTGTGTGCAGACCGCGCCCCCTATGTTGATCACAGCCAATCCATGAC
HSV-1 GACCTGTGTGCGGACCGCGCCCCCTACGTCGACCATAGCCAATCCATGAC
R1
200
HSV-2 TCTGTATGTCACAGAGAAGGCGGACGGGACGCTCCCCGCCTCCACCCTGG
HSV-1 CCTGTATGTCACGGAGAAGGCGGACGGGACCCTCCCAGCCTCCACCCTGG
R1' R2 TATA R2' 5' end
HSV-2 TCCGCCTTCTCGTCCACGCATATAAGCGCGGCCTGAAGACGGGGATGTAC
HSV-1 TCCGCCTTCTGGTCCACGCATATAAGCGCGGACTAAAAACAGGGATGTAC
300
HSV-2 TACTGCAAGGTTCGCAAGGCGACCAACAGCGGGGTGTTCGCCGGCGACGA
HSV-1 TACTGCAAGGTTCGCAAGGCGACCAACAGCGGGGTCTTTGGCGGCGACGA
C-term
HSV-2 CAACATCGTCTGCACAAGCTGCGCGCTGTAAGCAACA GCGCTCCGAT
HSV-1 CAACATTGTCTGCATGAGCTGCGCGCTGTGACCGACAAACCCCCTCCGCG
350
HSV-2 CGGGGTCAGGCGTCGCT CTCGGTCCCGCATATCGCCA
HSV-1 CCAGGCCCGCCGCCACTGTCGTCGCCGTCCCACGCTCTCCCCTGCTGCCA
N-term 400
HSV-2 TGGATCCCGCCGTCTC CCCCGCGAGC
HSV-1 TGGATTCCGCGGCCCCAGCCCTCTCCCCCGCT

Figure 32. A comparison between nucleotide sequences at the 5' termini of the HSV-2 and HSV-1 1.2kb mRNAs. The HSV-2 sequence is numbered as in Fig. 31 and identical bases are indicated by asterisks. The 5' terminus of the HSV-2 1.2kb mRNA is shown; the HSV-1 1.2kb mRNA 5' end is at an identical position in the HSV-1 sequence. TATA box homologues are underlined as are the inverted repeats (R1, R1', R2 and R2'). Codons which represent the N-termini of both Vmw38 proteins and the C-termini of Vmw138 and Vmw136 are marked above the HSV-2 sequence and below the HSV-1 sequence. The first ATG codon downstream from the 5' end of the HSV-1 1.2kb transcript, which is not conserved in HSV-2 DNA, is boxed.


A
200.
R1
R1'
TGTATGTCACAG-CGTCCACGCATATAAGCGCGGCCTGAAGACGGGGATGTAC
250.
R2
R2'
5'end

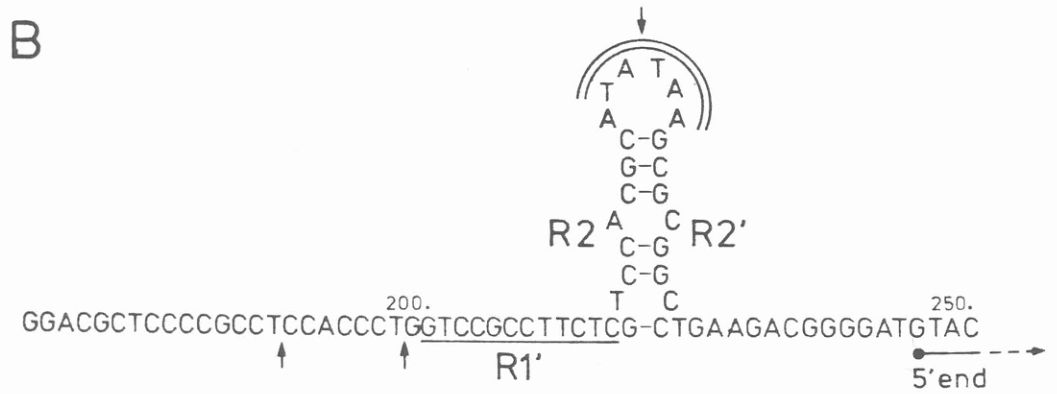

B
R2
R2'
200.
250.
GGACGCTCCCCGCCTCCACCCTGGTCCGCCTTCTCG-CTGAAGACGGGGATGTAC
R1'
5'end

Figure 33. Proposed stem and loop structures formed by inverted repeat sequences upstream from the 5' end of the HSV-2 1.2kb mRNA. The structure shown in (A) would be formed by the R1 and R1' inverted repeats and that in (B) by the R2 and R2' inverted repeats indicated in Fig. 31. Arrows locate minor nuclease S1 products indicated in Fig. 30. The TATA box homologue is overlined and the position of the mRNA 5' terminus is indicated. Sequences are numbered as in Fig. 31.

31 and 32) with the potential to form two alternative stem and loop structures in HSV-2 and HSV-1 DNA (Fig. 33). Inverted repeats R1 and R1' could form a structure comprising a stem of 12bp with a loop of 24b whereas the structure formed by R2 and R2' would consist of a stem of 8bp containing two mismatches and a loop of 6b. Under the DNA/RNA hybridisation conditions used in transcript mapping experiments, a minor proportion of molecules were generated in which the stem and loop structures were formed by intra-molecular reassociation of the DNA probe. In these molecules, sequences within the stems and loops were unavailable for hybridisation to the 4.5kb mRNA, and the intra-molecular loops were sensitive to nuclease S1. The remainder of the DNA probe was available for DNA/RNA hybridisation. Minor bands indicated in Fig. 30, track 2 correspond to nuclease S1 sensitive sites located within both loops; no nuclease S1 cleavage sites were detected at the stem regions. This indicates that the stem and loop structures can be generated by intra-molecular reassociation but does not provide evidence for their natural occurrence in virus DNA.

The 5' portions of the HSV-2 and HSV-1 1.2kb mRNAs lie within the C-terminal regions of Vmw138 and Vmw136 respectively. In both mRNAs, polypeptide stop codons are located 82nuc downstream from the 5' terminus (Fig. 32); the DNA sequences for the 1.2kb mRNAs are highly conserved (>90% homology) and colinear in this region. However, beyond the polypeptide stop codons, homology decreases to a

level of approximately 55% and the DNA sequences are no longer colinear. The distance between the Vmw138/Vmw136 protein stop codons and the initiation codons for the Vmw38 polypeptides is 16nuc shorter in HSV-2.

e) 3' co-terminus of the 6.4kb, 4.5kb and 1.2kb mRNAs.

The 3' co-terminus for the 6.4kb, 4.5kb and 1.2kb mRNAs was positioned within a Dde I/Xho I fragment (Fig. 34, track 2) which had been uniquely 3'-labelled at the Dde I site. The major nuclease S1 product located the 3' co-terminus within the sequence TTGTT (Fig. 31, positions 1452 to 1456), 20nuc downstream from the hexanucleotide sequence AATAAA. Minor nuclease S1 products were detected reproducibly which placed other 3' termini at preferred positions up to 70nuc downstream from the major 3' location (Fig. 31); the significance of these additional sites is unknown.

Draper et al. (1982) reported that the major 3' co-terminus in this region of HSV-1 strain KOS preceded the AATAAA signal. Data for both HSV-2 strain HG52 and HSV-1 strain 17^+ unambiguously locate the 3' co-termini downstream from the poly A signal which is consistent with the position of 3' ends of other eukaryotic mRNAs. A second AATAAA sequence which could serve as a poly A signal for a leftwards-transcribed mRNA, is present in the strand complementary to that specifying the 3' co-terminal transcripts (Fig. 31, positions 1605 to 1610); a poly A signal for a late leftwards-transcribed mRNA is present at

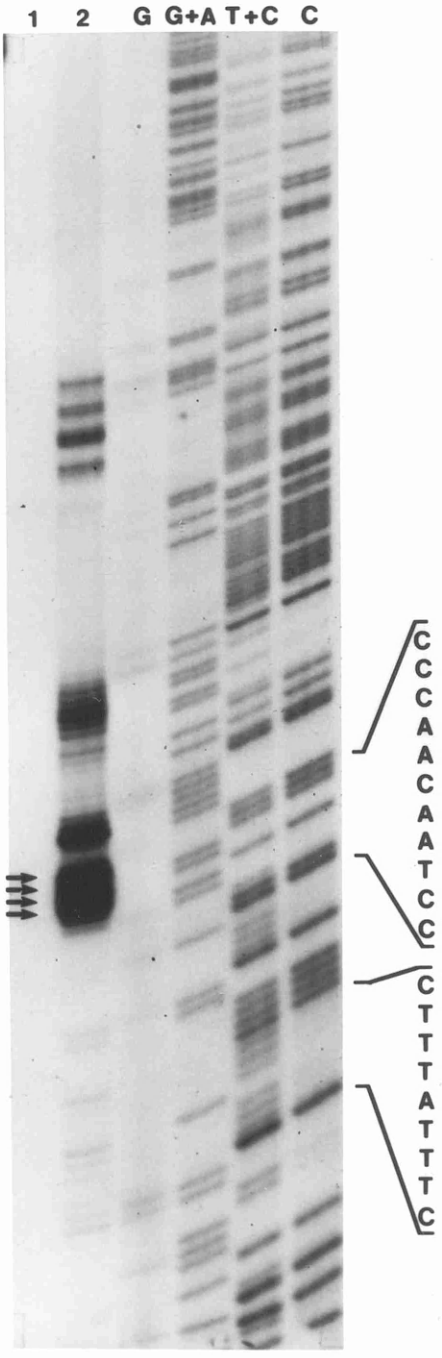

1
2
G
G+A
T+C
C
CCCAACAATCCC
CTTTATTTC

Figure 34. Location of the mRNA 3' co-terminus. From Bam HI t, a Dde I/Xho I fragment was uniquely 3'-labelled at the Dde I site and hybridised to early infected cell cytoplasmic RNA (track 2) and mock-infected cytoplasmic RNA (track 1). The Dde I site is located at position 1311 in Fig. 31 and the location of the Xho I site in Bam HI t is shown in Fig. 15. After hybridisation at 56°C, samples were treated with nuclease S1 and the digestion products were electrophoresed on an 8% denaturing polyacrylamide gel alongside the sequence reaction products of the DNA probe. Arrows indicate the major locations of the 3' co-terminus; the locations of additional nuclease S1 products are shown on the nucleotide sequence in Fig. 31.

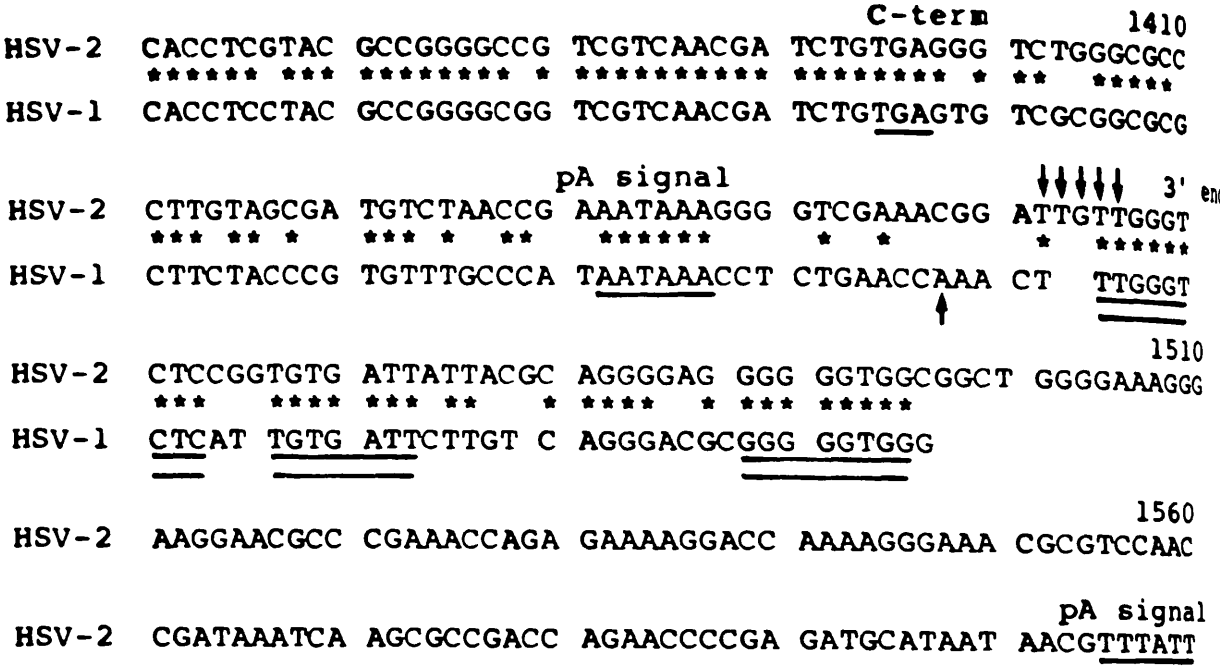


Figure 35. A comparison between the sequences at the HSV-2 3' co-terminus and a portion of those at the HSV-1 3' co-terminus. Asterisks denote conserved bases and the HSV-2 sequence is numbered as in Fig. 31. Locations of the 3' co-termini are arrowed and the AATAAA signals for the 3' co-terminal mRNAs are underlined as is the AATAAA signal for a leftwards-transcribed mRNA. Translational stop codons for HSV-1 and HSV-2 Vmw38 proteins are also underlined. Regions of homology downstream from the 3' co-termini are indicated by two solid lines.

the equivalent HSV-1 locus (McLauchlan and Clements, 1982). The distance between the poly A signals on opposite strands is 168nuc for HSV-2 and 125nuc for HSV-1.

HSV-2/HSV-1 nucleotide sequence comparisons reveal that upstream from the 3' co-terminus, in the 1.2kb mRNA untranslated region, the HSV-2 sequence has greater than 60% homology with the corresponding HSV-1 segment as far as the AATAAA signal and both sequences are colinear (Fig. 35). Beyond the AATAAA signal, homology decreases markedly with the exception of three blocks of highly conserved sequences, downstream from the mRNA poly A sites. These conserved sequences, as shown in Fig. 35, are; TTGGGTCTC (positions 1455 to 1463); TGTGATT (positions 1467 to 1473); GGGGGTGG (positions 1488 to 1495). The highly conserved sequence between positions 1467 to 1473 is related to a consensus sequence YGTGTTYY (Y = pyrimidine) which is located approximately 24 to 35nuc downstream from the poly A signals of many viral and eukaryotic mRNAs. The role of this sequence motif in mRNA 3' end formation was analysed further and the results are presented in Section E (Page 174). DNA sequences downstream from the poly A signal of the leftwards-transcribed mRNA share low homology with the corresponding HSV-1 region and do not contain a YGTGTTYY sequence motif.

DISCUSSION.

6. Organisation of Overlapping HSV mRNAs.

Transcripts which span the junction between HSV-2 transforming fragments, Bgl II c and Bgl II n, form a family of four unspliced, overlapping mRNAs which have an identical arrangement to transcripts in the corresponding region of HSV-1 DNA. Three mRNAs, with sizes of 6.4kb, 4.5kb and 1.2kb, share a common 3' terminus and the 6.4kb and 1.7kb transcripts are 5' co-terminal. As a result of this mRNA arrangement, DNA sequences required for polyadenylation of the 1.7kb transcript and for initiation of synthesis of the 4.5kb and 1.2kb mRNAs are located within the transcribed regions of overlapping messages (Fig. 36). The size differences between the larger HSV-2 and HSV-1 transcripts (6.4kb and 7.0kb; 4.5kb and 5.0kb) appear not to result from extensive insertions in the HSV-1 DNA sequence but may reflect inaccuracy of the methods used to determine the lengths of transcripts.

The 5' termini of three late mRNAs transcribed in the opposite orientation from the 6.4kb species have been mapped and the 5' terminal region of one of these mRNAs overlaps with the 6.4kb trancript. Overlapping regions between 5' terminal portions of HSV-1 U_S genes-2 and -3 (Fig. 3), which are divergently transcribed, have been identified (Rixon and McGeoch, 1985) and a family of transcripts are specified by the DNA strand complementary to that encoding an HSV-1 spliced mRNA (Fig. 4; Costa et

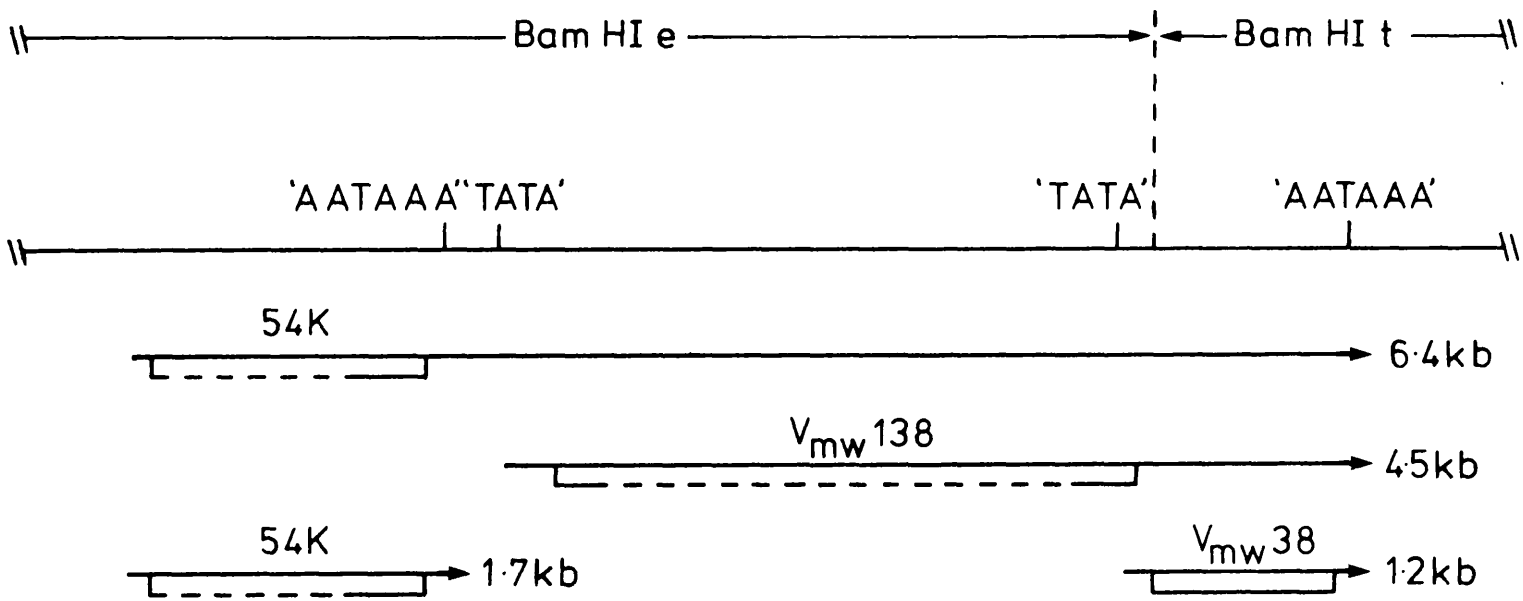

Bam HI e
Bam HI t
'AATAAA'
'TATA'
'TATA'
'AATAAA'
54K
6·4kb
V_{mw}138
4·5kb
54K
1·7kb
V_{mw}38
1·2kb

Figure 36. A summary of the locations of mRNA termini with respect to transcription control signals. Boxed areas indicate proposed protein coding regions within the transcripts and the dotted lines within these regions represent areas for which nucleotide sequence data are not available.

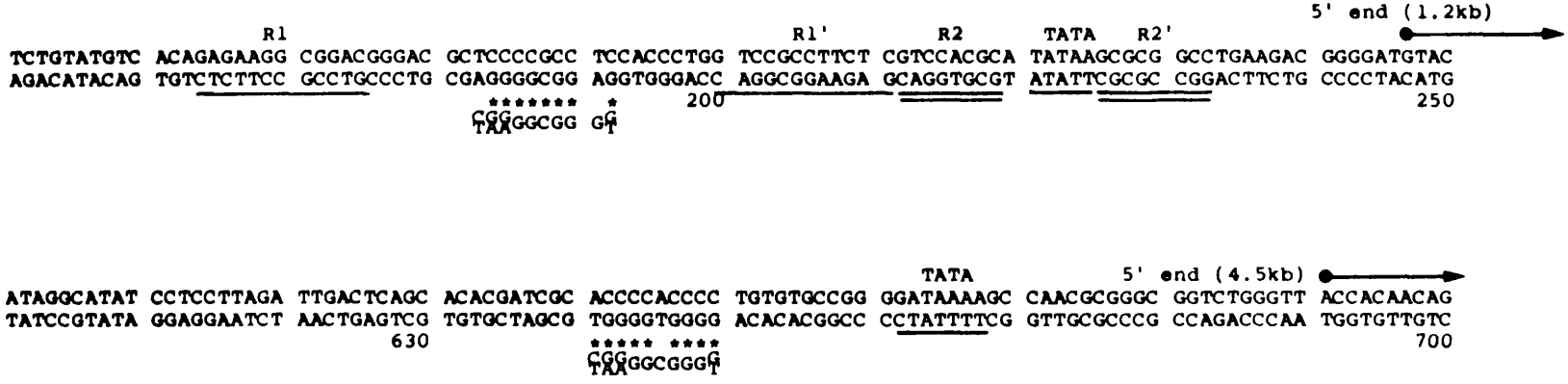

5' end (1.2kb)
R1
R1'
R2
TATA
R2'
TCTGTATGTC ACAGAGAAGG CGGACGGGAC GCTCCCCGCC TCCACCCTGG TCCGCCTTCT CGTCCACGCA TATAAGCGCG GCCTGAAGAC GGGGATGTAC
AGACATACAG TGTCTCTTCC GCCTGCCCTG CGAGGGGCGG AGGTGGGACC AGGCGGAAGA GCAGGTGCGT ATATTCGCGC CGGACTTCTG CCCCTACATG
200
250
CGG
TAAGGCGG GG
T
TATA
5' end (4.5kb)
ATAGGCATAT CCTCCTTAGA TTGACTCAGC ACACGATCGC ACCCCACCCC TGTGTGCCGG GGATAAAAGC CAACGCGGGC GGTCTGGGTT ACCACAACAG
TATCCGTATA GGAGGAATCT AACTGAGTCG TGTGCTAGCG TGGGGTGGGG ACACACGGCC CCTATTTTCG GTTGCGCCCG CCAGACCCAA TGGTGTTGTC
630
700
CGG
TAAGGCGGGG
T

Figure 37. Locations of putative Sp1 binding sites upstream from the 5' termini of the HSV-2 4.5kb and 1.2kb mRNAs. The nucleotide sequences of both DNA strands is shown and the positions of mRNA 5' termini are indicated. TATA box homologues and the inverted repeats, R1, R1', R2 and R2', described in Fig. 31, are underlined. Residues on the mRNA coding strand which have homology to the consensus sequence for Sp1 binding are indicated by asterisks. The Sp1 binding consensus is written in the 3' to 5' direction.

al., 1985). Therefore, in HSV, the DNA strand complementary to that specifying a transcript may contain control signals required for the expression of mRNA synthesised in the opposite orientation. Similar arrangements of genes occurs on the Ad genome where a number of transcripts are produced from opposite DNA strands (reviewed in Shenk and Williams, 1984).

7. Signals Involved in Initiation of mRNA Synthesis.

The DNA sequences upstream from the 5' termini of the 4.5kb and 1.2kb mRNAs contain conserved TATA box homologues and C-rich motifs, elements which are present in the promoter regions of many eukaryotic mRNAs. Conserved C-rich sequences positioned approximately 40 to 70nuc upstream from the 5' termini of the HSV-2 and HSV-1 mRNAs resemble C- and G-rich tracts of sequence which have been identified as components required for promoter activity in a number of genes (see Introduction, Page 62). The HSV-1 gD promoter contains two tracts of G-rich sequences which are required for mRNA synthesis and function in either orientation (Everett, 1983). In SV40, C-rich motifs influence the level of mRNA synthesis in either orientation (Everett et al., 1983) and these sequences are binding sites for the transcriptional factor Sp1 (Dynan and Tjian, 1983). From comparisons of strong binding sites, a consensus sequence, $\overset{G}{T}GGGCGG\overset{GGC}{AAT}$ has been derived for Sp1 sites (Dynan and Tjian, 1985). Inspection of the complement of C-rich sequences conserved upstream from the 4.5kb and

5.0kb mRNAs (Fig. 37) reveals reasonable homology with this consensus sequence, thus, these C-rich elements may represent Sp1 binding sites. Recent data indicates that Sp1 binds to G+C-rich elements in the promoter regions of HSV-1 IE (Jones and Tjian, 1985) and TK genes (Jones et al., 1985).

Tracts of sequence with homology to the second C-rich motif (ACACCCCC) at 110nuc from the 4.5kb mRNA 5' terminus are located at a similar distance from the cap sites of a number of early HSV mRNAs (Fig. 38; Wagner, 1985). These A+C-rich elements appear to be absent from equivalent positions in most late promoters, although, not all early promoters contain these sequences. However, the highly conserved nature of the ACACCCCC segment would indicate that this sequence may be a functional component of the promoters for the 4.5kb and 5.0kb mRNAs.

In conclusion, the promoter regions for the 4.5kb and 1.2kb mRNAs, which are located within the transcribed regions of larger, overlapping transcripts, appear to contain similar control sequences to those identified in other HSV and in eukaryotic genes.

8. Stem and Loop Structures may Influence Transcription of the 1.2kb mRNA.

The 5' end of the 1.2kb mRNA is located within the C-terminal portion of Vmw138, the protein encoded by the 4.5kb transcript. Thus, transcription control elements involved in specifying the 1.2kb mRNA 5' terminus form part


HSV Early mRNA Promoter Sequences
-100
-50
HSV-2 CATGGACATGTTACACCCCCCTGACTCAGGAGATAGGCATATCCTCCTTAGATTGACTCAGCACACGATCGCACCCCACCCCTGTGTGCCGGGGATAAAAGCCAACGCGGGCGGTCTGGG 1
4.5kb mRNA
-100
-50
HSV-2 GTTGATCACAGCCAATCCATGACTCTGTATGTCACAGAGAAGGCGGACGGGACGCTCCCCGCCTCCACCCTGGTCCGCCTTCTCGTCCACGCATATAAGCGCGGCCTGAAGACGGGGATG 1
1.2kb mRNA
-100
-50
HSV-1 CTATGATGACACACAAACCCCGCCCAGCGTCTTGTCATTGGCGAATTCGAACACGCAGATGCAGTCGGGGCGGCGCGGTCCCAGGTCCACTTCGCATATTAAGGTGACGCGTGTGGCCTCGA 2
TK mRNA
-100
-50
HSV-2 CCAGGATGACGCACACACCTCCCAACGTTTTGTCATTGGCGAATTCGAACACGCAGATGCAGTCTGGGCGGCGCGGCCCGAGGTCCACTTCGCATATTAAGGTGACGCGCGTGGCCTCGA 3
TK mRNA
-100
-50
HSV-1 GCCGTGTGACACTATCGTCCATACCGACCACACCGACGAATCCCCCAAGGGGGAGGGGCCATTTTACGAGGAGGAGGGGTATAACAAAGTCTGTCTTTAAAAAGCAGGGGTTAGGGAGTT 4
gD mRNA
-100
-50
HSV-1 GGACAAAGTGCGAACGCTTCGCGTTCTCACTTTTTTTATAATAGCGGCCACGCCCACCGGCTACGTCACTCTCCTGTCGGCCGCCGGCGGTCCATAAGCCCGGCCGGCCGGGCCGACGCG 5
DNA Pol. mRNA
-100
-50
HSV-1 CTTTTTTTATAATAGCGGCCACGCCCACCGGCTGATGACGCGCGGGGCGTGGGAGGGGCTGGGGCGGACCCGGCACGCCCCCAGGTAAAGTGTACATATACCAACCGCATACCAGACGCA 5
DBP mRNA
-100
-50
HSV-1 CCAACACCCACGGCCTGGCGTATGACGTCCCAGAGGGCATCCGGCGCCACCTCCGCAATCCCAAGATTCGGCGCGCGTTTACGGATCGGTGTATAAATTACCAGCACACACACAAGGCGA 6
Exo mRNA
HSV Late mRNA Promoter Sequences
-100
-50
HSV-1 TTGATATATTTTTCAATAAAAGGCATTAGTCCCGAAGACCGCCGGTGTGTGATGATTTCGCCATAACACCCAAACCCCGGATGGGGCCCGGGTATAAATTCCGGAAGGGGACACGGGCTA 7
gC mRNA
-100
-50
HSV-2 TACCACCGGCGTGTGATGATTTCGCCCTACCGCTCCGATCCCCGGGGGGAGGGGGGAAGGAAATGGGGGCGGGGGTGCCGTGGACGGGTATAAAGGCCAGGGGGGCAGGCGGGCCCATCA 8
gC mRNA
-100
-50
HSV-1 GCTTTGCCGCCTCTGCCAATTTCTTCCTGCACGCTTTTGGACCAGGGCCATCTTGAATGCACCCGTCGCGTTCTAACGGGGGTGGGGCGGGGGGGGTATATAAGGCCTGGGATCCCACGT 9
MCP mRNA
HSV IE mRNA Promoter Sequences
-100
-50
HSV-1 TTGGTTCTTGGGGACGGCGGGGGCGGGGGCGGTGACGCCCGACGGGGAGGGACAAGGAGGAGTTTCGGAAAGCCGGCCCCGGTCGTGCGGGTATAAGGGCAGCCACCGGCCCACTGGGCG 10
IE mRNA-2
-100
-50
HSV-1 CGTGCATAATGGAATTCCGTTCGGGGTGGGCCCGCCGGGGGGGCGGGGGGCCGGCGGCCTCCGCTGCTCCTCCTTCCCGCCGGCCCCTGGGACTATATGAGCCCGAGGACGCCCCGATCG 11
IE mRNA-3
-100
-50
HSV-1 ATTAGGGCGAAGTGCGAGCACTGGCGCCGTGCCCGACTCCGCGCCGGCCCCGGGGGCGGGCCCGGGCGGCGGGGGGCGGGTCTCTCCGGCGCACATAAAGGCCCGGCGCGACCGACGCCG 11
IE mRNAs-4/-5

Figure 38. Sequence comparison of regions upstream from the 5' termini of HSV mRNAs from different temporal classes. Locations of mRNA 5' termini are arrowed and TATA box homologues are underlined. Proposed A+C-rich sequences are underlined twice and a TAATGARAT motif in the promoter region of IE mRNA-3 is indicated by a dashed line. The key to abbreviations is given in the legend to Fig. 4. Nucleotide sequence data were taken from the following references:

1. this study; 2. McKnight, 1980; Wagner et al., 1981;
3. Swain and Galloway, 1983; 4. Watson et al., 1982;
5. Quinn, 1984; 6. Costa et al., 1983;
7. Frink et al., 1983; 8. Dowbenko and Lasky, 1984, Swain et al., 1985;
9. Costa et al., 1984; 10. Whitton et al., 1983;
11. Murchie and McGeoch, 1982.

of the Vmw138 protein coding region. The 5' terminus of the equivalent HSV-1 1.2kb mRNA is similarly located within the Vmw136 protein coding sequences. An additional feature of DNA sequences upstream from the 1.2kb mRNA is the presence of inverted repeats which may form stem and loop structures (Fig. 33). The nucleotide sequences which form these structures are present also in the 4.5kb RNA; from calculations on the energy of double-strand formation in RNA, the larger stem and loop (ΔG = -10.6kcal/mol.) has a greater stability than the smaller structure (ΔG = -0.9kcal/mol.). These structures may have no biological significance, however, they could act as promoter elements in the synthesis of the 1.2kb mRNA.

Transcription of the overlapping 4.5kb mRNA through this locus may generate a single-stranded DNA region which could facilitate the intra-strand formation of the stem and loop structures and thereby potentiate initiation of the 1.2kb mRNA. As a consequence of this proposal, synthesis of the 1.2kb mRNA would be dependent upon transcription of the 4.5kb mRNA and the structures would function essentially as transcription-activator elements. However, a plasmid construction consisting of the 1.2kb mRNA promoter fused to the bacterial CAT gene efficiently produces CAT enzyme activity in the absence of additional promoter sequences (O'Hare and Hayward, 1984). Thus, the 1.2kb mRNA promoter appears not to require an upstream promoter to function and probably contains all the transcription control signals necessary to initiate mRNA synthesis.

Alternatively, the stem and loop structures may be stabilised by the presence of DNA binding proteins and act as a pause site for RNA polymerase molecules which are actively transcribing the 4.5kb mRNA. Studies with avian leukosis virus (ALV), a retrovirus which integrates into the cellular genome, have shown that the promoter in a downstream long terminal repeat (LTR) inefficiently expresses flanking cellular sequences in the presence of a transcriptionally active promoter located in an upstream LTR; removal of this LTR increases promoter activity from a downstream LTR (Cullen et al., 1984). In HSV, a structure which prevents transcription through the promoter region of the 1.2kb mRNA may facilitate initiation of this mRNA.

9. Mechanisms for Generating Transcripts with Alternative 3' Termini.

The 5' co-terminal 6.4kb and 1.7kb mRNAs both encode a 54,000 mol. wt. protein but have different 3' termini. The mechanism for 3' end selection is unclear, however, there are a number of possible mechanisms for generating mRNAs with alternative 3' termini. Firstly, the 1.7kb mRNA 3' terminus may be an inefficient site for either termination or cleavage/polyadenylation of primary transcripts. DNA sequence analysis in this region indicates that the poly A site of the 1.7kb mRNA is flanked by a polyadenylation signal and a YGTGTTYY motif; these sequences also occur at the majority of mammalian and viral poly A sites. Moreover, plasmids carrying the 3' terminal

sequences of the 1.7kb mRNA produce CAT activities comparable with those obtained from plasmids containing poly A site sequences which are not within an overlapping transcript; this suggests that there is no inherent inefficiency in the processing signals at the 1.7kb mRNA 3' terminus as compared to those at other HSV poly A sites. It is also unlikely that initiation at the 4.5kb mRNA 5' terminus hinders polyadenylation of the 1.7kb mRNA; the distance between the 1.7kb mRNA 3' terminus and the 4.5kb mRNA 5' terminus is 178nuc which is greater than the spacing between termini of tandemly-arranged, non-overlapping transcripts in the HSV-1 U_S region (McGeoch et al., 1985).

A second possibility is that a virus-associated effect on the cellular transcription machinery results in read-through of the 1.7kb mRNA 3' terminus by RNA polymerase II. HSV infection inhibits host protein synthesis rapidly and suppresses cellular RNA transcription (see Introduction, Page 21). Potentially, effects on the cellular polyadenylation process may result in inefficient polyadenylation of viral as well as host transcripts. Extended transcripts of up to 10kb which contain internal poly A sites have been detected at a number of locations on the HSV genome (Wagner, 1985), thus transcription through functional poly A sites may be a general feature of HSV mRNA expression. Moreover, virus infection may induce a reduction in the levels of cellular components necessary for mRNA 3' end formation. Therefore, the mechanism for

synthesis of 5' co-terminal HSV mRNAs with different 3' termini may be similar to that for the selection of poly A sites for IgD and IgG mRNAs which is thought to be controlled by the level of an 'endase' enzyme (see Introduction, Page 75; Blattner and Tucker, 1984). A reduction in the concentration of cellular factors required for mRNA 3' end formation as a result of virus infection may decrease the number of cleavage events at potential poly A sites in a primary transcript, thereby increasing the probability of generating a mRNA which is polyadenylated at a distal 3' site.

A alanine

R arginine

N asparagine

D aspartic acid

C cysteine

E glutamic acid

Q glutamine

G glycine

H histidine

I isoleucine

L leucine

K lysine

M methionine

F phenylalanine

P proline

S serine

T threonine

W tryptophan

Y tyrosine

V valine

Figure 39. Key to amino acids represented by the single letter code.

SECTION B.

10. Analysis of the Polypeptide Coding Regions within the Overlapping mRNAs.

In this section, predicted amino acids specified by nucleotide sequences are presented in the single letter code; the key to this code is given in Fig. 39.

a) C-terminal region of the polypeptide encoded by the 6.4kb and 1.7kb mRNAs.

The HSV-1 7.0kb and 1.9kb 5' co-terminal mRNAs both specify a protein with a mol. wt. of 54,000 (see Page 124). The 3' portion of the HSV-2 1.7kb mRNA was analysed to identify possible polypeptide coding sequences of the equivalent HSV-2 protein. Computer analysis using the PTRANS program indicated that the longest open reading frame contains 122 amino acids and terminates at position 369 (Fig. 40) which is 142nuc upstream from the 1.7kb mRNA 3' terminus. Examination of the codon usage of HSV proteins has shown that, in most cases, the third base position of amino acid codons has a higher G+C content compared with the G+C content of the first and second positions (McGeoch, 1984). The G+C content across the proposed 54,000 mol. wt. protein coding region is 67% and in the third base position the G+C content is markedly higher (82%); these data therefore suggest that the longest reading frame in this region encodes the C-terminal portion of a HSV-2 protein equivalent to the HSV-1 54,000 mol. wt. polypeptide.

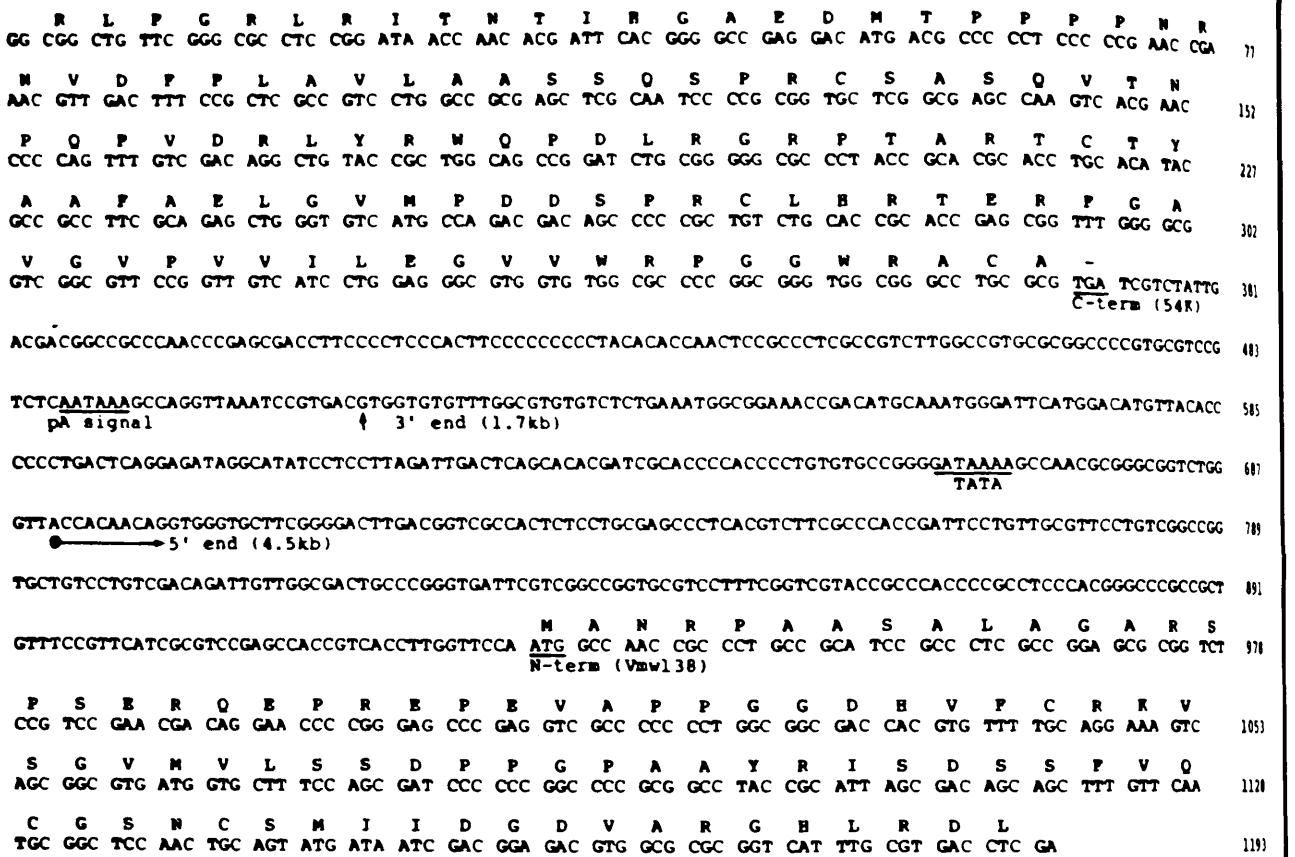


Figure 40. Locations of proposed amino acid sequences at the 3' terminal portion of the 1.7kb mRNA and the 5' terminal region of the 4.5kb transcript. The nucleotide sequence is numbered as in Fig. 25. Predicted amino acid residues are shown above the nucleotide sequence and proposed locations of the C-terminus of a 54,000 mol. wt. protein and the N-terminus of Vmw138 are underlined. The 5' and 3' termini of the 4.5kb and 1.7kb transcripts are indicated as are polyadenylation and TATA box signals for these transcripts.

Supportive evidence for this proposal is the high homology between amino acids in this reading frame and those of a putative VZV protein (see Page 163).

b) Homology between the proposed N-terminal regions of the HSV-2 Vmw138 and HSV-1 Vmw136 proteins.

The first ATG codon within the 4.5kb mRNA is located at position 934 (Fig. 40) which is 243nuc downstream from the 5' terminus. DNA sequences flanking the ATG triplet (CCAATGG) resemble those at other functional initiation codons (Kozak, 1981). This triplet precedes an open reading frame which extends to the limit of the DNA sequence and contains 86 codons; it is probable that this reading frame represents the N-terminal portion of Vmw138.

A comparison between this region and the equivalent N-terminal portion of HSV-1 Vmw136 reveals areas of high and low homology between both the nucleotide and the predicted amino acid sequences (Fig. 41). Alignment of amino acid sequences to maximise homology indicates that the N-terminal eight amino acids are identical in both serotypes apart from one alteration and the sequences share high amino acid homology (75%) between residues 40 to 80. These conserved regions are separated by an area of considerably lower homology (61%) and computer analysis predicts that the HSV-1 polypeptide contains an additional nine amino acids between residues 31 and 32 of the HSV-2 protein. These different levels of amino acid conservation

```
          M   A   N   R   P   A   A   S   A   L   A   G
HSV-2     ATG GCC AAC CGC CCT GCC GCA TCC GCC CTC GCC GGA   969
          *** *** * * *** **  *** *** *** *   * * * * * *
HSV-1     ATG GCC AGC CGC CCA GCC GCA TCC TCT CCC GTC GAA
                  S                       S   P   V   E

          A   R   S   P   S   E   R   Q   E   P   R   E
HSV-2     GCG CGG TCT CCG TCC GAA CGA CAG GAA CCC CGG GAG  1005
          *** *** *   *** *** *   **  *** **  **  *   *
HSV-1     GCG CGG GCC CCG GTT GGG GGA CAG GAG GCC GGC GGC
                  A   V   G   G               A   G   G

          P   E   V   A   P   P   G
HSV-2     CCC GAG GTC GCC CCC CCT GGC                      1026
          *** *   *   *** ** *   **
HSV-1     CCC AGC GCA GCC ACC CAG GGG GAG GCC GCC GGG GCC
                S   A       T   Q       E   A   A   G   A

                          G   D   H   V   F   C   R   K
HSV-2                     GGC GAC CAC GTG TTT TGC AGG AAA  1050
                          *** **  *** *** *   *** *   *
HSV-1     CCT CTC GCC CAC GGC CAC CAC GTG TAC TGC CAG CGA
            P   L   A   H     H           Y     Q   R

          V   S   G   V   M   V   L   S   S   D   P   P
HSV-2     GTC AGC GGC GTG ATG GTG CTT TCC AGC GAT CCC CCC  1086
          *** *   *** *** *** *** *** *** *   *   *   ***
HSV-1     GTC AAT GGC GTG ATG GTG CTT TCC GAC AAG ACG CCC
              N                       D   K   T

          G   P   A   A   Y   R   I   S   D   S   S   F
HSV-2     GGC CCC GCG GCC TAC CGC ATT AGC GAC AGC AGC TTT  1122
          **  **  *** **  *** *** **  *** **  *** * * ***
HSV-1     GGG TCC GCG TCC TAC CGC ATC AGC GAT AGC AAC TTT
                  S   S                       N

          V   Q   C   G   S   N   C   S   M   I   I   D
HSV-2     GTT CAA TGC GGC TCC AAC TGC AGT ATG ATA ATC GAC  1158
          **  *** **  **  *** *** *** *   *** **  *** ***
HSV-1     GTC CAA TGT GGT TCC AAC TGC ACC ATG ATC ATC GAC
                                      T

          G   D   V   A   R   G   H   L   R   D   L
HSV-2     GGA GAC GTG GCG CGC GGT CAT TTG CGT GAC CTC      1191
          *** *** *** * * *** **  *       *   * *
HSV-1     GGA GAC GTG GTG CGC GGG CGC CCC AGG ACC CGG
                      V       R   P       T   R
```

Figure 41. Alignment of the DNA sequences and predicted amino acid residues at the N-termini of HSV-2 Vmw138 and HSV-1 Vmw136. The HSV-2 DNA sequence is numbered as in Fig. 25 and conserved bases are indicated by asterisks. The predicted amino acid sequence for Vmw138 is shown above the HSV-2 DNA sequence and amino acid residues which are not conserved in Vmw136 are indicated below the HSV-1 DNA sequence.

reflect the nucleotide sequences in this region which show a similar pattern of variable homology (Fig. 41).

c) Homology between the proposed C-terminal regions of the Vmw138 and Vmw136 polypeptides.

Analysis of the equivalent HSV-1 region indicates that the C-terminal portion of Vmw136 extends into the untranslated leader sequence of the 1.2kb mRNA (see Page 138). The HSV-2 sequence also contains a single open reading frame which terminates 82nuc downstream from the 1.2kb mRNA 5' end (Fig. 42, position 329). The same open reading frame within equivalent HSV-2 and HSV-1 genome regions strongly suggests that the HSV-2 coding region is organised identically to that of HSV-1. This implies that transcription control signals required for synthesis of the HSV-2 1.2kb mRNA are located within the translated region of the 4.5kb mRNA. In this region, the HSV-2 and HSV-1 sequences are colinear and show a high degree of homology (Fig. 43). There are a total of 35 nucleotide sequence differences, 30 of which occur in the third base position of amino acid codons; these nucleotide differences result in five amino acid changes within a region encoding 109 amino acids. Recently, the DNA sequence of this region in HSV-2 strain 333 has been published (Galloway and Swain, 1984). These authors have also identified an open reading frame which contains an identical amino acid sequence to that predicted here for HSV-2 strain HG52.

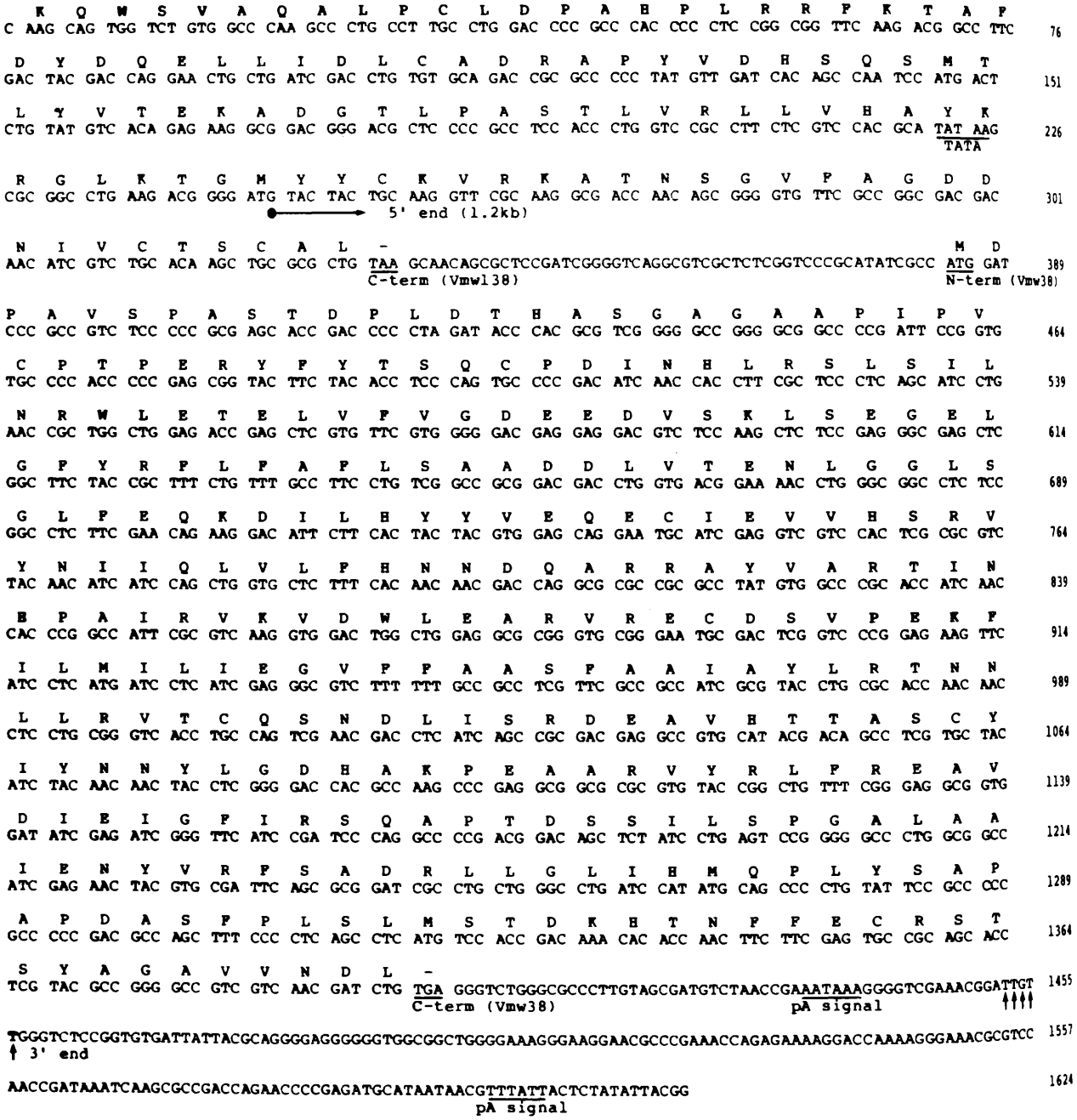

K Q W S V A Q A L P C L D P A H P L R R F K T A F
C AAG CAG TGG TCT GTG GCC CAA GCC CTG CCT TGC CTG GAC CCC GCC CAC CCC CTC CGG CGG TTC AAG ACG GCC TTC 76
D Y D Q E L L I D L C A D R A P Y V D H S Q S M T
GAC TAC GAC CAG GAA CTG CTG ATC GAC CTG TGT GCA GAC CGC GCC CCC TAT GTT GAT CAC AGC CAA TCC ATG ACT 151
L Y V T E K A D G T L P A S T L V R L L V H A Y K
CTG TAT GTC ACA GAG AAG GCG GAC GGG ACG CTC CCC GCC TCC ACC CTG GTC CGC CTT CTC GTC CAC GCA TAT AAG 226
TATA
R G L K T G M Y Y C K V R K A T N S G V F A G D D
CGC GGC CTG AAG ACG GGG ATG TAC TAC TGC AAG GTT CGC AAG GCG ACC AAC AGC GGG GTG TTC GCC GGC GAC GAC 301
5' end (1.2kb)
N I V C T S C A L -
AAC ATC GTC TGC ACA AGC TGC GCG CTG TAA GCAACAGCGCTCCGATCGGGGTCAGGCGTCGCTCTCGGTCCCGCATATCGCC ATG GAT 389
C-term (Vmw138)
M D
N-term (Vmw38)
P A V S P A S T D P L D T H A S G A G A A P I P V
CCC GCC GTC TCC CCC GCG AGC ACC GAC CCC CTA GAT ACC CAC GCG TCG GGG GCC GGG GCG GCC CCG ATT CCG GTG 464
C P T P E R Y F Y T S Q C P D I N H L R S L S I L
TGC CCC ACC CCC GAG CGG TAC TTC TAC ACC TCC CAG TGC CCC GAC ATC AAC CAC CTT CGC TCC CTC AGC ATC CTG 539
N R W L E T E L V F V G D E E D V S K L S E G E L
AAC CGC TGG CTG GAG ACC GAG CTC GTG TTC GTG GGG GAC GAG GAG GAC GTC TCC AAG CTC TCC GAG GGC GAG CTC 614
G F Y R F L F A F L S A A D D L V T E N L G G L S
GGC TTC TAC CGC TTT CTG TTT GCC TTC CTG TCG GCC GCG GAC GAC CTG GTG ACG GAA AAC CTG GGC GGC CTC TCC 689
G L F E Q K D I L H Y Y V E Q E C I E V V H S R V
GGC CTC TTC GAA CAG AAG GAC ATT CTT CAC TAC TAC GTG GAG CAG GAA TGC ATC GAG GTC GTC CAC TCG CGC GTC 764
Y N I I Q L V L F H N N D Q A R R A Y V A R T I N
TAC AAC ATC ATC CAG CTG GTG CTC TTT CAC AAC AAC GAC CAG GCG CGC CGC GCC TAT GTG GCC CGC ACC ATC AAC 839
H P A I R V K V D W L E A R V R E C D S V P E K F
CAC CCG GCC ATT CGC GTC AAG GTG GAC TGG CTG GAG GCG CGG GTG CGG GAA TGC GAC TCG GTC CCG GAG AAG TTC 914
I L M I L I E G V F F A A S F A A I A Y L R T N N
ATC CTC ATG ATC CTC ATC GAG GGC GTC TTT TTT GCC GCC TCG TTC GCC GCC ATC GCG TAC CTG CGC ACC AAC AAC 989
L L R V T C Q S N D L I S R D E A V H T T A S C Y
CTC CTG CGG GTC ACC TGC CAG TCG AAC GAC CTC ATC AGC CGC GAC GAG GCC GTG CAT ACG ACA GCC TCG TGC TAC 1064
I Y N N Y L G D H A K P E A A R V Y R L F R E A V
ATC TAC AAC AAC TAC CTC GGG GAC CAC GCC AAG CCC GAG GCG GCG CGC GTG TAC CGG CTG TTT CGG GAG GCG GTG 1139
D I E I G F I R S Q A P T D S S I L S P G A L A A
GAT ATC GAG ATC GGG TTC ATC CGA TCC CAG GCC CCG ACG GAC AGC TCT ATC CTG AGT CCG GGG GCC CTG GCG GCC 1214
I E N Y V R F S A D R L L G L I H M Q P L Y S A P
ATC GAG AAC TAC GTG CGA TTC AGC GCG GAT CGC CTG CTG GGC CTG ATC CAT ATG CAG CCC CTG TAT TCC GCC CCC 1289
A P D A S F P L S L M S T D K H T N F F E C R S T
GCC CCC GAC GCC AGC TTT CCC CTC AGC CTC ATG TCC ACC GAC AAA CAC ACC AAC TTC TTC GAG TGC CGC AGC ACC 1364
S Y A G A V V N D L -
TCG TAC GCC GGG GCC GTC GTC AAC GAT CTG TGA GGGTCTGGGCGCCCTTGTAGCGATGTCTAACCGAAATAAAGGGGTCGAAACGGATTGT 1455
C-term (Vmw38)
pA signal
TGGGTCTCCGGTGTGATTATTACGCAGGGGAGGGGGGTGGCGGCTGGGGAAAGGGAAGGAACGCCCGAAACCAGAGAAAAGGACCAAAAGGGAAACGCGTCC 1557
3' end
AACCGATAAATCAAGCGCCGACCAGAACCCCGAGATGCATAATAACGTTTATTACTCTATATTACGG 1624
pA signal

Figure 42. Locations of the proposed protein coding regions for Vmw38 and the C-terminal portion of Vmw138. The HSV-2 nucleotide sequence is numbered as in Fig. 31 and predicted amino acid residues are shown above the nucleotide sequence. The positions of the C-terminus of Vmw138 and the N- and C-termini of Vmw38 are underlined. Locations of the 5' and 3' termini of the 1.2kb mRNA are indicated as are the TATA box homology and AATAAA signal for this transcript; the AATAAA signal for a leftwards-transcribed mRNA also is shown.

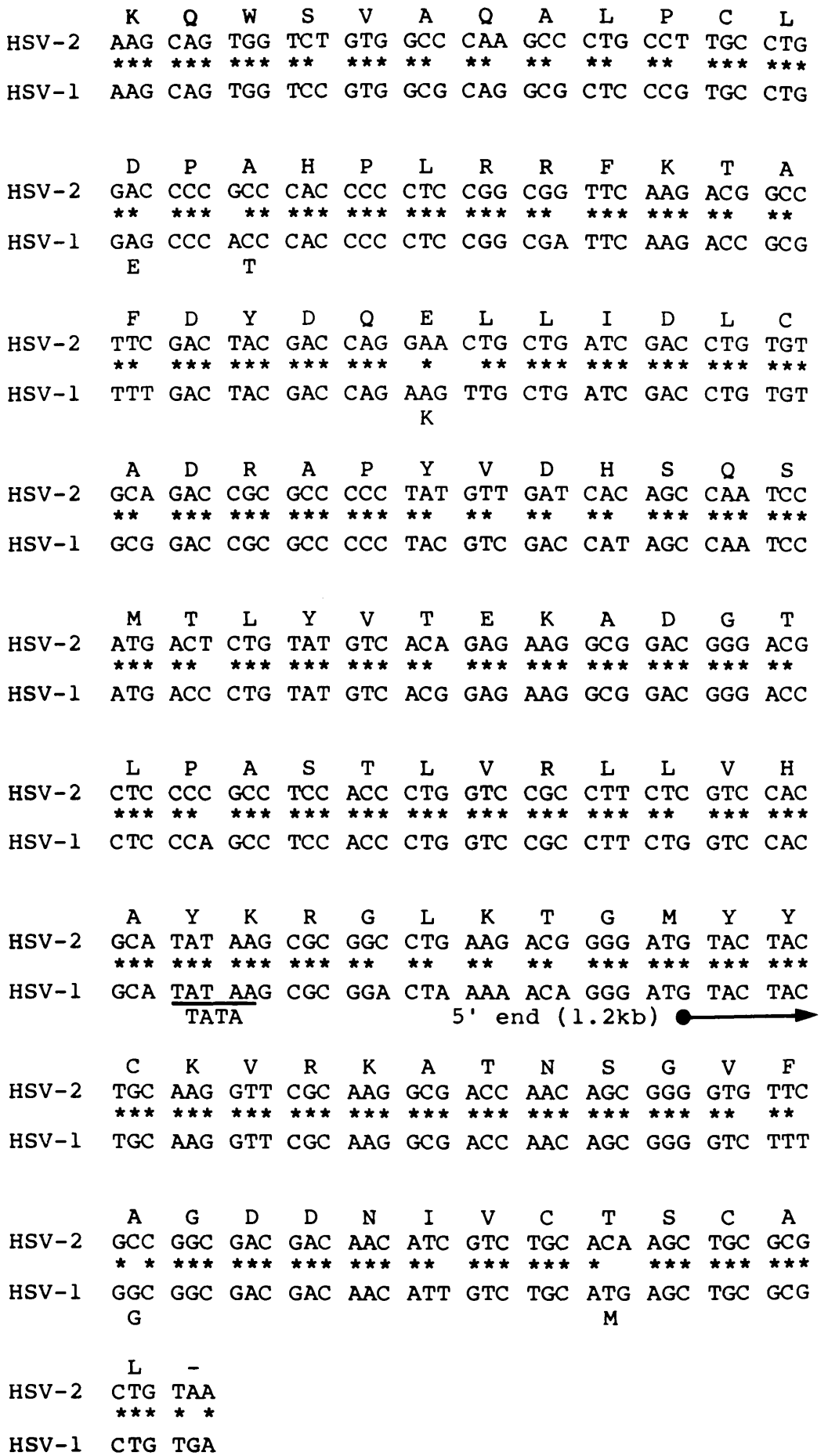

K Q W S V A Q A L P C L
HSV-2 AAG CAG TGG TCT GTG GCC CAA GCC CTG CCT TGC CTG 37
HSV-1 AAG CAG TGG TCC GTG GCG CAG GCG CTC CCG TGC CTG
D P A H P L R R F K T A
HSV-2 GAC CCC GCC CAC CCC CTC CGG CGG TTC AAG ACG GCC 73
HSV-1 GAG CCC ACC CAC CCC CTC CGG CGA TTC AAG ACC GCG
E T
F D Y D Q E L L I D L C
HSV-2 TTC GAC TAC GAC CAG GAA CTG CTG ATC GAC CTG TGT 109
HSV-1 TTT GAC TAC GAC CAG AAG TTG CTG ATC GAC CTG TGT
K
A D R A P Y V D H S Q S
HSV-2 GCA GAC CGC GCC CCC TAT GTT GAT CAC AGC CAA TCC 145
HSV-1 GCG GAC CGC GCC CCC TAC GTC GAC CAT AGC CAA TCC
M T L Y V T E K A D G T
HSV-2 ATG ACT CTG TAT GTC ACA GAG AAG GCG GAC GGG ACG 181
HSV-1 ATG ACC CTG TAT GTC ACG GAG AAG GCG GAC GGG ACC
L P A S T L V R L L V H
HSV-2 CTC CCC GCC TCC ACC CTG GTC CGC CTT CTC GTC CAC 217
HSV-1 CTC CCA GCC TCC ACC CTG GTC CGC CTT CTG GTC CAC
A Y K R G L K T G M Y Y
HSV-2 GCA TAT AAG CGC GGC CTG AAG ACG GGG ATG TAC TAC 253
HSV-1 GCA TAT AAG CGC GGA CTA AAA ACA GGG ATG TAC TAC
TATA
5' end (1.2kb)
C K V R K A T N S G V F
HSV-2 TGC AAG GTT CGC AAG GCG ACC AAC AGC GGG GTG TTC 289
HSV-1 TGC AAG GTT CGC AAG GCG ACC AAC AGC GGG GTC TTT
A G D D N I V C T S C A
HSV-2 GCC GGC GAC GAC AAC ATC GTC TGC ACA AGC TGC GCG 325
HSV-1 GGC GGC GAC GAC AAC ATT GTC TGC ATG AGC TGC GCG
G M
L -
HSV-2 CTG TAA
HSV-1 CTG TGA

Figure 43. A comparison between the DNA sequences and predicted amino acid residues at the C-termini of HSV-2 Vmw138 and HSV-1 Vmw136. The HSV-2 nucleotide sequence is numbered as in Fig. 31 and conserved bases are indicated by asterisks. The amino acid sequence of Vmw138 is shown above the HSV-2 DNA sequence and residues which are not conserved in Vmw136 are indicated below the HSV-1 DNA sequence. The locations of the 5' terminus of the HSV-1 1.2kb mRNA and TATA box homologues are shown.

------ AMINO ACID COMPOSITION ------

Ala	35	10.4%	Val	24	7.1%	Leu	35	10.4%	Ile	21	6.2%
Gly	14	4.2%	Pro	17	5.0%	Cys	7	2.1%	Met	4	1.2%
His	10	3.0%	Phe	18	5.3%	Tyr	15	4.5%	Trp	2	0.6%
Asn	15	4.5%	Gln	8	2.4%	Ser	27	8.0%	Thr	15	4.5%
Lys	6	1.8%	Arg	21	6.2%	Asp	21	6.2%	Glu	22	6.5%

Approximate Molecular Weight = 37667.52

------ CODON USAGE ------

TTT	Phe	7	2.1%	TCT	Ser	1	0.3%	TAT	Tyr	2	0.6%	TGT	Cys	0	0.0%
TTC	Phe	11	3.3%	TCC	Ser	9	2.7%	TAC	Tyr	13	3.8%	TGC	Cys	7	2.1%
TTA	Leu	0	0.0%	TCA	Ser	0	0.0%	TAA	---	0	0.0%	TGA	---	1	0.3%
TTG	Leu	0	0.0%	TCG	Ser	8	2.4%	TAG	---	0	0.0%	TGG	Trp	2	0.6%
CTT	Leu	2	0.6%	CCT	Pro	0	0.0%	CAT	His	2	0.6%	CGT	Arg	0	0.0%
CTC	Leu	14	4.1%	CCC	Pro	11	3.3%	CAC	His	8	2.4%	CGC	Arg	13	3.8%
CTA	Leu	1	0.3%	CCA	Pro	0	0.0%	CAA	Gln	0	0.0%	CGA	Arg	2	0.6%
CTG	Leu	18	5.3%	CCG	Pro	6	1.8%	CAG	Gln	8	2.4%	CGG	Arg	6	1.8%
ATT	Ile	3	0.9%	ACT	Thr	0	0.0%	AAT	Asn	0	0.0%	AGT	Ser	1	0.3%
ATC	Ile	18	5.3%	ACC	Thr	11	3.3%	AAC	Asn	15	4.4%	AGC	Ser	8	2.4%
ATA	Ile	0	0.0%	ACA	Thr	1	0.3%	AAA	Lys	1	0.3%	AGA	Arg	0	0.0%
ATG	Met	4	1.2%	ACG	Thr	3	0.9%	AAG	Lys	5	1.5%	AGG	Arg	0	0.0%
GTT	Val	0	0.0%	GCT	Ala	0	0.0%	GAT	Asp	5	1.5%	GGT	Gly	0	0.0%
GTC	Val	11	3.3%	GCC	Ala	23	6.8%	GAC	Asp	16	4.7%	GGC	Gly	7	2.1%
GTA	Val	0	0.0%	GCA	Ala	0	0.0%	GAA	Glu	4	1.2%	GGA	Gly	0	0.0%
GTG	Val	13	3.8%	GCG	Ala	12	3.6%	GAG	Glu	18	5.3%	GGG	Gly	7	2.1%

------ BASE COMPOSITION ------

	Y	T	C	A	G	R	ALL
NUMBER	557	186	371	177	280	457	1014
PERCENT	54.9	18.3	36.6	17.5	27.6	45.1	100.0

Figure 44. Predicted amino acid composition of HSV-2 Vmw38. Also shown are the frequency of usage for codons specifying the amino acids and the G+C content in the protein coding region.

d) Homology between HSV-2 and HSV-1 Vmw38 proteins.

Previous reports have mapped Vmw38 to the region examined here (Docherty et al., 1981; Galloway et al., 1982). The first ATG triplet at position 384 (Fig. 42), within the sequences specifying the 1.2kb mRNA, precedes a single open reading frame which extends for 337 codons and terminates with a TGA stop signal at position 1395 (Fig. 42). This sequence would specify a 37,667 mol. wt. polypeptide, which is close to the size of the 37,800 mol. wt. product obtained by in vitro translation of mRNA selected with HSV-2 Bgl II n (Docherty et al., 1981). The amino acid composition of Vmw38 is shown in Fig. 44; the G+C content in the third base position of the codons is 90% within a region whose total G+C content is 64%.

For the HSV-1 1.2kb mRNA, the first ATG triplet was followed by a stop codon in the same reading frame (Fig. 32, positions 314 and 329), which also marked the C-terminus of Vmw136. This ATG triplet therefore does not encode the N-terminus of HSV-1 Vmw38 and is absent in the HSV-2 sequence (Fig. 32). DNA sequences flanking the first ATG codon of the HSV-2 1.2kb mRNA are identical to those at the second ATG triplet of the HSV-1 1.2kb mRNA which reinforces the view that this second ATG codon is used in the HSV-1 1.2kb mRNA.

Alignment of HSV-2 and HSV-1 sequences by the HOMOL program predicts that the HSV-1 protein contains an additional three amino acid residues between positions 2 and 3 in the HSV-2 polypeptide (Fig. 45). Further

```
          M   D                       P   A   V   S   P   A   S
HSV-2     ATG GAT                     CCC GCC GTC TCC CCC GCG AGC   410
          *** ***                     **  *** **  *** *** **
HSV-1     ATG GAT TCC GCG GCC CCA GCC CTC TCC CCC GCT CTG
                  S   A   A           L                   L

          T   D   P   L   D   T   H   A   S   G   A   G
HSV-2     ACC GAC CCC CTA GAT ACC CAC GCG TCG GGG GCC GGG   446
          **  * * *     **      *  *** **  * * * * *  *
HSV-1     ACG GCC CTT ACG GAC CAG AGC GCG ACG GCG GAC CTG
              A   L   T       Q   S       T   A   D   L

          A   A   P   I   P   V   C   P   T   P   E   R
HSV-2     GCG GCC CCG ATT CCG GTG TGC CCC ACC CCC GAG CGG   482
          ***   * * * *** **    * *** ***   * *** *** **
HSV-1     GCG ATC CAG ATT CCA AAG TGC CCC GAC CCC GAG AGG
              I   Q           K           D

          Y   F   Y   T   S   Q   C   P   D   I   N   H
HSV-2     TAC TTC TAC ACC TCC CAG TGC CCC GAC ATC AAC CAC   518
          *** *** *** *** *** *** **  *** *** **  *** ***
HSV-1     TAC TTC TAC ACC TCC CAG TGT CCC GAC ATT AAC CAC

          L   R   S   L   S   I   L   N   R   W   L   E
HSV-2     CTT CGC TCC CTC AGC ATC CTG AAC CGC TGG CTG GAG   554
          **  *** *** *** *** *** **  *** *** *** *** **
HSV-1     CTG CGC TCC CTC AGC ATC CTT AAC CGC TGG CTG GAA

          T   E   L   V   F   V   G   D   E   E   D
HSV-2     ACC GAG CTC GTG TTC GTG GGG GAC GAG GAG GAC   587
          *** *** **  **  *** *** *** *** *** *** ***
HSV-1     ACC GAG CTT GTT TTC GTG GGG GAC GAG GAG GAC
```

Figure 45. A comparison between the DNA sequences and predicted amino acid residues at the N-termini of HSV-2 and HSV-1 Vmw38. The HSV-2 DNA sequence is numbered as in Fig. 31 and identical bases are indicated by asterisks. The amino acid sequence for HSV-2 Vmw38 is shown above the HSV-2 DNA sequence and residues which are not conserved in the HSV-1 protein are indicated below the HSV-1 DNA sequence.

downstream, there is a discrete region of low homology which results in 14 amino acid differences within a stretch of 22 amino acid residues (Fig. 45, positions 408 to 473). Following this non-homologous region, from positions 474 to 587 (Fig. 45), which is the limit of HSV-1 strain 17+ data, there are few differences in the nucleotide sequences and the amino acid residues are identical in both serotypes. Draper et al. (1982) have published the DNA sequence for the 1.2kb mRNA of HSV-1 strain KOS which has subsequently been amended (E.K. Wagner, personal communication). The predicted N-terminal amino acid sequence of strain 17+ differs from strain KOS between residues 7 and 19; there are five amino acid differences and the strain 17+ sequence contains an additional leucine residue (Fig. 46). A comparison of the HSV-2 amino acid sequence with the amended HSV-1 strain KOS sequence from residue 30 to the TGA stop codon reveals that the Vmw38 polypeptides are colinear and are highly conserved (94%) apart from the 30 amino acid region at the N-terminus (Fig. 46). Within the highly conserved region, there are only 19 amino acid differences; most nucleotide differences occur in the third base position of codons and both HSV-2 and HSV-1 reading frames terminate at the same stop codon. The HSV-2 strain HG52 amino acid sequence is identical to that of HSV-2 strain 333 (Galloway and Swain, 1984) apart from two differences at residues 173 (valine ⟶ isoleucine) and 235 (aspartic acid ⟶ glycine; Fig. 46).

```
HSV-1[17+] MDSAAPALSPALTALTDQSATADLAIQIPKCPDPERYFYTSQCPDINHL
           **    ** *** *   *  *     *  ** ** ****************
HSV-2      MD    PAVSPASTDPLDTHASGAGAAPIPVCPTPERYFYTSQCPDINHL
           **     * **          *     *  ** ** ****************
HSV-1KOS   MDSAAPA SPPLWAHTGHSATADLAIQIPKCPDPERYFYTSQCPDINHL

             50
HSV-2      RSLSILNRWLETELVFVGDEEDVSKLSEGELGFYRFLFAFLSAADDLVTE
           ********************************* ******************
HSV-1      RSLSILNRWLETELVFVGDEEDVSKLSEGELSFYRFLFAFLSAADDLVTE

            100
HSV-2      NLGGLSGLFEQKDILHYYVEQECIEVVHSRVYNIIQLVLFHNNDQARRAY
           ************************** ********************** *
HSV-1      NLGGLSGLFEQKDILHYYVEQECIEVAHSRVYNIIQLVLFHNNDQARREY

            150
HSV-2      VARTINHPAIRVKVDWLEARVRECDSVPEKFILMILIEGVFFAASFAAIA
           ** ******** ************ ************** **********
HSV-1      VAGTINHPAIRAKVDWLEARVRECASVPEKFILMILIEGIFFAASFAAIA

            200
HSV-2      YLRTNNLLRVTCQSNDLISRDEAVHTTASCYIYNNYLGDHAKPEAARVYR
           ************************************** ****   ***
HSV-1      YLRTNNLLRVTCQSNDLISRDEAVHTTASCYIYNNYLGGHAKPPPDRVYG

            250
HSV-2      LFREAVDIEIGFIRSQAPTDSSILSPGALAAIENYVRFSADRLLGLIHMQ
           *** ** ************** **** **********************
HSV-1      LFRQAVEIEIGFIRSQAPTDSHILSPAALAAIENYVRFSADRLLGLIHMK

            300
HSV-2      PLYSAPAPDASFPLSLMSTDKHTNFFECRSTSYAGAVVNDL-
           ** *** ************************************
HSV-1      PLFSAPPPDASFPLSLMSTDKHTNFFECRSTSYAGAVVNDL-
```

Figure 46. A comparison between the amino acid sequences of HSV-2 Vmw38 and the amended sequence for the Vmw38 polypeptide of HSV-1 strain KOS. The HSV-2 sequence is numbered and conserved residues are indicated by asterisks. Amino acid sequences between residues 5 to 15 are not identical in HSV-1 17^+ and HSV-1 KOS strains, therefore, the N-terminal 50 residues of HSV-1 strain 17^+ Vmw38 are also compared with the HSV-2 sequence. Amino acid differences between the two HSV-1 strains are underlined.

DISCUSSION.

11. Sequence Conservation between Equivalent HSV Proteins.

Proposed protein coding regions within the overlapping mRNAs have been predicted by DNA sequence analysis and comparisons between equivalent HSV-2 and HSV-1 protein coding sequences reveal homologous and non-homologous regions. Most of the Vmw38 protein sequences are well conserved in both serotypes except for an area of low homology at the N-termini where the HSV-1 protein contains an insertion of three amino acids. The C-terminal portions of HSV-2 Vmw138 and HSV-1 Vmw136 are highly conserved, however, the N-termini of these proteins have areas of high homology separated by a poorly conserved region in which the HSV-1 protein contains an additional nine amino acids.

The DNA sequences of the gC (Frink et al., 1983; Dowbenko and Lasky, 1984; Draper et al., 1984a; Swain et al., 1985), gD (Watson et al., 1982; Watson, 1983; McGeoch et al., 1985) and TK genes (McKnight, 1980; Wagner et al., 1981; Swain and Galloway, 1983) have been determined for both serotypes. Comparisons between these equivalent HSV-2 and HSV-1 protein coding regions have indicated that, in every case, there is an increased variation in amino acid sequences at the N-termini compared with other areas of the polypeptides. Furthermore, the N-terminal region of the HSV-1 gC polypeptide contains an insertion of 27 amino acids as compared with its HSV-2 counterpart. N-terminal

portions of the HSV-2 IE polypeptides, for which only limited sequence data are available, also share low homology with corresponding regions of HSV-1 IE proteins (Whitton, 1984; Whitton and Clements, 1984b) except for the sequences of IE Vmw65 which are highly conserved (Whitton et al., 1983).

It is unlikely that low conservation between short stretches of amino acids at the N-termini represents different functional domains in equivalent HSV proteins. A possible explanation for low sequence homology in these regions is that poorly conserved residues are not functionally important and do not play a significant role in determining protein structure. Supportive evidence for this proposal is that removal of the initiation codon from the HSV-1 TK gene does not significantly reduce viral TK activity (Halpern and Smiley, 1984). In this HSV-1 mutant, translation is probably initiated from two distal in-frame ATG codons which denote the N-termini of proteins with sizes of 39,000 and 38,000 mol. wt. (Preston and McGeoch, 1981; Marsden et al., 1983; Haarr et al., 1985); these polypeptides lack 45 and 59 N-terminal residues respectively as compared with the wild-type TK protein. Other comparisons between amino acid sequences of mammalian, yeast and bacillar alcohol dehydrogenases have shown that conserved residues can be correlated with the functional properties and tertiary structure of the enzymes (Eklund et al., 1976b). For example, residues involved in subunit interactions, ADP-ribose binding and enzymatic

activity are all well conserved. However, the N-terminal portions of these dehydrogenases share low homology and vary in length. Structural analysis of the horse liver alcohol dehydrogenase has indicated that the N-terminal region forms a loosely organised loop on the outside of the molecule and is not involved in enzyme function (Eklund et al., 1976a).

12. Components of HSV Ribonucleotide Reductase.

HSV-1 Vmw136 is a component of the virus-induced ribonucleotide reductase activity (Dutia, 1983; Preston et al., 1984b). In vitro translation of hybrid-selected mRNAs has mapped Vmw138 to the equivalent region of the HSV-2 genome (Galloway et al., 1982) and ribonucleotide reductase activity can be recovered from immunoprecipitates containing this protein (Huszar and Bacchetti, 1983). DNA sequence studies presented here show that Vmw136 and Vmw138 share amino acid homology, implying that Vmw138 is a component of HSV-2 ribonucleotide reductase activity. The properties of both HSV-2 and HSV-1 enzymes are identical (Averett et al., 1984), further suggesting that the enzymes from both serotypes share common structural features.

There is accumulating evidence that the HSV ribonucleotide reductase contains a second component which is the Vmw38 protein encoded by the 1.2kb mRNA. Biochemical studies on the structure of HSV ribonucleotide reductase indicate that the enzyme has similar properties to the

mammalian ribonucleotide reductase (see Introduction, Page 56) and comprises two non-identical subunits. Treatment of partially purified cell extracts with pyridoxal phosphate (Pyr.P) and 4-methyl-5-amino-1-formylisoquinoline thiosemicarbozone (MAIQ), both inhibitors of mammalian ribonucleotide reductase activity, inactivates the HSV enzyme (Cohen et al., 1985); Pyr.P interacts with the M1 mammalian subunit, possibly at a lysine residue in either the allosteric or catalytic sites (Cory and Mansell, 1975) and MAIQ treatment destroys the tyrosine free radical on the M2 subunit (Thelander and Graslund, 1983). Up to 77% of HSV ribonucleotide reductase activity can be recovered by combining extracts which had been treated separately with the two inhibitors suggesting that the sites of action of the inhibitors are not on the same HSV polypeptide and that inactivated subunits can associate to give enzyme activity.

Immunoprecipitation studies have shown that monoclonal antibodies directed against Vmw138 and Vmw136 also precipitated a 38,000 mol. wt. species (Galloway et al., 1982; Huszar and Bacchetti, 1983; Bacchetti et al., 1984; Preston et al., 1984b) shown to be the Vmw38 polypeptide encoded by the 1.2kb mRNA (Frame et al., 1985). There are three possible explanations for this observation: 1) firstly, it has been suggested that the coding sequences of Vmw138 and Vmw38 are at least partially colinear (Galloway et al., 1982), possibly as a result of mRNA splicing. This proposal conflicts with structural analyses of transcripts mapping within this region which indicate

that the mRNAs encoding these proteins are unspliced. In addition, the C-terminus of Vmw138 does not overlap with the N-terminal region of Vmw38, as also is the case for the equivalent HSV-1 polypeptides.

2) it is possible that monoclonal antibodies may recognise a common antigenic site in both proteins. A computer-assisted comparison between sequences of HSV-1 Vmw136 and Vmw38 has revealed no significant stretches of common amino acids; the longest stretch of homology between these proteins is four amino acids (I. Nikas, personal communication).

3) a third possibility is that Vmw138 and Vmw38 are associated in a protein-protein complex, hence an antibody directed against one protein could precipitate both species. Monoclonal antibodies which precipitated both polypeptides also inhibited HSV-2 ribonucleotide reductase activity (Huszar and Bacchetti, 1983) and enzyme activity was recovered also from the pelletted material of precipitated antibody-antigen complexes (Huszar and Bacchetti, 1983; Preston et al., 1984b). Furthermore, antiserum induced by a synthetic oligopeptide composed of the seven C-terminal amino acid residues of HSV-1 Vmw38 (McLauchlan and Clements, 1982) precipitated Vmw136 in addition to Vmw38; these amino acid residues are identical in the HSV-2 protein, hence the antiserum also precipitated HSV-2 Vmw138 and Vmw38 (Frame et al., 1985).

The HSV-1 mutant, 17tsVP1207, which contains a lesion in the protein coding region of Vmw136, fails to

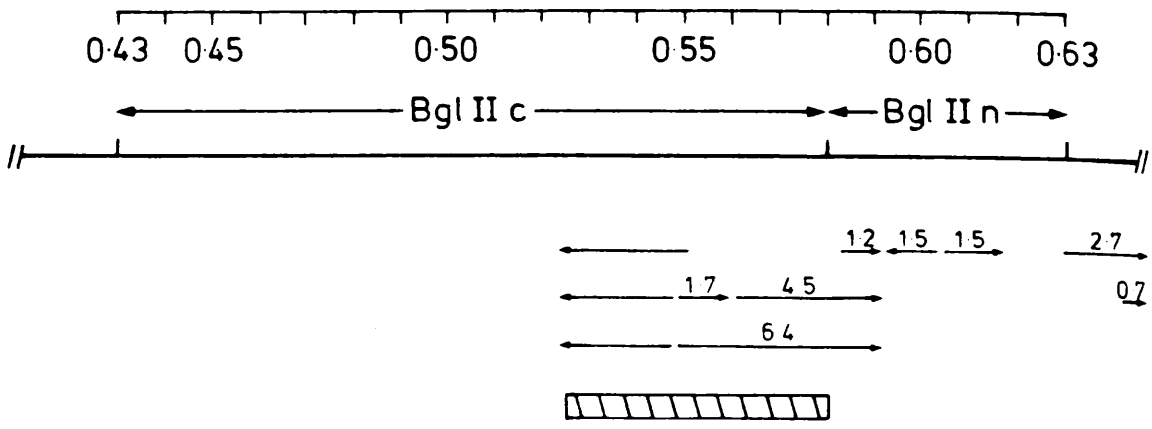


Figure 47. Locations of transcripts within the Bgl II c and Bgl II n transforming regions. Positions for the 6.4kb, 4.5kb, 1.7kb and 1.2kb mRNAs are taken from this study. The sizes of the leftwards-transcribed species, whose 5' ends map at the 5' co-terminal region of the 6.4kb and 1.7kb mRNAs, are not known, however, it is assumed that they are similar in size to a 3.8kb species which is located at the equivalent HSV-1 genomic region. The hatched area indicates the region within Bgl II c required for tumorigenicity. The 1.5kb mRNA to the right of the 3' co-terminal mRNA family encoding ribonucleotide reductase specifies a 61,000 mol. wt. polypeptide which contains those sequences implicated in morphological transformation. The 2.7kb species encodes gC-2 and only the 5' portion of this mRNA maps within Bgl II n. Transcripts to the right of the ribonucleotide reductase locus were taken from Draper et al., 1984b.

induce a viral ribonucleotide reductase activity at the non-permissive temperature (npt; Dutia, 1983; Preston et al., 1984b) Cells infected at the npt with 17tsVP1207 synthesise a non-functional Vmw136 and the oligopeptide-induced antiserum precipitates only Vmw38 (Frame et al., 1985). This suggests that, in cells infected with 17tsVP1207, a protein complex between Vmw136 and Vmw38 does not form at the npt and this may be the basis of the ribonucleotide reductase-negative phenotype of this mutant.

13. Location of the Ribonucleotide Reductase Gene within HSV-2 Transforming Regions.

Studies using cloned DNA have identified two HSV-2 fragments, Bgl II c and Bgl II n (mtr II), which are involved in cellular transformation (see Introduction, Page 46). The 6.4kb and 4.5kb mRNAs span both of these fragments while the 1.7kb and 1.2kb transcripts are contained entirely within Bgl II c and Bgl II n respectively (Fig. 47). The corresponding region of the HSV-1 genome, which specifies transcripts and proteins of similar sizes to those of HSV-2 DNA, does not have a comparable morphological transforming ability (Reyes et al., 1979); the ability of the HSV-1 region equivalent to Bgl II c to induce immortalisation of hamster cells has not been tested. There are no extensive insertions or deletions between the HSV-2 and HSV-1 DNA sequences examined here which may account for the different morphological transforming efficiences, however, there are areas of low

sequence homology. Within the left-hand portion of mtr II, there are regions of low homology downstream from the mRNA 3' co-terminus and at the N-terminus of Vmw38. As yet, it is not known whether these non-homologous segments contribute to the different transforming abilities of DNA from these HSV-2 and HSV-1 genomic regions. Recent evidence would suggest that a DNA fragment within mtr II which induces morphological transformation is located to the right of the HSV-2 ribonucleotide reductase locus, within the coding sequences of a 61,000 mol. wt. polypeptide (Fig. 47; Galloway et al., 1984).

Jariwalla et al. (1983) have proposed that immortalised cell lines require the right-hand 36% of Bgl II c for tumorigenicity (Fig. 47). This region contains most of the protein coding sequences of Vmw138 and probably all of the 54,000 mol. wt. polypeptide sequences. In addition, it is likely that major portions of the transcripts which map to the left of the 6.4kb and 1.7kb mRNAs (Fig. 47) are located within this proposed tumorigenic region; further studies are required to analyse the role of these sequences in tumorigenicity.

SECTION C.

14. Amino Acid Sequence Conservation between Proteins at the HSV-2 Ribonucleotide Reductase Locus and Other Herpesvirus Polypeptides.

The sequences of the entire genomes of EBV B95-8 strain (Baer et al., 1984) and VZV (Davison, 1983; A.J. Davison, personal communication) have been determined and analysed for potential protein coding regions. Comparisons of the predicted amino acid sequences of Vmw38 and the N- and C-terminal portions of Vmw138 with those of proposed EBV and VZV protein coding regions are presented in the following sections and reveal that both EBV and VZV have open reading frames corresponding to polypeptides which have homology with Vmw38 and Vmw138. The EBV proteins equivalent to Vmw138 and Vmw38 have sizes of 93,000 and 34,000 mol. wt. respectively while the equivalent VZV polypeptides have predicted sizes of 87,000 and 35,000 mol. wt. Comparisons with the proposed C-terminal portion of the HSV-2 54,000 mol. wt. protein with EBV and VZV sequences are also presented.

15. HSV-2 Vmw38 Shares Homology with Similarly-sized EBV and VZV Proteins.

Comparisons between the predicted amino acid sequence of HSV-2 Vmw38 and those of the EBV 34,000 and VZV 35,000 mol. wt. polypeptides are represented as CINTHOM plots in Figs. 48 and 49. Both comparisons indicate that

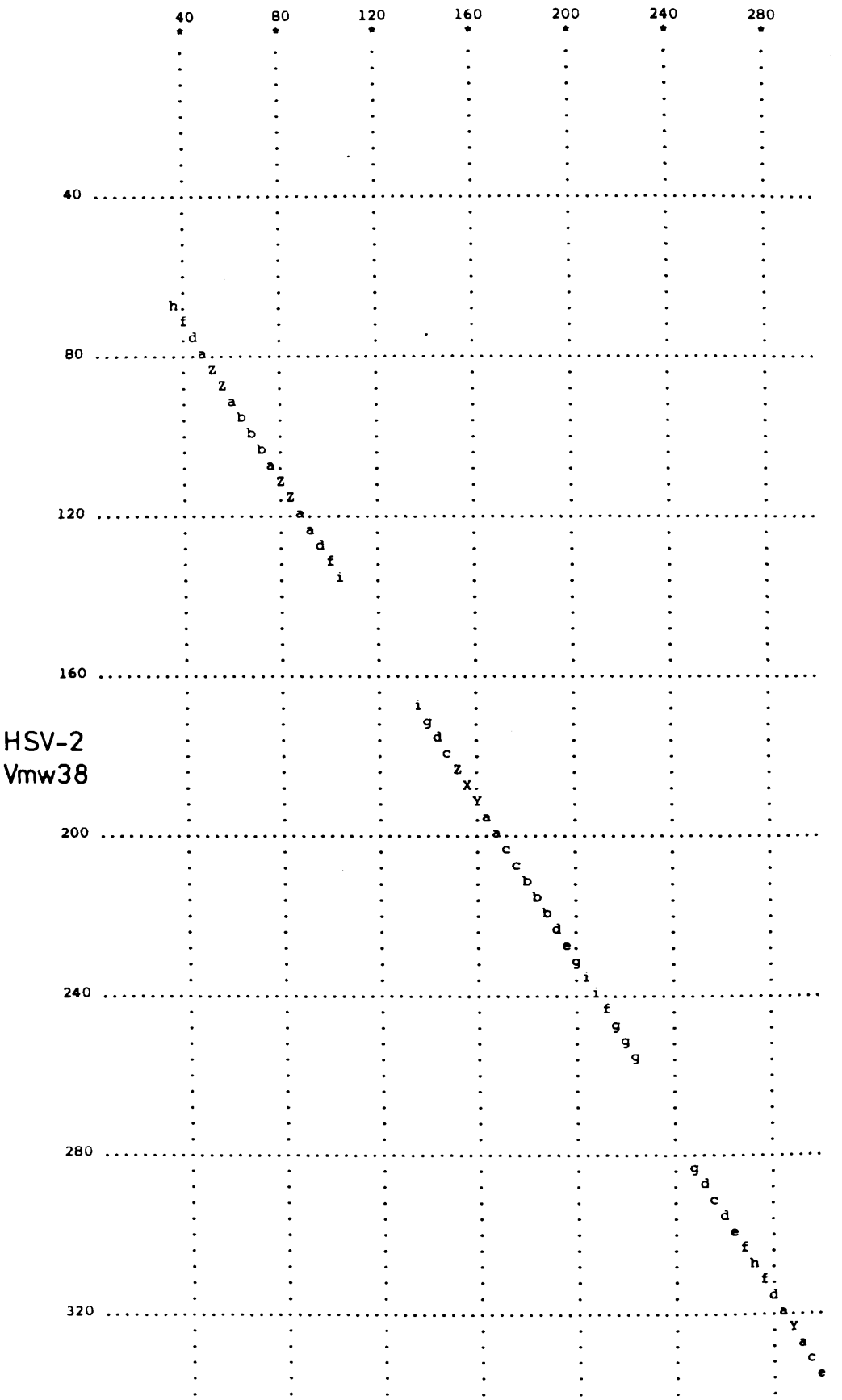

EBV 34K
40
80
120
160
200
240
280
320
HSV-2
Vmw38

Figure 48. A CINTHOM plot comparing amino acid residues of HSV-2 Vmw38 with those of the EBV 34,000 mol. wt. protein. Program settings were: range = 15; scale factor = 0.95; minimum value plotted = 30%; compressed x4.

VZV 35K

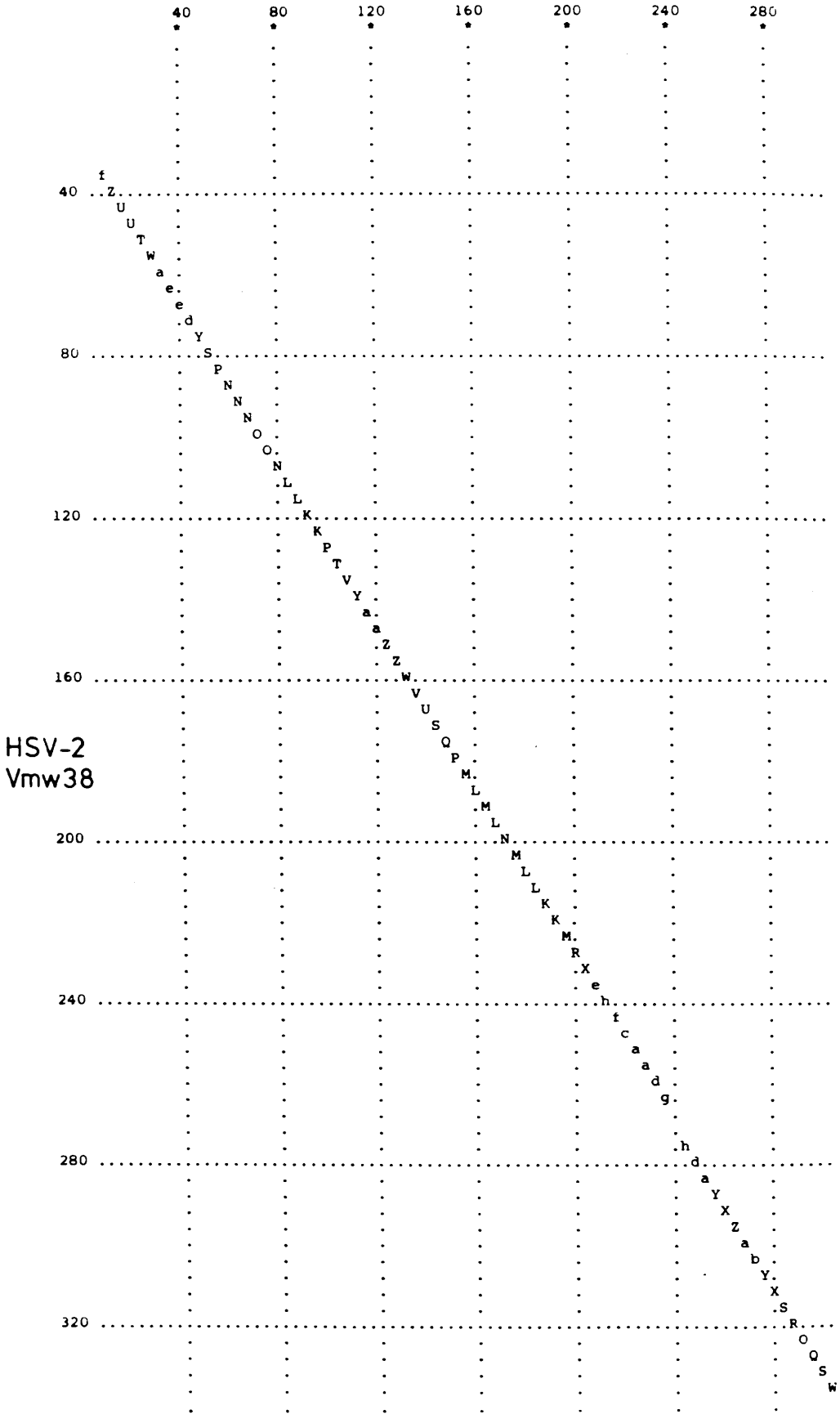


Figure 49. A CINTHOM plot comparing amino acid residues of HSV-2 Vmw38 with those of the VZV 35,000 mol. wt. protein. Program settings were: range = 15; scale factor = 0.95; minimum value plotted = 30%; compressed x4.

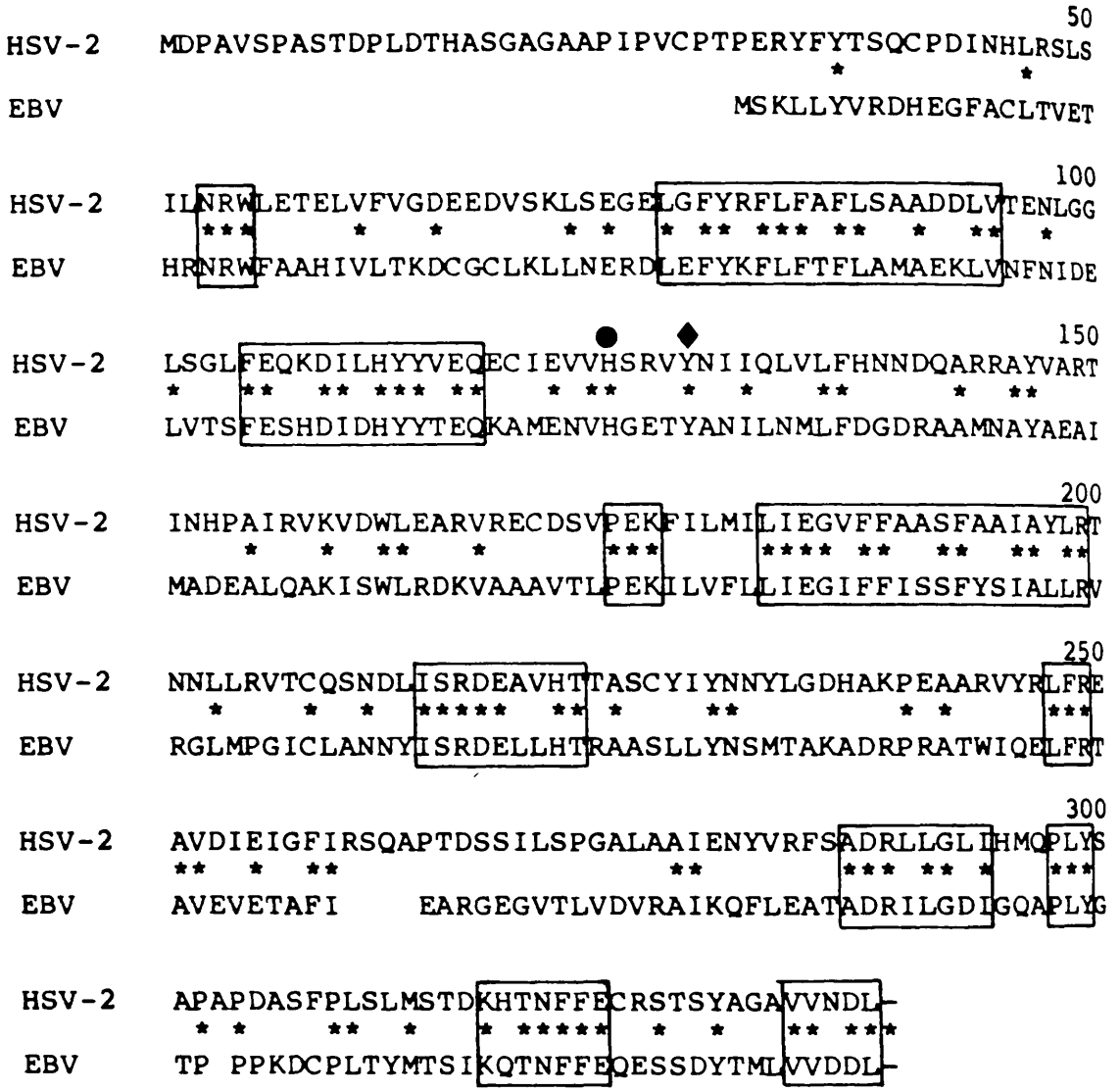


Figure 50. Alignment of the amino acid residues of HSV-2 Vmw38 with those of the EBV 34,000 mol. wt. protein. Asterisks denote conserved residues and the HSV-2 sequence is numbered. Blocks of higher homology are boxed and the proposed locations for the iron binding site (●) and tyrosine free radical (◆) are indicated.

homology with the EBV and VZV proteins extends across most of the Vmw38 coding region and that there are no large insertions/deletions. The HSV/EBV comparison contains two gaps in the diagonal (Fig. 48, between HSV-2 residues 136 to 168 and 256 to 284) and homology is not detectable at the N-termini (Fig. 48, HSV-2 residues 0 to 50). In contrast, the HSV/VZV homology is represented by an almost continuous diagonal which extends to the N-terminus of the VZV protein (Fig. 49) and there is greater homology with the VZV 35,000 mol. wt. protein than with the EBV 34,000 mol. wt. polypeptide. Alignment of protein coding sequences using the HOMOL program indicates that the HSV/VZV amino acid homology is 54% while the HSV/EBV homology is 33%. Alignment of the HSV-2 and EBV proteins reveals blocks of low and high homology (indicated in Fig. 50) and the N-terminal portion of HSV-2 Vmw38 is longer than that of the EBV 34,000 mol. wt. protein (Fig. 50).

Amino acid homology is reflected by the conservation between nucleotide sequences encoding the HSV-2, EBV, and VZV proteins (Figs. 51 and 52). Again, HSV-2 and VZV sequences share a greater degree of conservation than HSV-2 and EBV sequences, an interesting result considering the wide variation in G+C content of HSV (64%) and VZV DNA (38.5%) in this region. The G+C content of EBV DNA is 54.7% in the area encoding the 34,000 mol. wt. protein.

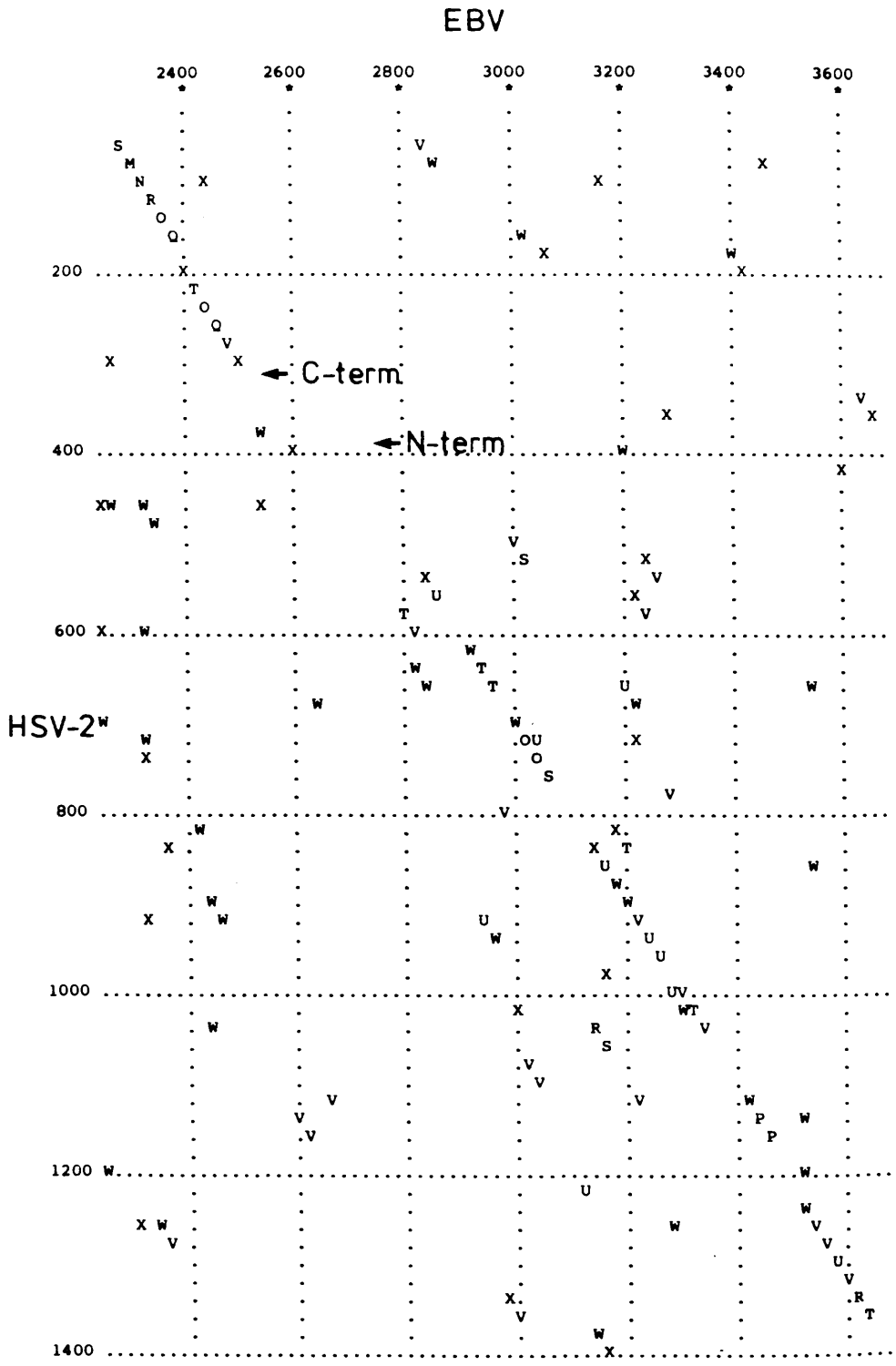

EBV
2400
2600
2800
3000
3200
3400
3600
200
400
600
800
1000
1200
1400
C-term
N-term
HSV-2

Figure 51. A CINTHOM plot comparing HSV-2 and EBV nucleotide sequences. The HSV-2 nucleotide sequence extends from bases 1 to 1400 as numbered in Fig. 31 and encodes the C-terminal portion of Vmw138 and the entire coding region of Vmw38. The EBV nucleotide sequence specifies equivalent regions of the 93,000 and 34,000 mol. wt. proteins. Locations of the C-terminus of Vmw138 and the N-terminus of Vmw38 are arrowed. Program settings were: range = 20; scale factor = 0.95; minimum value plotted = 54%; compressed x20.

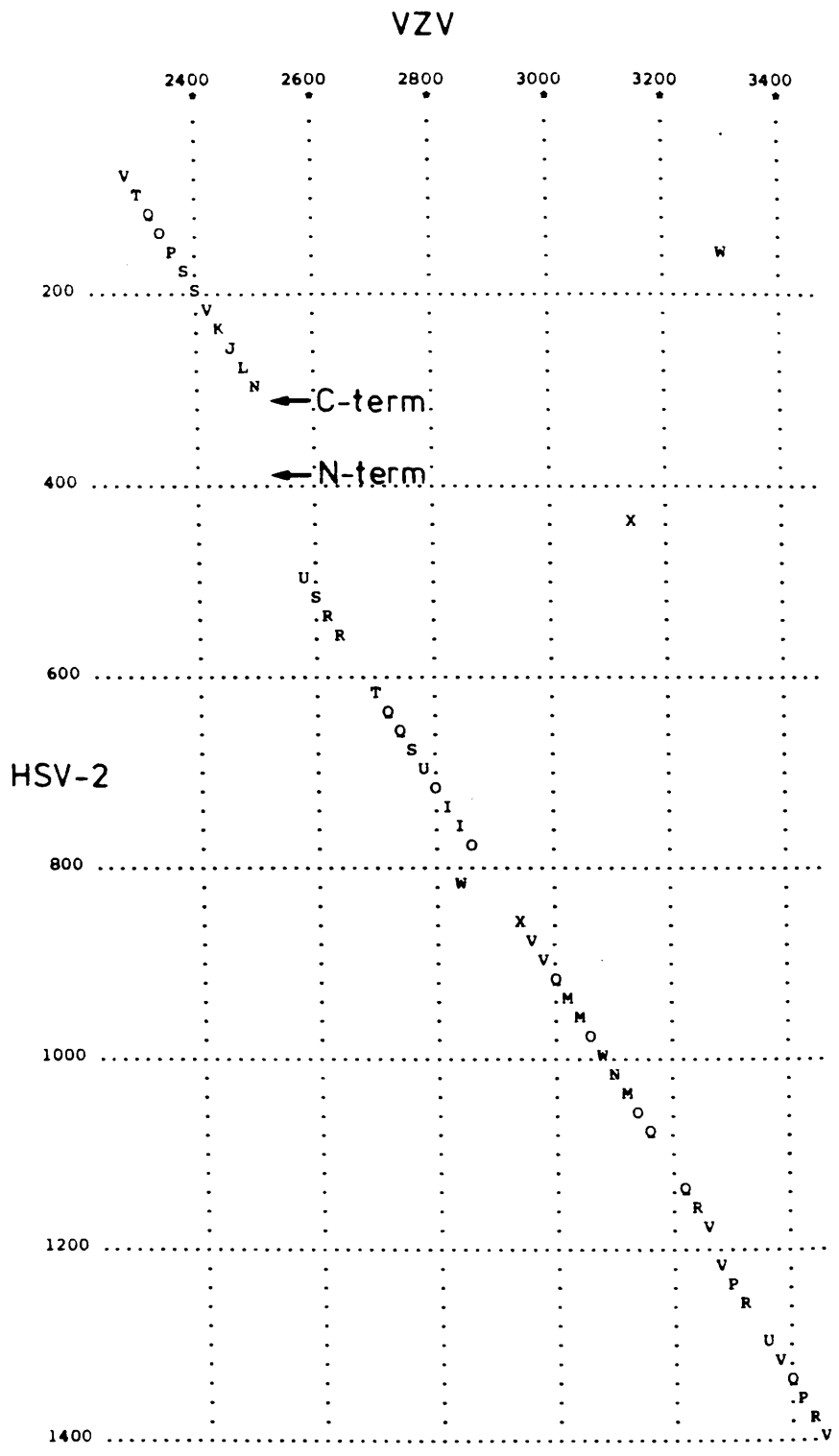

VZV
2400
2600
2800
3000
3200
3400
200
400
600
800
1000
1200
1400
C-term
N-term
HSV-2

Figure 52. A CINTHOM plot comparing HSV-2 and VZV nucleotide sequences. The HSV-2 nucleotide sequence extends from bases 1 to 1400 as numbered in Fig. 31 and encodes the C-terminal portion of Vmw138 and the entire coding region of Vmw38. The VZV nucleotide sequence specifies equivalent regions of the 87,000 and 35,000 mol. wt. proteins. Locations of the C-terminus of Vmw138 and the N-terminus of Vmw38 are arrowed. Program settings were: range = 20; scale factor = 0.95; minimum value plotted = 54%; compressed x20.

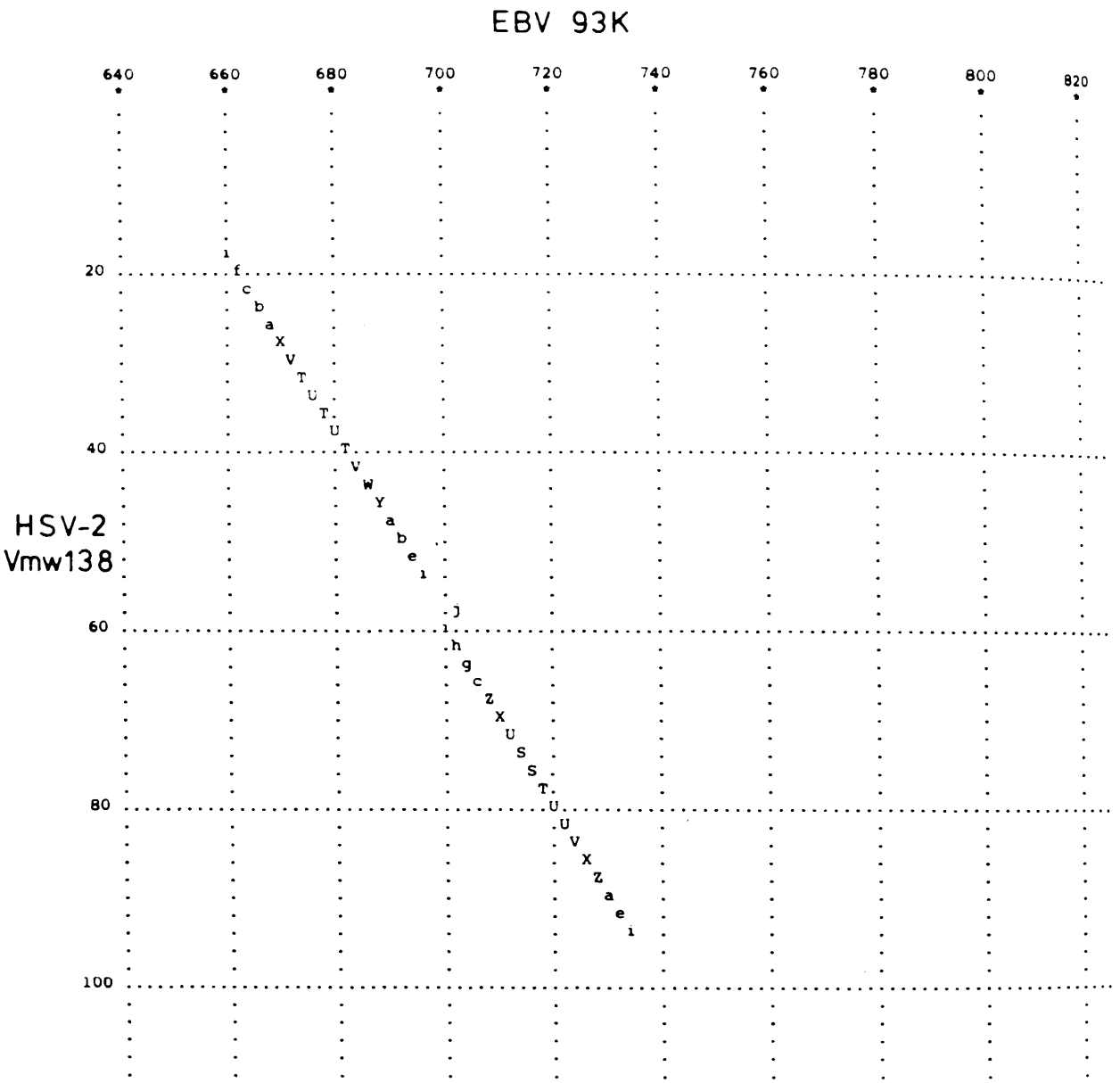


Figure 53. A CINTHOM plot comparing C-terminal amino acid residues of HSV-2 Vmw138 with the equivalent region of the EBV 93,000 mol. wt. polypeptide. Program settings were: range = 12; scale factor = 0.95; minimum value plotted = 30%; compressed x2.

16. Comparison of the HSV-2 Vmw138 Amino Acid Sequences with EBV and VZV Proteins.

a) C-terminal regions.

The 109 amino acid sequence of the C-terminal region of HSV-2 Vmw138 was compared with corresponding portions of the EBV 93,000 and VZV 87,000 mol. wt. polypeptides. Analysis using the CINTHOM program indicates that the HSV-2 and EBV sequences are essentially colinear, however, the EBV sequence contains a C-terminal region which is not present in the HSV-2 sequence (Fig. 53). The HSV-2 and VZV C-terminal regions are colinear with no extensive insertions or deletions (Fig. 54). Alignment of the sequences using the HOMOL program indicates that, in common with comparisons with HSV-2 Vmw38, the HSV-2/VZV homology is higher than that between HSV-2 and EBV sequences. Fig. 55 illustrates the HSV-2/EBV alignment and indicates that there are 75 additional residues at the C-terminus of the 93,000 mol. wt. polypeptide; homology between the sequences is principally associated with three regions shown in Fig. 55. Comparisons between nucleotide sequences in these polypeptide coding regions reflect the homology detected at amino acid level (Figs. 51 and 52).

b) N-terminal regions.

Comparisons between the N-terminal 86 amino acids of HSV-2 Vmw138 and the EBV 93,000 and VZV 87,000 mol. wt. proteins reveals no detectable homology (data not shown).

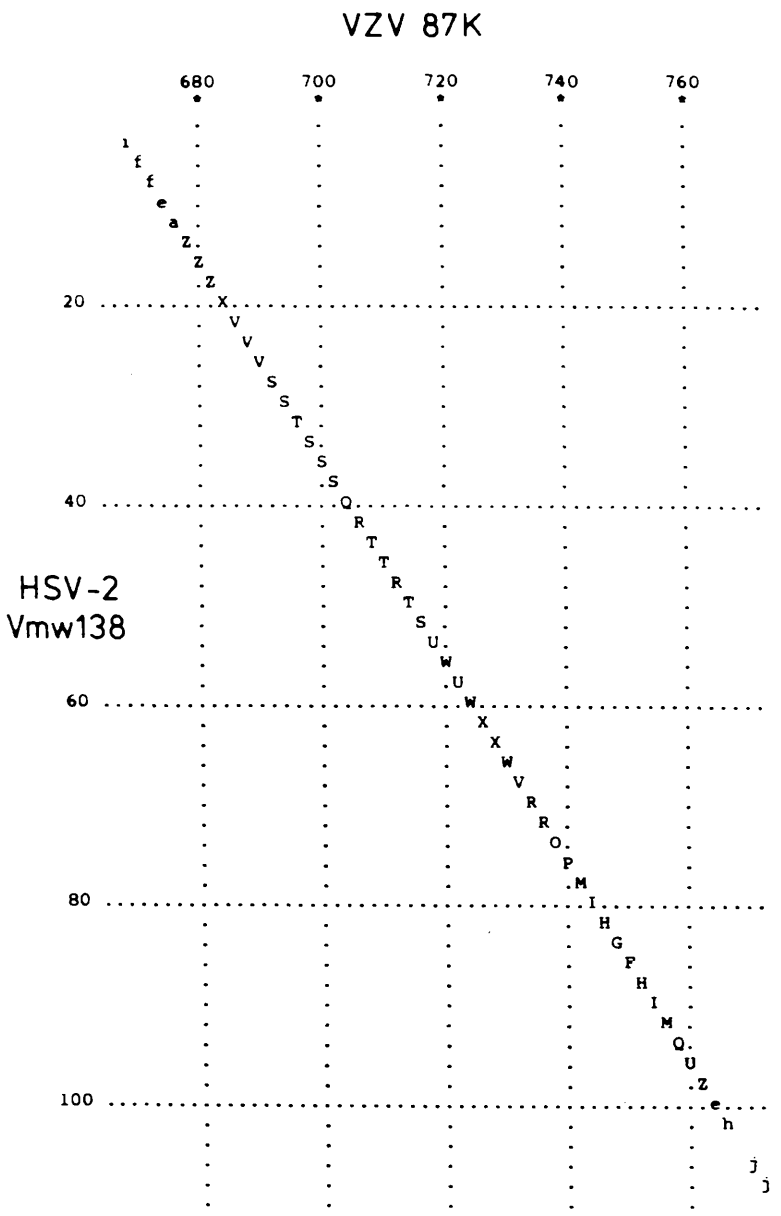


Figure 54. A CINTHOM plot comparing C-terminal amino acid residues of HSV-2 Vmw138 with the equivalent region of the VZV 87,000 mol. wt. protein. Program settings were: range = 12; scale factor = 0.95; minimum value plotted = 30%; compressed x2.

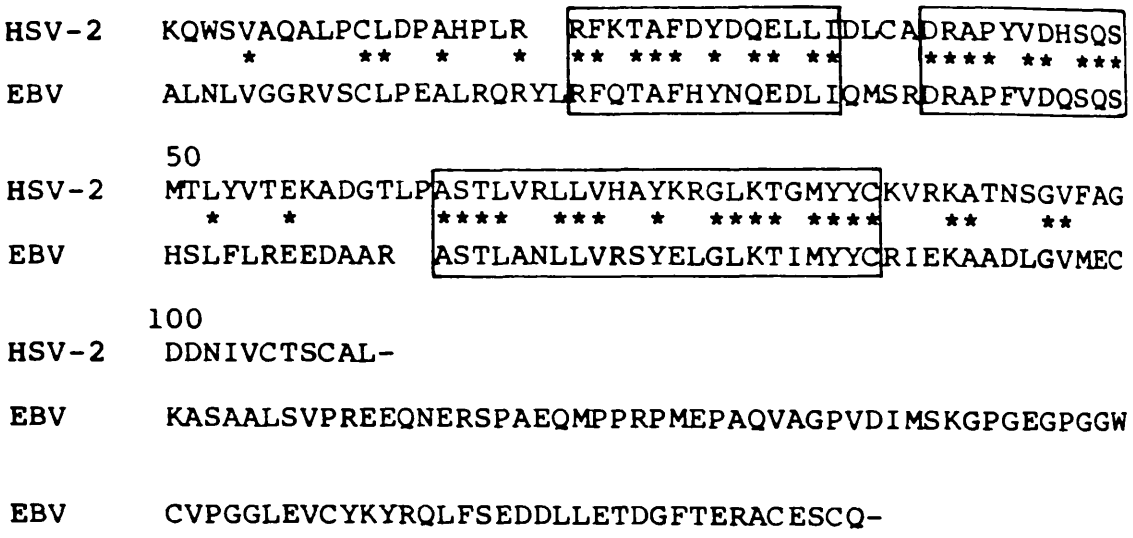


Figure 55. Alignment of the C-terminal amino acid residues of HSV-2 Vmw138 and the EBV 93,000 mol. wt. protein. Asterisks denote conserved residues and the HSV-2 sequence is numbered. Blocks of higher homology are boxed.

Other comparative studies indicate that N-terminal portions of the EBV and VZV proteins share homology with HSV-1 Vmw136 beginning at approximately amino acid residue 490 in the HSV-1 sequence, thus Vmw136 contains a large N-terminal region which is not present in either of the EBV or VZV proteins (I. Nikas, personal communication). Homology between the N-terminal regions of HSV-1 Vmw136 and HSV-2 Vmw138 (see Page 149) would indicate that the HSV-2 polypeptide also contains additional amino acid sequences as compared with the EBV and VZV proteins.

17. Comparison of the Proposed C-terminal Portion of the HSV-2 54,000 mol. wt. Protein with VZV and EBV Polypeptides.

The proposed C-terminal 122 amino acids of the HSV-2 54,000 mol. wt. protein were compared with the amino acid sequences of an open reading frame in VZV DNA which would specify a 54,000 mol. wt. protein (A.J. Davison, personal communication) located immediately upstream from the 87,000 mol. wt. protein. The HSV-2 sequence is homologous to the VZV sequence from approximately residue 360 to the C-terminus of the VZV protein (Fig. 56). This alignment of HSV-2 and VZV amino acid sequences suggests that the VZV and HSV-2 54,000 mol. wt. proteins are functionally equivalent. Comparisons of the HSV-2 and VZV amino acid sequences with proposed protein coding regions of EBV reveal no detectable homology, thus EBV appears not to specify a protein which is equivalent to the HSV-2 and VZV polypeptides.

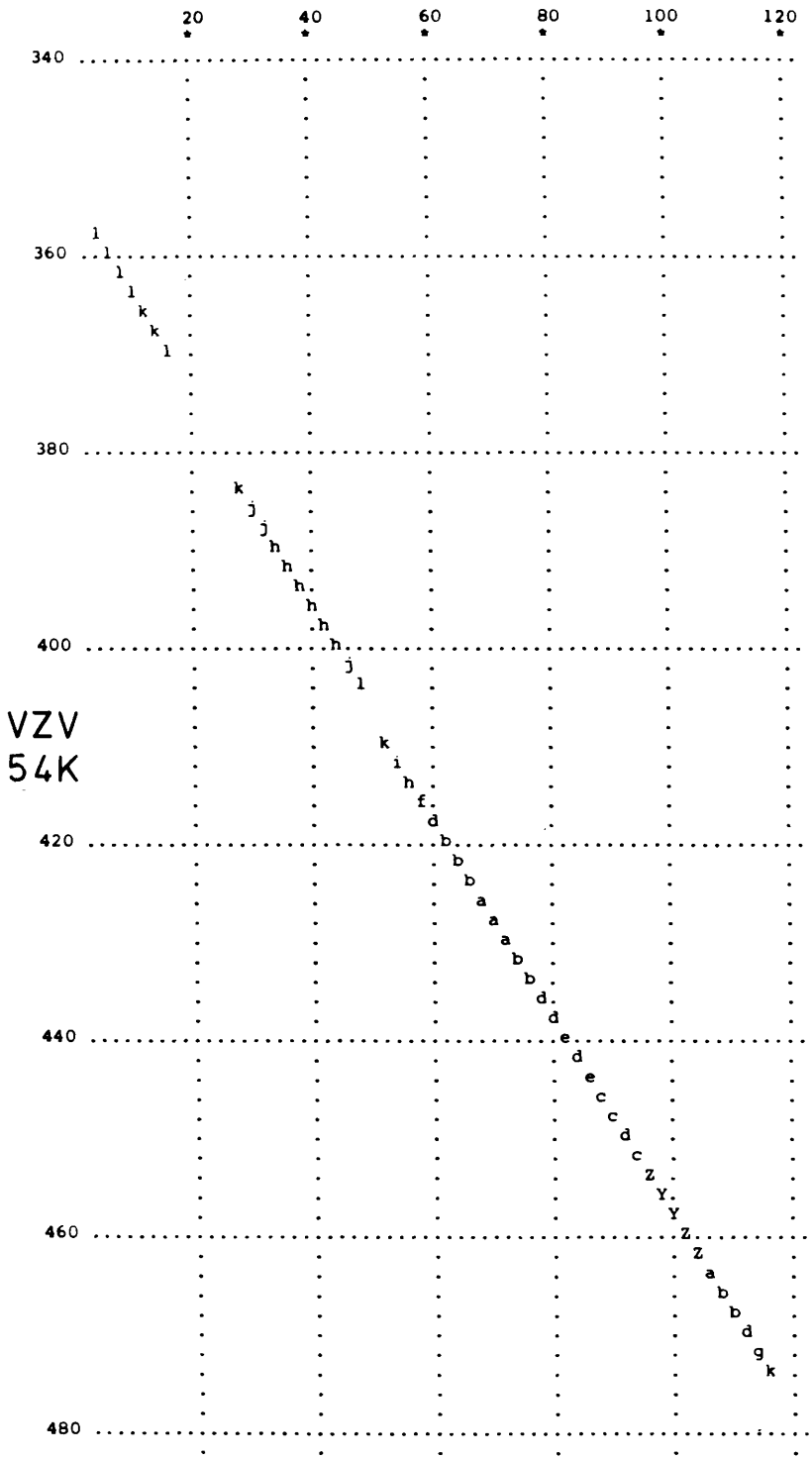

HSV-2 54K
20
40
60
80
100
120
340
360
380
400
420
440
460
480
VZV
54K

Figure 56. A CINTHOM plot comparing the C-terminal amino acid residues of the HSV-2 54,000 mol. wt. protein with the equivalent region of a VZV 54,000 mol. wt. polypeptide. Program settings were: range = 15; scale factor = 0.95; minimum value plotted = 25%; compressed x2.

DISCUSSION.

18. Conservation of Proteins between Herpesviruses.

The C-terminal portion of Vmw138, a component of the HSV-2-specified ribonucleotide reductase, has homology with VZV and EBV proteins with sizes of 87,000 and 93,000 mol. wt. respectively; this homology extends to the N-terminal regions of the VZV and EBV proteins (I. Nikas, personal communication). These results suggest that the 87,000 and 93,000 mol. wt. proteins may form components of ribonucleotide reductases specified by VZV and EBV. Furthermore, HSV-2 Vmw38, the proposed second component of HSV ribonucleotide reductase, has homology with VZV and EBV proteins with sizes of 35,000 and 34,000 mol. wt. respectively, a possible indication of the structural similarities between HSV and proposed VZV and EBV ribonucleotide reductases. EBV induces a novel ribonucleotide reductase activity during virus infection (Henry et al., 1978), although, at present, a virus-induced enzyme has not been detected in VZV-infected cells (W. Clark and M. Frame, personal communication). However, a 34,000 mol. wt. polypeptide is immunoprecipitated from extracts of VZV-infected cells by antiserum directed against the C-terminal seven amino acid residues of HSV-2 Vmw38 (Clark et al., 1985).

Amino acid and DNA sequence conservation between the HSV and VZV sequences is greater than that between the HSV and EBV sequences. DNA/DNA hybridisation studies

have demonstrated that the VZV genome is colinear with the I_L or I_{SL} arrangement of the HSV genome (see Fig. 2) and the region encoding the HSV ribonucleotide reductase has homology with VZV DNA (Davison and Wilkie, 1983a). Nucleotide sequence comparisons between the S segments of HSV and VZV DNA reveal that seven of the thirteen HSV-1 genes have homologues in VZV, however, the gene arrangements in the two S segments is not identical (McGeoch, 1984; A.J. Davison and D.J. McGeoch, personal communication). Similarly, the viral DNA polymerase (Quinn, 1984) and 65,000 mol. wt. structural protein (Dalrymple et al., 1985) involved in stimulation of IE mRNA synthesis, both of which are located in the HSV-1 L segment, are conserved in VZV.

The HSV-2 54,000 mol. wt. protein and equivalent VZV product have no detectable homology with any EBV protein coding sequences. A number of HSV genes, for example the 65,000 mol. wt. protein (Dalrymple et al., 1985) and genes encoded by the HSV-1 DNA S segment (D.J. McGeoch, personal communication), also do not appear to be conserved in the EBV genome, however, there is homology between the HSV-1 DNA polymerase gene and an EBV polypeptide (Quinn, 1984). These data further confirm that functions specified by the HSV and VZV genomes, both alphaherpesviruses with similar biological properties, are more closely related than with those of EBV, a gammaherpesvirus (see Introduction, Page 3).

19. Arrangements of the Transcripts Specifying the VZV and EBV Proteins Equivalent to Vmw138 and Vmw38.

The EBV and VZV polypeptides having homology with the Vmw138, Vmw38 and 54,000 mol. wt. proteins have the same genomic arrangement as the HSV-2 proteins. However, organisation of the EBV mRNAs encoding the 93,000 and 34,000 mol. wt. polypeptides is different from that of the HSV-2 4.5kb and 1.2kb 3' co-terminal transcripts. The EBV 93,000 mol. wt. protein is encoded by a 2.7kb transcript which has a 3' terminus located in the vicinity of the ATG triplet for the 34,000 mol. wt. product (Gibson et al., 1984); the AATAAA signal for this mRNA is partially encoded by the stop codon for the 93,000 mol. wt. protein. The EBV 34,000 mol. wt. protein is specified by a 3.3kb message which has a 5' terminus located within the 93,000 mol. wt. polypeptide coding region. Therefore, in common with the HSV-2 1.2kb message, the transcription control signals which direct synthesis of the 3.3kb mRNA form part of the sequences encoding the 93,000 mol. wt. protein. Sequences involved in transcription of the 3.3kb mRNA will be further described in the General Discussion (Page 195 and Fig. 70).

The mRNAs specifying the VZV polypeptides have not been analysed, however, the TATA box homology, upstream from the HSV-2 1.2kb mRNA and within the protein coding region of Vmw138, is retained in the VZV DNA sequence and an AATAAA sequence does not occur immediately downstream from the coding region of the VZV 87,000 mol. wt. protein (A.J. Davison, personal communication). A potential

polyadenylation signal is located downstream from the C-terminus of the 35,000 mol. wt. protein. These data suggest that the organisation of the VZV mRNAs specifying the 87,000 and 35,000 mol. wt. proteins is similar to that of the HSV-2 transcripts encoding Vmw138 and Vmw38.

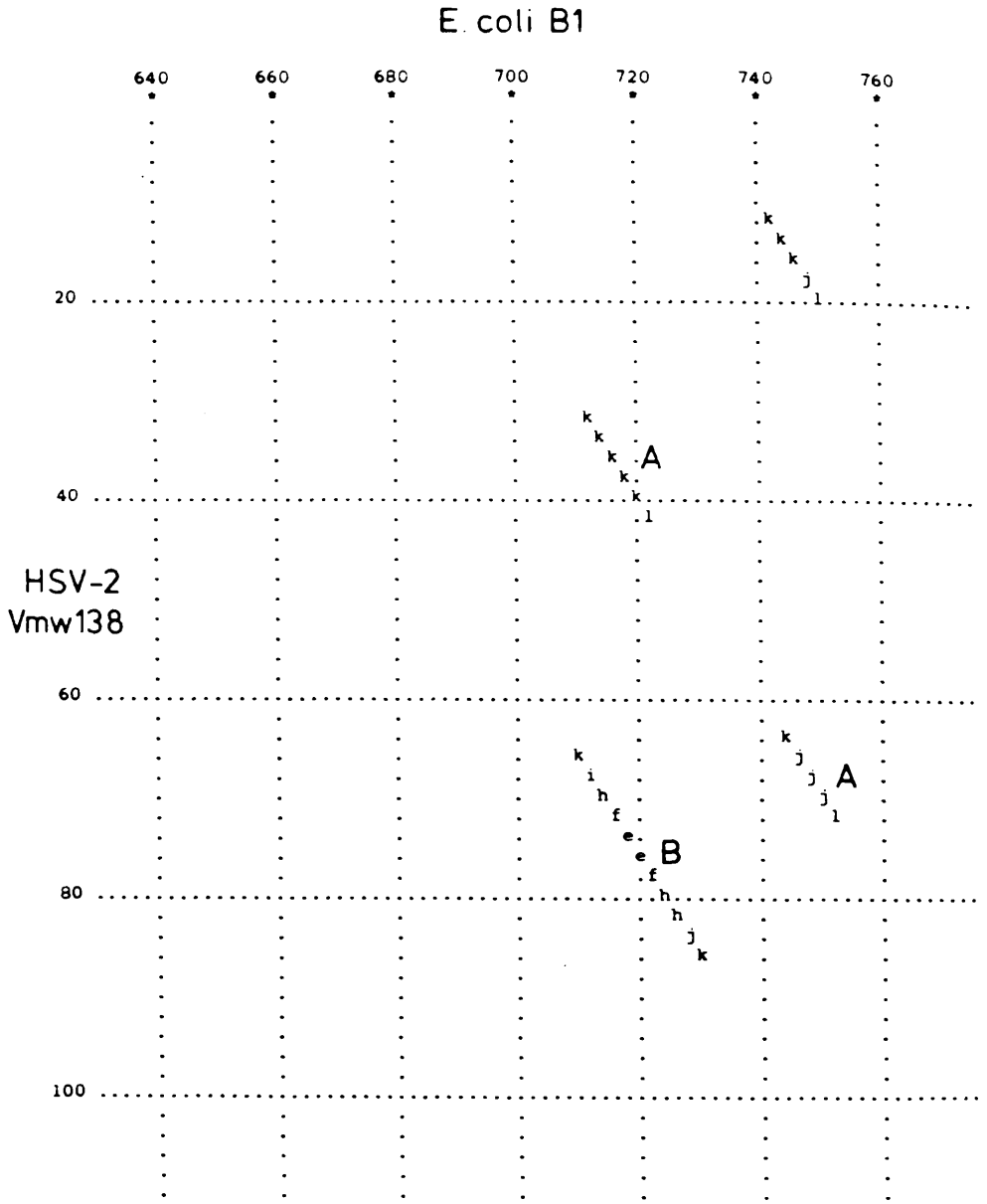


Figure 57. A CINTHOM plot comparing the C-terminal amino acid residues of HSV-2 Vmw138 with the equivalent region of the E. coli B1 polypeptide. Alternative alignments of the sequences are labelled A and B. Program settings were: range = 12; scale factor = 0.95; minimum value plotted = 26%; compressed x2.

SECTION D.

20. Amino Acid Conservation between HSV, Prokaryotic and Eukaryotic Ribonucleotide Reductases.

From the DNA sequences encoding bacterial and mammalian ribonucleotide reductases, amino acid sequences for the E. coli B1 (see Introduction, Page 53; Carlson et al., 1984) and mouse M1 subunits (see Introduction, Page 56; Caras et al., 1985) have been predicted and these were compared with the HSV-2 Vmw138 sequences. The N-terminus of Vmw138 had no homology with E. coli and mouse subunits (data not shown), however, a low level of conservation was detected between the Vmw138 C-terminus and the C-termini of the B1 and M1 polypeptides (Figs. 57 and 58). The CINTHOM plot predicts two possible alignments for the B1 and Vmw138 sequences (shown by A and B in Fig. 57); the alignment with the greater homology, marked B in Fig. 57, is shown in Fig. 59. The viral polypeptide has greater homology with the M1 subunit than with either alignment of the B1 sequences (Fig. 58). There are two regions conserved between viral and mammalian proteins which correspond to areas of homology in the HSV and EBV proteins; these areas are located between residues 44 to 48 and 77 to 84 (Figs. 55 and 59).

Low but detectable homology was observed principally between one area of the HSV-2 Vmw38 and E. coli B2 proteins (shown as A in Fig. 60). Vmw38 has more extensive homology with a surf clam (Spisula solidissima)

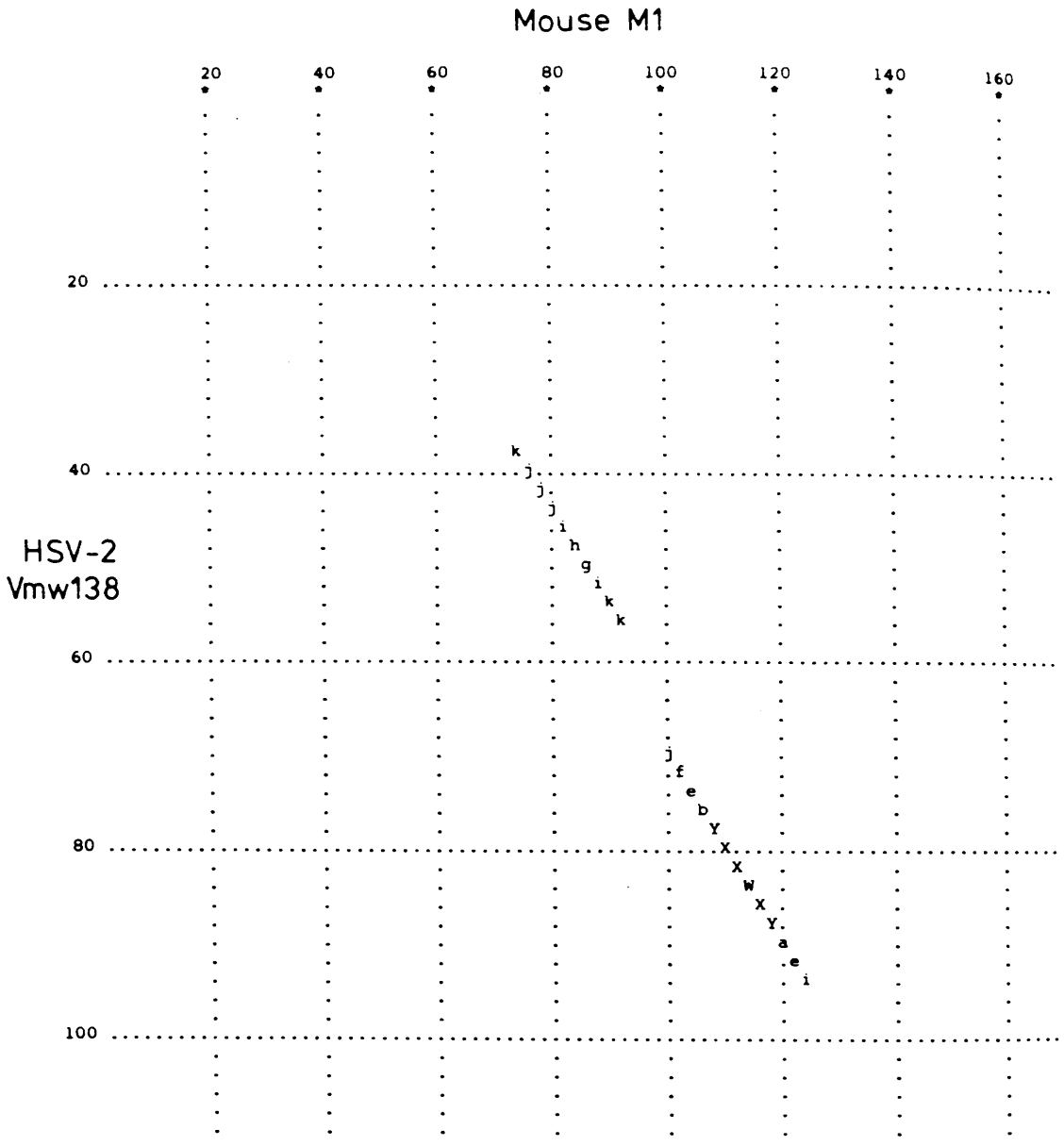


Figure 58. A CINTHOM plot comparing the C-terminal amino acid residues of HSV-2 Vmw138 with the equivalent region of the mouse M1 protein. Program settings were: range = 12; scale factor = 0.95; minimum value plotted = 28%; compressed x2.

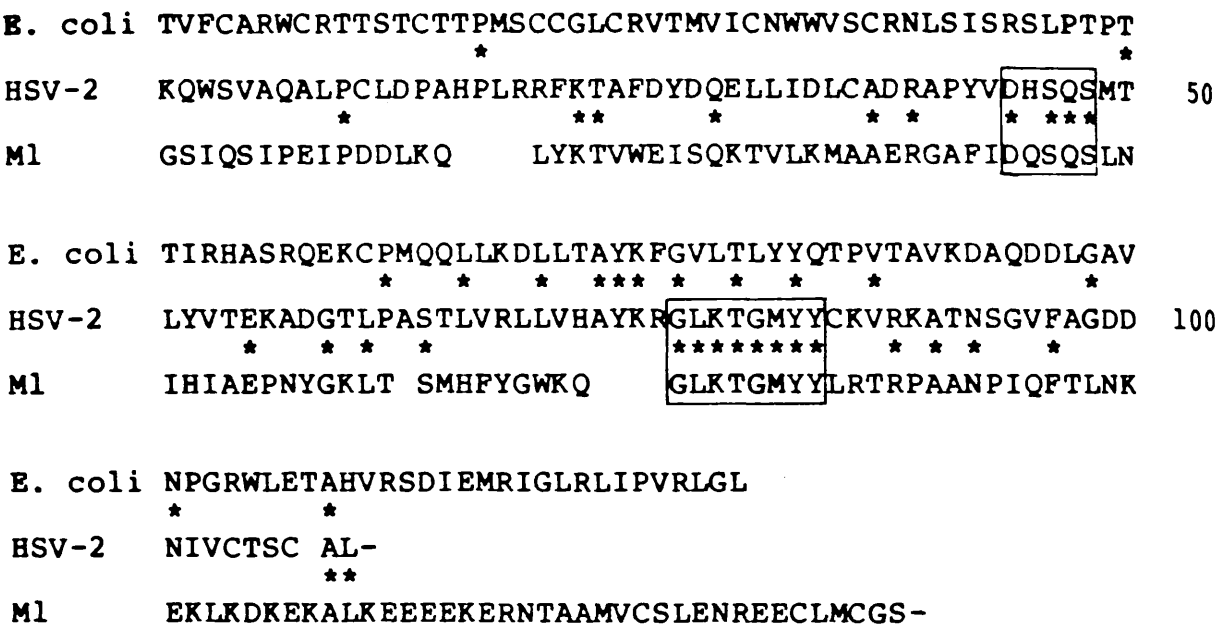


Figure 59. Alignment of the C-terminal amino acid residues of HSV-2 Vmw138 with equivalent regions of the mouse M1 and E. coli B1 proteins. Asterisks denote conserved residues and the HSV-2 sequence is numbered. The alignment of the HSV-2 and E. coli sequences is represented by B in Fig. 57. Blocks of homology are boxed.

protein with a mol. wt. of 40,000 (p40; Fig. 61) which is translated from stored, maternal mRNA following fertilisation of clam oocytes (Rosenthal et al., 1980 and 1983; Evans et al., 1983). Alignment of the sequences indicates that E. coli protein B2 contains two stretches of amino acid sequences not present in viral and clam polypeptides (Fig. 62, at residues 173 and 242). Regions conserved in all three proteins are located between residues 124 to 131, 191 to 193 and 213 to 221 (Fig. 62); two of these regions (residues 191 to 193 and 213 to 221) are also conserved in the EBV 34,000 mol. wt. protein. The overall homology of Vmw38 with clam p40 is greater than that with the B2 protein and five areas common to the HSV-2 and clam sequences are absent in the E. coli sequence (Fig. 62) These areas of homology between Vmw38 and p40 correspond to sequences conserved in the EBV 34,000 mol. wt. protein (Fig. 50). Interestingly, the amino acid sequence between residues 108 to 116 (Fig. 62) in the N-terminal region of the HSV-2 protein is almost completely conserved in the C-terminal portion of the E. coli B2 polypeptide (Fig. 62, residues 321 to 329).

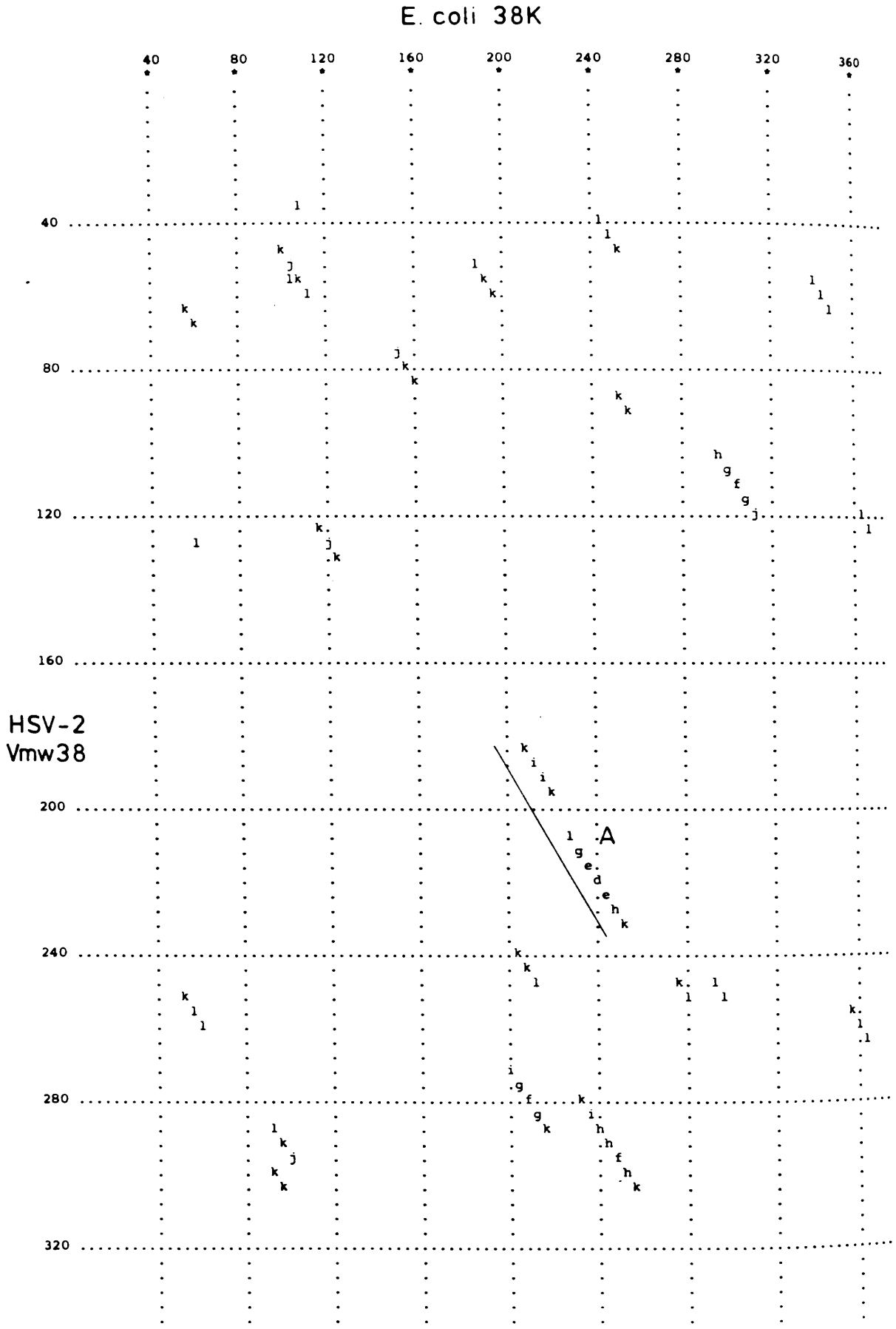

E. coli 38K
40
80
120
160
200
240
280
320
360
HSV-2
Vmw38
A

Figure 60. A CINTHOM plot comparing amino acid sequences of HSV-2 Vmw38 with those of the E. coli B2 protein. The region of highest homology between both proteins is labelled A. Program settings were: range = 12; scale factor = 0.95; minimum value plotted = 26%; compressed x4.

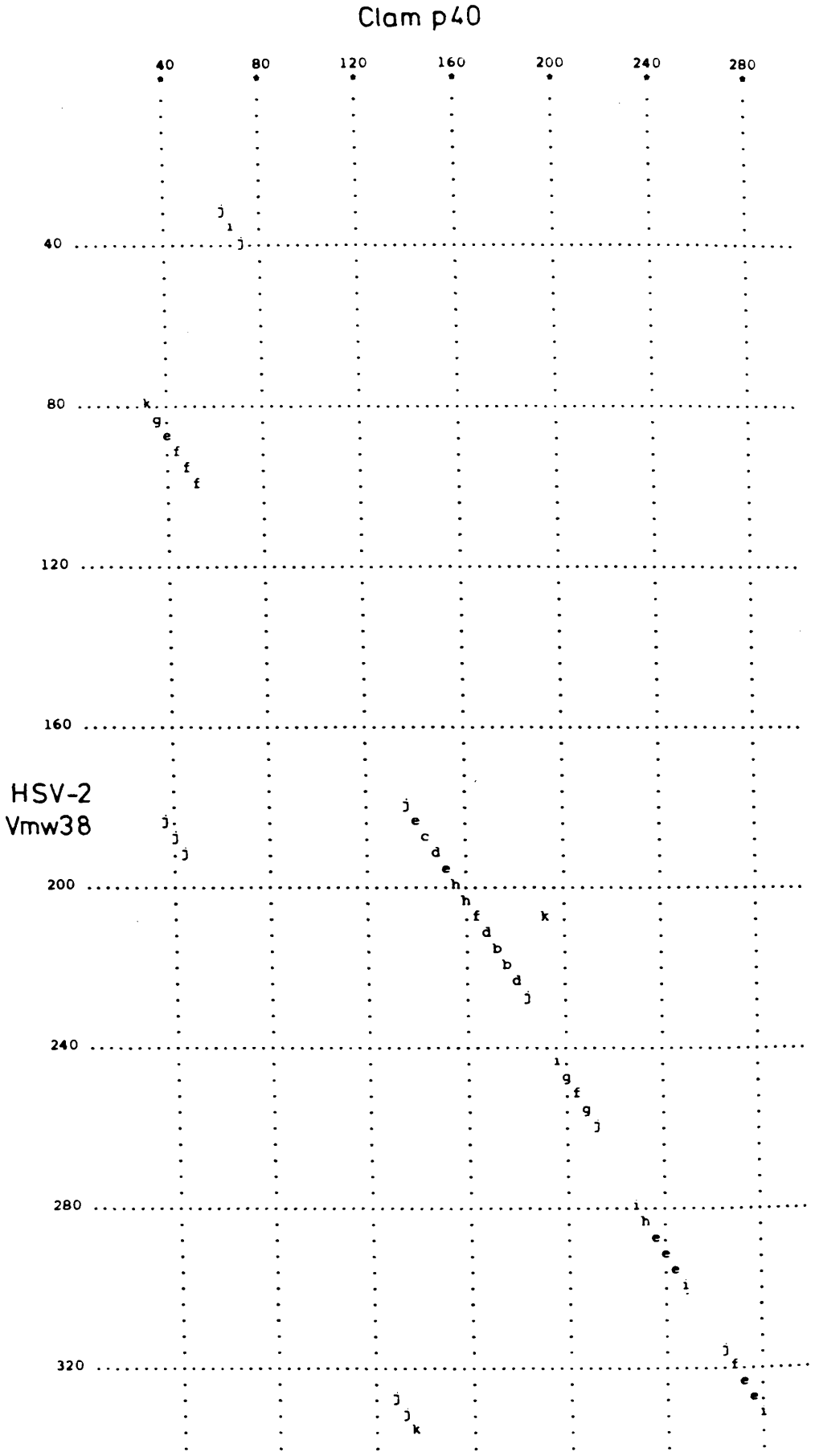

Clam p40
40
80
120
160
200
240
280
320
HSV-2
Vmw38

Figure 61. A CINTHOM plot comparing the amino acid sequences of HSV-2 Vmw38 with those of surf clam p40. Program settings were: range = 12; scale factor = 0.95; minimum value plotted = 28%; compressed x4.


Clam MYKK
HSV-2 MDPAVSPASTDPLDTHASGAGAAPIPVCPTPERYFYTSQCPDINHLRSLS 50
E.coli MAYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK
Clam AEASFWTAEEVDLSKDMAHWESLKKEEKHFISHVLAFFAASDGIVNENLV
HSV-2 ILNRWLETELVFVG DEEDVSKLSEGELGFYRFLFAFLSAADDLVTENL 98
E.coli QLSFFWRPEEVDVSRDRIDYQALPEHEKHIFISNLKYQTLLDSIQGRSPN
Clam ERFSKEVQVTEARCFYGFQIAMENIHSEMYSLLIDTYIKDPQERDFLFNA
HSV-2 GGLSGLFEQKDILHYYVEQECIEVVHSRVYN IIQLVLFHNNDQARRAYV 147
E.coli VALLPLISIPELETWVETWAFSETIHSRSYTHIIRNIVNDPSVVFDDIVT
Clam IETMPCVKEKA DWAMRWINDDSSSY
HSV-2 ARTINHPAIRVKVDWLEARVRECDSV 173
E.coli NEQIQKRAEGISSYYDELIEMTSYWHLLGEGTHTVNGKTVTVSLRELKKK
Clam AERVVAFAAVEGIFFSGSFASIFWLKKRGIMPGLTFSNELISRDEGLHC
HSV-2 PEKFILMILIEGVFFAASFAAIAYLRTNNLLRVTCQSNDLISRDEAVHT 222
E.coli LYLCLMSVNALEAIRLYVSFACSFAFAERELMEGNAKIIRLIARDEALHL
Clam DFAC LMFSHLVNKPSQE RIHQIIDEAVKIEQVFLTEALP
HSV-2 TASCYIYNNYLGDHAKPEAA RVYRLFREAVDIEIGFIRSQAP 264
E.coli TGTQHMLNLLRSGADDPEMAEIAEECKQACYDLFVQAAQQEKDWADYLFR
Clam CR LIGMNCDLMRQYIEFVADRLLLELKCDKLYNKENPFDFMEHISL
HSV-2 TDSSI LSPGALAAIENYVRFSADRLLGLIHMQPLYSAPAP DASFPLSL 312
E.coli DGSMIGLNKDILCQYVEYITNIRMQAVGLDLPFQTRSNPIPWINTWLVSD
Clam EGKTNFFEKRVGEYQKMGVMSGGNTGDSHAFTLDADF-
HSV-2 MSTDKHTNFFECRSTSYAGAVVNDL- 337
E.coli NVQVAPQEVEVSSYLVGQIDSEVDTDDLSNFQL-

Figure 62. Alignment of the amino acid sequences of HSV-2 Vmw38 with those of the surf clam p40 and E. coli B2 proteins. Asterisks denote conserved residues and the HSV-2 sequence is numbered. Regions of higher homology are boxed and dashed lines indicate the N-terminal region of Vmw38 which has homology with residues at the C-terminus of the E. coli polypeptide. Proposed locations for the iron binding site (●) and tyrosine free radical (◆) are indicated.

DISCUSSION.

21. Proposed Functional Domains within Ribonucleotide Reductases.

Sequence comparisons between the proposed subunits of the HSV-2 ribonucleotide reductase with those of eukaryotic and prokaryotic reductases have revealed regions of conserved amino acid residues. DNA hybridisation studies have shown that the nucleotide sequences encoding the mouse M1 subunit have homology with DNA sequences from a number of mammalian species, including humans (Caras et al., 1985). These data suggest that ribonucleotide reductases from a wide variety of species and organisms are evolutionarily related and regions of conserved amino acid residues may represent domains which are necessary for enzyme function.

Studies on the evolution of haemoglobins suggest that functionally critical regions are the most conserved, for example, there is no structural variation in the histidine residues which bind the heme (Perutz et al., 1965). In contrast, amino acids located on the surface of the haemoglobin molecule have evolved ten times more rapidly than those involved in heme binding (Kimura and Ohta, 1974). Other comparisons between E. coli DNA polymerase and T7 DNA polymerase indicate that domains involved in DNA binding are highly conserved and appear to share common structural features (Ollis et al., 1985); areas of low homology between these polymerases are in

regions of less secondary structure which lie on the surfaces.

The longest stretch of homology between the proposed small subunits of viral and cellular reductases has the amino acid sequence LISRDE in Vmw38 (Fig. 62, residues 213 to 218), of which four are either charged (R, D and E) or hydrophilic (S); the charged amino acids are invariant and conserved in all proposed small subunit proteins. In general, charged amino acids, which are usually located on the surfaces of proteins, are the least well conserved residues in related proteins (Creighton, 1983). Highly conserved amino acids are generally those with non-polar groups which are located internally within the protein structure. Thus, retention of charged residues in the proteins with homology to Vmw38 suggests that these amino acids have an important functional role. A second region of lower homology is located upstream from the LISRDE segment between residues 191 to 193 (Fig. 62) in HSV-2 Vmw38. These conserved residues lie within a highly hydrophobic region (Fig. 63) which extends from residue 175 to residue 197 and is present at a similar location in all proposed small subunit polypeptides (data not shown). The role of these conserved areas is unknown, however, the B2 subunit supplies elements which form part of the active centre of the E. coli enzyme (von Dobeln and Reichard, 1976). One possibility is that either the LISRDE sequence or the conserved hydrophobic region contribute to the active centre.

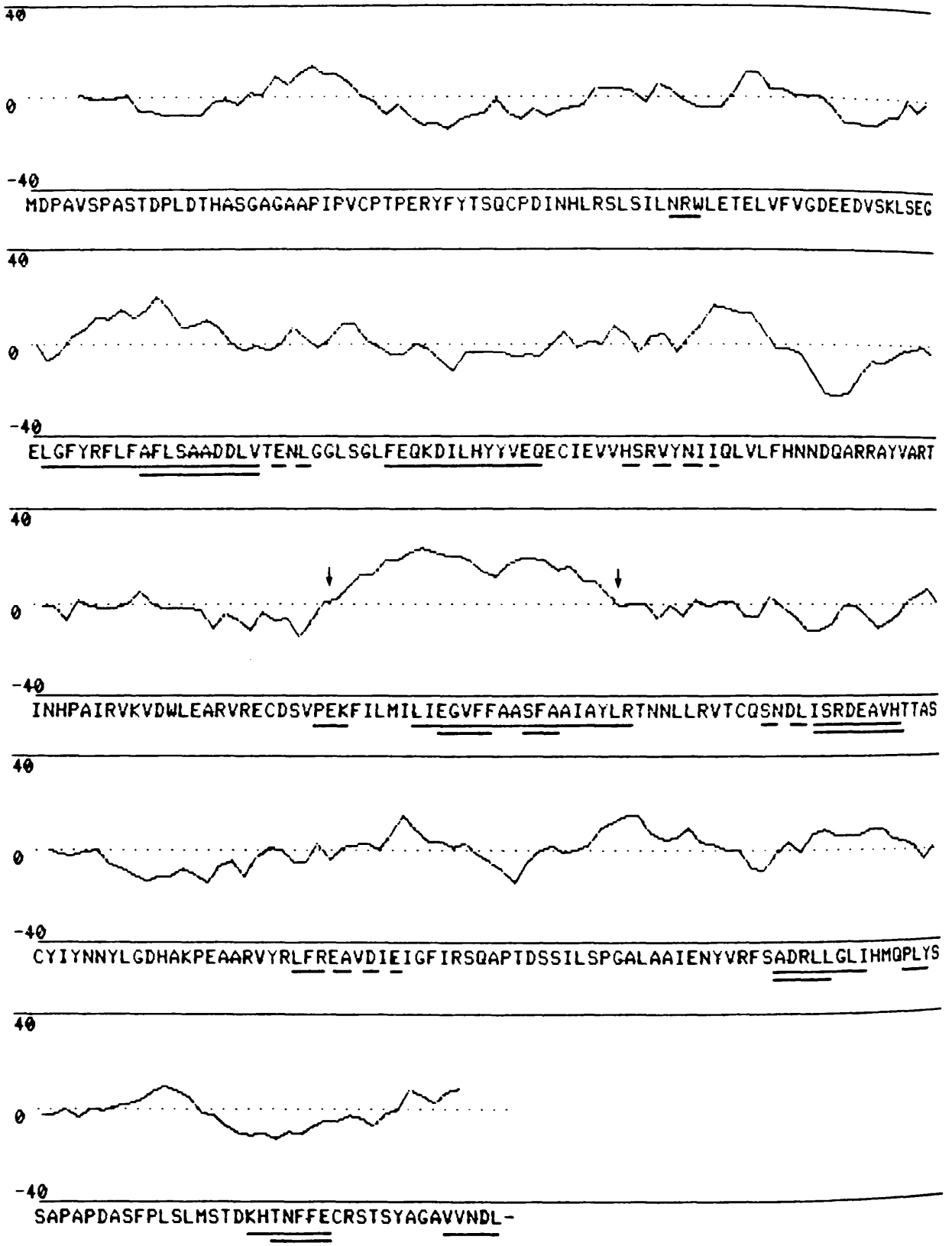

40
0
-40
MDPAVSPASTDPLDTHASGAGAAPIPVCPTPERYFYTSQCPDINHLRSLSILNRWLETELVFVGDEEDVSKLSEG
40
0
-40
ELGFYRFLFAFLSAADDLVTENLGGLSGLFEQKDILHYYVEQECIEVVHSRVYNIIQLVLFHNNDQARRAYVART
40
0
-40
INHPAIRVKVDWLEARVRECDSVPEKFILMILIEGVFFAASFAAIAYLRTNNLLRVTCQSNDLISRDEAVHTTAS
40
0
-40
CYIYNNYLGDHAKPEAARVYRLFREAVDIEIGFIRSQAPTDSSILSPGALAAIENYVRFSADRLLGLIHMQPLYS
40
0
-40
SAPAPDASFPLSLMSTDKHTNFFECRSTSYAGAVVNDL-

Figure 63. Hydropathicity plot of the HSV-2 Vmw38 amino acid sequence. Hydrophobic regions lie between 0 and 40 and hydrophilic regions lie between 0 and -40. The Vmw38 amino acid sequence is shown below the plot. Regions conserved between Vmw38, the EBV 34,000 mol. wt. and surf clam p40 polypeptides are shown by two solid lines. Underlined are areas of homology between Vmw38 and the EBV 34,000 mol. wt. polypeptide and dashed lines locate regions conserved between Vmw38 and surf clam p40. Arrows delimit the hydrophobic region conserved in the equivalent areas of EBV, VZV, surf clam and E. coli proteins. Program settings were: group length = 9; shift = 1; spacing of plotted points = 10.

In addition to supplying a portion of the enzyme active centre, each B2 subunit of E. coli ribonucleotide reductase contains a tyrosine free radical and two atoms of iron (see Introduction, Page 52). Iron is required to generate and stabilise the free radical by some unknown continued interaction, and removal of the iron leads to a loss of the free radical. Tyrosine free radicals have also been detected in the PRV (Lankinen et al., 1982) and mammalian reductases (Thelander et al., 1985), suggesting that the free radical may be ubiquitous in reductases which consist of two non-identical subunits. Locations for the iron binding site and tyrosine free radical within the small subunit polypeptide have been proposed from sequence comparisons (Sjoberg et al., 1985; Standart et al., 1985); the positions of these sites corresponds to the conserved region between residues 121 to 131 (Fig. 62). The iron binding site has been positioned on a histidine residue (Fig. 62, residue 134) and a further four amino acids downstream, a conserved tyrosine residue has been proposed as the location of the free radical. This region is not highly conserved in EBV (Fig. 50), however, both histidine and tyrosine residues are retained.

The amino acid sequence (KDILHYYVE) in the N-terminal portion of Vmw38 between residues 108 to 116 (Fig. 62) is almost perfectly conserved in the C-terminal region of the E. coli B2 protein (KDILCQYVE; Fig. 62, residues 272 to 280). These amino acids are poorly conserved in the equivalent region of the clam sequence,

however, most of these residues are retained in EBV (Figs. 50 and 62); the function of these amino acids is unknown. The presence of homologous sequences at different locations in Vmw38 and the B2 polypeptide is unlikely to have arisen by chance and may be a consequence of evolutionary pressure. Firstly, these N- and C-terminal positions in the Vmw38 and B2 proteins may be structurally related and the similarity in the sequences could be a result of convergent evolution. A second but less likely possibility is that the proteins have acquired these sequences following their divergence from a common ancestral gene. Thirdly, the different locations of homologous residues may be due to movement of protein domains. Recombination of exons into introns has been proposed as a mechanism for 'shuffling' domains and thereby increasing genetic complexity (Blake, 1978; Gilbert, 1978). By implication, precursor polypeptides to the B2 and Vmw38 proteins would have contained introns which have subsequently been deleted; it has been suggested that splicing did occur in primordial systems (Doolittle, 1978).

Although there is homology between the viral and cellular components of ribonucleotide reductases, viral and cellular subunits appear not to bind to a significant extent. In cells infected at the npt with the ribonucleotide reductase-negative mutant, 17tsVP1207, which fails to produce a complex between Vmw136 and Vmw38 (Page 158), antiserum directed against the seven C-terminal amino acids of Vmw38 does not immunoprecipitate a host protein corresponding to the size of M1 (Frame et al., 1985).

SECTION E.

22. Conserved DNA Sequences Downstream from mRNA 3' Termini.

Efficient polyadenylation requires the sequence AATAAA, a highly conserved poly A signal which is present upstream from the 3' termini of higher eukaryotic polyadenylated mRNAs (see Introduction, Page 87). However, this DNA sequence occurs also in regions which do not contain poly A sites, hence, additional sequences must be involved in determining the location of poly A sites. Comparisons between HSV-2 and HSV-1 DNA sequences at the poly A sites of the 3' co-terminal mRNAs revealed three blocks of homology downstream from the poly A sites (see Page 140). DNA sequences flanking the poly A sites of four other equivalent HSV-2 and HSV-1 genes also contained regions of conserved G+T-rich residues at approximately 30bp downstream from the poly A signal (Fig. 64). In addition, G+T-rich motifs were present in the flanking sequences of a number of HSV-1 mRNA 3' termini for which no equivalent HSV-2 DNA sequence data are available (Fig. 64). The comparative analysis was extended to include a large number of mammalian and eukaryotic viral genes and this survey indicated that greater than 60% of 3' terminal regions possessed sequences similar to those conserved downstream from HSV mRNA 3' termini (Fig. 65). From a compilation of these data, the consensus sequence YGTGTTYY was derived (Table 2); the preferred distance for sequences

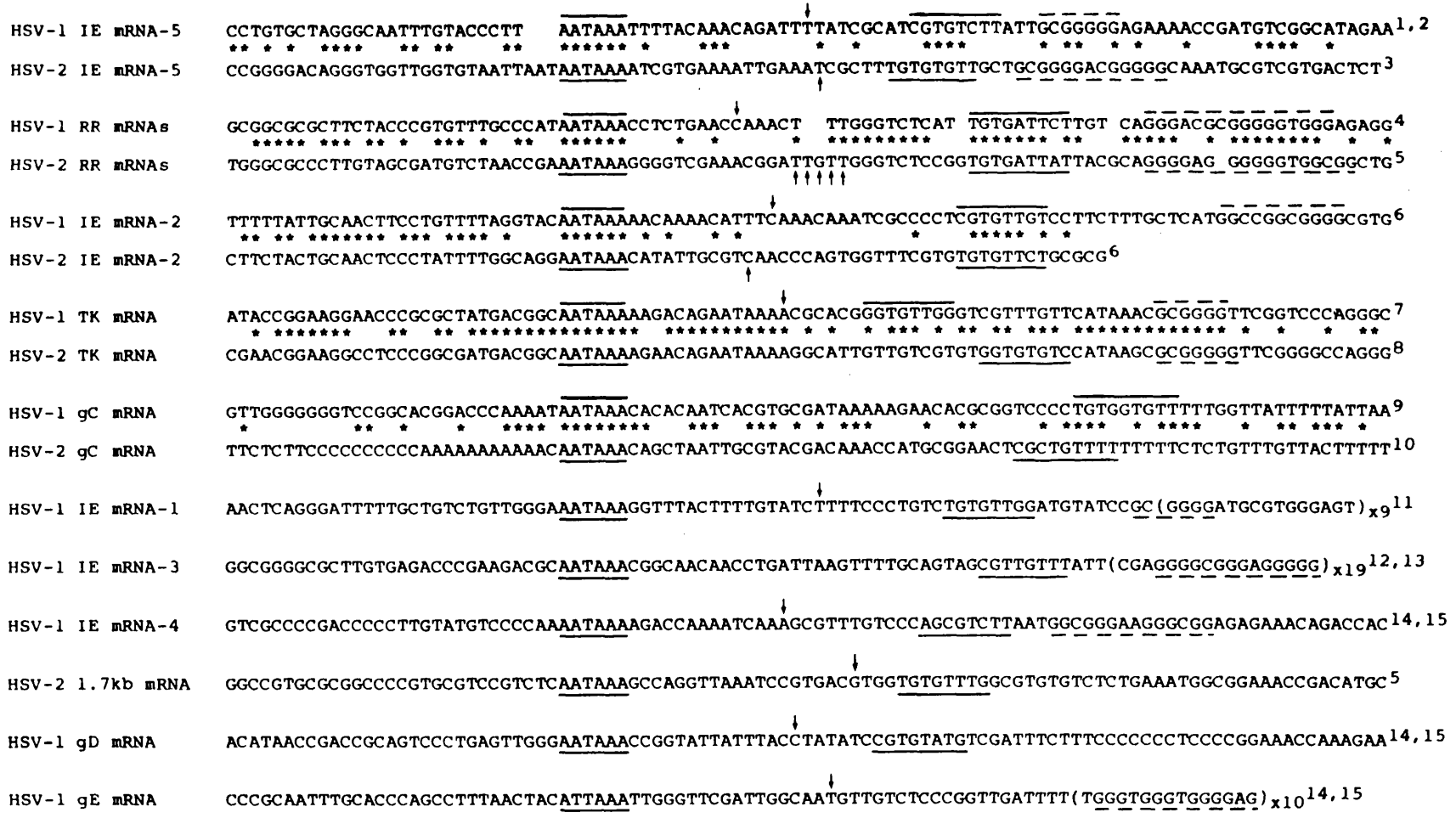

HSV-1 IE mRNA-5
CCTGTGCTAGGGCAATTTGTACCCTT AATAAATTTTACAAACAGATTTTATCGCATCGTGTCTTATTGCGGGGGAGAAAACCGATGTCGGCATAGAA 1,2
HSV-2 IE mRNA-5
CCGGGGACAGGGTGGTTGGTGTAATTAATAATAAAATCGTGAAAATTGAAATCGCTTTGTGTGTTGCTGCGGGGACGGGGGCAAATGCGTCGTGACTCT 3
HSV-1 RR mRNAs
GCGGCGCGCTTCTACCCGTGTTTGCCCATAATAAACCTCTGAACCAAACT TTGGGTCTCAT TGTGATTCTTGT CAGGGACGCGGGGGTGGGAGAGG 4
HSV-2 RR mRNAs
TGGGCGCCCTTGTAGCGATGTCTAACCGAAATAAAGGGGTCGAAACGGATTGTTGGGTCTCCGGTGTGATTATTACGCAGGGGAG GGGGGTGGCGGCTG 5
HSV-1 IE mRNA-2
TTTTTATTGCAACTTCCTGTTTTAGGTACAATAAAAACAAAACATTTCAAACAAATCGCCCCTCGTGTTGTCCTTCTTTGCTCATGGCCGGCGGGGCGTG 6
HSV-2 IE mRNA-2
CTTCTACTGCAACTCCCTATTTTGGCAGGAATAAACATATTGCGTCAACCCAGTGGTTTCGTGTGTGTTCTGCGCG 6
HSV-1 TK mRNA
ATACCGGAAGGAACCCGCGCTATGACGGCAATAAAAAGACAGAATAAAACGCACGGGTGTTGGGTCGTTTGTTCATAAACGCGGGGTTCGGTCCCAGGGC 7
HSV-2 TK mRNA
CGAACGGAAGGCCTCCCGGCGATGACGGCAATAAAAGAACAGAATAAAAGGCATTGTTGTCGTGTGGTGTGTCCATAAGCGCGGGGGTTCGGGGCCAGGG 8
HSV-1 gC mRNA
GTTGGGGGGGTCCGGCACGGACCCAAAATAATAAACACACAATCACGTGCGATAAAAAGAACACGCGGTCCCCTGTGGTGTTTTTGGTTATTTTTATTAA 9
HSV-2 gC mRNA
TTCTCTTCCCCCCCCCCAAAAAAAAAAAACAATAAACAGCTAATTGCGTACGACAAACCATGCGGAACTCGCTGTTTTTTTTTCTCTGTTTGTTACTTTTT 10
HSV-1 IE mRNA-1
AACTCAGGGATTTTTGCTGTCTGTTGGGAAATAAAGGTTTACTTTTGTATCTTTTCCCTGTCTGTGTTGGATGTATCCGC(GGGGATGCGTGGGAGT)x9 11
HSV-1 IE mRNA-3
GGCGGGGCGCTTGTGAGACCCGAAGACGCAATAAACGGCAACAACCTGATTAAGTTTTGCAGTAGCGTTGTTTATT(CGAGGGGCGGGAGGGGG)x19 12,13
HSV-1 IE mRNA-4
GTCGCCCCGACCCCCTTGTATGTCCCCAAAATAAAAGACCAAAATCAAAGCGTTTGTCCCAGCGTCTTAATGGCGGGAAGGGCGGAGAGAAACAGACCAC 14,15
HSV-2 1.7kb mRNA
GGCCGTGCGCGGCCCCGTGCGTCCGTCTCAATAAAGCCAGGTTAAATCCGTGACGTGGTGTGTTTGGCGTGTGTCTCTGAAATGGCGGAAACCGACATGC 5
HSV-1 gD mRNA
ACATAACCGACCGCAGTCCCTGAGTTGGGAATAAACCGGTATTATTTACCTATATCCGTGTATGTCGATTTCTTTCCCCCCCTCCCCGGAAACCAAAGAA 14,15
HSV-1 gE mRNA
CCCGCAATTTGCACCCAGCCTTTAACTACATTAAATTGGGTTCGATTGGCAATGTTGTCTCCCGGTTGATTTT(TGGGTGGGTGGGGAG)x10 14,15

Figure 64. DNA sequence comparison at the mRNA 3' termini of a number of HSV-2 and HSV-1 genes. The genes for which DNA sequence data are available for both serotypes are compared. Asterisks denote nucleotides which are conserved and the locations of poly A sites are arrowed. The AATAAA signals and sequences which fit the YGTGTTYY consensus are indicated by a solid line. Dashed lines locate the G-rich sequences present downstream from the YGTGTTYY consensus. Tandem repetitions in the 3' flanking regions of HSV mRNAs are bracketed and the number of copies of each repeat indicated. DNA sequence data were taken from the following references:

1. Watson and Vande Woude, 1982; 2. Rixon and McGeoch, 1984;

3. Whitton and Clements, 1984b; 4. McLauchlan and Clements, 1982;

5. this study; 6. Whitton et al., 1983; 7. McKnight, 1980;

8. Swain and Galloway, 1983; 9. Frink et al., 1983;

10. Swain et al., 1985; 11. Rixon et al., 1984;

12. Davison and Wilkie, 1981; 13. Rixon et al., 1982;

14. McGeoch et al., 1985; 15. Rixon and McGeoch, 1985.

A

Mammalian 3' termini

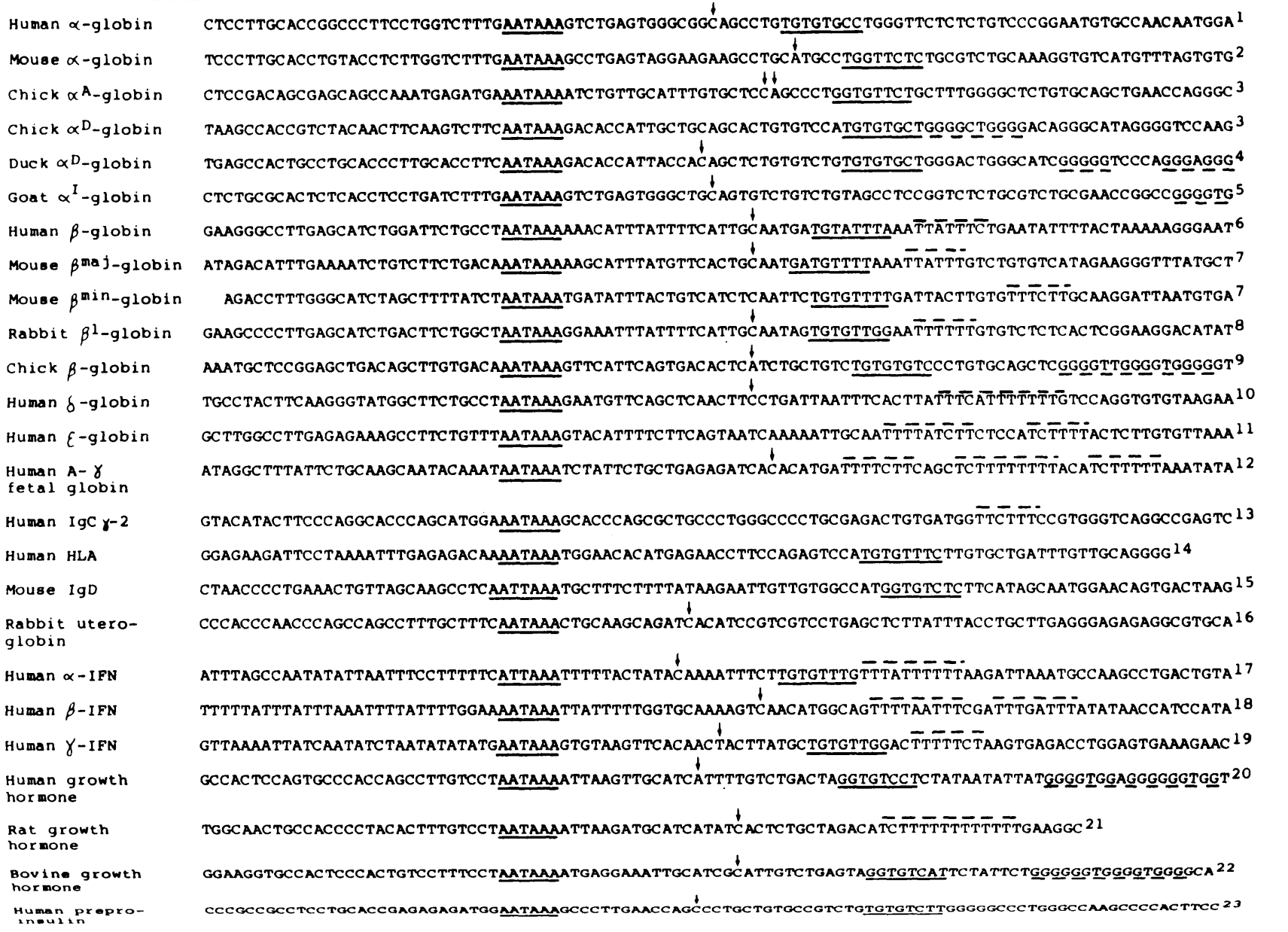


Figure 65. Compilation of DNA sequences at the 3' termini of mammalian and eukaryotic viral mRNAs. Sequences are aligned using the AATAAA signals which are underlined. Locations of poly A sites are arrowed and sequences which resemble the YGTGTTYY consensus are underlined. G-rich tracts beyond mRNA 3' termini are indicated by a dashed line below the sequence and T-rich tracts by a dashed line above the sequence.

A) IFN is the abbreviation for interferon. DNA sequence data were taken from the following references:

1. Liebhaber et al., 1980; 2. Nishioka and Leder, 1979;
3. Dodgson et al., 1983; 4. Erbil and Niessing, 1983;
5. Schon et al., 1982; 6. Lawn et al., 1980;
7. Konkel et al., 1979; 8. van Ooyen et al., 1979;
9. Dolan et al., 1983; 10. Spritz et al., 1980;
11. Baralle et al., 1980; 12. Shen et al., 1981;
13. Ellison et al., 1982; 14. Malissen et al., 1982;
15. Cheng et al., 1982; 16. Suske et al., 1983;
17. Ullrich et al., 1982; 18. Ohno and Taniguchi, 1981;
19. Taya et al., 1982; 20. Denoto et al., 1981;
21. Page et al., 1981; 22. Woychik et al., 1982;
23. Bell et al., 1980.

B

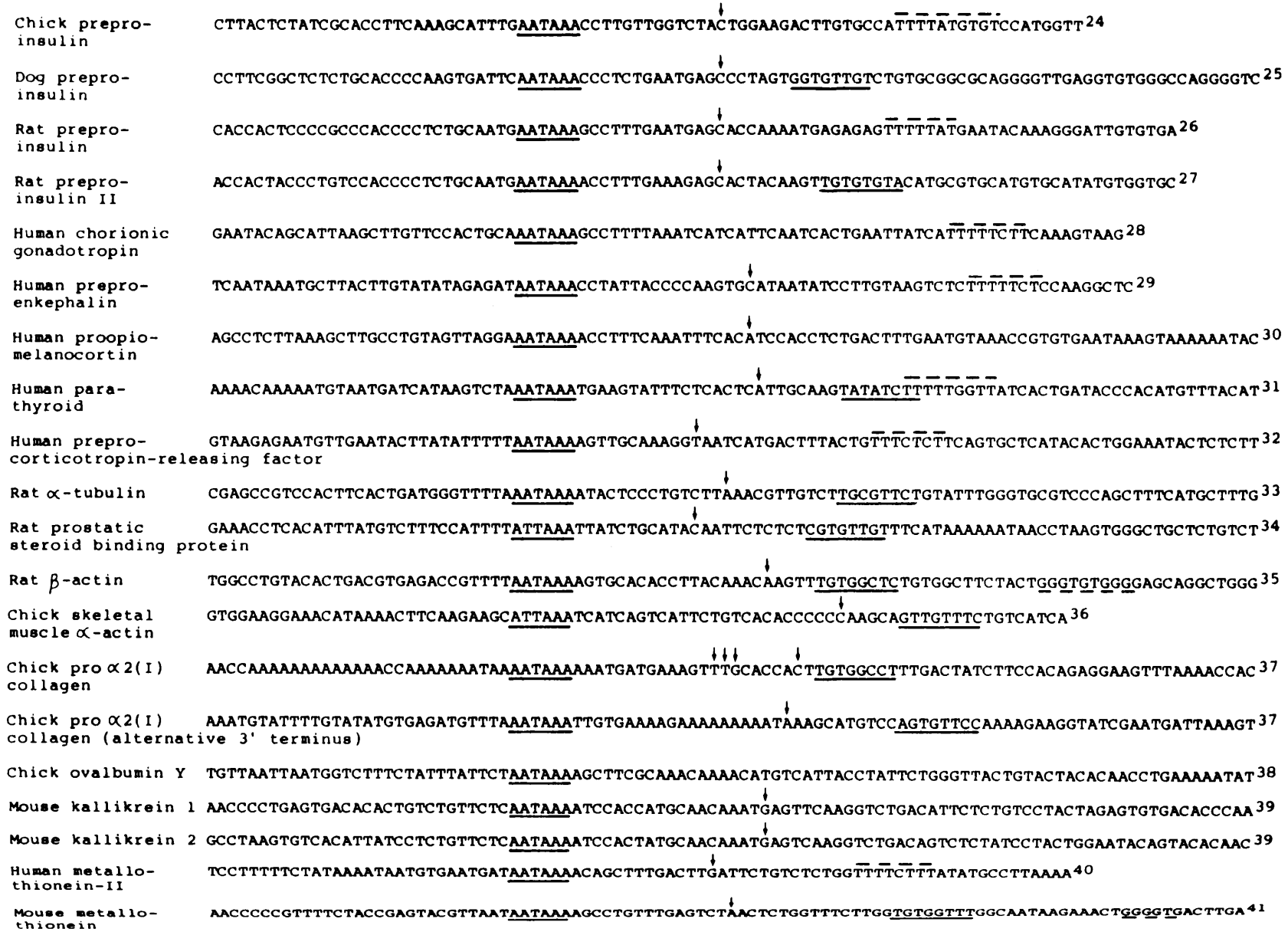

Chick prepro-insulin CTTACTCTATCGCACCTTCAAAGCATTTGAATAAACCTTGTTGGTCTACTGGAAGACTTGTGCCATTTTATGTGTCCATGGTT 24
Dog prepro-insulin CCTTCGGCTCTCTGCACCCCAAGTGATTCAATAAACCCTCTGAATGAGCCCTAGTGGTGTTGTCTGTGCGGCGCAGGGGTTGAGGTGTGGGCCAGGGGTC 25
Rat prepro-insulin CACCACTCCCCGCCCACCCCTCTGCAATGAATAAAGCCTTTGAATGAGCACCAAAATGAGAGAGTTTTTATGAATACAAAGGGATTGTGTGA 26
Rat prepro-insulin II ACCACTACCCTGTCCACCCCTCTGCAATGAATAAAACCTTTGAAAGAGCACTACAAGTTGTGTGTACATGCGTGCATGTGCATATGTGGTGC 27
Human chorionic gonadotropin GAATACAGCATTAAGCTTGTTCCACTGCAAATAAAGCCTTTTAAATCATCATTCAATCACTGAATTATCATTTTTCTTCAAAGTAAG 28
Human prepro-enkephalin TCAATAAATGCTTACTTGTATATAGAGATAATAAACCTATTACCCCAAGTGCATAATATCCTTGTAAGTCTCTTTTTCTCCAAGGCTC 29
Human proopio-melanocortin AGCCTCTTAAAGCTTGCCTGTAGTTAGGAAATAAAACCTTTCAAATTTCACATCCACCTCTGACTTTGAATGTAAACCGTGTGAATAAAGTAAAAAATAC 30
Human para-thyroid AAAACAAAAATGTAATGATCATAAGTCTAAATAAATGAAGTATTTCTCACTCATTGCAAGTATATCTTTTTGGTTATCACTGATACCCACATGTTTACAT 31
Human prepro-corticotropin-releasing factor GTAAGAGAATGTTGAATACTTATATTTTTAATAAAAGTTGCAAAGGTAATCATGACTTTACTGTTTCTCTTCAGTGCTCATACACTGGAAATACTCTCTT 32
Rat α-tubulin CGAGCCGTCCACTTCACTGATGGGTTTTAAATAAAATACTCCCTGTCTTAAACGTTGTCTTGCGTTCTGTATTTGGGTGCGTCCCAGCTTTCATGCTTTG 33
Rat prostatic steroid binding protein GAAACCTCACATTTATGTCTTTCCATTTTATTAAATTATCTGCATACAATTCTCTCTCGTGTTGTTTCATAAAAAATAACCTAAGTGGGCTGCTCTGTCT 34
Rat β-actin TGGCCTGTACACTGACGTGAGACCGTTTTAATAAAAGTGCACACCTTACAAACAAGTTTGTGGCTCTGTGGCTTCTACTGGGTGTGGGGAGCAGGCTGGG 35
Chick skeletal muscle α-actin GTGGAAGGAAACATAAAACTTCAAGAAGCATTAAATCATCAGTCATTCTGTCACACCCCCCAAGCAGTTGTTTCTGTCATCA 36
Chick pro α2(I) collagen AACCAAAAAAAAAAAAAACCAAAAAAAATAAAATAAAAAATGATGAAAGTTTGCACCACTTGTGGCCTTTGACTATCTTCCACAGAGGAAGTTTAAAACCAC 37
Chick pro α2(I) collagen (alternative 3' terminus) AAATGTATTTTGTATATGTGAGATGTTTAAATAAATTGTGAAAAGAAAAAAAAAATAAAGCATGTCCAGTGTTCCAAAAGAAGGTATCGAATGATTAAAGT 37
Chick ovalbumin Y TGTTAATTAATGGTCTTTCTATTTATTCTAATAAAAGCTTCGCAAACAAAACATGTCATTACCTATTCTGGGTTACTGTACTACACAACCTGAAAAATAT 38
Mouse kallikrein 1 AACCCCTGAGTGACACACTGTCTGTTCTCAATAAAATCCACCATGCAACAAATGAGTTCAAGGTCTGACATTCTCTGTCCTACTAGAGTGTGACACCCAA 39
Mouse kallikrein 2 GCCTAAGTGTCACATTATCCTCTGTTCTCAATAAAATCCACTATGCAACAAATGAGTCAAGGTCTGACAGTCTCTATCCTACTGGAATACAGTACACAAC 39
Human metallo-thionein-II TCCTTTTTCTATAAAATAATGTGAATGATAATAAAACAGCTTTGACTTGATTCTGTCTCTGGTTTTCTTTATATGCCTTAAAA 40
Mouse metallo-thionein AACCCCCGTTTTCTACCGAGTACGTTAATAATAAAAGCCTGTTTGAGTCTAACTCTGGTTTCTTGGTGTGGTTTGGCAATAAGAAACTGGGGTGACTTGA 41

Figure 65.

B) DNA sequence data were taken from the following references:

24. Perler et al., 1980; 25. Kwok et al., 1983;

26. Cordell et al., 1979; 27. Lomedico et al., 1979;

28. Fiddes and Goodman, 1981; 29. Noda et al., 1982;

30. Chang et al., 1980; 31. Vasicek et al., 1983;

32. Shibahara et al., 1983; 33. Lemischka and Sharp, 1982;

34. Hurst and Parker, 1983; 35. Nudel et al., 1983;

36. Fornwald et al., 1982; 37. Aho et al., 1983;

38. Heilig et al., 1982; 39. Mason et al., 1983;

40. Karin and Richards, 1982; 41. Glanville et al., 1981.

C


Adenovirus 3' termini
Ad2 E1a
Ad7 E1a
Ad12 E1a
Ad2 E1b
Ad7 E1b
Ad12 E1b
Ad2 E2a
Ad2 E2b
Ad7 E2b
Ad12 E2b
Ad2 E3
Ad2 E4
Ad2 L2
Ad2 L3
Ad2 L4
Ad2 L5
Papovavirus 3' termini
SV40 late
Polyoma late
BK late
SV40 early
Polyoma early
BK early
BPV-1 (E mRNAs)
HPV-1a (E mRNAs)
BPV-1 (L mRNAs)
HPV-1a (L mRNAs)

Figure 65.

C) The key to abbreviations is: BPV-1, bovine papillomavirus type 1; HPV-1, human papillomavirus type 1. DNA sequence data were taken from the following references:

42. Gingeras et al., 1982; 43. Dijkema et al., 1982;
44. Sugisaki et al., 1980; 45. Broker, 1980;
46. Akusjarvi et al., 1981; 47. Herisse and Galibert, 1981;
48. Akusjarvi and Persson, 1981b; 49. Herisse et al., 1980;
50. Fiers et al., 1978; 51. Soeda et al., 1980;
52. Seif et al., 1979; 53. Chen et al., 1982;
54. Danos et al., 1982.

D

RNA Tumour Viruses

A-MuLV	GACTGAGTCGCCCGGGTACCCGTGTATCCAATAAACCCTCTTGCAGTTGCATCCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGCGT	55
AMV (MC29)	ATGTTGTATTTAAGTGCCTAGCTCGTAACAATAAACGCCATTTTACCATCCACCACATTGGTCTGCACCTGGGTAGATGGACAGACCGTTGAGTCCCTAA	56
SSV	GACTGAGTCGCCCGGGTACCTGTGTGTTCAATAAAACCTCTTGCTATTTGCATCCGAAGCCGTGGTCTCGTTGTTCCTTGGGAGGGTCTCTCCTAACTGA	57
GA-FeSV	GACTTGACCGCCCCGGGTACCCGTGTACGAATAAACCTCTTGCTGTTTGCATCTGACTCGTGGTCTCGGTGTTCCGTGGGTACGGGGTCTCATCGCCGAG	58
MSV (strain 124)	GACTGCGTCGCCCGGGTACCCGTATTCCCAATAAAGCCTCTTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGAT	59
Avian endogenous provirus ev-2	TACGATTATATAAGCTGTTGCCACCATCAAATAAACGCCATTTTACCATTCACCACATTGGTGTGCACCTGGGTAGATGGACAGACCGTTGAGTCCCTAA	60
BaEV	ATTCCTTAGAGGGAGGGCCCTGGTGCACCAGTAAACGACTTTCTGCCGAAATTTGTGTGTTGGGTGGTTCCTCGCGCCGACTCTCAAACCCTAAGGAACA	61
SNV	TGTGACGTGCGGCCCAGATTCGAATCTGTAATAAAACTTTTTCTTCTGAATCCTCAGATTGGCAGTGAGAGGAGATTTTGTTCGTGGTGTTGGCTGGCCT	62
GaLV	ACTGAGTCGCCCGGGTACCCGTGTGTACGAATAAAACCTCTTGCTATTTGCATCCGAAGTCGTGGACTCGTTGTTCCTTGGGAGGGTCTTTCCCAACTGA	63
ASV	ATATTGTATTTAAGTGCCTAGCTCGATACAATAAACGCCATTTGACCATTCACCACATTGGTGTGTGCACCTGGGTTCATGGCTGGACCGTCGATTCCCT	64
MTV	CCATAATATAAAAGAGTGCTGATTTTTTGAGTAAACTTGCAACAGTCCTAACATTCTTCTCTCGTGTGTTTGTGTCTGTTCGCCATCCCGTCTCCGCTCG	65

Figure 65.

D) The key to abbreviations is: A-MuLV, Abelson murine leukaemia virus; AMV, avian myelocytomatosis virus; SSV, simian sarcoma virus; GA-FeSV, Gardner-Arnstein feline sarcoma virus; MSV, murine sarcoma virus; BaEV, baboon endogenous virus; SNV, spleen necrosis virus; GaLV, gibbon ape leukaemia virus; ASV, avian sarcoma virus; MTV, mammary tumor virus.

DNA sequence data were taken from the following references:

55. Reddy et al., 1983a; 56. Reddy et al., 1983b; 57. Devare et al., 1983; 58. Hampe et al., 1983; 59. Van Beveren et al., 1981; 60. Scholl et al., 1983; 61. Tamura et al., 1981; 62. Shimotohno et al., 1980; 63. Lovinger and Schochetman, 1979; 64. Czernilofsky et al., 1980; 65. Donehower et al., 1981.

A	8	6	0	3	3	1	4	9
G	15	57	5	65	11	9	9	13
C	12	1	1	0	0	15	18	14
T	36	7	65	3	57	46	40	35
	Y	G	T	G	T	T	Y	Y
	(67%)	(80%)	(91%)	(91%)	(80%)	(65%)	(82%)	(69%)

Table 2. Derivation of the YGTGTTYY consensus from sequences located 24bp to 38bp downstream from the AATAAA signal. This represents an analysis of 71 of those sequences listed in Figs. 64 and 65 which contain the signal. Numbers indicate the frequency of occurrence for each nucleotide and the percentage conservation for each residue within the consensus is given.

which fit this consensus is 24 to 35nuc beyond the poly A signal. A second conserved segment in certain HSV genes is a G-rich sequence downstream from the YGTGTTYY motif (Fig. 64). A number of eukaryotic genes contain similar tracts of G-rich residues, however, in other 3' terminal regions a T-rich segment is present at a location equivalent to the position of these G-rich sequences (Fig. 65).

23. Functional Analysis of the YGTGTTYY Sequence.

To test the effect of removing sequences downstream from the poly A site, plasmid constructions were made using the CAT gene. An HSV-2 IE promoter fragment was inserted upstream from the CAT gene giving plasmid pLW1 (see Methods, Page 99, Fig. 10), and a 'terminator' fragment from the same IE gene was inserted downstream to give plasmid pTER5.

The HSV-2 IE promoter was a 210bp Sma I/Sau 3A fragment from IE gene-4/-5 (see Fig. 3 for locations of IE mRNAs-4 and -5) which comprised 91bp of 5' flanking sequence and 119bp of leader sequence. This promoter fragment contains all the sequences required for initiation of transcription and the location of the mRNA 5' terminus is identical to that determined for HSV-2 IE mRNAs-4/-5 synthesised in vivo (Gaffney et al., 1985). Plasmid pLW1 has no ATG in the HSV-2 leader sequence; the first initiation codon in the mRNA is the CAT gene ATG.

The 'terminator' was a 100bp Sma I/Xba I fragment from the 3' terminus of HSV-2 IE gene-5. The 3' flanking sequences in this fragment extend 48bp beyond the poly A

site. The 'terminator' fragment contains an AATAAA signal, and located 8nuc downstream from the poly A site is the sequence TGTGTTGC which fits the YGTGTTYY consensus except in one position (Fig. 64). A G-rich segment (GGGGACGGGGG) is located a further 4nuc downstream from the YGTGTTYY sequence. Comparison of HSV-2 and HSV-1 DNA sequences from the 3' terminal portions of IE gene-5 (Fig. 64) reveals conservation of the AATAAA signal together with two blocks of homology downstream from the poly A site. The conserved block between positions 62 to 65 (Fig. 64) forms part of the sequence corresponding to the YGTGTTYY consensus.

a) Construction of 'terminator' plasmids.

Fig. 66 details the construction of plasmids used for short-term expression assays. The 'terminator' fragment was blunt-ended (see Methods, Page 112) and initially cloned into the Hinc II site of pUC9 to give plasmid pTER3. A Hind III/Xmn I fragment containing the 'terminator' was then ligated to a Hind III/Xmn I fragment of pLW1 containing the promoter/CAT modules to give plasmid pTER5. In pTER5, the 'terminator' is orientated with the AATAAA signal on the mRNA sense strand.

A series of deletions was made in plasmid pTER3 from a Bam HI site located in the multiple cloning site downstream from the 'terminator' (Fig. 66). These deletions removed sequences from the 3' portion of the 'terminator' and extended for different lengths through the YGTGTTYY

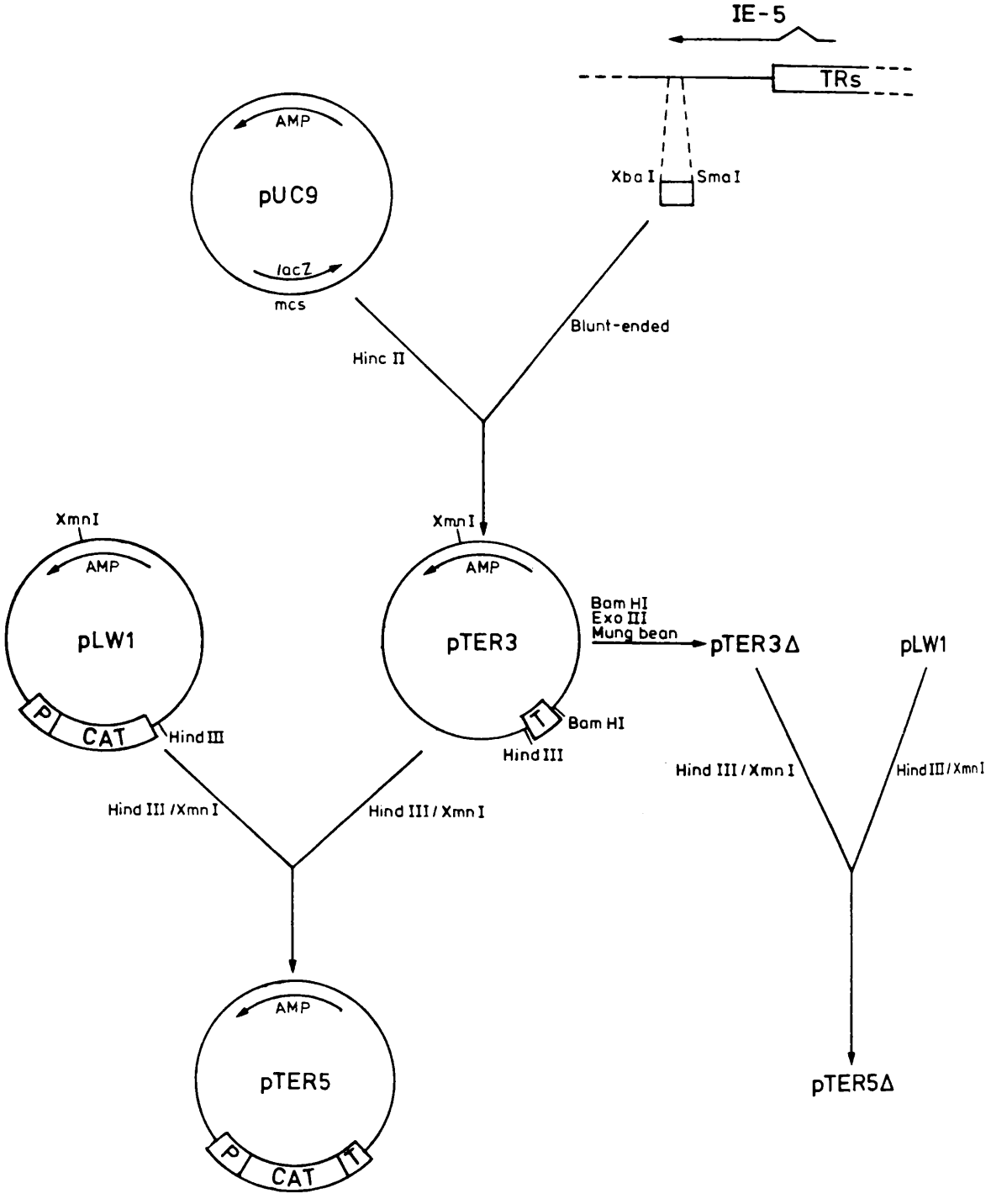

IE-5
TRs
AMP
pUC9
lacZ
mcs
Xba I
Sma I
Blunt-ended
Hinc II
Xmn I
AMP
pLW1
P
CAT
Hind III
Xmn I
AMP
pTER3
T
Bam HI
Hind III
Bam HI
Exo III
Mung bean
pTER3Δ
pLW1
Hind III / Xmn I
Hind III / Xmn I
Hind III / Xmn I
Hind III / Xmn I
AMP
pTER5
P
CAT
T
pTER5Δ

Figure 66. Construction of pTER plasmids. The Sma I/Xba I fragment (see Fig. 67) was derived from the HSV-2 plasmid pBam b' (kindly provided by J.L. Whitton) and blunt-ended using large fragment DNA polymerase I with all four dNTPs (see Methods, Page 112). The multiple cloning site of pUC9 (see Fig. 10) is designated mcs. The locations of restriction enzyme sites used for cloning purposes and deletion mutant analysis are indicated on plasmids. Deletions were produced in pTER3 as described in Methods, Page 110.

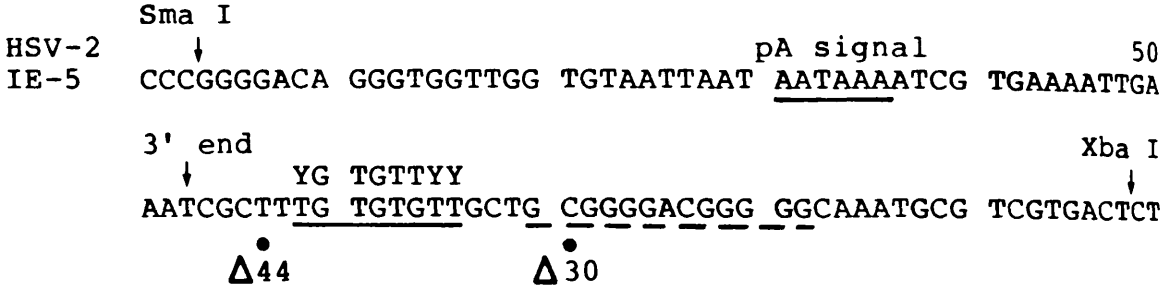


Figure 67. Nucleotide sequence at the 3' terminal region of HSV-2 IE mRNA-5. The AATAAA signal and sequences which fit the YGTGTTYY consensus are underlined; G-rich sequences are indicated by a dashed line. The location of the mRNA 3' terminus is arrowed and the positions of deletion end-points are indicated. Sma I and Xba I restriction enzyme sites, used for cloning, are also shown.

	1	2	3	4	poly A site	YGTGTTYY
pTER5	100(4.01)	100	100	100	+	+
pTER5 Δ30	35(1.41)	34	32	39	+	+
pTER5 Δ44	3(0.12)	4	1	1	+	-
pTER6	1.5(0.025)	2	0.6	1	inverted	
pLW1	0.4(0.013)	U	U	U	-	-
pβTER1	110(4.45)				+ β-globin	+

Table 3. CAT activities obtained with various plasmid constructions are shown for four separate transfection experiments. Activities are expressed as a percentage of the pTER5 value; U indicates undetectable CAT activity. The figures in brackets from experiment 1 are typical CAT values expressed as nmoles converted chloramphenicol/ μg protein/ h. Plasmids pTERΔ30 and pTER5Δ44 contain deletions in the HSV 'terminator' fragment; pTER6 has an inverted 'terminator' ; pLW1 has no 'terminator' fragment; pβTER1 contains a 1250bp rabbit β-globin fragment. The presence (+) or absence (-) of the poly A site and YGTGTTYY signal in each plasmid is indicated.

homology towards the AATAAA sequence to give plasmid series pTER3Δ. Hind III/Xmn I fragments containing these 'terminator' deletions were linked to the Hind III/Xmn I fragment of pLW1 containing the promoter/CAT modules to give expression plasmid series pTER5 Δ (Fig. 66).

b) Removal of 'terminator' sequences reduces CAT activity.

Plasmid pLW1 lacks the pTER5 'terminator' and Table 3 shows that Hela cells transfected with pLW1 produced little detectable CAT activity whereas high levels of activity were obtained with pTER5. The 'terminator' fragment therefore was required for detectable CAT gene expression. To determine the effect on CAT activity of an alternative 'terminator', a 1250bp Bam HI/Xho I fragment from the rabbit β-globin gene, containing all the IVS 2 sequences and some 600bp of flanking DNA (Banerji et al., 1981), was inserted into the Hind III site of pLW1 (see Methods, Page 101, Fig. 11). Plasmid pβTER1 contained this rabbit β-globin fragment with the AATAAA signal on the mRNA sense strand. Table 3 shows that CAT activity obtained with pβTER1 was comparable to that obtained with pTER5.

The CAT levels obtained with the pTER5Δ series were compared with pTER5 CAT activity. Two deleted plasmids with reduced CAT activities were analysed, and deletion end-points were determined by DNA sequencing. These two plasmids pTER5Δ30 and pTER5Δ44 had 30bp and 44bp respectively deleted from the 3' portion of the

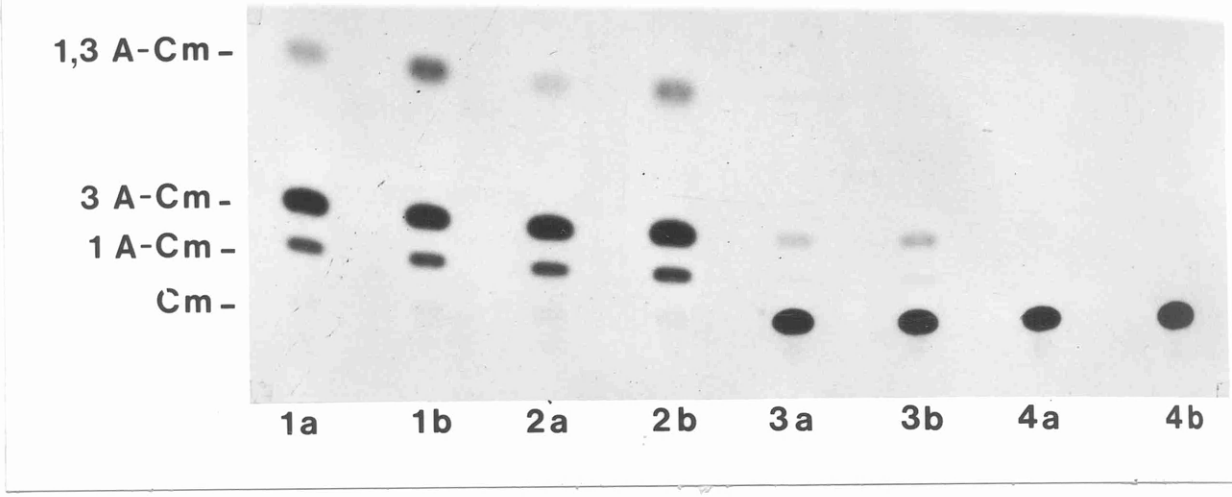


Figure 68. CAT activities produced by the pTER plasmids as determined by thin layer chromatography. Aliquots were removed from each reaction after (a) 20min and (b) 40min. The positions of acetylated (A-Cm) and non-acetylated forms (Cm) of ^{14}C chloramphenicol are indicated. HeLa cells were transfected with the following plasmids: tracks 1a, 1b pTER5; tracks 2a, 2b pTER5 Δ 30; tracks 3a, 3b pTER5 Δ 44; tracks 4a, 4b pTER6.

'terminator'. Locations of the deletion end-points are shown in Fig. 67; pTER5 Δ30 has a deletion which removes the G-rich sequences and extends to 3bp from the YGTGTTYY signal. In pTER5 Δ44, the YGTGTTYY signal is removed but the poly A site is retained.

Fig. 68 shows an autoradiograph of CAT assays obtained with the pTER5 plasmid series and the relative CAT activities from a number of separate experiments are listed in Table 3. CAT levels produced by pTER5 Δ30 and pTER5 Δ44 were low and trans-activation by co-transfection with the PRV plasmid, pPRVKpnh (see Methods, Page 100, Fig. 11), was required to boost activities to levels which could be accurately estimated. Plasmid pPRVKpnh expresses the PRV IE regulatory protein (Ihara et al., 1983) which increases transcription from other plasmid-borne eukaryotic promoters and from certain cellular promoters (Green et al., 1983; Imperiale et al., 1983; Everett, 1984b).

Deletion of 30bp from the 3' portion of the 'terminator' fragment reduced CAT levels to approximately 35% of the pTER5 value while removal of a further 14bp markedly reduced activity to a level of 1% to 4%; pTER5 Δ44 produced CAT activities only just above the background levels obtained with plasmids pLW1 and pTER6 (Table 3). Plasmid pLW1 has no 'terminator' fragment and pTER6 has the 'terminator' inserted in the opposite orientation to that of pTER5.

c) Levels of mRNA 3' and 5' termini produced by the pTER5 plasmid series.

Cytoplasmic RNA isolated from cells transfected with the pTER5 plasmid series was hybridised to strand-separated DNA probes which were 3'-labelled at the Hinf I site within the CAT coding sequences (Fig. 69). RNAs produced by pTER5 and pTER5Δ30 were hybridised to a probe produced from pTER5 DNA. As the deletion end-point of pTER5Δ44 is only 4bp from the poly A site (Fig. 67), hybrids produced by RNA transcribed across the poly A site could not be resolved from correctly-terminated RNAs, therefore RNA from pTER5Δ44 was hybridised to a probe from pTER5Δ44 DNA. The probes from pTER5 and pTER5Δ44 DNAs were end-labelled to similar specific activities. After treatment with nuclease S1, the protected DNA fragments were electrophoresed on a 6% denaturing polyacrylamide gel (Fig. 69).

RNA produced from pTER5 protected a fragment of 247nuc (Fig. 69, track 8) and this positioned the 3' end at position 53 (Fig. 67), the same location as the mRNA 3' terminus for IE mRNA-5 from HSV-2-infected cells (Whitton and Clements, 1984b). The 3' end of mRNA produced by pTER5Δ30 was located at the same position as that obtained with pTER5 but the amount of protected DNA fragment was reduced to approximately 30% of the pTER5 value (Fig. 69, track 9). No discrete 3' end was detected with pTER5Δ44 RNA (Fig. 69, track 11). Hence, the levels of mRNA 3' termini obtained with the pTER5 series reflected CAT activities obtained with these plasmids.

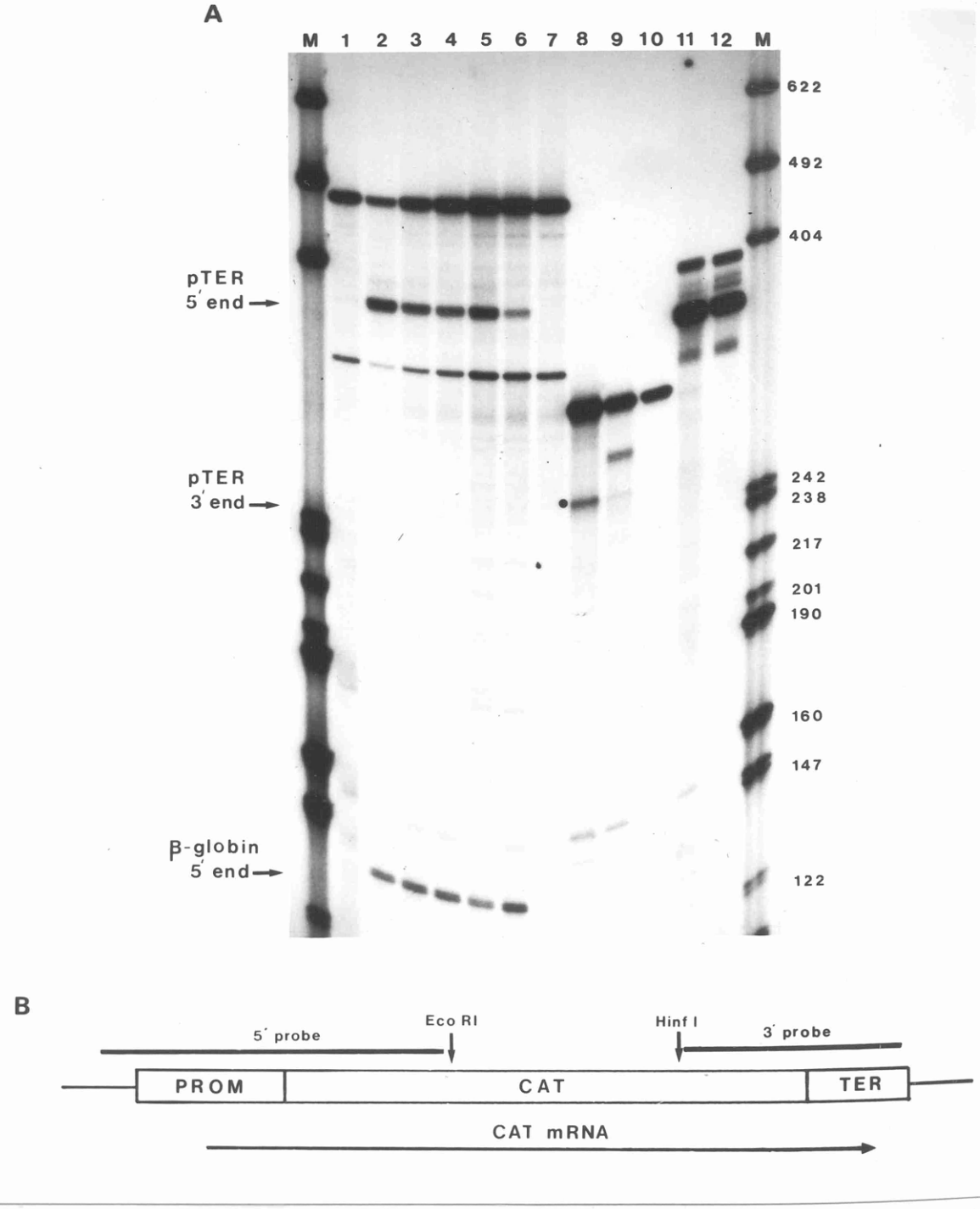

A
M 1 2 3 4 5 6 7 8 9 10 11 12 M
622
492
404
pTER
5′ end
pTER
3′ end
242
238
217
201
190
160
147
β-globin
5′ end
122
B
5′ probe
Eco RI
Hinf I
3′ probe
PROM
CAT
TER
CAT mRNA

Figure 69.

A) The abundance of mRNA 5' and 3' termini produced by the pTER plasmids. The DNA probes used in tracks 1 to 7 were: 1) an Eco RI/Sma I fragment 5'-labelled at the Eco RI site to determine the pTER CAT mRNA 5' termini and 2) a 5'-labelled Bst NI fragment derived from plasmid pRED4 to detect the β-globin mRNA 5' termini (see Fig. 11).

The single-stranded DNA probes used to detect pTER CAT mRNA 3' termini were a Hinf I fragment from pTER5, 3'-labelled at the Hinf I site in the CAT gene coding sequences (tracks 8 to 10), and a Hinf I/Bgl I fragment from pTER5Δ44, again 3'-labelled at the Hinf I site (tracks 11 and 12). RNA samples used were: tracks 1, 7, 10 and 12, mock-infected RNA; tracks 2 and 8, pTER5 RNA; tracks 3 and 9, pTER5Δ30 RNA; tracks 4 and 11, pTER5Δ44 RNA; track 5, pTER6 RNA; track 6, pLW1 RNA. Size markers (M) were pAT153 DNA cleaved with Hpa II.

B) Approximate locations for the 5' and 3' termini of pTER CAT mRNA are shown. Restriction enzyme sites within the CAT coding sequences which were used for making end-labelled DNA probes are indicated.

Plasmids pTER5 and pTER5Δ30 have identical sequences up to 18bp beyond the poly A site then the DNA sequences are dissimilar due to the 'terminator' deletion (30bp) in plasmid pTER5Δ30. Therefore, the band below the fully protected probe (Fig. 69, track 9) represents hybridisation of the pTER5 DNA probe to pTER5Δ30 RNA transcribed through the poly A site. Hence the size of this band reflects the sequence homology between pTER5Δ30 RNA and the pTER5 DNA probe.

To assess the validity of the comparative 3' terminal data, plasmid, pβ(244+)β, containing the β-globin promoter and the polyoma virus enhancer (see Methods, Page 100; de Villiers and Schaffner, 1981), was co-transfected with the various pTER5 plasmids. A strand-separated DNA probe from plasmid pRED4, 5'-labelled at a Bst NI site approximately 130bp from the β-globin mRNA 5' terminus was used to estimate β-globin mRNA levels (Fig. 11; Everett, 1983). As Fig. 69, tracks 2 to 6 indicate, β-globin mRNA levels were similar in all RNA samples.

Levels of mRNA 5' termini were estimated using a strand-separated DNA probe 5'-labelled at an EcoRI site within the CAT coding sequences of pTER5 (Fig. 69). Using this probe, a protected DNA fragment of 369nuc was obtained with RNA produced by all the pTER5 series plasmids and by pLW1 (Fig. 69, tracks 2 to 6). A similar amount of labelled probe was protected with the different RNA samples. Therefore, unlike the situation with the 3'-labelled probe, the levels of protected 5' probe

obtained with pTER5Δ30 and pTER5Δ44 were similar to that obtained with pTER5 although the CAT activities were lower than the pTER5 values. Indeed, detectable levels of 5' termini were obtained with pLW1 RNA in marked contrast to the inability of this plasmid to produce CAT activity. This suggests that cytoplasmic RNA with a normal 5' terminus but an incorrect 3' terminus does not produce functional protein.

DISCUSSION.

24. The Sequence YGTGTTYY is a Component of the Downstream Sequences Required for Efficient 3' End Formation.

Comparison of the sequences flanking HSV mRNA 3' termini has identified at least one element present also at the poly A sites of a large number of eukaryotic and viral genes. From conserved sequences located between 24 to 35bp downstream from AATAAA signals, a consensus sequence, YGTGTTYY, has been derived and greater than 60% of the genes examined contain sequences which resemble this consensus. A possible second element comprises either a G- or T-rich sequence, however, these tracts of residues occur less frequently than sequences which fit the YGTGTTYY motif.

Removal of sequences downstream from the poly A site of HSV-2 IE gene-5 reduces the level of correctly-processed mRNA 3' termini, and this effect is reflected in lower levels of CAT activity. Deletion plasmid

pTER5 Δ 30 which lacks the G-rich sequences gives reduced mRNA 3' termini and CAT levels are 35% of the values obtained with pTER5 which contains all the 'terminator' sequences. Removal of a further 14bp, which includes the YGTGTTYY signal but not the poly A site, markedly reduces CAT activity to 1% to 4% of the pTER5 value, and no correct 3' termini were detected. Therefore, it is likely that these conserved elements, in particular the YGTGTTYY signal, play an important role in the formation of mRNA 3' termini for a wide variety of mammalian genes and deletion analysis at other poly A sites would support this proposal.

McDevitt et al. (1984) have shown that T-rich sequences between 20 and 35nuc downstream from the poly A site, beyond the YGTGTTYY element, are required for production of functional Ad2 E2A mRNA, however correctly-terminated mRNA was produced at detectable levels. The data from the pTER5 Δ 30 plasmid also show that removal of sequences downstream from the YGTGTTYY signal, which included a G-rich tract, reduces the level of 3' end formation. Comparison of HSV-2 and Ad2 sequences in this region reveal no significant homology, however, analysis of the sequences at similar locations in other genes reveal that a number of 3' terminal sequences contain either a T- or G-rich segment (Figs. 64 and 65). A deletion which removes the poly A site of Ad2 E2A mRNA and the sequence AGTGTCTC resulted in no detectable mRNA; this sequence is 26nuc from the AATAAA signal and fits the YGTGTTYY consensus. Similarly, a 52bp deletion which removes the

poly A site together with sequences containing TGTGTTGG and a T-rich tract also abolishes rabbit β-globin 3' end formation (Gil and Proudfoot, 1984). Deletion studies on the SV40 late mRNA 3' terminus indicate that there are a minimum of two elements which affect the efficiency of cleavage at the correct poly A site (Sadofsky and Alwine, 1984; Conway and Wickens, 1985). At least part of one element is positioned between 21 and 38 residues from the AATAAA signal and this 17bp segment contains the sequence TATGTTTC (Conway and Wickens, 1985).

Approximately 33% of the 3' terminal DNA sequences examined do not contain a YGTGTTYY sequence motif. Absence of this sequence may be overcome by alternative downstream signals, and many of these 3' terminal regions contain T-rich residues (Fig. 65). In the HSV-1 TK gene, a T-rich segment between 32 and 37nuc from the AATAAA signal is required for efficient processing and polyadenylation (Cole and Stacy, 1985). The DNA sequences in this region (GTTTGTT) do not closely resemble the YGTGTTYY motif, however, these results do demonstrate that sequences approximately 30nuc from the AATAAA signal are functionally important in polyadenylation. The downstream YGTGTTYY sequence could therefore be analagous to transcription regulatory signals present upstream of the TATA box, for example the G+C-rich segments present in many but not all promoters (see Introduction, Page 62). Removal of G+C-rich sequences reduces the level of initiation (Everett, 1983;

McKnight and Kingsbury, 1982; Everett et al., 1983; Dierks et al., 1983), however, as these sequences are not present in the upstream regions of every promoter presumably they are not obligatory, or other sequences mimic their function.

25. Role of the YGTGTTYY Motif.

There is no evidence as to whether the YGTGTTYY signal either in DNA or precursor RNA is involved in mRNA termination or processing. To date, transcription termination signals for RNA polymerase II genes have not been identified in higher eukaryotic systems, however, mouse β-globin mRNA appears to terminate at random sites between 700 to 2000nuc downstream from the poly A site (see Introduction, Page 85; Citron et al., 1984). The presence of a YGTGTTYY motif approximately 30nuc downstream from the β-globin AATAAA signal would suggest that this sequence is not involved in termination within a region which is distant from the poly A site. Moreover, in the nucleus of infected cells, late polyoma RNAs have been detected which consist of tandem repeats of the viral genome (Buetti, 1974; Acheson, 1976). The flanking sequences of the polyoma late poly A site contain sequences which resemble the YGTGTTYY signal, again suggesting that this element is not a termination signal.

It would also appear that the YGTGTTYY motif is not a regulatory signal involved in poly A site selection. Analysis of alternative 3' termini which are used with differing efficiencies in the pro α2(I) collagen (Aho et

al., 1983), β2-microglobulin (Parnes et al., 1983) and mouse ∝amylase genes (Tosi et al., 1981) indicates that the presence of the YGTGTTYY signal is not associated with 3' end selection.

A likely role for these downstream sequences is in the cleavage/processing of nuclear RNAs which extend beyond the poly A site. Cleavage/processing of primary transcripts is thought to be the mechanism whereby 3' termini are generated for both poly A mRNAs and histone messages (see Introduction, Page 84; reviewed in Birnstiel et al., 1985). Histone genes generally do not contain AATAAA signals and a palindromic sequence frequently is present close to mRNA 3' termini (Kedes, 1979; Hentschel and Birnstiel, 1981). Point mutations within this palindromic sequence abolish the ability to form a correctly-terminated mRNA and sequences at least 24nuc downstream from the 3' terminus are required for efficient processing (Birchmeier et al., 1982). Within this region, a consensus sequence (CAAGAAAGA) has been derived from conserved sequences in the 3' flanking sequences of vertebrate and invertebrate histone genes (Busslinger et al., 1979; Hentschel and Birnstiel, 1981; Turner and Woodland, 1982; Birnstiel et al., 1985), however, the histone consensus sequence does not resemble the YGTGTTYY motif. This is not surprising since, although the mechanism for generating 3' termini in poly A and histone mRNAs may be similar, the processing proteins could be different and therefore alternative signals may be used.

26. Correctly-terminated mRNA is Required to Produce Functional Protein.

Deletion of the YGTGTTYY signal reduced 3' end formation but had no detectable effect on initiation at the 5' terminus. Therefore, it is likely that transcription has continued beyond the poly A site in plasmid pTER5 Δ 44, however, the location of the 3' termini of these transcripts has not been analysed. It is also not known whether the RNAs are polyadenylated. The 3' termini of extended transcripts produced as a result of the deletion of sequences at the HSV-1 TK poly A site map at approximately residue 4165 in the pBR322 plasmid sequences (Cole and Stacy, 1985). These RNAs are detectable in the poly A^+ fraction of cytoplasmic RNA preparations, suggesting that cryptic polyadenylation signals in the plasmid sequences may function following removal of the TK AATAAA sequence. The pTER5 plasmid series contain pBR322 sequences from residues 2067 to 4362 and the HSV/CAT gene is orientated in the same direction as the TK construct with respect to pBR322 sequences. Thus, it is possible that the 3' termini of extended transcripts generated by pTER5 Δ 44 also map at approximately position 4165 in the pBR322 sequences. An examination of the nucleotide sequences in this region indicates that two potential poly A signals are located between positions 4193 to 4198 (AATAAC) and 4172 to 4177 (AATAAT). Neither of these sequences fits exactly the AATAAA signal, although the sequence AATAAC does occur at the poly A site of one gene

(Wickens and Stephenson, 1985). Moreover, 36nuc downstream from the AATAAT sequence is a segment CGTGTCGC which resembles the YGTGTTYY motif. The 5' termini of the transcripts produced by pTER5Δ44 are detected in the cytoplasm, however, these RNAs do not produce functional CAT enzyme. It is possible that the pTER5Δ44 RNAs adopt a secondary structure which prevents ribosome binding and thereby the translation of protein.

Alternatively, transcripts which extend beyond the poly A site may not be polyadenylated. The inability to translate these RNAs may imply that their mode of transport to the cytoplasm differs from that of poly A mRNAs thus suggesting a coupling of mRNA polyadenylation, transport and translation. Further analysis is required to determine whether pTER5Δ44 cytoplasmic RNA is either polyadenylated or associated with polysomes.

GENERAL DISCUSSION

The Results and Discussion section has highlighted three areas concerning the expression and sequence content of HSV genes, namely:

1) organisation and regulation of mRNAs encoding the viral ribonucleotide reductase

2) conservation between the sequences specifying the HSV ribonucleotide reductase and other cellular and putative herpesvirus ribonucleotide reductases

3) identification of a consensus sequence, YGTGTTYY, which is located at a conserved distance downstream from most mRNA 3' termini and is required for efficient 3' end formation.

The General Discussion section proposes mechanisms for regulating HSV ribonucleotide reductase activity and suggests that expression of this function may be related to the mutagenic effect of HSV. From an interpretation of sequence comparisons, the possible evolution of the N- and C-terminal portions of the Vmw138 ribonucleotide reductase polypeptide is discussed. Finally, evidence is presented on DNA sequences involved in efficient mRNA 3' end formation and alternative functions for the YGTGTTYY signal in mRNA 3' end processing are suggested.

1. Regulation of HSV Ribonucleotide Reductase Activity.

Ribonucleotide reductase is essential for virus growth, as a mutant with a ts lesion in Vmw136 fails to synthesise viral DNA at the npt in resting cells (Preston et al., 1984b). Hence, the early synthesis of ribonucleotide reductase represents a critical step in HSV DNA biosynthesis. Previously, cellular ribonucleotide reductase activity has been proposed as a rate limiting step in DNA replication (see Introduction, Page 51); recent evidence indicates that changes in the concentration of dNTPs correlate with alterations in ribonucleotide reductase activity (Leeds et al., 1985). Synthesis of a viral reductase is necessary for at least two reasons:

1) Firstly, the level of cell ribonucleotide reductase activity is cell cycle-dependent (Eriksson et al., 1984; Engstrom et al., 1985). Enzyme activity appears to be controlled by the level of M2 protein which has a half-life of 3h and increases in abundance by three- to seven-fold prior to mitosis (Eriksson et al., 1984); by contrast, the M1 protein is produced constitutively and has a half-life of 15h (Engstrom et al., 1985). Thus, infected cells which are not undergoing DNA replication may possess a low cellular ribonucleotide reductase activity which is insufficient to meet the requirements for HSV DNA synthesis.

2) Secondly, in contrast to the mammalian enzyme, HSV ribonucleotide reductase is not regulated by allosteric effectors (see Introduction, Pages 56 and 58); hence, the

higher levels of dNTPs detected in infected cells (Cheng et al., 1975; Jamieson and Bjursell, 1976) would not inhibit the viral enzyme.

The following sections describe possible mechanisms for regulating this viral enzyme activity.

a) Regulation of mRNA synthesis.

The mRNAs specifying the proposed components of HSV ribonucleotide reductase overlap and, by nuclease S1 analysis, are synthesised at early times p.i. In cells infected with HSV-1 tsk, a ts mutant containing a lesion in the Vmw175 protein coding sequences which produces only the IE gene products at the npt, HSV-1 Vmw136 is synthesised. Thus, Vmw136, the equivalent HSV-1 protein to HSV-2 Vmw138, was classified as an IE product (Preston, 1979; Watson et al., 1979). A characteristic of HSV IE genes is the presence of multiple copies of the activator sequence, TAATGARAT, upstream from IE mRNA 5' termini (see Introduction, Page 28), however, TAATGARAT motifs are absent from the promoter sequences of the HSV-1 5.0kb and HSV-2 4.5kb transcripts. Moreover, cells infected in the presence of the protein synthesis inhibitor cycloheximide produce IE polypeptides but synthesise a markedly reduced level of Vmw136 compared with that in tsk-infected cells (H.S. Marsden, personal communication). While these data are consistent with the notion that the mRNAs encoding Vmw136 and Vmw138 are not IE products, Southern blot analysis with HSV-infected cell RNA indicates that transcripts from the ribonucleotide reductase locus are

detectable at low levels in the presence of cycloheximide (Clements et al., 1977). Moreover, following the removal of a cycloheximide block, transcripts from this genome region are rapidly and abundantly synthesised and their synthesis precedes that of other early mRNA species including those encoding DNA polymerase and TK functions (Clements et al., 1977). Hence, the transcripts encoding ribonucleotide reductase appear to possess promoters which are in a different temporal category from those of other early mRNAs. One possibility is that the promoters for the ribonucleotide reductase mRNAs are activated by proteins other than IE Vmw175 (see Introduction, Page 30). A candidate for mediating this activation is IE Vmw110 which is produced in tsk-infected cells at the npt and therefore could stimulate synthesis of the mRNA encoding Vmw136. Evidence from transient expression assays using plasmid constructions indicates that the promoter for the HSV-2 1.2kb mRNA encoding Vmw38 is activated by Vmw110 in the absence of Vmw175 (O'Hare and Hayward, 1985).

b) Translational control of Vmw38.

As a result of the overlapping nature of HSV transcripts, 3' co-terminal mRNAs which are unspliced carry the coding sequences for more than one protein. For example, the 4.5kb mRNA, which specifies Vmw138, also contains the coding sequences for Vmw38. In most cases, translation appears to begin at the 5' proximal initiation codon and subsequent ATG codons are utilised either with significantly reduced efficiency or not at all (Kozak,

1981). *In vitro* translation of the HSV-1 5.0kb mRNA, which encodes Vmw136 but also carries the Vmw38 coding sequences, produces a single polypeptide with a size of 140,000 mol. wt. (Anderson *et al.*, 1981). However, these results do not preclude the possibility that a proportion of Vmw38 produced *in vivo* is translated from the 5.0kb mRNA.

Within the HSV-1 TK gene, internal initiation occurs at two ATG codons downstream from the 5' proximal start codon giving rise to two minor species of 39,000 and 38,000 mol. wt. (see Introduction, Page 43). Moreover, insertion of ATG triplets upstream from the normal initiation codons of HSV-1 TK and hepatitis B surface antigen (HBsAg) reduces the level of initiation at authentic start codons (Bandyopadhay and Temin, 1984; Liu *et al.*, 1984). This effect can be partially suppressed by the presence of stop codons in phase with the inserted upstream ATG codons (Liu *et al.*, 1984); such stop codons may allow ribosomes to rapidly reinitiate at a second ATG on the same RNA molecule (Kozak, 1981). The initiation codon for Vmw38 is 53nuc downstream from the C-terminus of Vmw138, hence, the close proximity of these N- and C-termini may allow ribosomes which have translated Vmw138 from the 4.5kb mRNA to reinitiate at the ATG codon for Vmw38.

The first ATG in the HSV-1 1.2kb mRNA opens a reading frame which extends for only five amino acids and thus does not specify the N-terminus of Vmw38. The presence of this ATG may reduce the level of initiation at the

second ATG which marks the N-terminus of Vmw38. Vmw38 is almost certainly a component of ribonucleotide reductase (see Results and Discussion, Page 155) and enzyme activity in HSV-1-infected cells is less than that detected in HSV-2-infected cells (B. Dutia, personal communication). A lower level of initiation at the ATG for Vmw38 within the HSV-1 1.2kb mRNA as compared to that in the HSV-2 1.2kb transcript may contribute to this reduced ribonucleotide reductase activity.

2. Association of Ribonucleotide Reductase Activity with Cellular Transformation.

In vitro studies using separated restriction enzyme fragments have indicated that sequences adjacent to the ribonucleotide reductase locus are involved in morphological transformation (see Introduction, Page 48). However, the ability of HSV to transform cells may be related to its action as a mutagen. For example, increased mutation rates at the hypoxanthine guanine phosphoribosyltransferase (HGPRT) locus results from infection of permissive cells with partially-inactivated HSV-1 (Schlehofer and zur Hausen, 1982) and of non-permissive cells with HSV-2 (Pilon et al., 1985). Moreover, HSV induces chromosomal aberrations (Hampar and Ellison, 1963; Stich et al., 1964) which may be mediated by an early viral function (Waubke et al., 1968; O'Neill and Rapp, 1971).

In HSV-infected cells, chromosomal damage is repaired by cellular repair mechanisms (Nishiyama et al., 1983) whose ability to faithfully replicate DNA is influenced by changes in dNTP pools (Ben-Hur and Ben-Oshai, 1971; Collins et al., 1977). As infected cells contain higher levels of dNTPS than uninfected cells (Cheng et al., 1975; Jamieson and Bjursell, 1976), alterations in dNTP pool sizes may be the basis of mutagenic changes introduced into cell DNA following infection by HSV. Analysis of cells with a mutator phenotype have shown that the production of increased levels of dNTPs is associated with 10- to 100-fold increases in spontaneous mutation rate, possibly through competition between correct and incorrect dNTPs at sites of base addition (Kunkel and Loeb, 1979). In certain cases, such imbalances in dNTP pools result from the activity of cellular ribonucleotide reductases with altered allosteric properties such that enzyme activity is refractory to feedback inhibition (Ullman et al., 1980; Chan et al., 1981; Weinberg et al., 1981; Ayusawa et al., 1983); these data establish a correlation between mutation rate and ribonucleotide reductase activity. Thus, HSV ribonucleotide reductase, an early viral function, may act as a mutagen in infected cells by altering dNTP pool sizes. Of interest would be the mutation rate in cells possessing HSV ribonucleotide reductase, integrated into the cell genome, under the control of an inducible promoter. In such cells, stimulation of HSV ribonucleotide reductase activity may increase spontaneous mutation rate; these experiments are currently in progress.

3. Evolution of the Promoter for the Proposed Small Subunit of the HSV Ribonucleotide Reductase.

The 5' termini and flanking sequences of the HSV-2 1.2kb mRNA encoding Vmw38 and the EBV 3.3kb mRNA encoding the 34,000 mol. wt. protein which has homology with Vmw38, lie within the C-terminal regions of the HSV-2 Vmw138 and EBV 93,000 mol. wt. polypeptides respectively (Fig. 70; Gibson et al., 1984). The C-terminal regions of Vmw138 and the 93,000 mol. wt. protein share homology (see Results and Discussion, Page 162), but the locations of the 1.2kb and 3.3kb mRNA 5' termini are not at equivalent positions in these polypeptides (Fig. 70).

Alignment of the HSV-2 Vmw138 C-terminal residues with those of the mouse M1 polypeptide sequence (Caras et al., 1985) indicates that the 1.2kb mRNA 5' terminus lies within an area of high homology which is also conserved in the EBV 93,000 mol. wt. protein (Fig. 70). The HSV-2 amino acid sequence immediately upstream from the 5' end of the 1.2kb mRNA has low homology with the M1 protein and contains additional residues (Fig. 59). While the precise locations of the additional residues in the HSV-2 protein is unclear, the DNA sequences encoding these amino acids may represent transcription control signals involved in synthesis of the 1.2kb mRNA, for example the TATA box homology.

The equivalent region of the EBV 93,000 mol. wt. polypeptide also appears to contain additional residues compared with the M1 sequences. However, the TATAA box

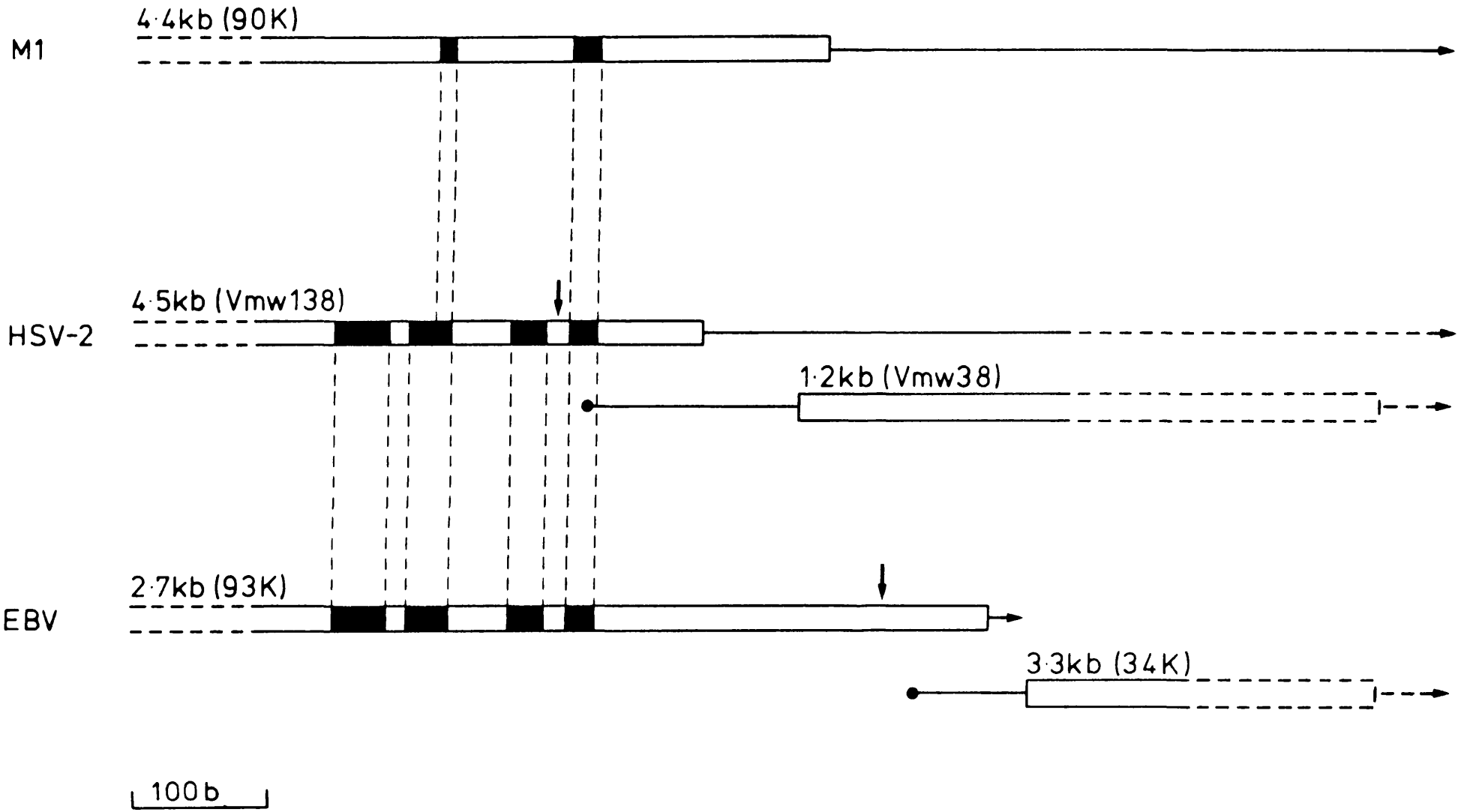

4·4kb (90K)
M1
4·5kb (Vmw138)
HSV-2
1·2kb (Vmw38)
2·7kb (93K)
EBV
3·3kb (34K)
100b

Figure 70. Alignment of conserved amino acid residues specified by the transcripts encoding C-terminal portions of HSV-2 Vmw138, the mouse M1 and EBV 93,000 mol. wt. proteins. Sizes of mRNAs are shown and the proteins specified by transcripts are given in brackets. Shaded areas represent regions of conserved amino acids and vertical dashed lines indicate their relative locations within equivalent polypeptides. The relative positions of the HSV-2 1.2kb and EBV 3.3kb mRNA 5' termini within protein coding regions are shown with the locations of TATA box homologues for these transcripts indicated by vertical arrows.

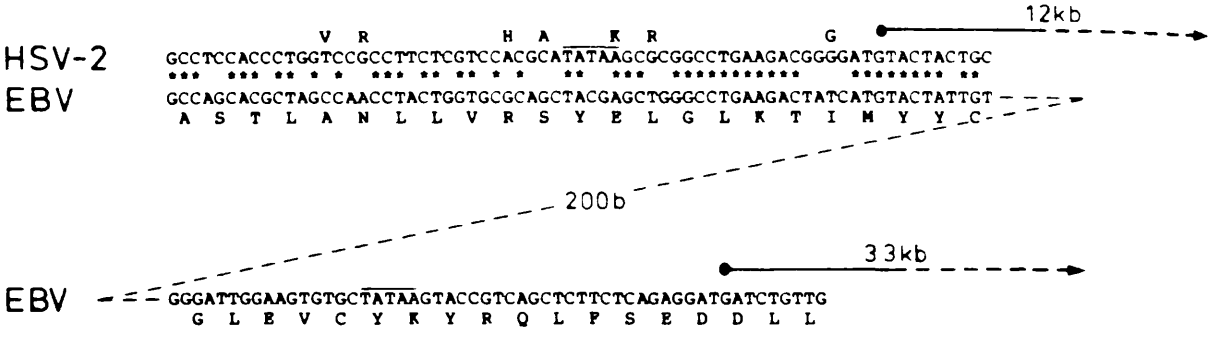


Figure 71. Alignment of nucleotide sequences specifying the C-terminal portions of HSV-2 Vmw138 and the EBV 93,000 mol. wt. protein. Asterisks denote conserved residues and predicted amino acid residues for the 93,000 mol. wt. protein are shown below the EBV nucleotide sequence. Amino acids which are not conserved in HSV-2 Vmw138 are shown above the HSV-2 nucleotide sequence. Positions of the HSV-2 1.2kb and EBV 3.3kb mRNA 5' termini are indicated and TATA box homologues for these transcripts are overlined.

motif in HSV-2 corresponds to a TACGA sequence in EBV and the cap site of the 3.3kb transcript is not positioned within 30nuc from this sequence (Fig. 71). Rather, the 3.3kb mRNA 5' terminus is located within a C-terminal portion of the 93,000 mol. wt. protein which is not present in either of the Vmw138 or M1 polypeptides(Fig. 71); the additional amino acid sequences in the EBV protein extend for 75 and 50 residues beyond the C-termini of Vmw138 and M1 polypeptides respectively. Thus, the promoters for the HSV-2 1.2kb and EBV 3.3kb transcripts appear to have arisen by different evolutionary events.

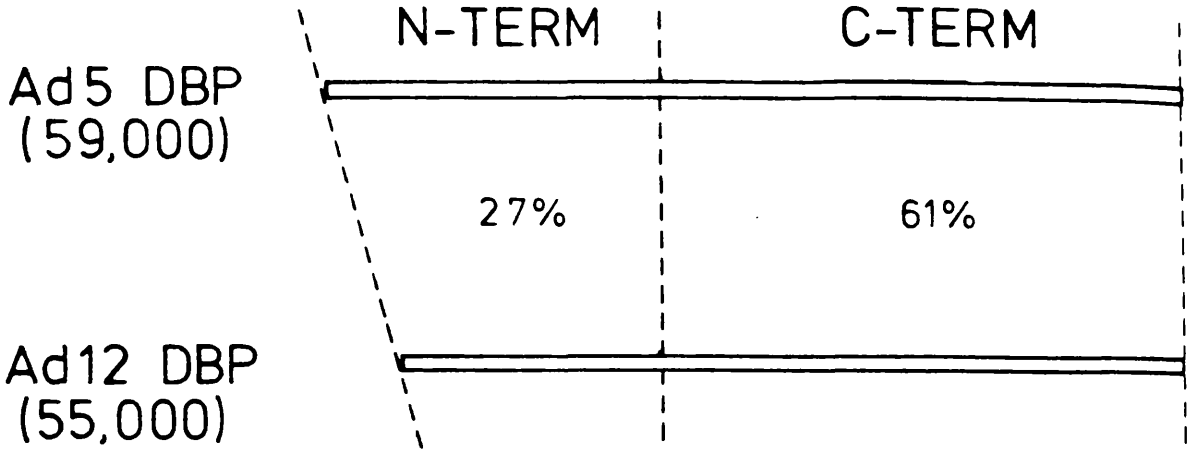
Additional amino acids at the C-terminus of the 93,000 mol. wt. polypeptide may be a result of a frame-shift in a precursor EBV protein, thereby incorporating untranslated sequences into a protein coding region. However, comparison of nucleotide sequences downstream from the C-termini of M1 and B1 proteins with those encoding the additional EBV amino acids reveals no detectable homology. Therefore, promoter sequences for the 3.3kb transcript appear not to be related to the sequences of other reductase genes. Alternatively, novel sequences containing transcription control signals may have been inserted fortuitously into the polypeptide coding region of a precursor EBV 93,000 mol. wt. protein. Insertion of a complete promoter downstream from a primitive promoter has been proposed for the major human chorionic somatomammotropin gene (hCS gene; Selby et al., 1984); these authors further suggest that the inserted promoter

sequences may have included transcriptional signals which altered the expression of a precursor hCS gene.

The promoter for the HSV-2 1.2kb mRNA does not appear to have arisen from the insertion of extensive nucleotide sequences into a protein coding region. Sequences upstream from the cap site of the 1.2kb RNA, which are absent in the M1 sequence, are present in EBV DNA, although this region does not contain a promoter in EBV (Fig. 71). Thus, the 1.2kb mRNA promoter signals may have evolved by mutation of pre-existing polypeptide coding sequences.

4. Is Vmw138 a Product of Gene Fusion?

The proposed components of HSV-2 ribonucleotide reductase (Vmw138 and Vmw38) share homology with the subunits comprising both mammalian and bacterial enzymes, however, Vmw138 is considerably greater in size than either the mouse M1 (90,000 mol. wt.; Thelander and Reichard, 1979) or E. coli B1 (80,000 mol. wt.; Thelander, 1973) proteins. Moreover, the EBV and VZV polypeptides which share homology with Vmw138 are of similar size to the M1 and B1 proteins (Gibson et al., 1984; A. J. Davison personal communication). Sequence comparisons between the various polypeptides indicate that HSV-1 Vmw136 contains approximately 480 additional amino acids at the N-terminus compared with equivalent EBV, VZV, mouse and E. coli proteins (I. Nikas, personal communication). Homology between the N-termini of HSV-1 Vmw136 and HSV-2 Vmw138



N-TERM: late gene expression

C-TERM: early gene expression; DNA replication

Figure 72. Amino acid homology between the N- and C-terminal portions of the Ad5 and Ad12 DNA binding proteins. The functions of these regions are also listed.

suggests that the HSV-2 protein also contains an additional N-terminal domain. The function of these N-terminal domains in Vmw136 and Vmw138 is unknown, however, these regions may not be directly involved in HSV-1 and HSV-2 ribonucleotide reductase activity. One process which could have generated the additional sequences is gene fusion whereby a precursor of Vmw136 or Vmw138 may have been linked to a second protein with different functional properties. Thus, Vmw136 and Vmw138 may comprise two functional units, the C-terminal portion which has a role in ribonucleotide reductase activity and the N-terminal region which has a separate but unknown function.

An example of a viral protein comprising two domains with different functions is the Ad DNA binding protein. Comparisons between coding sequences of the DNA binding proteins of Ad5 and Ad12 (Kruijer et al., 1981 and 1983) reveal low homology (27%) for one third of the protein at the N-terminus; the remainder of the protein sequences are essentially colinear and highly conserved (61%; Fig. 72). Kruijer et al. (1983) have suggested that these areas of different homology represent two domains in the Ad DNA binding protein, each with a different function. In support of this proposal, a 46Kd fragment from the C-terminal portion of the Ad5 DNA binding protein binds to single-stranded DNA (Horwitz, 1978) and mutations within this region affect DNA replication and early gene expression (Kruijer et al., 1981, 1982 and 1983; Klessig and Quinlan, 1982). However, a 26Kd fragment from

Vmw38	89%
TK	73%
gC	74%
gD	85%

Figure 73. Amino acid homology between equivalent HSV-2 and HSV-1 proteins. These data were taken from the following references:

HSV-1 Vmw38, Draper et al., 1982, McLauchlan and Clements, 1983a;

HSV-2 Vmw38, this study, McLauchlan and Clements, 1983b;

HSV-1 TK, McKnight, 1980, Wagner et al., 1981;

HSV-2 TK, Swain and Galloway, 1983;

HSV-1 gC, Frink et al., 1983, Draper et al., 1984a;

HSV-2 gC, Dowbenko and Lasky, 1984, Swain et al., 1985;

HSV-1 gD, Watson et al., 1982, McGeoch et al., 1985;

HSV-2 gD, Watson, 1983.

the N-terminal region does not bind to single-stranded DNA (Klein et al., 1979; Linne and Philipson, 1980) and mutations within this fragment influence late gene expression (Kruijer et al., 1981). Thus, different homologies between the N- and C-terminal regions may be a result of different evolutionary pressures on protein domains with separate functions.

Homology between equivalent HSV-1 and HSV-2 proteins varies from 73% to 89% (Fig. 73). Presumably, this variation in the degree of amino acid conservation reflects the evolutionary pressure on proteins which have different functions. If the additional N-terminal sequences in Vmw136 and Vmw138 are not directly involved in ribonucleotide reductase activity, then the level of homology may not be identical at the N- and C-terminal domains. The N-terminal 86 amino acids in Vmw136 and Vmw138 share lower homology (62%) than the C-terminal region (95%), however, a more extensive comparison is necessary to evaluate this possibility of different evolutionary rates for both domains. The function of these N-terminal domains in ribonucleotide reductase activity could be further tested by removing sequences at the N-termini of Vmw136 and Vmw138 and examining the effect on virus enzyme activity.

5. Limits of the DNA Sequences Involved in mRNA 3' End Formation.

The tandem reiterations located in intergenic regions of HSV DNA are putative sites for high rates of recombination (see Introduction, Page 15). Reiterations are present at a short distance, between 2 and 8nuc, beyond the YGTGTTYY motifs of three genes, IE genes-1 and -3 and gE (Fig. 64), hence, these repeats may delimit those sequences which function in 3' end formation. The repeats have a high G+C content and their sequences resemble the G-rich tracts found at a similar distance downstream from other 3' termini (Fig. 64). Therefore, it is possible that the first copy of each repeat may play a role in 3' end formation while the remaining copies are non-functional. The effect on CAT activity following the insertion of one or more tandem repeats downstream from the YGTGTTYY motif in plasmid pTERΔ30 (see Results and Discussion, Page 177) could test this prediction.

Analysis of gene duplications also suggests that downstream sequences contain signals necessary for the generation of 3' termini. The human α1 and α2 globin genes have arisen from a gene duplication event and show overall high sequence homology (Proudfoot and Maniatis, 1980). Sequence conservation at the 3' portions of α1 and α2 genes is markedly reduced beyond 15bp downstream from the poly A site; this 15bp segment has the nucleotide sequence GCCTGTGTGTGCCTG. Proudfoot and Maniatis (1980) have suggested that this segment delimits the signals

required for either recombination or the production of mRNA. The 15bp segment contains the sequence TGTGTGCC which fits closely the consensus YGTGTTYY, hence favouring the proposal that the 3' boundary of the sequences involved in the duplication event were determined by mRNA 3' end processing signals. The absence of G-rich and T-rich segments beyond the TGTGTGCC sequence suggests that these tracts are not an absolute requirement for 3' end formation.

6. Mechanisms of mRNA 3' End Formation.

The YGTGTTYY sequence may function as a processing signal either at the level of DNA or RNA. One mechanism for recognising a signal in RNA is through interaction with snRNPs since polyadenylation is inhibited *in vitro* by antisera which recognise such particles, indicating a role for these complexes in 3' end processing (Moore and Sharp, 1984). Berget (1984) has suggested that U4 snRNPs mediate 3' end formation through the binding of U4 snRNA to complementary sequences at mRNA 3' termini, principally the AATAAA signal and a conserved CAYUG pentamer (Y = pyrimidine) which lies at a variable distance (6 to 26nuc) beyond the polyadenylation signal. U4 RNA does not possess sequences complementary to the YGTGTTYY consensus, suggesting that either other U-type RNAs may be involved in the generation of poly A sites or an alternative function for YGTGTTYY.

SnRNPs have also been implicated in the formation of histone mRNA 3' termini where U7 RNA is required in oocytes for the processing of extended primary transcripts (see Introduction, Page 93). U7 RNA contains sequences which could base pair with the stem and loop structure and CAAGAAAGA motif (Strub et al., 1984) both of which are essential components for 3' end processing of histone mRNAs (see Introduction, Page 90). Altering the distance between the palindrome and the CAAGAAAGA sequence reduces the level of 3' end formation, possibly by disrupting the hybrid structure formed between U7 and the processing signals (Georgiev and Birnstiel, 1985). The distance between the poly A signal and the YGTGTTYY motif is relatively well conserved, thus the insertion or deletion of sequences between these elements may result in inefficient processing at the poly A site. This possibility could be examined by analysing the level of mRNA 3' end formation in plasmid constructions containing an altered spacing between the AATAAA and YGTGTTYY signals.

Alternatively, YGTGTTYY may act as a recognition site for proteins by altering the local configuration in DNA; the following studies suggest a critical role for the core of the consensus in producing changes in DNA configuration. The trinucleotide GTG occurs at a high frequency in eukaryotic and prokaryotic transcription control signals and at sites involved in DNA replication and recombination such as consensus sequences for phage replication origins, the LTRs of retroviruses and the

consensus for recombination sites in immunoglobulins (Cheung et al., 1984). Furthermore, GTG forms part of the binding site for a gene-specific transcription factor which interacts with the Ad major late promoter (Sawadogo and Roeder, 1985). Interestingly, GTG is the most highly conserved segment within the YGTGTTYY consensus (Table 2) and recent studies confirm this high frequency of G+T-rich residues downstream from mRNA 3' termini (Nussinov, 1985). From NMR studies using synthetic oligomers, the sequence GTG and its complement CAC exhibit faster opening kinetics in double-stranded DNA (Lu et al., 1983; Cheung et al., 1984). Moreover, the YGTGTTYY consensus contains alternating purine and pyrimidine residues which favour Z-DNA formation (Wang et al., 1979). The transcriptional enhancer region of SV40 also possesses alternating purine/pyrimidine segments which form Z-DNA and bind proteins (Nordheim and Rich, 1983). These analyses have led to the proposal that the GTG/CAC sequence may form a detent in DNA, thus acting as a pause site for proteins moving along DNA (Lu et al., 1983). This suggests that the GTG segment within YGTGTTYY may form part of the targetting signal for complexes by arresting proteins in the vicinity of the poly A site; the stable interaction of proteins at this site could then be determined by an examination of flanking sequences.

Studies using antibodies directed against both Z-DNA and snRNPs would be useful to determine whether YGTGTTYY functions in DNA or RNA. Protection experiments

using DNase I footprint methods would examine DNA regions at poly A sites which bind proteins. These binding studies could be extended to purify proteins from cell extracts which interact with 3' terminal sequences. A necessary part of this analysis would be the effect of point mutations within YGTGTTYY, in particular the effects of alteration of the GTG sequence on protein binding and DNA configuration. Analysis of a bacterial promoter region indicates that, in negatively-supercoiled regions of DNA, nuclease S1 sensitive sites are positioned at TpA dinucleotides due to local unwinding of the DNA (Drew et al., 1985). TpG doublets also have a low thermal stability (Gotoh and Tagashira, 1981) and a single alteration in the -10 region of a bacterial promoter from TATGATG to TATGAAG, which removes a TpG dinucleotide, reduces the ability of RNA polymerase to form a stable initiation complex (Travers et al., 1983). Therefore, the precise locations of nuclease S1 sensitive sites at poly A sites is relevant to future studies.

REFERENCES

Acheson, A. (1976). Polyoma virus giant RNAs contain tandem repeats of the nucleotide sequence of the entire viral genome. Proc. Natl. Acad. Sci. U.S.A. 75; 4754-4758.

Adam, E., Kaufman, R.H., Melnick, J.L., Levy, A.H. and Rawls, W.E. (1972). Sero-epidemiological studies of herpesvirus type 2 and carcinoma of the cervix. IV. Dysplasia and carcinoma in situ. Amer. J. Epidem. 96; 427-442.

Adesnik, M. and Darnell, J.E. (1972). Biogenesis and characterisation of histone messenger RNA in HeLa cells. J. Mol. Biol. 67; 397-406.

Aho, S., Tate, U. and Boedtker, H. (1983). Multiple 3' ends of the chicken pro α^2(1) collagen gene. Nucl. Acids Res. 11; 5443-5450.

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

Akusjarvi, G. and Persson, H. (1981b). Gene and mRNA for precursor polypeptide VI from adenovirus type 2. J. Virol. 38; 469-482.

Akusjarvi, G., Zabrielski, J., Perricaudet, M. and Pettersson, U. (1981). The sequence of the 3' non-coding region of the hexon mRNA discloses a novel adenovirus gene. Nucl. Acids Res. 9; 1-17.

Allen, G.P., Cohen, J.C., Randall, C.C. and O'Callaghan, D.J. (1978). Replication of equine herpesvirus type 1 and type 3: resistance to hydroxyurea and thymidine. Intervirology 9; 276-285.

Allen, J.R., Reddy, G.P.V., Lasser, G.W. and Mathews, C.K. (1980). T4 ribonucleotide reductase: physical and kinetic linkage to other enzymes of deoxyribonucleotide biosynthesis. J. Biol. Chem. 255; 7583-7588.

Alt, F.W., Bothwell, A.L.M., Knapp, M., Siden, E., Mather, A., Koshland, M. and Baltimore, D. (1980). Synthesis of secreted and membrane-bound immunoglobulin mu heavy chains is directed by mRNAs that differ at their 3' ends. Cell 20; 293-301.

Alwine, J.C., Steinhart, W.L. and Hill, C.W. (1974). Transcription of herpes simplex virus type 1 DNA in nuclei isolated from infected HEp-2 and KB cells. Virology 60; 302-307.

Alwine, J.C., Kemp, D.J., and Stark, G.R. (1977). Method for the detection of specific RNAs in agarose gels by transfer to diazobenzoyloxymethyl-paper and hybridisation with DNA probes. Proc. Natl. Acad. Sci. U.S.A. 74; 5350-5354.

Amara, S.G., Jonas, V., Rosenfeld, M.G., Ong, E.S. and Evans, R.M. (1982). Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. Nature 298; 240-244.

Amara, S.G., Evans, R.M. and Rosenfeld, M.G. (1984). Calcitonin/calcitonin gene-related peptide transcription unit: tissue-specific expression involves selective use of alternative polyadenylation sites. Mol. Cell. Biol. 4; 2151-2160.

Amin, J., Mestril, R., Lawson, R., Klapper, M. and Voellmy, R. (1985). The heat shock consensus sequence is not sufficient for hsp70 gene expression in Drosophila melanogaster. Mol. Cell. Biol. 5; 197-203.

Anderson, K.P., Costa, R.H., Holland, L.E. and Wagner, E.K. (1980). Characterisation of herpes simplex virus type 1 RNA present in the absence of de novo protein synthesis. J. Virol. 34; 9-27.

Anderson, K.P., Frink, R.J., Devi, G.B., Gaylord, B.H., Costa, R.H. and Wagner, E.K. (1981). Detailed characterisation of the mRNA mapping in the HindIII fragment k region of the herpes simplex virus type 1 genome. J. Virol. 37; 1011-1027.

Ashburner, M. and Bonner, J.J. (1979). The induction of gene activity in Drosophila by heat shock. Cell 17; 241-254.

Atkin, C.L., Thelander, L., Reichard, P. and Lang, G. (1973). Iron and free radical in ribonucleotide reductase: exchange of iron and Mossbauer spectroscopy of the protein B2 subunit of the Escherichia coli enzyme. J. Biol. Chem. 248; 7464-7472.

Aurelian, L. and Roizman, B. (1964). The host range of herpes simplex virus. Interferon, viral DNA and antigen synthesis in abortive infection of dog kidney cells. Virology 22; 452-461.

Averett, D.R., Lubbers, C., Elion, G.B. and Spector, T. (1983). Ribonucleotide reductase induced by herpes simplex virus type 1. Characterisation of a distinct enzyme. J. Biol. Chem. 258; 9831-9838.

Averett, D.R., Furman, P.A. and Spector, T. (1984). Ribonucleotide reductase of HSV-2 resembles that of HSV-1. J. Virol. 52; 981-983.

Avvedimento, V.E., Vogeli, G., Yamada, Y., Maizel, J.V.Jr., Pastan, I. and de Crombrugghe, B. (1980). Correlation between splicing sites within an intron and their sequence complementarity with U1 RNA. Cell 21; 689-696.

Ayusawa, D., Iwata, K. and Seno, T. (1983). Unusual sensitivity to bleomycin and joint resistance to 9-β-D-arabinofuranosyladenine and 1-β-D-arabinofuranosylcytosine of mouse FM3A cell mutants with altered ribonucleotide reductase thymidylate synthase. Cancer Res. 43; 814-818.

Bacchetti, S., Evelegh, M.J., Muirhead, B., Sartori, C.S. and Huszar, D. (1984). Immunological characterisation of HSV-1 and -2 polypeptide(s) involved in viral ribonucleotide reductase activity. J. Virol. 49; 591-593.

Bachenheimer, S.L. and Roizman, B. (1972). Ribonucleic acid synthesis in cells infected with herpes simplex virus. VI. Polyadenylic acid sequences in viral messenger ribonucleic acid. J. Virol. 10; 875-879.

Bachmann, B.J., Low, K.B. and Taylor, A.L. (1976). Recalibrated linkage map of Escherichia coli K12. Bacteriol. Rev. 40; 116-167.

Baer, R., Bankier, A.T., Biggin, M.D., Deininger, P.L., Farrell, P.J., Gibson, T.J., Hatfull, G., Hudson, G.S., Satchwell, S.C., Seguin, C., Tuffnell, P.S. and Barrell, B.G. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature 310; 207-211.

Baker, C.C. and Ziff, E.B. (1981). Promoters and heterogeneous 5' termini of the messenger RNAs of adenovirus serotype 2. J. Mol. Biol. 149; 189-221.

Bakken, A., Morgan, G., Sollner-Webb, B., Roan, J., Busby, S. and Roeder, R.H. (1982). Mapping of transcription initiation and termination signals on X. laevis ribosomal DNA. Proc. Natl. Acad. Sci. U.S.A. 79; 56-60.

Balachandran, N., Harnish, D., Rawls, W.E. and Bacchetti, S. (1982). Glycoproteins of herpes simplex virus type 2 as defined by monoclonal antibodies. J. Virol. 44; 344-355.

Bandyopadhyay, P.K. and Temin, H.M. (1984). Expression from an internal AUG codon of herpes simplex virus thymidine kinase gene inserted in a retrovirus vector. Mol. Cell. Biol. 4; 743-748.

Banerjee, A.K. (1980). 5'-terminal cap structure in eukaryotic messenger ribonucleic acids. Microbiol. Rev. 44; 175-205.

Banerji, J., Rusconi, S. and Schaffner, W. (1981). Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. Cell 27; 299-308.

Banerji, J., Olson, L. and Shaffner, W. (1983). A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell 33; 729-740.

Bankier, A.T., Dietrich, W., Baer, R., Barrell, B.G., Colbere-Garapin, F., Fleckenstein, B. and Bodemer, W. (1985). Terminal repetitive sequences in herpesvirus saimiri virion DNA. J. Virol. 55; 133-139.

Banks, L., Purifoy, D.M., Hurst, P.F., Killington, R.A. and Powell, K.L. (1983). Herpes simplex virus non-structural proteins. IV. Purification of the virus induced deoxyribonuclease and characterisation of the enzyme using monoclonal antibodies. J. Gen. Virol. 64; 2249-2260.

Barahona, H.H., Melendez, L.V., King, N.W., Daniel, M.D., Fraser, C.E.O. and Preville, A.C. (1973). Herpesvirus aotus type 2: a new viral agent from owl monkeys (Aotus trivirgatus). J. Infect. Dis. 127; 171-178.

Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980). The primary structure of the human ζ -globin gene. Cell 21; 621-626.

Baringer, J.R. (1974). Recovery of herpes simplex virus from human sacral ganglions. New England J. Med. 291; 828-830.

Bartkoski, M. and Roizman, B. (1976). RNA synthesis in cells infected with herpes simplex virus. XIII. Differences in the methylation patterns of viral RNA during the reproductive cycle. J. Virol. 20; 583-588.

Bastian, F.O., Rabson, A.S., Yee, C.L. and Tralka, T.S. (1972). Herpesvirus hominis: isolation from the human trigeminal ganglion. Science 178; 306-307.

Baty, D., Barrera-Saldana, H.A., Everett, R.D., Vigneron, M. and Chambon, P. (1984). Mutational dissection of the 21bp repeat region of the SV40 early promoter reveals that it contains overlapping elements of the early-early and late-early promoters. Nucl. Acids Res. 12; 915-932.

Baucke, R.B. and Spear, P.G. (1979). Membrane proteins specified by herpes simplex virus. V. Identification of an Fc-binding glycoprotein. J. Virol. 32; 779-789.

Becker, Y., Dym, H. and Sarov, I. (1968). Herpes simplex virus DNA. Virology 36; 184-192.

Bell, G.I., Pictet, R.L., Rutter, W.J., Cordell, B., Tischer, E. and Goodman, H.M. (1980). Sequence of the human insulin gene. Nature 284; 26-32.

Ben-Hur, E. and Ben-Oshai, R. (1971). DNA repair in ultraviolet light irradiated HeLa cells and its irreversible inhibition by hydroxyurea. Photochem. Photobiol. 13; 337-345.

Bennetzen, J.L. and Hall, B.D. (1982). The primary structure of the Saccharomyces cerevisae gene for alcohol dehydrogenase 1. J. Biol. Chem. 257; 3018-3025.

Benoist, C. and Chambon, P. (1981). In vivo sequence requirements of the SV40 early promoter region. Nature 290; 304-310.

Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980). The ovalbumin gene-sequence of putative control regions. Nucl. Acids Res. 8; 127-142.

Ben-Porat, T. and Kaplan, A.S. (1962). The chemical composition of herpes simplex and pseudorabies viruses. Virology 16; 261-266.

Ben-Porat, T. and Kaplan, A.S. (1985). Molecular biology of pseudorabies virus. In 'The Herpesviruses.' vol 3, pp. 105-173. (B. Roizman, ed.) Plenum Press, New York.

Ben-Porat, T., Rixon, F.J. and Blankenship, M.L. (1979). Analysis of the structure of the genome of pseudorabies virus. Virology 95; 285-294.

Ben-Zeev, A. and Becker, Y. (1977). Requirement of host cell RNA polymerase II in the replication of herpes simplex virus in α-amanitin-sensitive and -resistant cell lines. Virology 76; 246-253.

Ben-Zeev, A., Asher, Y. and Becker, Y. (1976). Synthesis of herpes simplex virus-specified RNA by an RNA polymerase II in isolated nuclei in vitro. Virology 71; 302-311.

Berget, S. (1984). Are U4 small nuclear ribonucleoproteins involved in polyadenylation? Nature 309; 179-182.

Berglund, O. (1972). Ribonucleotide diphosphate reductase induced by bacteriophage T4. II. Allosteric regulation of substrate specificity and catalytic activity. J. Biol. Chem. 247; 7276-7281.

Berglund, O. (1975). Ribonucleotide diphosphate reductase induced by bacteriophage T4. III. Isolation and characterisation of proteins B1 and B2. J. Biol. Chem. 250; 7450-7455.

Berglund, O. and Sjoberg, B.-M. (1970). Thioredoxin induced by bacteriophage T4. II. Purification and characterisation. J. Biol. Chem. 245; 6030-6035.

Berglund, O., Karlstrom, O. and Reichard, P. (1969). A new ribonucleotide reductase system after infection with phage T4. Proc. Natl. Acad. Sci. U.S.A. 62; 829-835.

Berk, A.J. and Sharp, P.A. (1978a). Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. 75; 1274-1278.

Berk, A.J. and Sharp, P.A. (1978b). Structure of the adenovirus 2 early mRNAs. Cell 14; 695-711.

Birchmeier, C., Grosschedl, R. and Birnstiel, M.L. (1982). Generation of authentic 3' termini of an H2A mRNA in vivo is dependent on a short inverted DNA repeat and on spacer sequences. Cell 28; 739-745.

Birchmeier, C., Folk, W. and Birnstiel, M.L. (1983). The terminal RNA stem-loop structure and 80bp of spacer DNA are required for the formation of 3' termini of sea urchin H2A mRNA. Cell 35; 433-440.

Birchmeier, C., Schumperli, D., Sconzo, G. and Birnstiel, M.L. (1984). 3' editing of mRNAs: sequence requirements and involvement of a 60 nucleotide RNA in maturation of histone mRNA precursors. Proc. Natl. Acad. Sci. U.S.A. 81; 1057-1061.

Birkenmeier, E.H., Brown, D.D. and Jordan, E. (1978). A nuclear extract of Xenopus laevis oocytes that accurately transcribes 5S RNA genes. Cell 15; 1077-1086.

Birnstiel, M.L., Busslinger, M. and Strub, K. (1985). Transcription termination and 3' processing: the end is in site! Cell 41; 349-359.

Biswal, N., Feldan, P. and Levy, C.C. (1983). A DNA topoisomerase activity copurifies with the DNA polymerase induced by herpes simplex virus. Biochim. Biophys. Acta 740; 379-389.

Biswas, C., Hardy, J. and Beck, W.S. (1965). Release of repressor control of ribonucleotide reductase by thymine starvation. J. Biol. Chem. 240; 3631-3640.

Blake, C.C.F. (1978). Do genes-in-pieces imply proteins-in-pieces? Nature 273; 267-268.

Blattner, F.R. and Tucker, P.W. (1984). The molecular biology of immunoglobulin D. Nature 307; 417-422.

Blue, W.T. and Stobbs, D.G. (1981). Isolation of a protein kinase induced by herpes simplex virus type 1. J. Virol. 38; 383-388.

Bogenhagen, D.F. and Brown, D.D. (1981). Nucleotide sequences in Xenopus 5S DNA required for transcription termination. Cell 24; 261-270.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977). Construction and characterisation of new cloning vehicles. II. A multi-purpose cloning system. Gene 2; 95-113.

Bornkamm, G.W., Delius, H., Fleckenstein, B., Werner, F.J. and Mulder, C. (1976). Structure of herpesvirus saimiri genomes: arrangement of heavy and light sequences in the M genome. J. Virol. 19; 154-161.

Borrelli, E., Hen, R. and Chambon, P. (1984). Adenovirus-2 E1A products repress enhancer-induced stimulation of transcription. Nature 312; 608-612.

Both, G.W., Banerjee, A.K. and Shatkin, A.J. (1975). Methylation-dependent translation of viral messenger RNAs in vitro. Proc. Natl. Acad. Sci. U.S.A. 72; 1189-1193.

Bowman, L.H., Rabin, B. and Schlessinger, D. (1981). Multiple ribosomal RNA cleavage pathways in mammalian cells. Nucl. Acids Res. 9; 4951-4966.

Boyd, A.L., Derge, J.G. and Hampar, B. (1978). Activation of endogenous type C virus in Balb/c mouse cells by herpesvirus DNA. Proc. Natl. Acad. Sci. U.S.A. 75; 4558-4562.

Boyd, A.L., Enquist, L., Vande Woude, G.F. and Hampar, B. (1980). Activation of mouse retrovirus by herpes simplex virus type 1 cloned DNA fragments. Virology 103; 228-231.

Boyer, H.W. and Roulland-Dussoix, D. (1969). A complementation analysis of the restriction modification of DNA in Escherichia coli. J. Mol. Biol. 41; 459-472.

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

Braun, D.K., Batterson, W. and Roizman, B. (1984). Identification and genetic mapping of a HSV capsid protein that binds DNA. J. Virol. 50; 645-648.

Brawerman, G. (1981). The role of poly(A) sequence in mammalian messenger RNA. C.R.C. Crit. Rev. Biochem. 10; 1-38.

Breathnach, R. and Chambon, P. (1981). Organisation and expression of eukaryotic split genes coding for proteins. Ann. Rev. Biochem. 50; 349-383.

Broker, T.R. (1980). Nucleotide sequences, transcription and translation analyses, and restriction endonuclease cleavage map of group-A human adenoviruses. In 'DNA Tumour Viruses.' pp. 1023-1036. (J. Tooze, ed.) Cold Spring Harbor Laboratory.

Brown, N.C. and Reichard, P. (1969a). Ribonucleoside diphosphate reductase. Formation of active and inactive complexes of proteins B1 and B2. J. Mol. Biol. 46; 25-38.

Brown, N.C. and Reichard, P. (1969b). Role of effector binding in allosteric control of ribonucleoside diphosphate reductase. J. Mol. Biol. 46; 39-55.

Brown, N.C., Eliasson, R., Reichard, P. and Thelander, L. (1969a). Spectrum and iron content of protein B2 from ribonucleoside diphosphate reductase. Eur. J. Biochem. 9; 512-518.

Brown, N.C., Canellakis, Z.N., Lundin, B., Reichard, P. and Thelander, L. (1969b). Ribonucleotide diphosphate reductase. Purification of the two subunits, proteins B1 and B2. Eur. J. Biochem. 9; 561-573.

Brown, R.D. and Brown, D.D. (1976). The nucleotide sequence adjoining the 3' end of the genes coding for oocyte-type 5S ribosomal RNA in Xenopus. J. Mol. Biol. 102; 1-14.

Brown, S.M., Ritchie, D.A. and Subak-Sharpe, J.H. (1973). Genetic studies with herpes simplex virus type 1. The isolation of temperature sensitive mutants, their arrangement into complementation groups and recombination analysis leading to a linkage map. J. Gen. Virol. 18; 329-346.

Buckmaster, E.A., Gompels, U. and Minson, A. (1984). Characterisation and physical mapping of an HSV-1 glycoprotein of approximately $115x10^3$ molecular weight. Virology, 139; 408-413.

Buetti, E. (1974). Characterisation of late polyoma mRNA. J. Virol. 14; 249-260.

Busby, D.W.G., House, W. and McDonald, J.R. (1964). In 'Virological Techniques.' Churchill, London.

Busch, H., Reddy, R., Rothblum, L. and Choi, Y.C. (1982). SnRNAs, snRNPs and RNA processing. Ann. Rev. Biochem. 51; 617-654.

Busslinger, M., Portman, R. and Birnstiel, M.L. (1979). A regulatory sequence near the 3' end of sea urchin histone genes. Nucl. Acids Res. 6; 2997-3008.

Busslinger, M., Portman, R., Irminger, J.C. and Birnstiel, M.L. (1980). Ubiquitous and gene-specific regulatory 5' sequences in a sea urchin histone DNA clone coding for histone protein variants. Nucl. Acids Res. 8; 957-977.

Busslinger, M., Moschonas, N. and Flavell, R.A. (1981). β^+-thalassaemia: aberrant splicing results from a single point mutation in an intron. Cell 27; 289-298.

Bzik, D.J., Fox, B.A., DeLuca, N.A. and Person, S. (1984). Nucleotide sequence specifying the glycoprotein gene, gB, of herpes simplex virus type 1. Virology 133; 301-314.

Camacho, A. and Spear, P. (1978). Transformation of hamster embryo fibroblasts by a specific fragment of the herpes simplex virus genome. Cell 15; 993-1002.

Cameron, I.R. and Macnab, J.C.M. (1980). Transformation studies using defined fragments of herpes simplex virus type 2. In 'The Human Herpesviruses.' p. 634. (A.J. Nahmias, W.R. Dowdle, and R.F. Schinazi, eds.) Elsevier, North-Holland, New York.

Cameron, I.R., Park, M., Dutia, B.M., Orr, A. and Macnab, J.C.M. (1985). Herpes simplex virus sequences involved in the initiation of oncogenic morphological transformation of rat cells are not required for maintenance of the transformed state. J. Gen. Virol. 66; 517-527.

Campbell, M.E.M., Palfreyman, J.W. and Preston, C.M. (1984). Identification of herpes simplex virus DNA sequences which encode a trans-acting polypeptide responsible for stimulation of immediate early transcription. J. Mol. Biol. 180; 1-19.

Caradonna, S.J. and Cheng, Y.C. (1981). Induction of uracil-DNA glycosylase and dUTP nucleotidylhydrolase activity in herpes simplex virus-infected human cells. J. Biol. Chem. 256; 9834-9837.

Caras, I.W., Jones, T., Eriksson, S. and Martin, D.W.Jr. (1983). Direct photoaffinity labelling of the catalytic site of mouse ribonucleotide reductase by CDP. J. Biol. Chem. 258; 3064-3068.

Caras, I.W., Levinson, B.B., Fabry, M., Williams, S.R. and Martin, D.W.Jr. (1985). Cloned mouse ribonucleotide reductase subunit M1 cDNA reveals amino acid sequence homology with Escherichia coli and herpesvirus ribonucleotide reductases. J. Biol. Chem. 260; 7015-7022.

Carlson, J., Fuchs, J.A. and Messing, J. (1984). The primary structure of the Escherichia coli ribonucleotide diphosphate reductase operon. Proc. Natl. Acad. Sci. U.S.A. 81; 4294-4297.

Chai, L. (1971). Architecture of the DNA molecule within Epstein-Barr virus (EBV). Proc. Amer. Assoc. for Cancer Res. 12, abstract no. 71.

Chan, V.L., Guttman, S. and Juranka, P. (1981). Mutator genes of baby hamster kidney cells. Mol. Cell. Biol. 1; 568-571.

Chaney, S.M.J., Warren, K.G., Kettyls, J., Zbitnue, A. and Subak-Sharpe, J.H. (1983). A comparative analysis of restriction enzyme digests of the DNA of herpes simplex virus isolated from genital and facial lesions. J. Gen. Virol. 64; 357-371.

Chang, A.C.Y., Cochet, M. and Cohen, S.N. (1980). Structural organisation of human genomic DNA encoding the pro-opiomelanocortin peptide. Proc. Natl. Acad. Sci. U.S.A. 77; 4890-4894.

Chang, C.-H. and Cheng, Y. (1979). Effects of nucleoside triphosphates on human ribonucleotide reductase from Molt-4F cells. Cancer Res. 39; 5087-5092.

Charnay, P., Treisman, R., Mellon, P., Chao, M., Axel, R. and Maniatis, T. (1984). Differences in human α - and β-globin gene expression in mouse erythroleukaemia cells: the role of intragenic sequences. Cell 38; 251-263.

Chartrand, P., Stow, N.D., Timbury, M.C. and Wilkie, N.M. (1979). Physical mapping of paa^r mutations of herpes simplex viruses types 1 and 2 by intertypic marker rescue. J. Virol. 31; 265-276.

Chartrand, P., Crumpacker, C.S., Schaffer, P.A. and Wilkie, N.M. (1980). Physical and genetic analysis of the herpes simplex virus DNA polymerase locus. Virology 103; 311-326.

Chase, J.W. and Richardson, C.C. (1974). Exonuclease VII of Escherichia coli. J. Biol. Chem. 249; 4553-4561.

Chen, A.K., Bhan, A., Hopper, S., Abrams, R. and Franzen, J.S. (1974). Substrate and effector binding to ribonucleoside triphosphate reductase of Lactobacillus leichmannii. Biochemistry 13; 654-661.

Chen, E.Y., Howley, P.M., Levinson, A.D. and Seeburg, P.H. (1982). The primary structure and genetic organisation of the bovine papillomavirus type 1 genome. Nature 299; 529-534.

Cheng, H.-L., Blattner, F.R., Fitzmaurice, L., Mushinski, J.F. and Tucker, P.W. (1982). Structure of genes for membrane and secreted murine IgD heavy chains. Nature 296; 410-415.

Cheng, Y.C., Goz, B. and Prusoff, W.H. (1975). Deoxyribonucleotide metabolism in herpes simplex virus infected HeLa cells. Biochim. Biophys. Acta 390; 253-263.

Cheung, A. and Kieff, E. (1981). Epstein-Barr virus DNA. X. A direct repeat within the internal direct repeat of EBV DNA. J. Virol. 40; 501-507.

Cheung, A. and Kieff, E. (1982). Epstein-Barr virus DNA. XI. The nucleotide sequence of the large internal repeat in EBV DNA. J. Virol. 44; 286-294.

Cheung, S., Arndt, K. and Lu, P. (1984). Correlation of lac operator DNA imino protein exchange rates with its function. Proc. Natl. Acad. Sci. U.S.A. 81; 3665-3669.

Childs, G., Nocente-McGrath, C., Lieber, T., Holt, C. and Knowles, J.A. (1982). Sea urchin (Lytechinus pictus) late-stage histone H3 and H4 genes: characterisation and mapping of a clustered but non-tandemly linked multi-gene family. Cell 31; 383-393.

Chiu, C.-S., Cox, S.M. and Greenberg, G.R. (1980). Effect of bacteriophage T4 nrd mutants on deoxyribonucleotide synthesis in vivo. J. Biol. Chem. 255; 2747-2751.

Chiu, C.-S., Cook, K.S. and Greenberg, G.R. (1982). Characteristics of a bacteriophage T4-induced complex synthesising deoxyribonucleotides. J. Biol. Chem. 257; 15087-15097.

Chou, J. and Roizman, B. (1985). Isomerisation of the herpes simplex virus 1 genome: identification of the cis-acting and recombination sites within the domain of the 'a' sequence. Cell 41; 803-811.

Chousterman, S., Lacasa, M. and Sheldrick, P. (1979). Physical map of the channel catfish virus genome: location of sites for restriction endonucleases EcoRI, HindIII, HpaI and XbaI. J. Virol. 31; 73-85.

Chow, L.T., Broker, T. and Lewis, J.B. (1979). The complex splicing patterns of RNA from the early regions of adenovirus. J. Mol. Biol. 134; 265-303.

Churchill, A.E. and Biggs, P.M. (1967). Agent of Marek's disease in tissue culture. Nature 215; 528-530.

Citron, B., Falck-Pedersen, E., Salditt-Georgieff, M. and Darnell, J.E. (1984). Transcription termination occurs within a 1000 base pair region downstream from the poly(A) site of the mouse β-globin (major) gene. Nucl. Acids Res. 12; 8723-8731.

Clark, W., Frame, M.C. and Dutia, B.M. (1985). manuscript in preparation.

Clements, J.B., Cortini, R. and Wilkie, N.M. (1976). Analysis of herpesvirus substructure by means of restriction endonucleases. J. Gen. Virol. 30; 243-256.

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

Clements, J.B., McLauchlan, J. and McGeoch, D.J. (1979). Orientation of herpes simplex virus type 1 immediate-early mRNAs. Nucl. Acids Res. 7; 77-91.

Clewell, D.B. (1972). Nature of ColE1 plasmid replication in Escherichia coli in the presence of chloramphenicol. J. Bacteriol. 110; 667-676.

Coffin, J. (1982). Structure of the retroviral genome. In 'RNA Tumour Viruses.' 2nd ed. pp. 261-368. (J. Tooze, ed.) Cold Spring Harbor Laboratory.

Cohen, E.A, Charron, J., Perret, J. and Langelier, Y. (1985). Herpes simplex virus ribonucleotide reductase induced in infected BHK-21/C13 cells: biochemical evidence for the existence of two non-identical subunits, H1 and H2. J. Gen. Virol. 66; 733-745.

Cohen, G.H. (1972). Ribonucleotide reductase activity of synchronised KB cells infected with herpes simplex virus. J. Virol. 9; 408-418.

Cohen, G.H., Factor, M.N. and Ponce de Leon, M. (1974). Inhibition of herpes simplex virus type 2 replication by thymidine. J. Virol. 14; 20-25.

Cohen, G.H., Ponce De Leon, M., Diggelman, H., Lawrence, W.C., Vernon, S.K. and Eisenberg, R.J. (1980). Structural analysis of the capsid polypeptides of herpes simplex virus types 1 and 2. J. Virol. 34; 521-531.

Cohen, J.C., Henry, B.E., Randall, C.C. and O'Callaghan, D.J. (1977). Ribonucleotide reductase activity in hydroxyurea-resistant herpesvirus replication. Proc. Soc. Exp. Biol. Med. 155; 395-399.

Cohen, L.H., Westerhuis, L.W., de Jong, W.W. and Bloemendal, H. (1978). Rat α-crystallin A chain with an insertion of 22 residues. Eur. J. Biochem. 89; 259-266.

Cohen, S.S. and Barner, H.D. (1962). Spermidine in the extraction of the deoxyribosyl-synthesising system from T6r$^+$-infected Escherichia coli. J. Biol. Chem. 237; 1376-1378.

Cole, C.N. and Stacy, T.P. (1985). Identification of sequences in the herpes simplex virus thymidine kinase gene required for efficient processing and polyadenylation. Mol. Cell. Biol. 5; 2104-2113.

Collins, A.R., Schor, S. and Johnson, R.T. (1977). The inhibition of repair in UV irradiated human cells. Mutat. Res. 42; 413-432.

Compton, J.G., Schrader, W.T. and O'Malley, B.W. (1983). DNA sequence preference of the progesterone receptor. Proc. Natl. Acad. Sci. U.S.A. 80; 16-20.

Conway, L. and Wickens, M. (1985). A sequence downstream of AAUAAA is required for formation of simian virus 40 late mRNA 3' termini in frog oocytes. Proc. Natl. Acad. Sci. U.S.A. 82; 3949-3953.

Cook, K.S. and Greenberg, G.R. (1983). Properties of bacteriophage T4 ribonucleoside diphosphate reductase subunits coded by nrdA and nrdB mutants. J. Biol. Chem. 258; 6064-6072.

Cooper, J.A., Wittek, R. and Moss, B. (1981). Extension of the transcriptional and translational map of the left end of the vaccinia virus genome to 21 kilobase pairs. J. Virol. 39; 733-745.

Cooper, M.D., Kearney, J.F., Lydward, P.M., Grossi, C.E. and Lawton, A.R. (1976). Studies of generation of B-cell diversity in mouse, man and chicken. Cold Spring Harbor Symp. Quant. Biol. 41; 139-147.

Cordell, B., Bell, G., Tischer, E., DeNoto, F.M., Ullrich, A., Pictet, R., Rutter, W.J. and Goodman, H.M. (1979). Isolation and characterisation of a cloned rat insulin gene. Cell 18; 533-543.

Cordingley, M.G., Campbell, M.E. and Preston, C.M. (1983). Functional analysis of a herpes simplex type 1 promoter: identification of far-upstream regulatory sequences. Nucl. Acids Res. 11; 2347-2365.

Corsalo, C.M. and Pearson, M.L. (1981). Enhancing the efficiency of DNA-mediated gene transfer in mammalian cells. Genetics 7; 603-616.

Cortini, R. and Wilkie, N.M. (1978). Physical maps for HSV type 2 DNA with restriction endonucleases. J. Gen. Virol. 39; 259-280.

Cory, J.G. and Mansell, M.M. (1975). Studies on mammalian ribonucleotide reductase inhibition by pyridoxal phosphate and the dialdehyde derivatives of adenosine, adenosine 5'-monophosphate and adenosine 5'-triphosphate. Cancer Res. 35; 390-396.

Costa, R.H., Devi, G., Anderson, K.P., Gaylord, B.H. and Wagner, E.K. (1981). Characterisation of a major late herpes simplex virus type 1 mRNA. J. Virol. 38; 483-496.

Costa, R.H., Draper, K.G., Banks, L., Powell, K.L., Cohen, G., Eisenberg, R. and Wagner, E.K. (1983). High resolution characterisation of herpes simplex virus type 1 transcripts encoding alkaline exonuclease and a 50,000-dalton protein tentatively identified as a capsid protein. J. Virol. 48; 591-603.

Costa, R.H., Cohen, G., Eisenberg, R., Long, D. and Wagner, E.K. (1984). Direct demonstration that the abundant 6-kilobase herpes simplex virus type 1 mRNA mapping between 0.23 and 0.27 map units enodes the major capsid protein VP5. J. Virol. 49; 287-292.

Costa, R.H., Draper, K.G., Kelly, T.J. and Wagner, E.K. (1985). An unusual spliced herpes simplex virus type 1 transcript with sequence homology to Epstein-Barr virus DNA. J. Virol. 54; 317-328.

Costanzo, F., Campadelli-Fiume, G., Foa-Tomasi, L. and Cassai, E. (1977). Evidence that herpes simplex virus DNA is transcribed by cellular RNA polymerase B. J. Virol. 21; 996-1001.

Cozzarelli, N.R., Gerrard, S.P., Schlissel, M., Brown, D.D. and Bogenhagen, D.F. (1983). Purified RNA polymerase III accurately and efficiently terminates transcription of 5S RNA genes. Cell 34; 829-835.

Crawford, L.V. and Lee, A.J. (1964). The nucleic acid of human cytomegalovirus. Virology 23; 105-107.

Creighton, T.E. (1983). 'Proteins. Structures and Molecular Properties.' W.H. Freeman and Co., New York.

Cullen, B.R., Lomedico, P.T. and Ju, G. (1984). Transcriptional interference in avian retroviruses - implications for the promoter insertion model of leukaemogenesis. Nature 307; 241-245.

Czernilofsky, A.P., DeLorbe, W., Swanstrom, R., Varmus, H.E., Bishop, J.M., Tischer, E. and Goodman, H.M. (1980). The nucleotide sequence of an untranslated but conserved domain at the 3' end of the avian sarcoma virus genome. Nucl. Acids Res. 8; 2967-2984.

Dalrymple, M.A., McGeoch, D.J., Davison, A.J. and Preston, C.M. (1985). DNA sequence of the herpes simplex virus type 1 gene whose product is responsible for transcriptional activation of immediate early promoters. Nucl. Acids Res. 13; 7865-7879.

Dalziel, R.G. and Marsden, H.S. (1984). Identification of two herpes simplex virus type 1-induced proteins (21K and 22K) which interact specifically with the 'a' sequence of herpes simplex virus DNA. J. Gen. Virol. 65; 1467-1475.

Danos, O., Katinka, M. and Yaniv, M. (1982). Human papillomavirus 1a complete DNA sequence: a novel type of genome organisation among Papovaviridae. EMBO J. 1; 231-236.

Darlington, R.W. and Randall, C.C. (1963). The nucleic acid content of equine abortion virus. Virology 19; 322-327.

Darnell, J.E. (1982). Variety in the level of gene control in eukaryotic cells. Nature 297; 365-371.

Davidson, I. and Stow, N.D. (1985). Expression of an immediate early polypeptide and activation of a viral origin of DNA replication in cells containing a fragment of herpes simplex virus DNA. Virology 141; 77-88.

Davison, A.J. (1983). DNA sequence of the U_S component of the varicella-zoster virus genome. EMBO J. 2; 2203-2209.

Davison, A.J. (1984). Structure of the genome termini of varicella-zoster virus. J. Gen. Virol. 65; 1969-1977.

Davison, A.J. and Wilkie, N.M. (1981). Nucleotide sequences of the joint between the L and S segments of herpes simplex virus types 1 and 2. J. Gen. Virol. 55; 315-331.

Davison, A.J and Wilkie, N.M. (1983a). Location and orientation of homologous sequences in the genomes of five herpesviruses. J. Gen. Virol. 64; 1927-1942.

Davison, A.J and Wilkie, N.M. (1983b). Inversion of the two segments of the herpes simplex virus genome in intertypic recombinants. J. Gen. Virol. 64; 1-18.

Debroy, C., Pederson, N., and Person, S. (1985). Nucleotide sequence of a herpes simplex virus type 1 gene that causes cell fusion. Virology 145; 36-48.

de Jong, W.W., Cohen, L.H., Leunissen, J.A.M. and Zweers, A. (1980). Internally elongated rodent α-crystallin A chain: resulting from incomplete RNA splicing. Biochem. Biophys. Res. Commun. 96; 648-655.

Delius, H. and Clements, J.B. (1976). A partial denaturation map of herpes simplex virus type 1 DNA: evidence for inversions of the unique DNA regions. J. Gen. Virol. 33; 125-133.

DeLuca, N., Bzik, D.J., Bond, V.C., Person, S. and Snipes, W. (1982). Nucleotide sequences of herpes simplex virus type 1 (HSV-1) affecting virus entry, cell fusion, and production of glycoprotein gB (VP7). Virology 122; 411-423.

DeLuca, N.A., Courtney, M.A. and Schaffer, P.A. (1984). Temperature-sensitive mutants in herpes simplex virus type 1 ICP4 permissive for early gene expression. J. Virol. 52; 767-776.

DeMarchi, J.M., Blankenship, M.L., Brown, G.D. and Kaplan, A.S. (1978). Size and complexity of human cytomegalovirus DNA. Virology 89; 643-646.

Denhardt, D.T. (1966). A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Comm. 23; 641-646.

Dennis, D. and Smiley, J.R. (1984). Transactivation of a late herpes simplex virus promoter. Mol. Cell. Biol. 4; 544-551.

DeNoto, F.M., Moore, D.D. and Goodman, H.M. (1981). Human growth hormone DNA sequence and mRNA structure: possible alternative splicing. Nucl. Acids Res. 9; 3719-3730.

Derse, D., Bastow, K.F. and Cheng, Y.-C. (1982). Characterisation of the DNA polymerases induced by a group of herpes simplex virus type 1 variants selected for growth in the presence of phosphonoformic acid. J. Biol. Chem. 257; 10251-10260.

de The, G., Geser, A., Day, N.E., Tukei, P.M., Williams, E.H., Beri, D.P., Smith, P.G., Dean, A.G., Bornkamm, G.W., Feorino, P. and Henle, W. (1978). Epidemiological evidence for a causal relationship between Epstein-Barr virus and Burkitt's lymphoma: results of the Ugandan prospective study. Nature 274; 756-761.

Devare, S.G., Reddy, E.P., Law, J.D., Robbins, K.C. and Aaronson, S.A. (1983). Nucleotide sequence of the simian sarcoma virus genome: demonstration that its acquired cellular sequences encode the transforming gene product p28sis. Proc. Natl. Acad. Sci. U.S.A. 80; 731-735.

de Villiers, J. and Schaffner, W. (1981). A small segment of polyoma virus DNA enhances the expression of a cloned β-globin gene over a distance of 1400 base pairs. Nucl. Acids Res. 9; 6251-6264.

de Villiers, J., Olson, J., Tyndall, C. and Schaffner, W. (1982). Transcriptional enhancers from SV40 and polyoma virus show a cell type preference. Nucl. Acids Res. 10; 7965-7976.

Dierks, P., van Ooyen, A., Mantei, N. and Weissman, C. (1981). DNA sequences preceding the rabbit β-globin gene are required for formation in mouse L cells of β-globin RNA with the correct 5' terminus. Proc. Natl. Acad. Sci. U.S.A. 78; 1411-1415.

Dierks, P., van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. and Weissmann, C. (1983). Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit β-globin gene in mouse 3T6 cells. Cell 32; 695-706.

Dijkema, R., Dekker, B.M.M. and van Ormondt, H. (1982). Gene organisation of the transforming region of adenovirus type 7 DNA. Gene 18; 143-156.

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

Docherty, J.J., Subak-Sharpe, J.H. and Preston, C.M. (1981). Identification of a virus-specific polypeptide associated with a transforming fragment (Bgl II-N) of herpes simplex virus type 2 DNA. J. Virol. 40; 126-132.

Dodgson, J.B. and Engel, J.D. (1983). The nucleotide sequence of the adult chicken α-globin genes. J. Biol. Chem. 258; 4623-4629.

Dolan, M., Dodgson, J.B. and Engel, J.D. (1983). Analysis of the adult chicken β-globin gene: nucleotide sequence of the locus, microheterogeneity at the 5' end of β-globin mRNA and aberrant nuclear RNA species. J. Biol. Chem. 258; 3983-3990.

Domdey, H., Apostol, B., Lin, R.-J., Newman, A., Brody, E. and Abelson, J. (1984). Lariat structures are in vivo intermediates in yeast pre-mRNA splicing. Cell 39; 611-621.

Donehower, L.A., Huang, A.L. and Hager, G.L. (1981). Regulatory and coding potential of the mouse mammary tumor virus long terminal redundancy. J. Virol. 37; 226-238.

Doolittle, W.F. (1978). Genes in pieces: were they ever together? Nature 272; 581-582.

Dowbenko, D.J. and Lasky, L.A. (1984). Extensive homology between the HSV-2 gF gene and the HSV-1 gC gene. J. Virol. 52; 154-163.

Draper, K.G., Frink, R.J. and Wagner, E.K. (1982). Detailed characterisation of an apparently unspliced β herpes simplex virus type 1 gene mapping in the interior of another. J. Virol. 43; 1123-1128.

Draper, K.G., Costa, R.H., Lee, G. T.-Y., Spear, P.G. and Wagner, E.K. (1984a). Molecular basis of the gC-negative phenotype of HSV-1 macroplaque strain. J. Virol. 51; 578-585.

Draper, K.G., Frink, R.J., Devi, G.B., Swain, M., Galloway, D.A. and Wagner, E.K. (1984b). Herpes simplex virus types 1 and 2 homology in the region between 0.58 and 0.68 map units. J. Virol. 52; 615-623.

Dreesman, G.R., Burek, J., Adam, E., Kaufman, R.H., Melnick, J.L., Powell, K. and Purifoy, D.J.M. (1980). Expression of herpesvirus induced antigens in human cervical cancer. Nature 283; 591-593.

Drew, H.R., Weeks, J.R. and Travers, A.A. (1985). Negative supercoiling induces spontaneous unwinding of a bacterial promoter. EMBO J. 4; 1025-1032.

Duff, R. and Rapp, F. (1973). Oncogenic transformation of hamster embryo cells after exposure to inactivated herpes simplex virus type 1. J. Virol. 12; 209-217.

Dumas, A.M., Geelen, J.L.M.C., Westrate, M.W., Wertheim, P. and van der Noordaa, J. (1981). XbaI, PstI and BglII restriction endonuclease maps of the two orientations of the varicella-zoster genome. J. Virol. 39; 390-400.

Dutia, B.M. (1983). Ribonucleotide reductase induced by herpes simplex virus has a virus-specified constituent. J. Gen. Virol. 64; 513-521.

Dynan, W.S. and Tjian, R. (1983). The promoter specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell 35; 79-87.

Dynan, W.S. and Tjian, R. (1985). Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature 316; 774-778.

Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. and Hood, L. (1980). Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways. Cell 20; 313-319.

Easton, A.J. and Clements, J.B. (1980). Temporal regulation of herpes simplex virus type 2 transcription and characterisation of virus immediate-early mRNAs. Nucl. Acids Res. 8; 2627-2645.

Ecker, J.R. and Hyman, R.W. (1981). Analysis of interruptions in the phosphodiester backbone of herpes simplex virus DNA. Virology 110; 213-216.

Ecker, J.R. and Hyman, R.W. (1982). Varicella-zoster DNA exists as two isomers. Proc. Natl. Acad. Sci. U.S.A. 79; 156-160.

Edbrooke, M.R., Parker, D., McVey, J.H., Riley, J.H., Sorenson, G.D., Pettengill, O.S. and Craig, R.K. (1985). Expression of the human calcitonin/CGRP gene in lung and thyroid carcinoma. EMBO J. 4; 715-724.

Efstratiadis, A., Posakony, J., Maniatis, T., Lawn, R., O'Connell, C., Spritz, R., deRiel, J., Forget, B., Weissman, S., Slightom, J., Blechl, A., Smithies, O., Baralle, F., Shoulders, C. and Proudfoot, N.J. (1980). The structure and evolution of the β-globin gene family. Cell 21; 653-668.

Eglin, R.P., Sharp, F., MacLean, A.B., Cordiner, J.W., Kitchener, H.C., Macnab, J.C.M., Clements, J.B. and Wilkie, N.M. (1981a). Herpes simplex virus type 2 coded material in cervical neoplasia. In 'Pre-clinical Neoplasia of the Cervix.' pp. 31-46. (J. Jordan, F. Sharp and A. Singer, eds.) Royal College of Obstetricians and Gynaecologists, London.

Eglin, R.P., Sharp, F., MacLean, A.B., Macnab, J.C.M., Clements, J.B. and Wilkie, N.M. (1981b). Detection of RNA complementary to herpes simplex virus DNA in human cervical squamous cell neoplasms. Cancer Res. 41; 3597-3604.

Ehrenberg, A. and Reichard, P. (1972). Electron spin resonance of the iron-containing protein B2 from ribonucleotide reductase. J. Biol. Chem. 247; 3485-3488.

Eisenberg, S.P., Coen, D.M. and McKnight, S.L. (1985). Promoter domains required for expression of plasmid-borne copies of the herpes simplex virus thymidine kinase gene in virus-infected mouse fibroblasts and microinjected frog oocytes. Mol. Cell. Biol. 5; 1940-1947.

Eklund, H., Nordstrom, B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T., Soderberg, B.-O., Tapia, O., Branden, C.-I. and Akeson, A. (1976a). Three-dimensional structure of horse liver alcohol dehydrogenase at 2.4A resolution. J. Mol. Biol. 102; 27-59.

Eklund, H., Branden, C.-I. and Jornvall, H. (1976b). Structural comparisons of mammalian, yeast and bacillar alcohol dehydrogenases. J. Mol. Biol. 102; 61-73.

Elkaim, R., Goding, C. and Kedinger, C. (1983). The adenovirus-2 EIIa early gene promoter: sequences required for efficient in vitro and in vivo transcription. Nucl. Acids Res. 11; 7105-7117.

ElKareh, A., Murphy, A.J.M., Fichter, T., Efstratiadis, A. and Silverstein, S. (1985). 'Transactivation' control signals in the promoter of the herpesvirus thymidine kinase gene. Proc. Natl. Acad. Sci. U.S.A. 82; 1002-1006.

Ellison, J., and Hood, L. (1982). Linkage and sequence homology of two human immunoglobulin gamma heavy chain constant region genes. Proc. Natl. Acad. Sci. U.S.A. 79; 1984-1988.

Engstrom, Y., Eriksson, S., Thelander, L. and Akerman, M. (1979). Ribonucleotide reductase from calf thymus. Purification and properties. Biochemistry 18; 2941-2948.

Engstrom, Y., Rozell, B., Hansson, H.-A., Stemme, S. and Thelander, L. (1984). Localisation of ribonucleotide reductase in mammalian cells. EMBO J. 3; 863-867.

Engstrom, Y., Eriksson, S., Jildevik, I., Skog, S., Thelander, L. and Tribukait, B. (1985). Cell cycle-dependent expression of mammalian ribonucleotide reductase. Differential regulation of the two subunits. J. Biol. Chem. 260; 9114-9116.

Epstein, M.A. (1962). Observations on the fine structure of mature herpes simplex virus and on the composition of its nucleoid. J. Exp. Med. 115; 1-9.

Epstein, M.A., Achong, B.G. and Barr, Y.M. (1964). Virus particles in cultured lymphoblasts from Burkitt's lymphoma. Lancet 1; 759-761.

Epstein, M.A., Henle, W., Achong, B.G. and Barr, Y.M. (1965). Morphological and biological studies on a virus in cultured lymphoblasts from Burkitt's lymphoma. J. Exp. Med. 121; 761-770.

Erbil, C. and Niessing, J. (1983). The primary structure of the duck α^D-globin gene: an unusual 5' splice junction sequence. EMBO J. 2; 1339-1343.

Eriksson, S. and Berglund, O. (1974). Bacteriophage-induced ribonucleotide reductase systems: T5- and T6-specific ribonucleotide reductase and thioredoxin. Eur. J. Biochem. 46; 271-278.

Eriksson, S. and Martin, D.W.Jr. (1981). Ribonucleotide reductase in cultured mouse lymphoma cells. J. Biol. Chem. 256; 9436-9440.

Eriksson, S., Thelander, L. and Akerman, M. (1979). Allosteric regulation of calf thymus ribonucleotide diphosphate reductase. Biochemistry 18; 2948-2952.

Eriksson, S., Graslund, A., Skog, S., Thelander, L. and Tribukait, B. (1984). Cell cycle-dependent regulation of mammalian ribonucleotide reductase. The S phase-correlated increase in subunit M2 is regulated by de novo protein synthesis. J. Biol. Chem. 259; 11695-11700.

Esparza, J., Benyesh-Melnick, M. and Schaffer, P.A. (1976). Intertypic complementation and recombination between temperature sensitive mutants of herpes simplex virus types 1 and 2. Virology 70; 372-384.

Evans, A.S. and Niederman, J.C. (1976). Epstein-Barr virus. In 'Viral infections of Humans: Epidemiology and Control.' pp. 209-234. (A.S. Evans, ed.) Plenum Press, New York.

Evans, A.S., Niederman, J.C. and McCollum, R.W. (1968). Seroepidemiologic studies of infectious mononucleosis with EB virus. New England J. Med. 279; 1121-1127.

Evans, T., Rosenthal, E., Youngblom, J., Distel, D. and Hunt, T. (1983). Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. Cell 33; 389-396.

Everett, R.D. (1983). DNA sequence elements required for regulated expression of the HSV-1 glycoprotein D gene lie within 83 bp of the RNA capsites. Nucl. Acids Res. 11; 6647-6666.

Everett, R.D. (1984a). A detailed analysis of an HSV-1 early promoter: sequences involved in trans-activation by viral immediate-early gene products are not early-gene specific. Nucl. Acids Res. 12; 3037-3056.

Everett, R.D. (1984b). Trans activation of transcription by herpes virus products: requirement for two HSV-1 immediate-early polypeptides for maximum activity. EMBO J. 3; 3135-3141.

Everett, R.D. (1985). Activation of cellular promoters during herpes virus infection of biochemically transformed cells. EMBO J. 4; 1973-1980.

Everett, R.D. and Dunlop. M. (1984). Trans activation of plasmid-borne promoters by adenovirus and several herpes group viruses. Nucl. Acids Res. 12; 5969-5978.

Everett, R.D., Baty, D. and Chambon, P (1983). The repeated GC-rich motifs upstream from the TATA box are important elements of the SV40 early promoter. Nucl. Acids Res. 11; 2447-2464.

Fahrner, K., Yarger, J. and Hereford, L. (1980). Yeast histone mRNA is polyadenylated. Nucl. Acids Res. 8; 5725-5737.

Falck-Pedersen, E., Logan, J., Shenk, T. and Darnell, J.E.Jr. (1985). Transcription termination within the Ela gene of adenovirus induced by insertion of the mouse β -major globin terminator element. Cell 40; 897-905.

Falk, L., Deinhardt, F., Nonoyama, M., Wolfe, L.G., Bergholtz, C., Lapin, B., Yakovlera, L., Agrba, V., Henle, G. and Henle, W. (1976). Properties of a baboon lymphotropic herpesvirus related to Epstein-Barr virus. Int. J. Cancer 18; 798-807.

Felber, B.K., Orkin, S.H. and Hamer, D.H. (1982). Abnormal RNA splicing causes one form of α-thalassaemia. Cell 29; 895-902.

Feldman, L.T., Imperiale, M.J. and Nevins, J.R. (1982). Activation of early adenovirus transcription by the herpesvirus immediate early gene: evidence for a common cellular control factor. Proc. Natl. Acad. Sci. U.S.A. 79; 4952-4956.

Feller, W. and Follman, H. (1976). Ribonucleotide reductase activity in green algae. Biochem. Biophys. Res. Commun. 70; 752-758.

Fenner, F. (1976). Classification and nomenclature of viruses. Intervirology 7; 1-115.

Fenwick, M.L. and Walker, M.J. (1978). Suppression of the synthesis of cellular macromolecules by herpes simplex virus. J. Gen. Virol. 41; 37-51.

Fenwick, M.L., Morse, L.S. and Roizman, B. (1979). Anatomy of herpes simplex virus DNA. XI. Apparent clustering of functions effecting rapid inhibition of host DNA and protein synthesis. J. Virol. 29; 825-827.

Fetherston, J., Werner, E. and Patterson, R. (1984). Processing of the external transcribed spacer of murine rRNA and site of action of actinomycin D. Nucl. Acids Res. 12; 7187-7198.

Fiddes, J.C. and Goodman, H.M. (1981). The gene encoding the common alpha subunit of the four human glycoprotein hormones. J. Mol. Appl. Genet. 1; 3-18.

Fiddian, A.P., Yeo, J.M. and Clark, A.E. (1983). Treatment of herpes labialis. J. Infect. 6, Suppl. 1; pp. 41-48.

Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voode, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. and Ysehaert, M. (1978). Complete nucleotide sequence of SV40 DNA. Nature 273; 113-120.

Fitzgerald, M. and Shenk, T. (1981). The sequence 5'-AAUAAA-3' forms part of the recognition site for polyadenylation of late SV40 mRNAs. Cell 24; 251-260.

Flanagan, J.F. (1967). Virus-specified ribonucleic acid synthesis in KB cells infected with herpes simplex virus. J. Virol. 1; 583-590.

Flanegan, J.B. and Greenberg, G.R. (1977). Regulation of deoxyribonucleotide biosynthesis during *in vivo* bacteriophage T4 DNA replication: intrinsic control of synthesis of thymine and 5-hydroxymethylcytosine deoxyribonucleotides at precise ratio found in DNA. J. Biol. Chem. 252; 3019-3027.

Flannery, V.L., Courtney, R.J. and Schaffer, P.A. (1977). Expression of an early, nonstructural antigen of herpes simplex virus in cells transformed *in vitro* by herpes simplex virus. J. Virol. 21; 284-291.

Fleckenstein, B. and Wolf, H. (1974). Purification and properties of herpesvirus saimiri DNA. Virology 58; 55-64.

Fleckenstein, B. and Bornkamm, G.W. (1975). Structure and function of H. saimiri DNA. In 'Oncogenesis and Herpesviruses II.' part 1, pp. 145-150. (G. de The, H. zur Hausen and M.A. Epstein, eds.) M.A. IARC Press, Lyon.

Fleckenstein, B. and Mulder, C. (1980). Molecular biological aspects of herpesvirus saimiri and herpesvirus ateles. In 'Viral Oncology.' pp. 799-812. (G. Klein, ed.) Raven Press, New York.

Fleckenstein, B., Bornkamm, G.W., Mulder, C., Werner, F.-J., Daniel, M.D., Falk, L.A. and Delius, H. (1978). Herpesvirus ateles DNA and its homology with herpesvirus saimiri nucleic acid. J. Virol. 25; 361-373.

Ford, J.P. and Hsu, M.-T. (1978). Transcription pattern of in vivo-labelled late simian virus 40 RNA: equimolar transcription beyond the mRNA 3' terminus. J. Virol. 28; 795-801.

Fornwald, J.A., Kuncio, G., Peng, I. and Ordahl, C.P. (1982). The complete nucleotide sequence of the chick ∝-actin gene and its evolutionary relationship to the actin gene family. Nucl. Acids Res. 10; 3861-3876.

Forsdyke, D.R. and Scott, F.W. (1979). In 'Cell Compartmentation and Metabolic Channeling.' pp. 177-184. (F. Lynden, K. Mothes and L. Nover, eds.) Elsevier, Amsterdam.

Frame, M.C., Marsden, H.S. and Dutia, B.M. (1985). The ribonucleotide reductase induced by herpes simplex virus type 1 involves minimally a complex of two polypeptides (136K and 38K). J. Gen. Virol. 66; 1581-1587.

Fraser, N.W., Nevins, J.R., Ziff, E. and Darnell, J.E. (1979). The major late adenovirus type-2 transcription unit: termination is downstream from the last poly(A) site. J. Mol. Biol. 129; 643-656.

Frendewey, D., Dingermann, T., Cooley, L. and Soll, D. (1985). Processing of precursor tRNAs in Drosophila. Processing of the 3' end involves an endonucleolytic cleavage and occurs after 5' end maturation. J. Biol. Chem. 260; 449-454.

Frenkel, N. and Roizman, B. (1972). Separation of the herpesvirus deoxyribonucleic acid duplex into unique fragments and intact strands on sedimentation in alkaline gradients. J. Virol. 10; 565-572.

Friedmann, A., Shlomai, J. and Becker, Y. (1977). Electron microscopy of herpes simplex virus DNA molecules isolated from infected cells by centrifugation in CsCl density gradients. J. Gen. Virol. 34; 507-522.

Frink, R.J., Anderson, K.P. and Wagner, E.K. (1981). Herpes simplex virus type 1 HindIII fragment 1 encodes spliced and complementary mRNA species. J. Virol. 39; 559-572.

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

Fromm, M. and Berg, P. (1982). Deletion mapping of DNA required for SV40 early promoter in vivo. J. Mol. Appl. Genet. 1; 457-481.

Fuchs, J.A. (1977). Coordinate control of the synthesis of ribonucleoside diphosphate reductase components in Escherichia coli. J. Bacteriol. 130; 957-959.

Fuchs, J.A. and Karlstrom, H.O. (1976). Mapping of nrdA and nrdB in Escherichia coli K-12. J. Bacteriol. 128; 810-814.

Fuchs, P.G., Ruger, R., Pfister, H. and Fleckenstein, B. (1985). Genome organisation of herpesvirus aotus type 2. J. Virol. 53; 13-18.

Fukumaki, Y., Ghosh, P.K., Benz, E.J.Jr., Reddy, V.B., Lebowitz, P., Forget, B.G. and Weissman, S.M. (1982). Abnormally spliced messenger RNA in erythroid cells from patients with β^+ thalassaemia and monkey cells expressing a cloned β^+ thalassaemic gene. Cell 28; 585-593.

Furlong, D., Swift, H. and Roizman, B. (1972). Arrangement of the herpesvirus deoxyribonucleic acid in the core. J. Virol. 10; 1071-1074.

Gaffney, D.F., McLauchlan, J., Whitton, J.L. and Clements, J.B. (1985). A modular system for the assay of transcription regulatory signals: the sequence TAATGARAT is required for herpes simplex virus immediate early gene activation. Nucl. Acids Res. 13; 7847-7863.

Galli, G., Hofstetter, H., Stunnenberg, H.G. and Birnstiel, M.L. (1983). Biochemical complementation with RNA in the Xenopus oocyte: a small RNA is required for the generation of histone mRNA 3' termini. Cell 34; 823-828.

Galloway, D.A. and McDougall, J.K. (1981). Transformation of rodent cells by a cloned DNA fragment of herpes simplex virus type 2. J. Virol. 38; 749-760.

Galloway, D.A. and McDougall, J.K. (1983). The oncogenic potential of herpes simplex viruses: evidence for a 'hit and run' mechanism. Nature 302; 21-24.

Galloway, D.A. and Swain, M.A. (1984). Organisation of the left-hand end of the herpes simplex virus type 2 Bgl II n fragment. J. Virol. 49; 724-730.

Galloway, D.A., Goldstein, L.C. and Lewis, J.B. (1982). Identification of proteins encoded by a fragment of herpes simplex virus type 2 DNA that has transforming ability. J. Virol. 42; 530-537.

Galloway, D.A., Nelson, J.A. and McDougall, J.K. (1984). Small fragments of herpesvirus DNA with transforming activity contain insertion sequence-like structures. Proc. Natl Acad. Sci. U.S.A. 81; 4736-4740.

Gannon, F., O'Hare, K., Perrin, F., LePennec, J.P., Benoist, C., Cochet, M., Breathnach, R., Royal, A., Garapin, A., Cami, B. and Chambon, P. (1979). Organisation and sequences at the 5' end of a cloned complete ovalbumin gene. Nature 278; 428-434.

Geelen, J.L.M.C. and Westrate, M.W. (1982). Organisation of the human cytomegalovirus genome. In 'Herpesvirus DNA.' pp. 325-343. (Y. Becker, ed.) Martinus Nijhoff Medical Publishers, The Hague, Holland.

Geelen, J.L.M.C., Walig, C., Wertheim, P. and van der Noordaa, J. (1978). Human cytomegalovirus DNA. I. Molecular weight and infectivity. J. Virol. 26; 813-816.

Gelb, L.D., Huang, J.J. and Wellinghoff, W.J. (1980). Varicella-zoster virus transformation of hamster embryo cells. J. Gen. Virol. 51; 171-177.

Gelman, I.H. and Silverstein, S. (1985). Identification of immediate early genes from herpes simplex virus that transactivate the virus thymidine kinase gene. Proc. Natl. Acad. Sci. U.S.A. 82; 5265-5269.

Georgiev, O. and Birnstiel, M.L. (1985). The conserved CAAGAAAGA spacer sequence is an essential element for the formation of 3' termini of the sea urchin H3 histone mRNA by RNA processing. EMBO J. 4; 481-489.

Ghosh, P.K. and Lebowitz, P. (1981). Simian virus 40 mRNAs contain multiple 5' termini upstream and downstream from a Hogness-Goldberg sequence; a shift in 5' termini during the lytic cycle is mediated by large T antigen. J. Virol. 40; 224-240.

Gibson, T., Stockwell, P., Ginsburg, M. and Barrell, B. (1984). Homology between two EBV early genes and HSV ribonucleotide reductase and 38k genes. Nucl. Acids Res. 12; 5087-5099.

Gil, A. and Proudfoot, N.J. (1984). A sequence downstream of AAUAAA is required for rabbit β-globin mRNA 3' end formation. Nature. 312; 473-474.

Gilbert, W. (1978). Why genes in pieces? Nature 271; 501.

Gilden, D., Wroblewska, Z., Kindt, V., Warren, K. and Wolinsky, J. (1978). Varicella-zoster virus infection of human brain cells and ganglion cells in tissue culture. Arch. Virol. 56; 105-117.

Gillies, S.D., Morrison, S.L., Oi, V.T. and Tonegawa, S. (1983). A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. Cell 33; 717-728.

Gilman, S.C., Docherty, J.J., Clarke, A. and Rawls, W.E. (1980). Reaction patterns of herpes simplex virus type 1 and type 2 proteins with sera of patients with uterine cervical cancer and matched controls. Cancer Res. 40; 4640-4647.

Gingeras, T.R., Sciaky, D., Gelinas, R.E., Bing-Dong, J., Yen, C.E., Kelly, M.M., Bullock, P.A., Parsons, B.L., O'Neill, K.E. and Roberts, R.J. (1982). Nucleotide sequences from the adenovirus 2 genome. J. Biol. Chem. 257; 13475-13491.

Giraldo, G., Beth, E., Kourilsky, G., Henle, W., Mike, V., Huraux, J.M., Andersen, H.K., Gharbi, M.R., Kyalwazi, S.K. and Puissant, A. (1975). Antibody patterns to herpesvirus in Kaposi's sarcoma: serological association of European Kaposi's sarcoma with cytomegalovirus. Int. J. Cancer 15; 839-848.

Given, D. and Kieff, E. (1978). DNA of Epstein-Barr virus. IV. Linkage map for restriction enzyme fragments of the B95-8 and W91 strains of EBV. J. Virol. 28; 524-542.

Given, D. and Kieff, E. (1979). DNA of Epstein-Barr virus. VI. Mapping of the internal tandem reiteration. J. Virol. 31; 315-324.

Given, D., Yee, D., Griem, K. and Kieff, E. (1979). DNA of Epstein-Barr virus. V. Direct repeats at the ends of EBV DNA. J. Virol. 30; 852-862.

Glanville, N., Durnam, D.M. and Palmiter, R.D. (1981). Structure of the mouse metallothionein-I gene and its mRNA. Nature 292; 267-269.

Gleason, F.K. and Hogenkamp, H.P.C. (1972). 5'-deoxyadenosylcobalamin-dependent ribonucleotide reductase: a survey of its distribution. Biochim. Biophys. Acta 277; 466-470.

Gleason, F.K. and Wood, J.M. (1976). Ribonucleotide reductase in blue-green algae: dependence on adenosylcobalamin. Science 192; 1343-1344.

Godowski, P.J. and Knipe, D.M. (1985). Identification of a herpes simplex virus function that represses late gene expression from parental viral genomes. J. Virol. 55; 357-365.

Gold, E. and Nankervis, G.A. (1976). Cytomegalovirus. In 'Viral Infections of Humans: Epidemiology and Control.' pp. 143-161. (A.S. Evans, ed.) Plenum Publishing Corp., New York.

Gomez-Marquez, J., Puga, A. and Notkins, A.L. (1985). Regions of the terminal repetitions of the herpes simplex virus type 1 genome. Relationship to immunoglobulin switch-like DNA sequences. J. Biol. Chem. 260; 3490-3495.

Goodheart, C.R., Plummer, G. and Waner, J.L. (1968). Density differences of DNA of human herpes simplex viruses types 1 and 2. Virology 35; 473-475.

Gordin, H., Olshevsky, U., Rosenkrantz, H.S. and Becker, Y. (1973). Studies on herpes simplex virus DNA: denaturation properties. Virology 55; 280-284.

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

Gotoh, O. and Tagashira, Y. (1981). Stabilities of nearest-neighbor doublets in double-helical DNA determined by fitting calculated melting profiles to observed profiles. Biopolymers 20; 1033-1042.

Gourse, R.L. and Gerbi, S.A. (1980). Fine structure of ribosomal RNA. IV. Extraordinary evolutionary conservation in sequences that flank introns in rDNA. Nucl. Acids Res. 8; 3623-3637.

Grabowski, P.J., Padgett, R.A. and Sharp, P.A. (1984). Messenger RNA splicing in vitro: an excised intervening sequence and a potential intermediate. Cell 37; 415-427.

Grafstrom, R.H., Alwine, J.C., Steinhart, W.L. and Hill, C.W. (1974). Terminal repetitions in herpes simplex virus type 1 DNA. Cold Spring Harbor Symp. Quant. Biol. 39; 679-681.

Grafstrom, R.H., Alwine, J.C., Steinhart, W.L., Hill, C.W. and Hyman, R.W. (1975). The terminal repetition of herpes simplex virus DNA. Virology 67; 144-157.

Graslund, A., Ehrenberg, A. and Thelander, L. (1982). Characterisation of the free radical of mammalian ribonucleotide reductase. J. Biol. Chem. 257; 5711-5715.

Gray, C.P. and Kaerner, H.C. (1984). Sequence of the putative origin of replication in the U_L region of herpes simplex virus type 1 Ang DNA. J. Gen. Virol. 65; 2109-2119.

Green, M.R., Treisman, R. and Maniatis, T. (1983). Transcriptional activation of cloned human β-globin genes by viral immediate-early gene products. Cell 35; 137-148.

Greenberg, G.R. and Chiu, C.-S. (1978). In vivo and in vitro evidence for a phage T4-induced deoxyribonucleotide synthetase/DNA replication complex. Fed. Proc. vol 39, Abstract no. 768.

Greenberg, J.R. and Perry, R. (1972). Relative occurrence of polyadenylic acid in messenger and heterogeneous nuclear RNA of L-cells as determined by poly(U)-hydroxylapatite chromatography. J. Mol. Biol. 72; 91-96.

Grose, C., Perrotta, D., Brunell, P. and Smith, G. (1979). Cell-free varicella-zoster virus in cultured human melanoma cells. J. Gen. Virol. 43; 15-27.

Grosschedl, R. and Birnstiel, M.L. (1980a). Identification of regulatory sequences in the prelude sequences of an H2A histone gene by the study of specific deletion mutants in vivo. Proc. Natl. Acad. Sci. U.S.A. 77; 1432-1436.

Grosschedl, R. and Birnstiel, M.L. (1980b). Spacer DNA sequences upstream of the T-A-T-A-A-A-T-A sequence are essential for promotion of H2A histone gene transcription in vivo. Proc. Natl. Acad. Sci. U.S.A. 77; 7102-7106.

Grosschedl, R., Wasylyk, B., Chambon, P. and Birnstiel, M.L. (1981). Point mutation in the TATA box curtails expression of sea urchin H2A histone gene in vivo. Nature 294; 178-180.

Grosveld, G.C., de Boer, E., Shewmaker, C.K. and Flavell, R.A. (1982a). DNA sequences necessary for transcription of the rabbit β-globin gene in vivo. Nature 295; 120-126.

Grosveld, G.C., Rosenthal, A. and Flavell, R.A. (1982b). Sequence requirements for the transcription of the rabbit β-globin gene in vivo. Nucl. Acids Res. 10; 4951-4971.

Gustafsohn, D.P. (1970). Pseudorabies in Diseases of Swine. 3rd ed. pp. 337-355, Iowa State Univ.

Haarr, L. and Marsden, H.S. (1981). Two dimensional gel analysis of HSV type 1 induced polypeptides and glycoprotein processing. J. Gen. Virol. 52; 77-92.

Haarr, L., Marsden, H.S., Preston, C.M., Smiley, J.R., Summers, W.C. and Summers, W.P. (1985). Utilisation of internal AUG codons for initiation of protein synthesis directed by mRNAs from normal and mutant genes encoding herpes simplex virus-specified thymidine kinase. J. Virol. 56; 512-519.

Haegeman, G. and Fiers, W. (1978). Characterisation of the 5' terminal cap structures of late simian virus 40-specific mRNA. J. Virol. 25; 824-830.

Haegeman, G. and Fiers, W. (1980). Characterisation of the 5' terminal cap structures of early simian virus 40 mRNA. J. Virol. 35; 955-961.

Hall, L.M., Draper, K.G., Frink, R.J., Costa, R.H. and Wagner, E.K. (1982). Herpes simplex virus mRNA species mapping in the Eco RI fragment I. J. Virol. 43; 594-607.

Halliburton, I.W. (1972). Biological comparison of type 1 and type 2 herpes simplex viruses. In 'Oncogenesis and Herpesviruses.' pp. 432-438. (P.M. Biggs, G. de The and L.M. Payne, eds.) IARC Scientific Publications, no. 2, Lyons.

Halliburton, I.W. and Timbury, M.C. (1973). Characterisation of temperature-sensitive mutants of herpes simplex virus type 2. Growth and DNA synthesis. Virology 54; 60-68.

Halliburton, I.W., Morse, L.S., Roizman, B. and Quinn, K.E. (1980). Mapping of the thymidine kinase genes of type 1 and type 2 herpes simplex virus using intertypic recombinants. J. Gen. Virol. 49; 235-254.

Halpern, M.E. and Smiley, J.R. (1984). Effects of deletions on expression of the HSV TK gene from the intact viral genome: the amino terminus of the enzyme is dispensable for catalytic activity. J. Virol. 50; 708-716.

Hamada, H., Muramatsu, M., Urano, Y., Onishi, T. and Kominami, R. (1979). In vitro synthesis of a 5S RNA precursor by isolated nuclei of rat liver and HeLa cells. Cell 17; 163-173.

Hampar, B.H. and Ellison, S.A. (1963). Cellular alterations in the MCH line of Chinese hamster cells following infection with herpes simplex virus. Proc. Natl. Acad. Sci. U.S.A. 49; 474-480.

Hampe, A., Gobet, M., Even, J., Sherr, C.J. and Galibert, F. (1983). Nucleotide sequences of feline sarcoma virus long terminal repeats and 5' leaders show extensive homology to those of other mammalian retroviruses. J. Virol. 45; 466-472.

Hanke, P.D. and Fuchs, J.A. (1983a). Regulation of ribonucleoside diphosphate reductase mRNA synthesis in Escherichia coli. J. Bacteriol. 154; 1040-1045.

Hanke, P.D. and Fuchs, J.A. (1983b). Characterisation of the mRNA coding for ribonucleoside diphosphate reductase in Escherichia coli. J. Bacteriol. 156; 1192-1197.

Hanke, P.D. and Fuchs, J.A. (1984). Requirement of protein synthesis for the induction of ribonucleoside diphosphate reductase mRNA in Escherichia coli. Mol. Gen. Genet. 193; 327-331.

Harris-Hamilton, E. and Bachenheimer, S.L. (1985). Accumulation of herpes simplex virus type 1 RNAs of different kinetic classes in the cytoplasm of infected cells. J. Virol. 53; 144-151.

Hay, J., Koteles, G.J., Keir, H.M. and Subak-Sharpe, H. (1966). Herpes virus specified ribonucleic acids. Nature 210; 387-390.

Haynes, S.R. and Jelinek, W.R. (1981). Low molecular weight RNAs transcribed in vitro by RNA polymerase III from Alu-type dispersed repeats in Chinese hamster DNA are also found in vivo. Proc. Natl. Acad. Sci. U.S.A. 78; 6130-6134.

Hayward, G.S., Jacob, R.J., Wadsworth, S.C. and Roizman, B. (1975). Anatomy of herpes simplex virus DNA: evidence for four populations of molecules that differ in the relative orientations of their long and short components. Proc. Natl. Acad. Sci. 72; 4243-4247.

Hayward, S.D., Lazarowitz, S.G. and Hayward, G.S. (1982). Organisation of the Epstein-Barr virus DNA molecule. II. Fine mapping of the boundaries of the internal repeat cluster of B95-8 and identification of additional small tandem repeats adjacent to the HR-1 deletion. J. Virol. 43; 201-212.

Heilig, R., Muraskowsky, R., Kloepfer, C. and Mandel, J.L. (1982). The ovalbumin gene family: complete sequence and structure of the Y gene. Nucl. Acids Res. 10; 4363-4382.

Heine, J.W., Honess, R.W., Cassai, E. and Roizman, B. (1974). Proteins specified by herpes simplex virus. XII. The virion polypeptide of type 1 strains. J. Virol. 14; 640-651.

Heintz, N., Zernik, M. and Roeder, R.G. (1981). The structure of the human histone genes: clustered but not tandemly repeated. Cell 24; 661-668.

Heller, M. and Kieff, E. (1981). Colinearity between the DNAs of Epstein-Barr virus and herpesvirus papio. J. Virol. 37; 821-826.

Heller, M., Gerber, P. and Kieff, E. (1981). Herpesvirus papio DNA is similar in organisation to Epstein-Barr virus DNA. J. Virol. 37; 698-709.

Henikoff, S. and Cohen, E.H. (1984). Sequences responsible for transcription termination on a gene segment in Saccharomyces cerevisae. Mol. Cell. Biol. 4; 1515-1520.

Henikoff, S., Kelly, J.D. and Cohen, E.H. (1983a). Transcription terminates in yeast distal to a control sequence. Cell 33; 607-614.

Henikoff, S., Sloan, J.S. and Kelly, J.D. (1983b). A Drosophila metabolic gene transcript is alternatively processed. Cell 34; 405-414.

Henle, G., Henle, W. and Diehl, V. (1968). Relation of Burkitt's tumor-associated herpes-type virus to infectious mononucleosis. Proc. Natl. Acad. Sci. U.S.A. 59; 94-101.

Henry, B.E., Glaser, R., Hewetson, J. and O'Callaghan, D.J. (1978). Expression of altered ribonucleotide reductase activity associated with the replication of Epstein-Barr virus. Virology 89; 262-271.

Henry, B.E., Robinson, R.A., Dauenhauer, S.A., Atherton, S.S., Hayward, G.S. and O'Callaghan, D.J. (1981). Structure of the genome of equine herpesvirus type 1. Virology 115; 97-114.

Hentschel, C.C. and Birnstiel, M.L. (1981). The organisation and expression of histone gene families. Cell 25; 301-313.

Hentschel, C., Irminger, J.-C., Bucher, P. and Birnstiel, M.L. (1980). Sea urchin histone mRNA termini are located in gene regions downstream from putative regulatory signals. Nature 285; 147-151.

Herisse, J. and Galibert, F. (1981). Nucleotide sequence of the Eco RI E fragment of adenovirus 2 genome. Nucl. Acids Res. 9; 1229-1240.

Herisse, J., Courtois, G. and Galibert, F. (1980). Nucleotide sequence of the Eco RI D fragment of adenovirus 2 genome. Nucl. Acids Res. 8; 2173-2192.

Higgs, D.R., Goodbourn, S.E.Y., Lamb, J., Clegg, J.B., Weatherall, D.J. and Proudfoot, N.J. (1983). α-thalassaemia caused by a polyadenylation signal mutation. Nature 306; 398-400.

Hirsch, I., Cabral, G., Patterson, M. and Biswal, N. (1977). Studies on the intracellular replicating DNA of herpes simplex virus type 1. Virology 81; 48-61.

Hofer, E. and Darnell, J.E. (1981). The primary transcription unit of the mouse β-major globin gene. Cell 23; 585-593.

Holland, L.E., Anderson, K.P., Stringer, J.R. and Wagner, E.K. (1979). Isolation and localisation of herpes simplex virus type 1 mRNA abundant before viral DNA synthesis. J. Virol. 31; 447-462.

Holland, L.E., Anderson, K.P., Shipman, C. and Wagner, E.K. (1980). Viral DNA synthesis is required for the efficient expression of specific herpes simplex virus type 1 mRNA species. Virology 101; 10-24.

Holland, L.E., Sandri-Goldin, R.M., Goldin, A.L., Glorioso, J.C. and Levine, M. (1984). Transcriptional and genetic analyses of the herpes simplex virus type 1 genome: coordinates 0.29 to 0.45. J. Virol. 49; 947-959.

Holmes, D.S. and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114; 193-197.

Holmgren, A. (1981). Regulation of ribonucleotide reductase. Current Topics in Cellular Regulation. 19; 47-76.

Honess, R.W. and Roizman, B. (1973). Proteins specified by herpes simplex virus. XI. Identification and relative molar rates of synthesis of structural and non structural herpes virus polypeptides in the infected cell. J. Virol. 12; 1347-1365.

Honess, R.W. and Roizman, B. (1974). Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J. Virol. 14; 8-19.

Honess, R.W. and Watson, D.H. (1974). Herpes simplex virus-specific polypeptides studied by polyacrylamide gel electrophoresis of immune precipitates. J. Gen. Virol. 22; 171-185.

Honess, R.W. and Watson, D.H. (1977). Unity and diversity in the herpesviruses. J. Gen. Virol. 37; 15-37.

Hope, R.G. and Marsden, H.S. (1983). Processing of glycoproteins induced by herpes simplex virus type 1: sulphation and nature of oligosaccharide linkage. J. Gen. Virol. 64; 1943-1953.

Hope, R.G., Palfreyman, J.W., Suh, M. and Marsden, H.S. (1982). Sulphated glycoproteins induced by herpes simplex virus. J. Gen. Virol. 58; 399-415.

Hopper, S. (1972). Ribonucleotide reductase of rabbit bone marrow. I. Purification, properties and separation into two protein fractions. J. Biol. Chem. 247; 3336-3340.

Hopper, S. (1978). Ribonucleotide reductase of rabbit bone marrow. Methods Enzymol. 51; 237-246.

Horwitz, M.S. (1978). Temperature-sensitive DNA replications of H5ts125 adenovirus DNA in vitro. Proc. Natl. Acad. Sci. U.S.A. 75; 4291-4295.

Hosbach, H.A., Silberklang, M. and McCarthy, B.J. (1980). Evolution of a Drosophila melanogaster glutamate tRNA gene cluster. Cell 21; 169-178.

Huang, E.-S., Chen, S.-T. and Pagano, J.S. (1973). Human cytomegalovirus. I. Purification and characterisation of viral DNA. J. Virol. 12; 1473-1481.

Huang, E.-S., Kilpatrick, B.A., Huang, Y.-T. and Pagano, J.S. (1976). Detection of human cytomegalovirus and analysis of strain variation. Yale J. Biol. Med. 49; 29-43.

Hubenthal-Voss, J. and Roizman, B. (1985). Herpes simplex virus 1 reiterated S component sequences (c_1) situated between the 'a' sequence and α4 gene are not essential for virus replication. J. Virol. 54; 509-514.

Hummeler, K., Tomassini, N. and Zajac, B. (1969). Early events in herpes simplex virus infection: a radioautographic study. J. Virol. 4; 67-74.

Hurst, H.C. and Parker, M.G. (1983). Rat prostatic steroid binding protein: DNA sequence and transcript maps of the two C3 genes. EMBO J. 2; 769-774.

Huszar, D. and Bacchetti, S. (1981). Partial purification and characterisation of the ribonucleotide reductase induced by herpes simplex virus infection of mammalian cells. J. Virol. 37; 580-588.

Huszar, D. and Bacchetti, S. (1983). Is ribonucleotide reductase the transforming function of herpes simplex virus 2? Nature 302; 76-79.

Hyman, R.W. (1980). Identification of proteins tightly bound to herpes simplex virus DNA. Virology 105; 254-255.

Ihara, S., Feldman, L., Watanabe, S. and Ben-Porat, T. (1983). Characterisation of the immediate-early functions of pseudorabies virus. Virology 131; 437-454.

Imperiale, M.J., Feldman, L.T. and Nevins, J.R. (1983). Activation of gene expression by adenovirus and herpesvirus regulatory genes acting in *trans* and by a *cis*-acting adenovirus enhancer element. Cell 35; 127-136.

Inglis, S.C. (1982). Inhibition of host protein synthesis and degradation of cellular mRNAs during infection by influenza and herpes simplex virus. Mol. Cell. Biol. 2; 1644-1648.

Isakson, P.C., Pure, E., Uhr, J.W. and Vitetta, E.S. (1981). Induction of proliferation and differentiation of neoplastic B cells by anti-immunoglobulin and T-cell factors. Proc. Natl. Acad. Sci. U.S.A. 78; 2507-2511.

Jacob, R.J. and Roizman, B. (1977). Anatomy of herpes simplex virus DNA. VIII. Properties of the replicating DNA. J. Virol. 23; 394-411.

Jacob, R.J., Morse, L.S. and Roizman, B. (1979). Anatomy of herpes simplex virus DNA. XII. Accumulation of head-to-tail concatemers in nuclei of infected cells and their role in the generation of the four isomeric arrangements of viral DNA. J. Virol. 29; 448-457.

Jacq, B., Jourdan, R. and Jourdan, B.R. (1977). Structure and processing of precursor 5S RNA in *Drosophila melanogaster.* J. Mol. Biol. 117; 785-795.

Jamieson, A.T. and Bjursell, G. (1976). Deoxyribonucleoside triphosphate pools in herpes simplex virus type 1 infected cells. J. Gen. Virol. 31; 101-113.

Jamieson, A.T. and Subak-Sharpe, J.H. (1978). Interallelic complementation of mutants of herpes simplex virus deficient in deoxypyrimidine kinase activity. Virology 85; 109-117.

Jamieson, A.T., Gentry, G.A. and Subak-Sharpe, J.H. (1974). Induction of both thymidine and deoxycytidine kinase activity by herpes simplex viruses. J. Gen. Virol. 24; 465-480.

Jamieson, A.T., Macnab, J.C.M., Perbal, B. and Clements, J.B. (1976). Virus specified enzyme activity and RNA species in herpes simplex virus type 1 transformed mouse cells. J. Gen. Virol. 32; 493-508.

Jariwalla, R.J., Aurelian, L. and Ts'o, P.O.P. (1979). Neoplastic transformation of cultured Syrian hamster embryo cells by DNA of herpes simplex virus type 2. J. Virol. 30; 404-409.

Jariwalla, R.J., Aurelian, L. and Ts'o, P.O.P. (1980). Tumorigenic transformation induced by a specific fragment of DNA from herpes simplex virus type 2. Proc. Natl. Acad. Sci. U.S.A. 77; 2279-2283.

Jariwalla, R.J., Aurelian, L. and Ts'o, P.O.P. (1983). Immortalisation and neoplastic transformation of normal diploid cells by defined cloned DNA fragments of herpes simplex virus type 2. Proc. Natl. Acad. Sci. U.S.A. 80; 5902-5906.

Jean, J.-H. and Ben-Porat, T. (1976). Appearance in vivo of single-stranded complementary ends on herpesvirus DNA. Proc. Natl. Acad. Sci. U.S.A. 73; 2674-2678.

Jenkins, F.J., Howett, M.K., Spector, D.J. and Rapp, F. (1982). Detection by RNA blot hybridisation of RNA sequences homologous to the Bgl II-N fragment of herpes simplex virus type 2 DNA. J. Virol. 44; 1092-1096.

Johnson, D.C. and Spear, P.G. (1983). O-linked oligosaccharides are acquired by herpes simplex virus glycoproteins in the Golgi apparatus. Cell 32; 987-997.

Johnson, D.C. and Spear, P.G. (1984). Evidence for translational regulation of HSV-1 gD expression J. Virol. 51; 389-394.

Jones, K.A. and Tjian, R. (1985). Spl binds to promoter sequences and activates herpes simplex virus 'immediate-early' gene transcription in vitro. Nature 317; 179-182.

Jones, K.A., Yamamoto, K.R. and Tjian, R. (1985). Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell 42; 559-572.

Jones, M.D. and Griffin, B.E. (1983). Clustered repeat sequences in the genome of Epstein-Barr virus. Nucl. Acids Res. 11; 3919-3937.

Jones, P.C. and Roizman, B. (1979). Regulation of herpesvirus macromolecular synthesis. VIII. The transcription program consists of three phases during which both extent of transcription and accumulation of RNA in the cytoplasm are regulated. J. Virol. 31; 299-314.

Jongeneel, C.V. and Bachenheimer, S.L. (1981). Structure of replicating herpes simplex virus DNA. J. Virol. 39; 656-660.

Josey, W.E., Nahmias, A.J. and Naib, Z.M. (1972). The epidemiology of type 2 (genital) herpes simplex virus infection. Obstet. Gynecol. Surv. 27; 295-302.

Kalinyak, J.E., Fleagle, G., and Docherty, J.J. (1977). Incidence and distribution of herpes simplex virus types 1 and 2 from genital lesions in college women. J. Med. Virol. 1; 173-181.

Karin, M. and Richards, R.I. (1982). Human metallothionein genes - primary structure of the metallothionein-II gene and a related processed gene. Nature 299; 797-802.

Karin, M., Haslinger, A., Holtgreve, H., Richards, R.I., Krauter, P., Westphal, H.M. and Beato, M. (1984). Characterisation of sequences through which cadmium and glucocorticoid hormones induce human metallothionein-II_A gene. Nature 308; 513-519.

Kazama, F.Y. and Schornstein, K.L. (1972). Herpes-type virus particles associated with a fungus. Science 177; 696-697.

Kedes, L.H. (1979). Histone genes and histone messengers. Ann. Rev. Biochem. 48; 837-870.

Keir, H.M. (1968). Virus-induced enzymes in mammalian cells infected with DNA-viruses. In 'The Molecular Biology of Viruses.' pp. 67-99. (L.V. Crawford and M.G.P. Stoker, eds.) Cambridge University Press, Cambridge.

Keir, H.M. and Gold, E. (1963). Deoxyribonucleic acid nucleotidyl transferase and deoxyribonuclease from cultured cells infected with herpes simplex virus. Biochim. Biophys. Acta 72; 263-276.

Keir, H.M., Hay, J., Morrison, J.M. and Subak-Sharpe, J.H. (1966). Altered properties of deoxyribonucleic acid nucleotidyltransferase after infection of mammalian cells with herpes simplex virus. Nature 210; 369-371.

Kemp, D.J., Morahan, G., Cowman, A.F. and Harris, A.W. (1983). Production of RNA for secreted immunoglobulin chains does not require transcriptional termination 5' to the mu exons. Nature 301; 84-86.

Kessler, I. (1977). Venereal factors in human cervical cancer. Cancer 39; 1912-1919.

Kieff, E.D., Bachenheimer, S.L. and Roizman, B. (1971). Size, composition and structure of the deoxyribonucleic acid of herpes simplex virus subtypes 1 and 2. J. Virol. 8; 125-132.

Kilpatrick, B.A. and Huang, E-S. (1977). Human cytomegalovirus genome: partial denaturation map and organisation of genome sequences. J. Virol. 24; 261-276.

Kimura, M. and Ohta, T. (1974). On some principles governing molecular evolution. Proc. Natl. Acad. Sci. U.S.A. 71; 2848-2852.

King, C.R. and Piatigorsky, J. (1983). Alternative RNA splicing of the murine ∝A-crystallin gene: protein coding information within an intron. Cell 32; 707-712.

Kintner, C.R. and Sugden, B. (1979). The structure of the termini of the DNA of Epstein-Barr virus. Cell 17; 661-671.

Kit, S. and Dubbs, D.R. (1963). Acquisition of thymidine kinase activity by herpes simplex infected mouse fibroblast cells. Biochem. Biophys. Res. Comm. 11; 55-59.

Klein, H., Maltzman, W. and Levine, A.J. (1979). Structure-function relationships of the adenoviral DNA binding protein. J. Biol. Chem. 254; 11051-11060.

Klemperer, H.G., Haynes, G.R., Sheddon, W.I.H. and Watson, D.H. (1967). A virus-specific thymidine kinase in BHK21 cells infected with herpes simplex virus. Virology 31; 120-128.

Klessig, D.F. and Quinlan, M.P. (1982). Genetic evidence for separate functional domains on human adenovirus specific 72 kd DNA binding protein. J. Mol. Appl. Genet. 1; 263-272.

Knapp, M.R., Liu, C.-P., Newell, N., Ward, R.B., Tucker, P.W., Strober, S. and Blattner, F. (1982). Simultaneous expression of immunoglobulin mu and delta heavy chains by a cloned B-cell lymphoma: a single copy of the V_H gene is shared by two adjacent C_H genes. Proc. Natl. Acad. Sci. U.S.A. 79; 2996-3000.

Knopf, K.W. (1979). Properties of herpes simplex virus DNA polymerase and characterisation of its associated exonuclease activity. Eur. J. Biochem. 98; 231-244.

Knust, E., Schirm, S., Dietrich, W., Bodemer, W., Kolb, E. and Fleckenstein, B. (1983). Cloning of herpesvirus saimiri DNA fragments representing the entire L-region of the genome. Gene 25; 281-289.

Konarska, M.M., Grabowski, P.J., Padgett, R.A. and Shar, P.A. (1985). Characterisation of the branch site in lariat RNAs produced by splicing of mRNA precursors. Nature 313; 552-557.

Konkel, D.A., Tilgham, S.M. and Leder, P. (1978). The sequence of the chromosomal mouse β-globin major gene: homologies in capping, splicing and poly A sites. Cell 15; 1125-1132.

Konkel, D.A., Maizel, J.V. and Leder, P. (1979). The evolution and sequence comparison of two recently diverged mouse chromosomal β-globin genes. Cell 18; 865-873.

King, C.R. and Piatigorsky, J. (1983). Alternative RNA splicing of the murine ∝A-crystallin gene: protein coding information within an intron. Cell 32; 707-712.

Kintner, C.R. and Sugden, B. (1979). The structure of the termini of the DNA of Epstein-Barr virus. Cell 17; 661-671.

Kit, S. and Dubbs, D.R. (1963). Acquisition of thymidine kinase activity by herpes simplex infected mouse fibroblast cells. Biochem. Biophys. Res. Comm. 11; 55-59.

Klein, H., Maltzman, W. and Levine, A.J. (1979). Structure-function relationships of the adenoviral DNA binding protein. J. Biol. Chem. 254; 11051-11060.

Klemperer, H.G., Haynes, G.R., Sheddon, W.I.H. and Watson, D.H. (1967). A virus-specific thymidine kinase in BHK21 cells infected with herpes simplex virus. Virology 31; 120-128.

Klessig, D.F. and Quinlan, M.P. (1982). Genetic evidence for separate functional domains on human adenovirus specific 72 kd DNA binding protein. J. Mol. Appl. Genet. 1; 263-272.

Knapp, M.R., Liu, C.-P., Newell, N., Ward, R.B., Tucker, P.W., Strober, S. and Blattner, F. (1982). Simultaneous expression of immunoglobulin mu and delta heavy chains by a cloned B-cell lymphoma: a single copy of the V_H gene is shared by two adjacent C_H genes. Proc. Natl. Acad. Sci. U.S.A. 79; 2996-3000.

Knopf, K.W. (1979). Properties of herpes simplex virus DNA polymerase and characterisation of its associated exonuclease activity. Eur. J. Biochem. 98; 231-244.

Knust, E., Schirm, S., Dietrich, W., Bodemer, W., Kolb, E. and Fleckenstein, B. (1983). Cloning of herpesvirus saimiri DNA fragments representing the entire L-region of the genome. Gene 25; 281-289.

Konarska, M.M., Grabowski, P.J., Padgett, R.A. and Shar, P.A. (1985). Characterisation of the branch site in lariat RNAs produced by splicing of mRNA precursors. Nature 313; 552-557.

Konkel, D.A., Tilgham, S.M. and Leder, P. (1978). The sequence of the chromosomal mouse β-globin major gene: homologies in capping, splicing and poly A sites. Cell 15; 1125-1132.

Konkel, D.A., Maizel, J.V. and Leder, P. (1979). The evolution and sequence comparison of two recently diverged mouse chromosomal β-globin genes. Cell 18; 865-873.

Korn, L.J. and Brown, D.D. (1978). Nucleotide sequence of Xenopus borealis oocyte 5S DNA: comparison of sequences that flank several related eucaryotic genes. Cell 15; 1145-1156.

Kousoulas, K.G., Bzik, D.J., DeLuca, N. and Persson, S. (1983). The effect of ammonium chloride and tunicamycin on the glycoprotein content and infectivity of herpes simplex virus type 1. Virology 125; 468-474.

Kozak, M. (1981). Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucl. Acids Res. 9; 5233-5252.

Kozak, M. and Roizman, B. (1974). Regulation of herpesvirus macromolecular synthesis: nuclear retention of nontranslated viral RNA sequences. Proc. Natl. Acad. Sci. U.S.A. 71; 4322-4326.

Krainer, A.R., Maniatis, T., Ruskin, B. and Green, M.R. (1984). Normal and mutant human β-globin pre-mRNAs are faithfully and efficiently spliced in vitro. Cell 36; 993-1005.

Kramer, A., Keller, W., Appel, B. and Luhrmann, R. (1984). The 5' terminus of the RNA moiety of U1 small ribonucleoprotein particles is required for the splicing of messenger RNA precursors. Cell 38; 299-307.

Krieg, P.A. and Melton, D.A. (1984). Formation of the 3' end of histone mRNA by post-transcriptional processing. Nature 308; 203-206.

Krieg, P.A., Robins, A.J., Colman, H. and Wells, J.R.E. (1983). Chicken histone H5 mRNA: the polyadenylated RNA lacks the conserved histone 3' terminator sequence. Nucl. Acids Res. 10; 6777-6785.

Kristie, T.M. and Roizman, B. (1984). Separation of sequences defining basal expression from those conferring α gene recognition within the regulatory domains of herpes simplex virus type 1 α genes. Proc. Natl. Acad. Sci. U.S.A. 81; 4065-4069.

Kruijer, W., van Schaik, F.M.A. and Sussenbach, J.S. (1981). Structure and organisation of the gene coding for the DNA binding protein of adenovirus type 5. Nucl. Acids Res. 9; 4439-4457.

Kruijer, W., van Schaik, F.M.A. and Sussenbach, J.S. (1982). Nucleotide sequence of the gene encoding adenovirus type 2 DNA binding protein. Nucl. Acids Res. 10; 4493-4500.

Kruijer, W., van Schaik, F.M.A., Speijer, J.G. and Sussenbach, J.S. (1983). Structure and function of adenovirus DNA binding protein: comparison of the amino acid sequences of the Ad5 and Ad12 proteins derived from the nucleotide sequence of the corresponding genes. Virology 128; 140-153.

Kudler, L. and Hyman, R.W. (1979). Exonuclease III digestion of herpes simplex virus DNA. Virology 92; 68-81.

Kuhne, T., Wieringa, B., Reiser, J. and Weissman, C. (1983). Evidence against a scanning model of RNA splicing. EMBO J. 2; 727-733.

Kumar, A. and Lindberg, U. (1972). Characterisation of messenger ribonucleoprotein and messenger RNA from KB cells. Proc. Natl. Acad. Sci. U.S.A 69; 681-685.

Kunkel, T. and Loeb, L.A. (1979). On the fidelity of DNA replication. Effect of divalent metal ion activators and deoxyribonucleoside triphosphate pools on in vitro mutagenesis. J. Biol. Chem. 254; 5718-5725.

Kwok, S.C.M., Chan, S.J. and Steiner, D.F. (1983). Cloning and nucleotide sequence analysis of the dog insulin gene: coded amino acid sequence of canine preproinsulin predicts an additional C-peptide fragment. J. Biol. Chem. 258; 2357-2363.

Ladin, B.F., Blankenship, M.L. and Ben-Porat, T. (1980). Replication of herpesvirus DNA. V. Maturation of concatemeric DNA of pseudorabies virus to genome length is related to capsid formation. J. Virol. 33; 1151-1164.

LaFemina, R.L. and Hayward, G.S. (1980). Structural organisation of the DNA molecules from human cytomegalovirus. In 'Animal Virus Genetics.' pp. 39-55. (B.N. Fields and R. Jaenisch, eds.) Academic Press, New York.

Landesman, R. and Gross, P. (1968). Patterns of macromolecular synthesis during development in X. laevis. Dev. Biol. 18; 571-589.

Lang, J.S., Spandidos, D.A. and Wilkie, N.M. (1984). Transcriptional regulation of a herpes simplex virus immediate early gene is mediated through an enhancer-type sequence. EMBO J. 3; 389-395.

Langelier, Y. and Buttin, G. (1981). Characterisation of ribonucleotide reductase induction in BHK-21/C13 Syrian hamster cell line upon infection by herpes simplex virus (HSV). J. Gen. Virol. 57; 21-31.

Langelier, Y., DeChamps, M. and Buttin, G. (1978). An analysis of dCMP deaminase and CDP reductase levels in hamster cells infected by herpes simplex virus. J. Virol. 26; 547-553.

Langford, C.J. and Gallwitz, D. (1983). Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell 33; 519-527.

Langford, C.J., Klinz, F.-J., Donath, C. and Gallwitz, D. (1984). Point mutations identify the conserved, intron-contained TATAAC box as an essential splicing signal sequence in yeast. Cell 36; 645-653.

Langley, K.E., Villarejo, M.R., Fowler, A.V., Zamenhof, P.J. and Zabin, I. (1975). Molecular basis of β-galactosidase complementation. Proc. Natl. Acad. Sci. U.S.A. 72; 1254-1257.

Lankinen, H., Graslund, A. and Thelander, L. (1982). Induction of a new ribonucleotide reductase after infection of mouse L cells with pseudorabies virus. J. Virol. 41; 893-900.

Larsson, A. and Reichard, P. (1966a). Enzymatic synthesis of deoxyribonucleotides. IX. Allosteric effects in the reduction of pyrimidine ribonucleotides by ribonucleoside diphosphate reductase system of Escherichia coli. J. Biol. Chem. 241; 2533-2539.

Larsson, A. and Reichard, P. (1966b). Enzymatic synthesis of deoxyribonucleotides. X. Reduction of purine ribonucleotides; allosteric behaviour and substrate specificity of the enzyme system from Escherichia coli B. J. Biol. Chem. 241; 2540-2549.

LaThangue, N.B., Shriver, K., Dawson, C. and Chan, W.L. (1984). Herpes simplex virus infection causes the accumulation of a heat-shock protein. EMBO J. 3; 267-277.

Lawn, R.M., Efstratiadis, A., O'Connell, C. and Maniatis, T. (1980). The nucleotide sequence of the human β-globin gene. Cell 21; 647-651.

Leary, K. and Francke, B. (1984). The interaction of a topoisomerase-like enzyme from herpes simplex virus type 1-infected cells with non-viral circular DNA. J. Gen. Virol. 65; 1341-1350.

Lebowitz, P. and Ghosh, P.K. (1982). Initiation and regulation of simian virus 40 early transcription in vitro. J. Virol. 41; 449-461.

Lee, Y.S., Tanaka, A., Lau, R.Y., Nonoyama, M. and Rabin, H. (1980). Comparative studies of herpesvirus papio (baboon herpesvirus) DNA and Epstein-Barr virus DNA. J. Gen. Virol. 51; 245-253.

Lee, Y.S., Nonoyama, M. and Rabin, H. (1981). Colinear relationships of herpesvirus papio DNA to Epstein-Barr virus DNA. Virology 110; 248-252.

Leeds, J.M., Slabaugh, M.B. and Mathews, C.K. (1985). DNA precursor pools and ribonucleotide reductase activity: distribution between the nucleus and cytoplasm of mammalian cells. Mol. Cell. Biol. 5; 3443-3450.

Lemischka, I. and Sharp, P.A. (1982). The sequences of an expressed rat α-tubulin gene and a pseudogene with an inserted repetitive element. Nature 300; 330-335.

Lewis, J.G., Kucera, L.S., Eberle, R. and Courtney, R.J. (1982). Detection of herpes simplex virus type 2 glycoproteins expressed in virus-transformed rat cells. J. Virol. 42; 275-282.

Liebhaber, S.A., Goossens, M.J. and Wai Kan, Y. (1980). Cloning and complete nucleotide sequence of human 5'-α-globin gene. Proc. Natl. Acad. Sci. U.S.A. 77; 7054-7058.

Liebhaber, S.A., Goosens, M. and Kan, Y.W. (1981). Homology and concerted evolution at the α1 and α2 loci of human α-globin. Nature 290; 26-29.

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

Linne, T. and Philipson, L. (1980). Further characterisation of the phosphate moiety of the adenovirus type 2 DNA binding protein. Eur. J. Biochem. 103; 259-270.

Little, S.P., Jofre, J.T., Courtney, R.J. and Schaffer, P.A. (1981). A virion-associated glycoprotein essential for infectivity of herpes simplex virus type 1. Virology 115; 149-160.

Liu, C.-C., Simonsen, C.C. and Levinson, A.D. (1984). Initiation of translation at internal AUG codons in mammalian cells. Nature 309; 82-85.

Lomedico, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R. and Tizard, R. (1979). The structure and evolution of the two nonallelic rat preproinsulin genes. Cell 18; 545-548.

Lonsdale, D.M., Brown, S.M., Subak-Sharpe, J.H., Warren, K.G. and Koprowski, H. (1979). The polypeptide and the DNA restriction enzyme profiles of spontaneous isolates of herpes simplex virus type 1 from explants of human trigeminal, superior cervical and vagus ganglia. J. Gen. Virol. 43; 151-171.

Lovinger, G.G. and Schochetman, G. (1979). 5' terminal nucleotide sequences of the Rauscher leukemia virus and gibbon ape leukemia virus genomes exhibit a high degree of correspondence. J. Virol. 32; 803-811.

Lowe, P.A. (1978). Levels of DNA-dependent RNA polymerase in herpes simplex virus-infected BHK 21 C13 cells. Virology 86; 577-580.

Lu, P., Cheung, S. and Arndt, K. (1983). Possible molecular detent in the DNA structure at regulatory sequences. J. Biomolec. Struc. Dynam. 1; 509-521.

Lusky, M., Berg, L., Weiher, H. and Botchan, M. (1983). Bovine papilloma virus contains an activator of gene expression at the distal end of the early transcription unit. Mol. Cell. Biol. 3; 1108-1122.

MacPherson, I. and Stoker, M. (1962). Polyoma transformation of hamster cell clones - an investigation of genetic factors affecting cell competence. Virology 16; 147-151.

McDevitt, M.A., Imperiale, M.J., Ali, H. and Nevins, J.R. (1984). Requirement of a downstream sequence for generation of a poly(A) addition site. Cell 37; 993-999.

McDougall, J.K., Galloway, D.A. and Fenoglio, C.M. (1980). Cervical carcinoma: detection of herpes simplex virus RNA in cells undergoing neoplastic change. Int. J. Cancer 25; 1-8.

McDougall, J.K., Crum, C.K., Fenoglio, C.M., Goldstein, L.C. and Galloway, D.A. (1982). Herpesvirus specific RNA and protein in carcinoma of the uterine cervix. Proc. Natl. Acad. Sci. U.S.A. 69; 3853-3857.

McGeoch, D.J. (1984). The nature of animal virus genetic material. In 'The Microbe 1984.' part 1, pp. 75-107. (B.W.J. Mahy and J.R. Pattison, eds.) Cambridge University Press, Cambridge.

McGeoch, D.J., Dolan, A., Donald, S. and Rixon, F.J. (1985). Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. J. Mol. Biol. 181; 1-13.

McKnight, S.L. (1980). The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. Nucl. Acids Res. 8; 5949-5964.

McKnight, S.L. and Kingsbury, R. (1982). Transcriptional control signals of a eukaryotic protein-coding gene. Science 217; 316-324.

McKnight, S.L., Gavis, E.R., Kingsbury, R. and Axel, R. (1981). Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: identification of an upstream control region. Cell 25; 385-398.

McKnight, S.L., Kingsbury, R.C., Spence, A. and Smith, M. (1984). The distal transcription signals of the herpesvirus tk gene share a common hexanucleotide control sequence. Cell 37; 253-262.

McLauchlan, J. and Clements, J.B. (1982). A 3' co-terminus of two early herpes simplex virus type 1 mRNAs. Nucl. Acids Res. 10; 501-512.

McLauchlan, J. and Clements, J.B. (1983a). Organisation of the herpes simplex virus type 1 transcription unit encoding two early proteins with molecular weights of 140000 and 40000. J. Gen. Virol. 64; 997-1006.

McLauchlan, J. and Clements, J.B. (1983b). DNA sequence homology between two colinear loci on the HSV genome which have different transforming abilities. EMBO J. 2; 1953-1961.

McReynolds, L. and Penman, S. (1974). A polymerase activity forming 5S and pre-4S RNA in isolated HeLa cell nuclei. Cell 1; 139-145.

Mackem, S. and Roizman, B. (1982a). Differentiation between ∝-promoter and regulator regions of herpes simplex virus 1: the functional domains and sequence of a movable ∝ regulator. Proc. Natl. Acad. Sci. U.S.A. 79; 4917-4921.

Mackem, S. and Roizman, B. (1982b). Structural features of the herpes simplex virus ∝ gene 4, 0, and 27 promoter-regulatory sequences which confer ∝ regulation on chimeric thymidine kinase genes. J. Virol. 44; 939-949.

Macnab, J.C.M. (1974). Transformation of rat embryo cells by temperature sensitive mutants of herpes simplex virus. J. Gen. Virol. 24; 143-153.

Macnab, J.C.M. (1979). Tumour production by HSV-2 transformed lines in rats and the varying response to immunosuppression. J. Gen. Virol. 43; 39-56.

Macnab, J.C.M., Visser, L., Jamieson, A.T. and Hay, J. (1980). Specific viral antigens in rat cells transformed by herpes simplex virus type 2 and in rat tumours induced by inoculation of transformed cells. Cancer Res. 40; 2074-2097.

Mahr, A. and Roberts, B.E. (1984a). Organisation of six early transcripts synthesised from a vaccinia virus Eco RI DNA fragment. J. Virol. 49; 497-509.

Mahr, A. and Roberts, B.E. (1984b). Arrangement of late RNAs transcribed from a 7.1kb Eco RI vaccinia virus DNA fragment. J. Virol. 49; 510-520.

Maki, R., Roeder, W., Traunecker, A., Sidman, C., Wahl, M., Raschke, W. and Tonegawa, S. (1981). The role of DNA rearrangement and alternative RNA processing in the expression of immunoglobulin delta genes. Cell 24; 353-365.

Malissen, M., Malissen, B., and Jordan, B.R., (1982). Exon/intron organisation and complete nucleotide sequence of an HLA gene. Proc. Natl. Acad. Sci. U.S.A. 79; 893-897.

Maniatis, T., Jeffrey, A. and van de Sande, H. (1975). Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. Biochemistry 14; 3787-3794.

Manley, J.L. (1983). Accurate and specific polyadenylation of mRNA precursors in a soluble whole-cell lysate. Cell 33; 595-605.

Manley, J.L., Sharp, P.A. and Gefter, M.L. (1982). RNA synthesis in isolated nuclei: processing of adenovirus serotype 2 late messenger RNA precursors. J. Mol. Biol. 159; 581-599.

Manley, J.L., Yu, H. and Ryner, L. (1985). RNA sequence containing hexanucleotide AAUAAA directs efficient mRNA polyadenylation in vitro. Mol. Cell. Biol. 5; 373-379.

Mao, J.C-H., Robishaw, E.E. and Overby, L.R. (1975). Inhibition of DNA polymerase from herpes simplex virus-infected Wi-38 cells by phosphonoacetic acid. J. Virol. 15; 1281-1283.

Marcu, K.B. (1982). Immunoglobulin heavy-chain constant-region genes. Cell 29; 719-721.

Marek, J. (1907). Multiple nervenentzundung bei huhnen. Deutsche Tierarztliche Wochenschrift. 15; 417-421.

Marsden, H.S., Crombie, I.K. and Subak-Sharpe, J.H. (1976). Control of protein synthesis in herpes virus-infected cells: analysis of the polypeptides induced by wild-type and sixteen temperature-sensitive mutants of strain 17. J. Gen. Virol. 31; 347-372.

Marsden, H.S., Stow, N.D., Preston, V.G., Timbury, M.C. and Wilkie, N.M. (1978). Physical mapping of herpes simplex virus-induced polypeptides. J. Virol. 28; 624-642.

Marsden, H.S., Lang, J., Davison, A.J., Hope, R.G. and McDonald, D.M. (1982). Genomic location and lack of phosphorylation of the HSV immediate-early polypeptide IE12. J. Gen. Virol. 62; 17-27.

Marsden, H.S., Haarr, L. and Preston, C.M. (1983). Processing of herpes simplex virus proteins and evidence that translation of thymidine kinase mRNA is initiated at three separate AUG codons. J. Virol. 46; 434-445.

Marsden, H.S., Buckmaster, A., Palfreyman, J.W., Hope, R.G. and Minson, A.C. (1984). Characterisation of the 92,000 dalton glycoprotein induced by herpes simplex virus type 2. J. Virol. 50; 547-554.

Marshall, R.D. (1974). The nature and metabolism of the carbohydrate-peptide linkage of glycoproteins. Biochem. Soc. Symp. 40; 17-26.

Marshall, R.D. and Neuberger, A. (1964). Carbohydrates in protein. VIII. The isolation of 2-acetamido-1(L- β -aspartamido)-1,2-dideoxy-β-D-glucose from hen's egg albumin. Biochemistry 3; 1596-1600.

Martin, W.B., Hay, D., Crawford, L.V., Lebouvier, G.L. and Crawford, E.M. (1966). Characteristics of bovine mammillitis virus. J. Gen. Microbiol. 45; 325-332.

Mason, A.J., Evans, B.A., Cox, D.R., Shine, J. and Richards, R.I. (1983). Structure of mouse kallikrein gene family suggests a role in specific processing of biologically active peptides. Nature 303; 300-307.

Mason, P.J., Jones, M.B., Elkington, J.A. and Williams, J.G. (1985). Polyadenylation of the Xenopus β1 globin mRNA at a downstream minor site in the absence of the major site and utilisation of an AAUACA polyadenylation signal. EMBO J. 4; 205-211.

Mathis, D.J. and Chambon, P. (1981). The SV40 early region TATA box is required for accurate *in vitro* initiation of transcription. Nature 290; 310-315.

Matthews, R.E.F. (1982). Classification and nomenclature of viruses. Intervirology 17; 47-51.

Maxam, A.M. and Gilbert, W. (1977). A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. 74; 560-564.

Maxam, A.M. and Gilbert, W. (1980). Sequencing end-labelled DNA with base specific chemical cleavage. Methods Enzymol. 65; 499-560.

Mayman, B.A. and Nishioka, Y. (1985). Differential stability of host mRNAs in Friend erythroleukaemia cells infected with herpes simplex virus type 1. J. Virol. 53; 1-6.

Mayo, K.E. and Palmiter, R.D. (1981). Glucocorticoid regulation of metallothionein-I mRNA synthesis in cultured mouse cells. J. Biol. Chem. 256; 2621-2624.

Melendez, L.V., Daniel, M.D., Hunt, R.D. and Garcia, F.G. (1968). An apparently new herpesvirus from primary kidney cultures of the squirrel monkey (Saimiri sciureus). Laboratory and Animal Care 18; 374-381.

Melendez, L.V., Hunt, R.D., King, N.W., Barahona, H.H., Daniel, M.D., Fraser, C.E.O. and Garcia, F.G. (1972). Herpesvirus ateles, a new lymphoma virus of monkeys. Nature New Biology 235; 182-184.

Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981). Identification of DNA sequences required for transcription of the human α1 globin gene in a new SV40 host-vector system. Cell 27; 279-288.

Melton, D.A., de Robertis, E.M. and Cortese, R. (1980). Order and intracellular location of the events involved in the maturation of a spliced tRNA. Nature 284; 143-148.

Messing, J. (1979). A multi-purpose cloning system based on the single-stranded DNA bacteriophage M13. Recombinant DNA Technical Bulletin, NIH Publication No. 79-99, 2; No. 2, 43-48.

Messing, J., Crea, R. and Seeburg, P.H. (1981). A system for shotgun DNA sequencing. Nucl. Acids Res. 9; 309-321.

Meuth, M., L'Heureux-Huard, N. and Trudel, M. (1979). Characterisation of a mutator gene in Chinese hamster ovary cells. Proc. Natl. Acad. Sci. U.S.A. 76; 6505-6509.

Miller, J.H. (1972). Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Minson, A.C., Thouless, M.E., Eglin, R.P. and Darby, G. (1976). The detection of virus DNA sequences in HSV-2 transformed hamster cell line (3338-3339). Int. J. Cancer. 17; 493-500.

Mocarski, E.S. and Roizman, B. (1981). Site specific inversion of the herpes simplex virus genome: domain and structural features. Proc. Natl. Acad. Sci. U.S.A. 78; 7047-7051.

Mocarski, E.S. and Roizman, B. (1982a). Structure and role of the herpes simplex virus DNA termini in inversion, circularisation and generation of virion DNA. Cell 31; 89-97.

Mocarski, E.S. and Roizman, B. (1982b). Herpesvirus dependent amplification and inversion of cell associated viral thymidine kinase gene flanked by 'a' sequences and linked to an origin of DNA replication. Proc. Natl. Acad. Sci. U.S.A. 79; 5626-5630.

Mocarski, E.S., Post, L.E. and Roizman, B. (1980). Molecular engineering of the herpes simplex virus genome: insertion of a second L-S junction into the genome causes additional genome inversions. Cell 22; 243-255.

Montell, C., Fisher, E.F., Caruthers, M.H. and Berk, A.J. (1982). Resolving the functions of overlapping viral genes by site-specific mutagenesis at a mRNA splice site. Nature 295; 380-384.

Montell, C., Fisher, E.F., Caruthers, M.H. and Berk, A.J. (1983). Inhibition of RNA cleavage but not polyadenylation by a point mutation in mRNA 3' consensus sequence AAUAAA. Nature 305; 600-605.

Moore, C.L. and Sharp, P.A. (1984). Site-specific polyadenylation in a cell-free reaction. Cell 36; 581-591.

Moore, E.C. (1977). Components and control of ribonucleotide reductase system of the rat. Adv. Enz. Reg. 15; 101-114.

Morelli, G., Nagy, F., Fraley, R.T., Rogers, S.G., and Chua, N.-H. (1985). A short conserved sequence is involved in the light-inducibility of a gene encoding ribulose 1,5-biphosphate carboxylase small subunit of pea. Nature 315; 200-204.

Morgan, J.R. and Roberts, B.E. (1984). Organisation of RNA transcripts from a vaccinia virus early gene cluster. J. Virol. 51; 283-297.

Morrison, J.M. and Keir, H.M. (1968). A new DNA-exonuclease in cells infected with herpes virus: partial purification and properties of the enzyme. J. Gen. Virol. 3; 337-347.

Morse, L.S., Pereira, L., Roizman, B. and Schaffer, P.A. (1978). Anatomy of herpes simplex virus (HSV) DNA. X. Mapping of viral genes by analysis of polypeptides and functions specified by HSV-1 x HSV-2 recombinants. J. Virol. 26; 389-410.

Moss, B., Gershowitz, A., Stringer, J.R., Holland, L.E. and Wagner, E.K. (1977). 5' terminal and internal methylated nucleosides in herpes simplex virus type 1 mRNA. J. Virol. 23; 234-239.

Moss, H., Chartrand, P., Timbury, M.C. and Hay, J. (1979). Mutant of herpes simplex virus type 2 with temperature sensitive lesions affecting virion thermostability and DNase activity: identification of the lethal mutation and physical mapping of the nuc⁻ lesion. J. Virol. 32; 140-146.

Mount, S.M. (1982). A catalogue of splice junction sequences. Nucl. Acids Res. 10; 459-472.

Mount, S.M., Petterson, I., Hinterberger, M., Karmas, A. and Steitz, J.A. (1983). The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. Cell 33; 509-518.

Muller, M.T., Bolles, C.S. and Parris, D.S. (1985). Association of type 1 DNA topoisomerase with herpes simplex virus. J. Gen. Virol. 66; 1565-1574.

Muller, W.E.G., Zahn, R.K., Arendes, J. and Falke, D. (1979). Oligoribonucleotide initiators for herpes simplex virus DNA synthesis in vivo and in vitro. Virology 98; 200-210.

Munk, K., and Sauer, G.Z. (1964). Relationship between cell DNA metabolism and nucleocytoplasmic alterations in herpes virus-infected cells. Virology 22; 153-154.

Murchie, M.-J. and McGeoch, D.J. (1982). DNA sequence analysis of an immediate-early gene region of the herpes simplex virus type 1 genome (map coordinates 0.950 to 0.978). J. Gen. Virol. 62; 1-15.

Nabeshima, Y., Fujii-Kuriyama, Y., Muramatsu, M. and Ogata, K. (1984). Alternative transcription and two modes of splicing result in two myosin light chains from one gene. Nature 308; 333-338.

Nagata, S., Mantei, N. and Weissmann, C. (1980). The structure of one of the eight or more distinct chromosomal genes for human interferon-α. Nature 287; 401-408.

Nahmias, A.J., Josey, W.E., Naib, Z.M., Luce, C.F. and Guest, B. (1970). Antibodies to herpesvirus hominis types 1 and 2 in humans. II. Women with cervical cancer. Amer. J. Epidem. 91; 547-552.

Naib, Z.M., Nahmias, A.J. and Joset, W.E. (1966). Cytology and histopathology of cervical herpes simplex infection. Cancer 19; 1026-1031.

Nakai, H., Maxwell, I.H. and Pizer, L.I. (1982). Herpesvirus infection alters the steady-state levels of cellular polyadenylated RNA in polyoma virus-transformed BHK cells. J. Virol. 42; 1131-1134.

Nayak, D.P. (1971). Isolation and characterisation of a herpesvirus from leukemic guinea pigs. J. Virol. 8; 579-588.

Nevins, J.R. (1982). Induction of the synthesis of a 70000 dalton mammalian heat shock protein by the adenovirus E1A gene product. Cell 29; 913-919.

Nevins, J.R. (1983). The pathway of eukaryotic mRNA formation. Ann. Rev. Biochem. 52; 441-466.

Nevins, J.R. and Darnell, J.E. (1978). Steps in the processing of Ad2 mRNA: poly $(A)^+$ nuclear sequences are conserved and poly (A) addition precedes splicing. Cell 15; 1477-1493.

Nevins, J.R., Blanchard, J.-M. and Darnell, J.E. (1980). Transcription units of adenovirus type 2: termination of transcription beyond the poly (A) addition site in early regions 2 and 4. J. Mol. Biol. 144; 377-386.

Nishioka, Y. and Silverstein, S. (1977). Degradation of cellular mRNA during infection by herpes simplex virus. Proc. Natl. Acad. Sci. U.S.A. 74; 2370-2374.

Nishioka, Y. and Silverstein, S. (1978). Requirement of protein synthesis for the degradation of host mRNA in Friend erythroleukaemia cells infected with herpes simplex virus type 1. J. Virol. 27; 619-627.

Nishioka, Y. and Leder, P. (1979). The complete sequence of a chromosomal mouse α-globin gene reveals elements conserved throughout vertebrate evolution. Cell 18; 875-882.

Nishioka, Y.G., Jones, G. and Silverstein, S. (1983). Inhibition by vesicular stomatitis virus of herpes simplex virus-directed protein synthesis. Virology 124; 238-250.

Nishiyama, Y., Tsurumi, T., Aoki, H. and Maeno, K. (1983). Identification of DNA polymerase(s) involved in the repair of viral and cellular DNA in herpes simplex virus type 2-infected cells. Virology 129; 524-528.

Noda, M., Teranishi, Y., Takahashi, H., Toyosato, M., Notake, M., Nakanishi, S. and Numa, S. (1982). Isolation and structural organisation of the human preproenkephalin gene. Nature 297; 431-434.

Nonoyama, M. and Pagano, J.S. (1973). Homology between Epstein-Barr virus DNA and viral DNA from Burkitt's lymphoma and nasopharyngeal carcinoma determined by DNA-DNA reassociation kinetics. Nature 242; 44-47.

Nordheim, A. and Rich, A. (1983). Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. Nature 303; 674-679.

Norrild, B. and Pedersen, B. (1982). Effect of tunicamycin on the synthesis of herpes simplex virus type 1 glycoproteins and their expression on the cell surface. J. Virol. 43; 395-402.

Notarianni, E.L. and Preston, C.M. (1982). Activation of cellular stress protein genes by herpes simplex virus temperature-sensitive mutants which overproduce immediate-early polypeptides. Virology 123; 113-122.

Nudel, U., Zakut, R., Shani, M., Neuman, S., Levy, Z. and Yaffe, E. (1983). The nucleotide sequence of the rat cytoplasmic β-actin gene. Nucl. Acids Res. 11; 1759-1771.

Nussinov, R. (1985). Sequence signals which may be required for efficient formation of mRNA 3' termini. manuscript submitted.

O'Hare, P. and Hayward, G.S. (1984). Expression of recombinant genes containing herpes simplex virus delayed-early and immediate-early regulatory regions and trans activation by herpesvirus infection. J. Virol. 52; 522-531.

O'Hare, P. and Hayward, G.S. (1985). Evidence for a direct role for both the 175,000- and 110,000-molecular-weight immediate-early proteins of herpes simplex virus in the trans activation of delayed-early promoters. J. Virol. 53; 751-760.

Ohno, S. and Taniguchi, T. (1981). Structure of a chromosomal gene for human interferon β. Proc. Natl. Acad. Sci. U.S.A. 78; 5305-5309.

Ollis, D.L., Kline, C. and Steitz, T.A. (1985). Domain of E. coli DNA polymerase I showing sequence homology to T7 DNA polymerase. Nature 313; 818-819.

O'Neill, F.J. and Rapp, F. (1971). Early events required for induction of chromosome abnormalities in human cells by herpes simplex virus. Virology 44; 544-553.

Orkin, S.H., Kazazian, H.H., Antonarakis, S.E., Goff, S.C., Boehm, C.D., Sexton, J.P., Waber, P.G. and Giardina, P.J.V. (1982). Linkage of β-thalassaemia mutations and β-globin gene polymorphisms with DNA polymorphisms in human β-globin gene cluster. Nature 296; 627-631.

Osborne, T.F., Goldstein, J.L. and Brown, M.S. (1985). 5' end of HMG CoA reductase gene contains sequences responsible for cholesterol-mediated inhibition of transcription. Cell 42; 203-212.

Ostrander, M. and Cheng, Y.-C. (1980). Properties of herpes simplex virus type 1 and type 2 DNA polymerase. Biochim. Biophys. Acta 609; 232-245.

Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F. and Sharp, P.A. (1984). Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. Science 225; 898-903.

Pagano, J.S. (1975). Diseases and mechanisms of persistent DNA virus infection: latency and cellular transformation. J. Infect. Dis. 132; 209-223.

Page, G.S., Smith, S. and Goodman, H.M. (1981). DNA sequence of the rat growth hormone gene: location of the 5' terminus of the growth hormone mRNA and identification of an internal transposon-like element. Nucl. Acids Res. 9; 2087-2104.

Palfreyman, J.W., Haarr, L., Cross, A., Hope, R.G. and Marsden, H.S. (1983). Processing of herpes simplex virus type 1 glycoproteins: two dimensional gel analysis using monoclonal antibodies. J. Gen. Virol. 64; 873-886.

Pan, J., Elder, J.T., Duncan, C.H. and Weissman, S.M. (1981). Structural analysis of interspersed repetitive polymerase III transcription units in human DNA. Nucl. Acids Res. 9; 1151-1170.

Panagou, D., Orr, M.D., Dunstone, J.R. and Blakley, R.L. (1972). A monomeric, allosteric enzyme with a single polypeptide chain. Ribonucleotide reductase of Lactobacillus leichmannii. Biochemistry 11; 2378-2388.

Para, M.F., Goldstein, L. and Spear, P.G. (1982). Similarities and differences in the Fc binding glycoprotein (gE) of herpes simplex types 1 and 2 and tentative mapping of the viral gene for this glycoprotein. J. Virol. 41; 137-144.

Para, M.F., Zezulak, K.M., Conley, A.J., Weinberger, M., Snitzer, K. and Spear, P.G. (1983). Use of monoclonal antibodies directed against two 75,000 molecular weight glycoproteins specified by herpes simplex virus type 2 in glycoprotein identification and gene mapping. J. Virol. 45; 1223-1227.

Park, M., Lonsdale, D.M., Timbury, M.C., Subak-Sharpe, J.H. and Macnab, J.C.M. (1980). Genetic retrieval of viral genome sequences from herpes simplex virus transformed cells. Nature 285; 412-415.

Park, M., Kitchener, H.C. and Macnab, J.C.M. (1983). Detection of herpes simplex virus type 2 DNA restriction fragments in human cervical carcinoma tissue. EMBO J. 2; 1029-1034.

Parker, D.C., Fothergill, J.J. and Wadsworth, D.C. (1979). B lymphocyte activation by insoluble anti-immunoglobulin: induction of immunoglobulin secretion by a T cell-dependent soluble factor. J. Immun. 123; 931-941.

Parnes, J.R., Robinson, R.R. and Seidman, J.G. (1983). Multiple mRNA species with distinct 3' termini are transcribed from the β 2-microglobulin gene. Nature 302; 449-452.

Pelham, H.R.B. (1982). A regulatory upstream promoter element in the Drosophila hsp 70 heat-shock gene. Cell 30; 517-528.

Pelham, H.R.B. and Bienz, M. (1982). A synthetic heat-shock promoter element confers heat-inducibility on the herpes simplex virus thymidine kinase gene. EMBO J. 1; 1473-1477.

Pellett, P.E., Kousoulas, K.G., Pereira, L. and Roizman, B. (1985a). Anatomy of the herpes simplex virus strain F glycoprotein B gene: primary sequence and predicted protein structure of the wild type and of monoclonal antibody-resistant mutants. J. Virol. 53; 243-253.

Pellett, P.E., McKnight, J.L.C., Jenkins, F.J. and Roizman, B. (1985b). Nucleotide sequence and predicted amino acid sequence of a protein encoded in a small herpes simplex virus DNA fragment capable of trans-inducing genes. Proc. Natl. Acad. Sci. U.S.A. 82; 5870-5874.

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

Pereira, L., Dondero, D., Norrild, B. and Roizman, B. (1981). Differential immunologic reactivity and processing of glycoproteins gA and gB of herpes simplex virus types 1 and 2 made in Vero and HEp-2 cells. Proc. Natl. Acad. Sci. U.S.A. 78; 5202-5206.

Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. and Dodgson, J. (1980). The evolution of genes: the chicken preproinsulin gene. Cell 20; 555-566.

Pernis, B., Forni, L. and Luzzati, A. (1976). Synthesis of multiple immunoglobulin classes by single lymphocytes. Cold Spring Harbor Symp. Quant. Biol. 41; 139-145.

Perricaudet, M., Akusjarvi, G., Virtanen, A. and Petterson, U. (1979). Structure of two spliced mRNAs from the transforming region of human subgroup C adenoviruses. Nature 281; 694-696.

Perricaudet, M., leMoullec, J.M., Tiollais, P. and Pettersson, U. (1980). Structure of two adenovirus type 12 transforming polypeptides and their evolutionary implications. Nature 288; 174-176.

Persson, R.H., Bacchetti, S. and Smiley, J.R. (1985). Cells that constitutively express the herpes simplex virus immediate-early protein ICP4 allow efficient activation of viral delayed-early genes in trans. J. Virol. 54; 414-421.

Perutz, M.F., Kendrao, J.C. and Watson, H.C. (1965). Structure and function of haemoglobin. II. Some relationships between polypeptide chain configuration and amino acid sequence. J. Mol. Biol. 13; 669-678.

Peutherer, J.F., Smith, I.W. and Hunter, J.M. (1981). Herpes simplex virus 1 infection of the cervix. Lancet 2; 1285.

Pikielny, C.W., Teem, J.L. and Rosbash, M. (1983). Evidence for the biochemical role of an internal sequence in yeast nuclear mRNA introns: implications for U1 RNA and metazoan mRNA splicing. Cell 34; 395-403.

Pilon, L., Royal, A. and Langelier, Y. (1985). Increased mutation frequency after herpes simplex virus type 2 infection in non-permissive XC cells. J. Gen. Virol. 66; 259-265.

Pizer, L., Cohen, G.H. and Eisenberg, R.J. (1980). Effect of tunicamycin on herpes simplex virus glycoproteins and infectious virus production. J. Virol. 34; 142-153.

Plummer, G. (1973). Cytomegalovirus of man and animals. Prog. Med. Virol. 15; 92-125.

Poffenberger, K.L. and Roizman, B. (1985). A noninverting genome of a viable herpes simplex virus 1: presence of head-to-tail linkages in packaged genomes and requirements for circularisation after infection. J. Virol. 53; 587-595.

Ponce de Leon, M., Eisenberg, R.J. and Cohen, G.H. (1977). Ribonucleotide reductase from herpes simplex virus (types 1 and 2) infected and uninfected KB cells: properties of the partially purified enzymes. J. Gen. Virol. 36; 163-173.

Post, L.E., Mackem, S. and Roizman, B. (1981). Regulation of ∝ genes of herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with ∝ gene promoters. Cell 24; 555-565.

Powell, K.L. and Courtney, R.J. (1975). Polypeptides synthesised in herpes simplex virus type 2-infected HEp-2 cells. Virology 66; 217-228.

Powell, K.L. and Purifoy, D.J.M. (1977). Nonstructural proteins of herpes simplex virus. I. Purification of the induced DNA polymerase. J. Virol. 24; 618-626.

Powell, K.L., Purifoy, D.J.M. and Courtney, R.J. (1975). The synthesis of herpes simplex virus proteins in the absence of virus DNA synthesis. Biochem. Biophys. Res. Comm. 66; 262-271.

Preston, C.M. (1977). The cell free synthesis of herpes-induced polypeptides. Virology 78; 349-353.

Preston, C.M. (1979). Abnormal properties of an immediate-early polypeptide in cells infected with the herpes simplex virus type 1 mutant tsK. J. Virol. 32; 357-359.

Preston, C.M. and McGeoch, D.J. (1981). Identification and mapping of two polypeptides encoded within the herpes simplex virus type 1 thymidine kinase gene sequences. J. Virol. 38; 593-605.

Preston, C.M. and Cordingley, M.G. (1982). mRNA and DNA-directed synthesis of herpes simplex virus-coded exonuclease in Xenopus laevis oocytes. J. Virol. 43; 386-394.

Preston, C.M., Cordingley, M.G. and Stow, N.D. (1984a). Analysis of DNA sequences which regulate the transcription of a herpes simplex virus immediate early gene. J. Virol. 50; 708-716.

Preston, V.G. and Fisher, F.B. (1984). Identification of the herpes simplex virus type 1 gene encoding the dUTPase. Virology 138; 58-68.

Preston, V.G., Coates, J.A.V. and Rixon, F.J. (1983). Identification and characterisation of a herpes simplex virus gene product required for encapsidation of virus DNA. J. Virol. 45; 1056-1064.

Preston, V.G., Palfreyman, J.W. and Dutia, B.M. (1984b). Identification of a herpes simplex virus type 1 polypeptide which is a component of the virus-induced ribonucleotide reductase. J. Gen. Virol. 65; 1457-1466.

Price, D.H. and Parker, C.S. (1984). The 3' end of Drosophila histone H3 mRNA is produced by a processing activity in vitro. Cell 38; 423-429.

Pritchett, R.F. (1980). DNA nucleotide sequence heterogeneity between the Towne and AD169 strains of cytomegalovirus. J. Virol. 36; 152-161.

Proudfoot, N.J. (1978). Eukaryotic promoters. Nature 279; 376.

Proudfoot, N.J. and Brownlee, G.G. (1976). 3' non-coding region sequences in eukaryotic mRNA. Nature (London) 263; 211-214.

Proudfoot, N.J. and Maniatis, T. (1980). The structure of a human α-globin pseudogene and its relationship to α-globin gene duplication. Cell 21; 537-544.

Quinlan, M.P. and Knipe, D.M. (1985). Stimulation of expression of a herpes simplex virus DNA-binding protein by two viral functions. J. Virol. 5; 957-963.

Quinn, J.P. (1984). Sequence of the DNA polymerase locus in the genome of herpes simplex virus type 1. Ph. D. Thesis, University of Glasgow.

Randall, C.C., Ryden, F.W., Doll, E.R. and Shell, F.S. (1953). Cultivation of equine abortion virus in foetal horse tissue in vitro. Amer. J. Pathol. 29; 139-153.

Rapp, F. (1980). Persistence and transmission of cytomegalovirus. In 'Virus-Host Interactions: Viral Invasion, Persistence and Disease.' vol 16, pp. 193-232. (H. Fraenkel-Conrat and R.R. Wagner, eds.) Plenum Publishing Corp., New York.

Rawls, W.E., Tompkins, W.A.F. and Melnick, J.L. (1969). The association of herpes virus type 2 and carcinoma of the uterine cervix. Amer. J. Epidem. 89; 547-554.

Rawls, W.E., Gardner, H.L., Flanders, R.W., Lowry, S.P., Kaufman, R.H. and Melnick, J.C. (1971). Genital herpes in two social groups. Amer. J. Obstet. Gynecol. 110; 682-689.

Read, G.S. and Frenkel, N. (1983). Herpes simplex virus mutants defective in the virion-associated shutoff of host polypeptide synthesis and exhibiting abnormal synthesis of α (immediate-early) viral polypeptides. J. Virol. 46; 498-512.

Rechavi, G., Givol, D. and Canaani, E. (1982). Activation of a cellular oncogene by DNA rearrangement: possible involvement of an IS-like element. Nature (London) 300; 607-611.

Reddy, E.P., Smith, M.J. and Srinivasan, A. (1983a). Nucleotide sequence of Abelson murine leukemia virus genome: structural similarity of its transforming gene product to other onc gene products with tyrosine-specific kinase activity. Proc. Natl. Acad. Sci. U.S.A. 80; 3623-3627.

Reddy, E.P., Reynolds, R.K., Watson, D.K., Schultz, R.A., Lautenberger, J. and Papas, T.S. (1983b). Nucleotide sequence analysis of the proviral genome of avian myelocytomatosis virus (MC29). Proc. Natl. Acad. Sci. U.S.A. 80; 2500-2504.

Reddy, G.P.V. and Pardee, A.B. (1980). Multienzyme complex for metabolic channeling in mammalian DNA replication. Proc. Natl. Acad. Sci. U.S.A. 77; 3312-3316.

Reddy, G.P.V. and Pardee, A.B. (1982). Coupled ribonucleoside diphosphate reduction, channeling and incorporation into DNA of mammalian cells. J. Biol. Chem. 257; 12526-12531.

Reed, C.L., Cohen, G.H. and Rapp, F. (1975). Detection of a virus-specific antigen on the surface of herpes simplex virus-transformed cells. J. Virol. 15; 668-670.

Reichard, P. and Ehrenberg, A. (1983). Ribonucleotide reductase - a radical enzyme. Science 221; 514-519.

Reyes, G.R., La Femina, R., Hayward, S.D. and Hayward, G.S. (1979). Morphological transformation by DNA fragments of human herpesviruses: evidence for two distinct transforming regions in HSV-1 and HSV-2 and lack of correlation with biochemical transfer of the thymidine kinase gene. Cold Spring Harbor Symp. Quant. Biol. 44; 629-641.

Ricciardi, R., Jones, R., Cepko, C., Sharp, P. and Roberts, B. (1981). Expression of early adenovirus genes requires a viral encoded acidic polypeptide. Proc. Natl. Acad. Sci. U.S.A. 78; 6121-6125.

Rifkind, D. (1966). The activation of varicella-zoster virus infections by immunosuppressive therapy. J. Lab. Clin. Med. 68; 463-474.

Rigby, P.W.J., Dieckermann, M., Rhodes, C. and Berg, P. (1977). Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. 113; 237-251.

Rixon, F.J. and Clements, J.B. (1982). Detailed structural analysis of two spliced HSV-1 immediate-early mRNAs. Nucl. Acids Res. 10; 2241-2256.

Rixon, F.J. and McGeoch, D.J. (1984). A 3' co-terminal family of mRNAs from the herpes simplex virus type 1 short region: two overlapping reading frames encode unrelated polypeptides one of which has a highly reiterated amino acid sequence. Nucl. Acids Res. 12; 2473-2487.

Rixon, F.J. and McGeoch, D.J. (1985). Detailed analysis of the mRNAs mapping in the short unique region of herpes simplex virus type 1. Nucl. Acids Res. 13; 953-973.

Rixon, F.J., Campbell, M.E. and Clements, J.B. (1982). The immediate-early mRNA that encodes the regulatory polypeptide Vmw175 of herpes simplex virus type 1 is unspliced. EMBO J. 1; 1273-1277.

Rixon, F.J., Campbell, M.E. and Clements, J.B. (1984). A tandemly reiterated DNA sequence in the long repeat region of herpes simplex virus type 1 found in close proximity to immediate early mRNA 1. J. Virol. 52; 715-718.

Rodriguez, J.R., Pikielny, C.W. and Rosbash, M. (1984). In vivo characterisation of yeast mRNA processing intermediates. Cell 39; 603-610.

Roizman, B. (1982). The family herpesviridae: general description, taxonomy and classification. In 'The Herpesviruses.' vol 1, pp. 1-23. (B. Roizman, ed.) Plenum Press, N.Y.

Roizman, B. and Roane, P.R.Jr. (1964). The multiplication of herpes simplex virus. II. The relation between protein synthesis and the duplication of viral DNA in infected HEp-2 cells. Virology 22; 262-269.

Roizman, B. and Furlong, D. (1974). The replication of herpesviruses. In 'Comprehensive Virology.' vol 3, pp. 229-403. (H. Fraenkel-Conrat and R.R. Wagner, eds.) Plenum Press, N.Y.

Roizman, B., Kozak, M., Honess, R.W., and Hayward, G. (1974). Regulation of herpesvirus macromolecular synthesis: evidence for multilevel regulation of herpes simplex 1 RNA and protein synthesis. Cold Spring Harbor Symp. Quant. Biol. 39; 687-701.

Roizman, B., Carmichael, S., de The, G.B., Masic, M., Nahmias, A.J., Plowright, W., Rapp, F., Sheldrick, P., Takahashi, M., Terni, M. and Wolfe, K. (1978). Provisional classification of herpesviruses. In 'Oncogenesis of Herpesviruses III.' Part 2, pp. 1079-1082. (G. de The, F. Rapp and W. Henle, eds.) IARC Scientific Publications, No 24, Lyon.

Roizman, B., Jacob, R.J., Knipe, D.M., Morse, L.S. and Ruyechan, W.T. (1979). On the structure, functional equivalence and replication of the four rearrangements of herpes simplex virus DNA. Cold Spring Harbor Symp. Quant. Biol. 43; 809-826.

Roizman, B., Norrild, B., Chan, C. and Pereira, L. (1984). Identification and preliminary mapping with monoclonal antibodies of a herpes simplex virus 2 glycoprotein lacking a known type-1 counterpart. Virology 133; 242-247.

Rosenfeld, M.G., Mermod, J.-J., Amara, S.G., Swanson, L.W., Sawchenko, P.E., Rivier, J., Vale, W.W. and Evans, R.M. (1983). Production of a novel neuropeptide encoded by the calcitonin gene via tissue-specific RNA processing. Nature 304; 129-135.

Rosenthal, E., Hunt, T. and Ruderman, J.V. (1980). Selective translation of mRNA controls the pattern of protein synthesis during early development of the surf clam, Spisula solidissima. Cell 20; 487-496.

Rosenthal, E.T., Tansey, T.R. and Ruderman, J.V. (1983). Sequence-specific adenylations and deadenylations accompany changes in the translation of maternal mRNA after fertilisation of Spisula oocytes. J. Mol. Biol. 166; 309-327.

Royston, I. and Aurelian, L. (1970). Immunofluorescent detection of herpesvirus antigens in exfoliated cells from human cervical carcinoma. Proc. Natl. Acad. Sci. U.S.A. 67; 204-212.

Rubenstein, A.S. and Kaplan, A.S. (1975). Electron microscopic studies of the DNA of defective and standard pseudorabies virions. Virology 66; 385-392.

Ruskin, B., Krainer, A.R., Maniatis, T. and Green, M.R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell 38; 317-331.

Ruskin, B., Greene, J.M. and Green, M.R. (1985). Cryptic branch point activation allows accurate in vitro splicing of human β-globin intron mutants. Cell 41; 833-844.

Russell, W.C. (1962). Herpes virus nucleic acid. Virology 16; 355-357.

Russell, W.C. and Crawford, L.V. (1963). Some characteristics of the deoxyribonucleic acid from herpes simplex virus. Virology 21; 353-361.

Russell, W.C. and Crawford, L.V. (1964). Properties of the nucleic acids from some herpes group viruses. Virology 22; 288-292.

Sacks, W.R., Greene, C.C., Aschman, D.P. and Schaffer, P.A. (1985). Herpes simplex virus type 1 ICP27 is an essential regulatory protein. J. Virol. 55; 796-805.

Sadofsky, M. and Alwine, J.C. (1984). Sequences on the 3' side of hexanucleotide AAUAAA affect efficiency of cleavage at the polyadenylation site. Mol. Cell. Biol. 4; 1460-1468.

Sandri-Goldin, R.M., Goldin, A.L., Holland, L.E., Glorioso, J.C. and Levine, M. (1983). Expression of herpes simplex virus β and γ genes integrated in mammalian cells and their induction by an α gene product. Mol. Cell. Biol. 3; 2028-2044.

Sawadogo, M. and Roeder, R.G. (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell 43; 165-175.

Schaak, J., Sharp, S., Dingermann, T. and Soll, D. (1983). Transcription of eukaryotic tRNA genes in vitro. II. Formation of stable complexes. J. Biol. Chem. 258; 2447-2453.

Schaak, J., Sharp, S., Dingermann, T., Burke, D.J., Cooley, L. and Soll, D. (1984). The extent of a eukaryotic tRNA gene. 5'- and 3'-flanking sequence dependence for transcription and stable complex formation. J. Biol. Chem. 259; 1461-1467.

Schaffer, P.A., Carter, V.C. and Timbury, M.C. (1978). Collaborative complementation study of temperature-sensitive mutants of herpes simplex virus types 1 and 2. J. Virol. 27; 490-504.

Scheidereit, C. and Beato, M. (1984). Contacts between hormone receptor and DNA double helix within a glucocorticoid regulatory element of mouse mammary tumor virus. Proc. Natl. Acad. Sci. U.S.A. 81; 3029-3033.

Schek, N. and Bachenheimer, S.L. (1985). Degradation of cellular mRNAs induced by a virion-associated factor during herpes simplex virus infection of Vero cells. J. Virol. 55; 601-610.

Schlehofer, J.R. and zur Hausen, H. (1982). Induction of mutations within the host cell genome by partially inactivated herpes simplex virus type 1. Virology 122; 471-475.

Scholl, D.R., Kahn, S., Malavarca, R., Astrin, S. and Skalka, A.M. (1983). Nucleotide sequence of the long terminal repeat and flanking cellular sequences of avian endogenous retrovirus ev-2: variation in Rous-associated virus-0 expression cannot be explained by differences in primary sequence. J. Virol. 45; 868-871.

Schon, E.A., Wernke, S.M. and Lingrel, J.B. (1982). Gene conversion of two functional goat α-globin genes preserves only minimal flanking sequences. J. Biol. Chem. 257; 6825-6835.

Schrader, W.T., Birnbaumer, M.E., Hughes, M.R., Weigel, N.L., Grody, W.W. and O'Malley, B.W. (1981). Studies on the structure and function of the chicken progesterone receptor. Recent Prog. Horm. Res. 37; 583-633.

Schulte-Holthausen, H. and zur Hausen, H. (1970). Partial purification of the Epstein-Barr virus and some properties of its DNA. Virology 40; 776-779.

Searle, P.F., Stuart, G.W. and Palmiter, R.D. (1985). Building a metal-responsive promoter with synthetic regulatory elements. Mol. Cell. Biol. 5; 1480-1489.

Seif, I., Khoury, G. and Dhar, R. (1979). The genome of human papovavirus BKV. Cell 18; 963-977.

Selby, M.J., Barta, A., Baxter, J.D., Bell, G.I. and Eberhardt, N.L. (1984). Analysis of a major human chorionic somatomammotropin gene. Evidence for two functional promoter elements. J. Biol. Chem. 259; 13131-13138.

Setzer, D.R., McGrogan, M., Nunberg, J.H. and Schimke, R.T. (1980). Size heterogeneity in the 3' end of dihydrofolate reductase messenger RNAs in mouse cells. Cell 22; 361-370.

Sharp, J.A., Wagner, M.J. and Summers, W.C. (1983). Transcription of herpes simplex virus genes *in vivo*: overlap of a late promoter with the 3' end of the early thymidine kinase gene. J. Virol. 45; 10-17.

Sharp, P.A. (1981). Speculations on RNA splicing. Cell 23; 643-646.

Shatkin, A.J. (1976). Capping of eukaryotic mRNAs. Cell 9; 645-653.

Shaw, G.D., Boll, W., Taira, H., Mantei, N., Lengyel, P. and Weissmann, C. (1983). Structure and expression of cloned murine IFN-α genes. Nucl. Acids Res. 11; 555-573.

Shaw, P., Sordat, B. and Schibler, U. (1985). The two promoters of the mouse α-amylase gene Amy-1α are differentially activated during parotid gland differentiation. Cell 40; 907-912.

Sheldrick, P. and Berthelot, N. (1974). Inverted repetitions in the chromosome of herpes simplex virus. Cold Spring Harbor Symp. Quant. Biol. U.S.A. 39; 667-678.

Shen, S., Slightom, J.L. and Smithies, O. (1981). A history of the human fetal globin gene duplication. Cell 26; 191-203.

Shenk, T. and Williams, J. (1984). Genetic analysis of adenoviruses. In 'The Molecular Biology of Adenoviruses.' Current Topics in Microbiology and Immunology. vol 3, pp. 1-39. (W. Doerfler, ed.) Springer-Verlag, Berlin and Heidelberg.

Shibahara, S., Morimoto, Y., Furutani, Y., Notake, M., Takahashi, H., Shimizu, S., Horikawa, S. and Numa, S. (1983). Isolation and sequence analysis of the human corticotropin-releasing factor precursor gene. EMBO J. 2; 775-779.

Shimizu, A., Takahashi, N., Yaoita, Y. and Honjo, T. (1982). Organisation of the constant-region gene family of the mouse immunoglobulin heavy chain. Cell 28; 499-506.

Shimotohno, K., Mizutani, S. and Temin, H.M. (1980). Sequence of retrovirus provirus resembles that of bacterial transposable elements. Nature 285; 550-554.

Shlomai, J., Friedmann, A. and Becker, Y. (1976). Replicative intermediates of herpes simplex virus DNA. Virology 69; 647-659.

Silver, S. and Roizman, B. (1985). γ_2 thymidine kinase chimeras are identically transcribed but regulated as γ_2 genes in herpes simplex virus genomes and as β genes in cell genomes. Mol. Cell. Biol. 5; 518-528.

Silverstein, S., Millette, R., Jones, P. and Roizman, B. (1976). RNA synthesis in cells infected with herpes simplex virus. XII. Sequence complexity and properties of RNA differing in extents of adenylation. J. Virol. 18; 977-991.

Simonds, J.A., Rabey, W.G., Graham, B.J., Oie, H. and Vande Woude, G.F. (1975). Purification of herpesvirus saimiri and properties of the viral DNA. Arch. Virol. 49; 249-259.

Singh, D., Tamao, Y. and Blakely, R.L. (1977). Allosterism, regulation and cooperativity: the case of ribonucleotide reductase of Lactobacillus leichmannii. Adv. Enz. Regul. 15; 81-100.

Sjoberg, B.-M., Reichard, P., Graslund, A. and Ehrenberg, A. (1977). Nature of the free radical in ribonucleotide reductase from Escherichia coli. J. Biol. Chem. 252; 536-541.

Sjoberg, B.-M., Loehr, T.M. and Sanders-Loehr, J. (1982). Raman spectral evidence for a μ-oxo bridge in the binuclear iron center of ribonucleotide reductase. Biochemistry 21; 96-102.

Sjoberg, B.-M., Eklund, H., Fuchs, J.A., Carlson, J., Standart, N.M., Ruderman, J.V., Bray, S.J. and Hunt, T. (1985). Identification of the stable free radical tyrosine residue in ribonucleotide reductase: a sequence comparison. FEBS Lett. 183; 99-102.

Skinner, G.B.R., Whitnex, J.E. and Hartley, C. (1977). Prevalence of type specific antibody against type 1 and type 2 herpes simplex virus in women with abnormal cervical cytology: evidence towards prepubertal vaccination of seronegative female subjects. Arch. Virol. 54; 211-221.

Slabaugh, M.B. and Mathews, C.K. (1984). Vaccinia virus-induced ribonucleotide reductase can be distinguished from host cell activity. J. Virol. 52; 501-506.

Slabaugh, M.B., Johnson, T.L. and Mathews, C.K. (1984). Vaccinia virus induces ribonucleotide reductase in primate cells. J. Virol. 52; 507-514.

Smiley, J.R., Fong, B.S. and Leung, W.-C. (1981). Construction of a double jointed herpes simplex viral DNA molecule: inverted repeats are required for segment inversion and direct repeats promote deletions. Virology 113; 345-362.

Smith, I.W., Peutherer, J.F. and Hunter, J.M. (1981). Cervical infection with herpes simplex virus. Lancet 1; 1051.

Smith, M.G. (1954). Propagation of salivary gland virus of the mouse in tissue cultures. Proc. Soc. Exp. Biol. Med. 86; 435-440.

Smith, M.G. (1956). Propagation in tissue cultures of a cytopathogenic virus from human salivary gland virus (SGV) disease (22498). Proc. Soc. Exp. Biol. Med. 92; 424-430.

Soderlund, H., Pettersson, U., Vennstrom, B., Philipson, L. and Mathews, M.B. (1976). A new species of virus-coded low molecular weight RNA from cells infected with adenovirus type 2. Cell 7; 585-593.

Soeda, E., Arrand, J.R., Smolar, N., Walsh, J.E. and Griffin, B.E. (1980). Coding potential and regulatory signals of the polyoma virus genome. Nature 283; 445-453.

Sollner-Webb, B. and Reeder, R.H. (1979). The nucleotide sequence of the initiation and termination sites for ribosomal RNA transcription in X. laevis. Cell 18; 485-499.

Solnick, D. (1981). An adenovirus mutant defective in splicing RNA from early region 1A. Nature 291; 508-510.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98; 503-517.

Spaete, R.R. and Frenkel, N. (1982). The herpes simplex virus amplicon: a new eucaryotic defective virus cloning-amplifying vector. Cell 30; 295-304.

Spaete, R.R. and Mocarski, E.S. (1985). The 'a' sequence of the cytomegalovirus genome functions as a cleaving/packaging signal for herpes simplex virus defective genomes. J. Virol. 54; 817-824.

Spandidos, D.A. and Paul, J. (1982). Transfer of human globin genes to erythroleukemic mouse cells. EMBO J. 1; 15-20.

Spear, P.G. (1976). Membrane proteins specified by herpes simplex virus. I. Identification of four glycoprotein precursors and their products in type 1 infected cells. J. Virol. 17; 991-1008.

Spear, P.G. and Roizman, B. (1972). Proteins specified by herpes simplex virus. V. Purification and structural proteins of the herpes virion. J. Virol. 9; 143-159.

Spear, P.G. and Roizman, B. (1980). Herpes simplex viruses. In 'DNA Tumour Viruses.' 2nd ed., pp. 615-746. (J. Tooze, ed.) Cold Spring Harbor Laboratory.

Spector, D.H., Hock, L. and Tamashiro, J.C. (1982). Cleavage maps for human cytomegalovirus DNA strain AD169 for restriction endonucleases EcoRI, BglII and HindIII. J. Virol. 42; 558-582.

Speirs, J. and Birnstiel, M. (1974). Arrangement of the 5.8S RNA cistrons in the genome of X. laevis. J. Mol. Biol. 87; 237-256.

Spring, S.B., Roizman, B. and Schwartz, J. (1968). Herpes simplex virus products in productive and abortive infection. II. Electron microscopic and immunological evidence for failure of virus envelopment as a cause of abortive infection. J. Virol. 2; 384-392.

Spritz, R.A., DeRiel, J.K., Forget, B.G. and Weissman, S.M. (1980). Complete nucleotide sequence of the human δ-globin gene. Cell 21; 639-646.

Standart, N.M., Bray, S.J., George, E.L., Hunt, T. and Ruderman, J.V. (1985). The small subunit of ribonucleotide reductase is encoded by one of the most abundant translationally regulated maternal RNAs in clam and sea urchin eggs. J. Cell Biol. 100; 1968-1976.

Stenberg, R.M. and Pizer, L. (1982). Herpes simplex virus-induced changes in cellular and adenovirus metabolism in an adenovirus type 5-transformed cell line. J. Virol. 42; 474-487.

Stevely, W.S. (1977). Inverted repetition in the chromosome of pseudorabies virus. J. Virol. 22; 232-234.

Stich, H.F., Hsu, T.C. and Rapp, F. (1964). Viruses and mammalian chromosomes. I. Localisation of chromosome aberrations after infection with herpes simplex virus. Virology 22; 439-445.

Stillman, B.W., Lewis, J.B., Chow, L.T., Mathews, M.B. and Smart, J.E. (1981). Identification of the gene and mRNA for the adenovirus terminal protein precursor. Cell 23; 497-508.

Storz, J., Ehlers, B., Todd, W.J. and Ludwig, H. (1984). Bovine cytomegaloviruses: identification and differential properties. J. Gen. Virol. 65; 697-706.

Stow, N.D. (1982). Localization of an origin of DNA replication within the TR_S/IR_S repeated region of the herpes simplex virus type 1 genome. EMBO J. 1; 863-867.

Stow, N.D. and McMonagle, E.C. (1983). Characterization of the TR_S/IR_S origin of DNA replication of herpes simplex virus type 1. Virology 130; 427-438.

Stow, N.D., McMonagle, E.C. and Davison, A.J. (1983). Fragments from both termini of the herpes simplex virus type 1 genome contain signals required for the encapsidation of viral DNA. Nucl. Acids Res. 11; 8205-8220.

Straus, S.E., Owens, J., Ruyechan, W.T., Takiff, H.E., Casey, T.A., Vande Woude, G.F. and Hay, J. (1982). Molecular cloning and physical mapping of varicella-zoster virus DNA. Proc. Natl. Acad. Sci. U.S.A. 79; 993-997.

Stringer, J.R., Holland, L.E., Swanstrom, R.I., Pivo, K. and Wagner, E.K. (1977). Quantitation of herpes simplex virus type 1 RNA in infected cells. J. Virol. 21; 889-901.

Strub, K., Galli, G., Busslinger, M. and Birnstiel, M.L. (1984). The cDNA sequences of the sea urchin U7 small nuclear RNA suggest specific contacts between histone mRNA precursor and U7 RNA during RNA processing. EMBO J. 3; 2801-2807.

Struck, D.K. and Lennarz, W.J. (1980). The function of saccharide-lipids in synthesis of glycoproteins. In 'The Biochemistry of Glycoproteins and Proteoglycans.' pp. 35-83. (W.J. Lennarz, ed.) Plenum Press, New York.

Sugisaki, H., Sugimoto, K., Takanami, M., Shiroki, K., Saito, I., Shimojo, H., Sawada, Y., Uemizu, Y., Uesugi, S.-I. and Fujinaga, K. (1980). Structure and gene organisation in the transforming Hind III-G fragment of Ad12. Cell 20; 777-786.

Suske, G., Wenz, M., Cato, A.C.B. and Beato, M. (1983). The uteroglobin gene region: hormonal regulation, repetitive elements and complete nucleotide sequence of the gene. Nucl. Acids Res. 11; 2257-2271.

Sutcliffe, J.G. (1978). Complete nucleotide sequence of the Escherichia coli plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. 43; 77-90.

Swain, M.A. and Galloway, D.A. (1983). Nucleotide sequence of the herpes simplex virus type 2 thymidine kinase gene. J. Virol. 46; 1045-1050.

Swain, M.A., Peet, R.W. and Galloway, D.A. (1985). Characterisation of the gene encoding herpes simplex virus type 2 glycoprotein C and comparison with the type 1 counterpart. J. Virol. 53; 561-569.

Swanstrom, R.I. and Wagner, E.K. (1974). Regulation of synthesis of herpes simplex virus type 1 virus mRNA during productive infection. Virology 60; 522-533.

Swanstrom, R.I., Pivo, K. and Wagner, E.K. (1975). Restricted transcription of the herpes simplex virus genome occurring early after infection and in the presence of metabolic inhibitors. Virology 66; 140-150.

Sydiskis, R.J. and Roizman, B. (1967). The disaggregation of host ribosomes in productive and abortive infection with herpes simplex virus. Virology 32; 678-686.

Sydiskis, R.J. and Roizman, B. (1968). The sedimentation profiles of cytoplasmic polyribosomes in mammalian cells productively and abortively infected with herpes simplex virus. Virology 34; 562-565.

Tamashiro, J.C., Filpula, D., Friedmann, T. and Spector, D.H. (1984). Structure of the heterogeneous L-S junction region of human cytomegalovirus strain AD169 DNA. J. Virol. 52; 541-548.

Tamura, T., Noda, M. and Takano, T. (1981). Structure of the baboon endogenous virus genome: nucleotide sequences of the long terminal repeat. Nucl. Acids Res. 9; 6615-6626.

Taya, Y., Devos, R., Tavernier, J., Cheroutre, H., Engler, G. and Fiers, W. (1982). Cloning and structure of the human immune interferon-γ chromosomal gene. EMBO J. 1; 953-958.

Teich, N., Wyke, J., Mak, T., Bernstein, A. and Hardy, W. (1982). Pathogenesis of retrovirus-induced disease. In 'RNA Tumor Viruses.' 2nd ed., pp. 785-998. (R. Weiss, N. Teich, H. Varmus and J. Coffin, eds.) Cold Spring Harbor Laboratory.

Tekamp, P.A., Garcea, R.L. and Rutter, W.J. (1980). Transcription and *in vitro* processing of yeast 5S RNA. J. Biol. Chem. 255; 9501-9506.

Thelander, L. (1973). Physiochemical characterisation of ribonucleoside diphosphate reductase from *Escherichia coli.* J. Biol. Chem. 248; 4591-4601.

Thelander, L. and Reichard. P. (1979). Reduction of ribonucleotides. Ann. Rev. Biochem. 48; 133-158.

Thelander, L. and Graslund, A. (1983). Mechanism of inhibition of the mammalian ribonucleotide reductase by the iron chelate of 1-formylisoquinoline thiosemicarbazone. J. Biol. Chem. 258; 4063-4066.

Thelander, L., Eriksson, S. and Akerman, M. (1980). Ribonucleotide reductase from calf thymus. Separation of the enzyme into two nonidentical subunits, proteins M1 and M2. J. Biol. Chem. 255; 7426-7432.

Thelander, M., Graslund, A. and Thelander, L. (1985). Subunit M2 of mammalian ribonucleotide reductase: characterisation of a homogeneous protein isolated from M2-overproducing mouse cells. J. Biol. Chem. 260; 2737-2741.

Timbury, M.C. (1971). Temperature-sensitive mutants of herpes simplex virus type 2. J. Gen. Virol. 13; 373-376.

Timbury, M.C. and Subak-Sharpe, J.H. (1973). Genetic interactions between temperature-sensitive mutants of types 1 and 2 herpes simplex virus. J. Gen. Virol. 18; 347-357.

Todd, W.J. and Storz, J. (1983). Morphogenesis of a cytomegalovirus from an American bison affected with malignant catarrhal fever. J. Gen. Virol. 64; 1025-1030.

Tosi, M., Young, R.A., Hagenbuchle, O. and Schibler, U. (1981). Multiple polyadenylation sites in a mouse α-amylase gene. Nucl. Acids Res. 9; 2313-2323.

Travers, A.A., Lamond, A.I., Mace, H.A.F. and Berman, M.L. (1983). RNA polymerase interactions with the upstream region of the E. coli tyrT promoter. Cell 35; 265-273.

Treisman, R., Orkin, S.H. and Maniatis, T. (1983). Specific transcription and RNA splicing defects in five cloned β-thalassaemia genes. Nature 302; 591-596.

Turner, P.C. and Woodland, H.R. (1982). H3 and H4 histone cDNA sequences from Xenopus: a sequence comparison of H4 genes. Nucl. Acids Res. 10; 3769-3780.

Ullman, B., Clift, S.M., Gudas, L.J., Levinson, B.B., Wormsted, M.A. and Martin, D.W.Jr. (1980). Alterations in deoxyribonucleotide metabolism in cultured cells with ribonucleotide reductase activities refractory to feedback inhibition by 2'-deoxyadenosine triphosphate. J. Biol. Chem. 255; 8308-8314.

Ullrich, A., Gray, A., Goeddel, D.V. and Dull, T.J. (1982). Nucleotide sequence of a portion of human chromosome 9 containing a leukocyte interferon gene cluster. J. Mol. Biol. 156; 467-486.

van Beveren, C., van Straaten, F., Galleshaw, J.A. and Verma, I.M. (1981). Nucleotide sequence of the genome of a murine sarcoma virus. Cell 27; 97-108.

van Ooyen, A., van den Berg, J., Mantei, N. and Weissman, C. (1979). Comparison of total sequence of a cloned rabbit β-globin gene and its flanking regions with a homologous mouse sequence. Science 206; 337-344.

Varmuza, S.L. and Smiley, J.R. (1985). Signals for site-specific cleavage of HSV DNA: maturation involves two separate cleavage events at sites distal to the recognition sequences. Cell 41; 793-802.

Vasicek, T.J., McDevitt, B.E., Freeman, M.W., Fennick, B.J., Hendy, G.N., Potts, J.T., Rich, A. and Kronenberg, H.M. (1983). Nucleotide sequence of the human parathyroid gene. Proc. Natl. Acad. Sci. U.S.A. 80; 2127-2131.

Vieira, J. and Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19; 259-268.

Vitols, E., Brownson, C., Gardiner, W. and Blakely, R.L. (1967). Cobamides and ribonucleotide reduction. V. A kinetic study of the ribonucleoside triphosphate reductase of Lactobacillus leichmannii. J. Biol. Chem. 242; 3035-3041.

Vlazny, D.A. and Frenkel, N. (1981). Replication of herpes simplex virus DNA: localisation of replicative recognition signals within defective virus genomes. Proc. Natl. Acad. Sci. U.S.A. 78; 742-746.

Vlazny, D.A., Kwong, A. and Frenkel, N. (1982). Site-specific cleavage/packaging of herpes simplex virus DNA and the selective maturation of nucleocapsids containing full-length viral DNA. Proc. Natl. Acad. Sci. U.S.A. 79; 1423-1427.

von der Ahe, D., Janich, S., Scheidereit, C., Renkawitz, R., Schutz, G. and Beato, M. (1985). Glucocorticoid and progesterone receptors bind to the same sites in two hormonally regulated promoters. Nature 313; 706-709.

von Dobeln, U. and Reichard, P. (1976). Binding of substrates to Escherichia coli ribonucleotide reductase. J. Biol. Chem. 251; 3616-3622.

Vogt, V.M. (1973). Purification and further properties of single strand specific nuclease from Aspergillus oryzae. Eur. J. Biochem. 33; 192-200.

Wadsworth, S., Hayward, G.S. and Roizman, B. (1976). Anatomy of herpes simplex virus DNA. V. Terminally repetitive sequences. J. Virol. 17; 503-512.

Wagner, E.K. (1972). Evidence for transcriptional control of the herpes simplex virus genome in infected human cells. Virology 47; 502-506.

Wagner, E.K. (1985). Individual HSV transcripts: characterisation of specific genes. In 'The Herpesviruses.' vol 3, pp. 45-104. (B. Roizman, ed.) Plenum Press, New York and London.

Wagner, E.K. and Summers, W.C. (1978). Structure of the joint region and the termini of the DNA of herpes simplex virus type 1. J. Virol. 27; 374-387.

Wagner, M.J., Sharp, J.A. and Summers, W.C. (1981). Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. Proc. Natl. Acad. Sci. U.S.A. 78; 1441-1445.

Wang, A.H.-J., Quigley, G.J., Kolpak, F.J., Crawford, J.L., van Boom, J.H., van der Marel, G. and Rich, A. (1979). Molecular structure of a left-handed double helical DNA fragment at atomic resolution. Nature 282; 680-686.

Wasylyk, B., Derbyshire, R., Guy, A., Molko, D., Roget, A., Teolue, R. and Chambon, P. (1980). Specific *in vitro* transcription of conalbumin gene is drastically decreased by a single-point mutation in TATA box homology sequence. Proc. Natl. Acad. Sci. U.S.A. 77; 7024-7028.

Watson, R.J. (1983). DNA sequence of the herpes simplex virus type 2 glycoprotein D gene. Gene 26; 307-312.

Watson, R.J. and Clements, J.B. (1978). Characterization of transcription-deficient temperature-sensitive mutants of herpes simplex virus type 1. Virology 91; 364-379.

Watson, R.J. and Clements, J.B. (1980). Identification of a herpes simplex virus type 1 function continuously required for synthesis of early and late virus RNAs. Nature 285; 329-330.

Watson, R.J. and Vande Woude, G.F. (1982). DNA sequence of an immediate-early gene (IEmRNA-5) of herpes simplex virus type 1. Nucl. Acids Res. 10; 979-991.

Watson, R.J., Preston, C.M. and Clements, J.B. (1979). Separation and characterization of herpes simplex virus type 1 immediate-early mRNAs. J. Virol. 31; 42-52.

Watson, R.J., Sullivan, M. and Vande Woude, G.F. (1981a). Structure of two spliced herpes simplex virus type 1 immediate-early mRNAs which map at the junctions of the unique and reiterated regions of the virus DNA S component. J. Virol. 37; 431-444.

Watson, R.J., Umene, K. and Enquist, L.W. (1981b). Reiterated sequences within the intron of an immediate-early gene of herpes simplex virus type 1. Nucl. Acids Res. 9; 4189-4199.

Watson, R.J., Weis, J.H., Salstrom, J.S. and Enquist, L.W. (1982). Herpes simplex virus type-1 glycoprotein D gene: nucleotide sequence and expression in *Escherichia coli.* Science 218; 381-384.

Waubke, R., zur Hausen, H. and Henle, W. (1968). Chromosomal and autoradiographic studies of cells infected with herpes simplex virus. J. Virol. 2; 1047-1054.

Weaver, R.F. and Weissmann, C. (1979). Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucl. Acids Res. 7; 1175-1193.

Weinberg, G., Ullman, B. and Martin, D.W.Jr. (1981). Mutator phenotypes in mammalian cell mutants with distinct biochemical defects and abnormal deoxyribonucleoside triphosphate pools. Proc. Natl. Acad. Sci. U.S.A. 78; 2447-2451.

Wellauer, P. and Dawid, I. (1974). Secondary structure maps of ribosomal RNA and DNA. J. Mol. Biol. 89; 379-395.

Weller, T.H. (1953). Serial propagation in vitro of agents producing inclusion bodies derived from varicella and herpes zoster. Proc. Soc. Exp. Biol. Med. 83; 3440-3446.

Weller, T.H. (1971). The cytomegalovirus: ubiquitous agents with protein clinical manifestation. New England J. Med. 285; 267-274.

Weller, T.H. (1976). Varicella-herpes zoster virus. In 'Viral Infections of Humans: Epidemiology and Control.' pp. 457-480. (A.S. Evans, ed.) Plenum Publishing Corp., New York.

Weller, T., Witton, H. and Bell, E. (1958). The etiologic agents of varicella and herpes zoster: isolation, propagation and cultural characteristics in vitro. J. Exp. Med. 108; 843-868.

Weller, S.K., Lee, K.J., Sabourin, D.J. and Schaffer, P.A. (1983). Genetic analysis of temperature sensitive mutants which define the gene for the major herpes simplex virus type 1 DNA binding protein. J. Virol. 45; 354-366.

Weller, S.K., Spadaro, A., Schaffer, J.E., Murray, A.W., Maxam, A.M. and Schaffer, P.A. (1985). Cloning, sequencing and functional analysis of ori_L, a herpes simplex type 1 origin of DNA synthesis. Mol. Cell. Biol. 5; 930-942.

Westrate, M.W., Geelen, J.L.M.C. and van der Noordaa, J. (1980). Human cytomegalovirus DNA: physical maps for the restriction endonucleases BglII, HindIII and XbaI. J. Gen. Virol. 49; 1-22.

Whalley, J.M., Robertson, G.R. and Davison, A.J. (1981). Analysis of the genome of equine herpesvirus type 1: arrangement of cleavage sites for restriction endonucleases Eco RI, Bgl II and Bam HI. J. Gen. Virol. 57; 307-323.

Whitley, R.J. (1985). Epidemiology of herpes simplex viruses. In 'The Herpesviruses.' vol 3, pp. 1-44. (B. Roizman, ed.) Plenum Press, New York.

Whitton, J.L. (1984). Structural elements involved in herpes simplex virus immediate-early transcription. Ph. D. Thesis, University of Glasgow.

Whitton, J.L. and Clements, J.B. (1984a). Replication origins and a sequence involved in coordinate induction of the immediate-early gene family are conserved in an intergenic region of herpes simplex virus. Nucl. Acids Res. 12; 2061-2079.

Whitton, J.L. and Clements, J.B. (1984b). The junctions between the repetitive and the short unique sequences of the herpes simplex virus genome are determined by the polypeptide coding regions of two spliced immediate-early mRNAs. J. Gen. Virol. 65; 451-466.

Whitton, J.L., Rixon, F.J., Easton, A.J. and Clements, J.B. (1983). Immediate-early mRNA-2 of herpes simplex viruses types 1 and 2 is unspliced: conserved sequences around the 5' and 3' termini correspond to transcription regulatory signals. Nucl. Acids Res. 11; 6271-6287.

Wickens, M. and Stephenson, P. (1984). Role of the conserved AAUAAA sequence: four AAUAAA point mutants prevent messenger RNA 3' end formation. Science 226; 1045-1051.

Wieringa, B., Meyer, F., Reiser, J. and Weissmann, C. (1983). Unusual sequence of cryptic splice sites utilized in the β-globin gene following inactivation of an authentic 5' splice site by site-directed mutagenesis. Nature 301; 38-43.

Wieringa, B., Hofer, E. and Weissmann, C. (1984). A minimum intron length but no specific internal sequence is required for splicing the large rabbit β-globin intron. Cell 37; 915-925.

Wigler, M., Pellicer, A., Silverstein, S. and Axel, R. (1978). Biochemical transfer of single-copy eucaryotic genes using total cellular DNA as donor. Cell 14; 725-731.

Wilcox, K.W., Kohn, A., Sklyanskaya, E. and Roizman, B. (1980). Herpes simplex virus phosphoproteins. I. Phosphate cycles on and off some viral polypeptides and can alter their affinity for DNA. J. Virol. 33; 167-182.

Wildy, P., Russell, W.C. and Horne, R.W. (1960). The morphology of herpesvirus. Virology 12; 204-222.

Wilkie, N.M. (1973). The synthesis and substructure of herpesvirus DNA: the distribution of alkali-labile single strand interruptions in HSV-1 DNA. J. Gen. Virol. 21; 453-467.

Wilkie, N.M. (1976). Physical maps for herpes simplex virus type 1 DNA for restriction endonucleases HindIII, HpaII and XbaI. J. Virol. 20; 222-233.

Wilkie, N.M. and Cortini, R. (1976). Sequence arrangements in herpes simplex virus type 1 DNA: identification of terminal fragments in restriction endonuclease digests as evidence for inversion in redundant and unique sequences. J. Virol. 20; 211-221.

Williams, M.V. (1984). Deoxyuridine triphosphate nucleotidylhydrolase induced by herpes simplex virus type 1: purification and characterisation of induced enzyme. J. Biol. Chem. 259; 10080-10084.

Wolf, K. and Darlington, R.W. (1971). Channel catfish virus: a new herpes virus of ictalurid fish. J. Virol. 8; 525-533.

Wolontis, S. and Jeansson, S. (1977). Correlation of herpes simplex virus types 1 and 2 with clinical features of infection. J. Infect. Dis. 135; 28-33.

Woychik, R.P., Camper, S.A., Lyons, R.H., Horowitz, S., Goodwin, E.C. and Rottman, F.M. (1982). Cloning and nucleotide sequencing of the bovine growth hormone gene. Nucl. Acids Res. 10; 7197-7210.

Wright, S., Rosenthal, A., Flavell, R. and Grosveld, F. (1984). DNA sequences required for regulated expression of β-globin genes in murine erythroleukemia cells. Cell 38; 265-273.

Wu, M., Hyman, R.W. and Davidson, N. (1979). Electron microscopic mapping of proteins bound to herpes simplex virus DNA. Nucl. Acids Res. 6; 3427-3441.

Yamada, M., Takeda, Y., Okamoto, K. and Hirota, Y. (1982). Physical map of the nrdA-nrdB-ftsB-glpT region of the chromosomal DNA of Escherichia coli. Gene 18; 309-318.

Yamanishi, K., Ogino, T. and Takahashi, M. (1975). Induction of celluar DNA synthesis by a temperature sensitive mutant of herpes simplex virus type 2. Virology 67; 450-462.

Yeh, Y.-C. and Tessman, I. (1972). Control of pyrimidine biosynthesis by phage T4. II. Complementation between ribonucleotide reductase mutants. Virology 47; 767-772.

Yeh, Y.-C., Dubovi, E.J. and Tessman, I. (1969). Control of pyrimidine biosynthesis by phage T4: mutants unable to catalyse the reduction of cytidine diphosphate. Virology 37; 615-623.

Young, R.A., Hagenbuchle, O. and Schibler, U. (1981). A single mouse ∝-amylase gene specifies two different tissue-specific mRNAs. Cell 23; 451-458.

Zaret, K.S. and Sherman, F. (1982). DNA sequence required for efficient transcription termination in yeast. Cell 28; 563-573.

Zaret, K.S. and Sherman, F. (1984). Mutationally altered 3' ends of yeast CYC1 mRNA affects transcript stability and translational efficiency. J. Mol. Biol. 177; 107-135.

Zeitlin, S. and Efstratiadis, A. (1984). In vivo splicing products of the rabbit β-globin pre-mRNA. Cell 39; 589-602.

Zezulak, K.M. and Spear, P.G. (1983). Characterisation of a herpes simplex virus type 2, 75,000 molecular weight glycoprotein antigenically related to herpes simplex virus type 1 glycoprotein C. J. Virol. 47; 553-562.

zur Hausen, H., Schulte-Holthausen, H., Klein, G., Henle, G., Clifford, P. and Santesson, L. (1970). EBV DNA in biopsies of Burkitt's tumours and anaplastic carcinomas of the nasopharynx. Nature 228; 1056-1058.

zur Hausen, H., Schulte-Holthausen, H., Wolf, H., Dorries, K. and Egger, H. (1974). Attempts to detect virus-specific DNA in human tumors. II. Nucleic acid hybridisations with complementary RNA of human herpes group viruses. Int. J. Cancer. 13; 657-664.

Zweig, M., Heilman, C.J.Jr. and Hampar, B. (1979). Identification of disulphide-linked protein complexes in the nucleocapsids of herpes simplex virus type 2. Virology 94; 442-450.

Zweig, M., Showalter, S.D., Bladen, S.V., Heilman, C.J. and Hampar, B. (1983). Herpes simplex virus type 2 glycoprotein gF and type 1 glycoprotein gC have related antigenic determinants. J. Virol. 47; 185-192.

The EMBO Journal Vo..2 No.11 pp.1953 – 1961, 1983

DNA sequence homology between two co-linear loci on the HSV genome which have different transforming abilities

John McLauchlan and J.Barklie Clements*

Institute of Virology, University of Glasgow, Church St., Glasgow G11 5JR, UK

Communicated by J.H.Subak-Sharpe
Received on 1 July 1983; revised on 1 August 1983

A transcription unit at the herpes simplex virus (HSV) type 2 transforming region, mtr-2 (map coordinates 0.580 – 0.625), comprises two early, unspliced mRNAs of 4.5 kb and 1.2 kb which are 3′ co-terminal; a region including that specifying the 1.2-kb mRNA has been sequenced. The putative translated portions of these two mRNAs do not overlap and this feature, together with the arrangement of the mRNAs, is similar to the apparently equivalent co-linear HSV-1 locus which however does not transform. A putative stem and loops structure containing a TATA box is located upstream from the 5′ terminus of the 1.2-kb mRNA within the translated portion of the 4.5-kb mRNA. Evidence for the generation of this structure by intra-strand reassociation under our hybridisation conditions has been obtained and possibilities are that it may function as a transcription-activated promoter or as an RNA polymerase pause site. A comparison of the equivalent HSV-2 and HSV-1 regions reveals a conserved sequence downsteam from the 3′ co-terminus which is present at a similar location in many eukaryotic genes (consensus sequence YGTGTTYY). The overall sequence conservation at this transcription unit is high except for regions located at: (1) the untranslated leaders of the 1.2-kb mRNAs; (2) the N termini of the polypeptides specified by the HSV-2/HSV-1 1.2-kb mRNAs; (3) the intergenic region beyond the 3′ co-termini. Regions (2) and (3) are located within a transforming fragment of HSV-2. The possible significance of these data for HSV-mediated cell transformation is discussed.

Key words: conserved 3′-terminal sequence/DNA sequence homology/HSV-2 transcription unit/protein-coding regions/stem and loops structure

Introduction

Two separate regions (*Bgl*II *c* and mtr-2) of the herpes simplex virus type 2 (HSV-2) genome (Figure 1) have been implicated in morphological and oncogenic transformation of rodent cells and hamster cells (Galloway *et al.*, 1982; Reyes *et al.*, 1979; Jariwalla *et al.*, 1980). An interesting feature of the mtr-2 region, represented by the HSV-2 *Bgl*II *n* fragment (Figure 1, map coordinates 0.580 – 0.625), is that the co-linear HSV-1 locus does not transform cells; a different HSV-1 region (mtr-1, Figure 1, map coordinates 0.31 – 0.42) was found positive for transformation (Camacho and Spear; 1978; Reyes *et al.*, 1979). However, the equivalent HSV-2 and HSV-1 regions between maps coordinates 0.58 and 0.62 show good homology by heteroduplex mapping except for a short stretch at map coordinate 0.61 (Kudler *et al.*, 1983).

We have shown that the HSV-1 region around map coordinate 0.6 specifies two early, 3′ co-terminal mRNAs of 5.0 kb and 1.2 kb which are unspliced and overlapping (McLauchlan and Clements, 1982, 1983). They encode polypeptides of 140 000 and 40 000 mol. wt., respectively (Anderson *et al.*, 1981). Upstream signals involved in transcription initiation of the 1.2-kb mRNA and sequences transcribed within the 5′ leader of the 1.2-kb mRNA appear to encode the C terminus of the 140 000 mol. wt. polypeptide (McLauchlan and Clements, 1983). The HSV-1-specified ribonucleotide reductase activity has been mapped to this location (Dutia, 1983; V.G.Preston, personal communication); this activity is therefore a function of the 140 000 mol. wt. polypeptide though it is possible that the 40 000 mol. wt. polypeptide is also involved. A similar enzymatic activity has been described for HSV-2 and it has been suggested that ribonucleotide reductase activity may be the basis of the transforming ability (Huszar and Bacchetti, 1983).

This paper deals with the mtr-2 region, wherein we locate the 5′ and 3′ termini of an unspliced, early 1.2-kb mRNA which is 3′ co-terminal with, and overlapped by, a 4.5-kb mRNA. The genomic DNA sequence specifying the 1.2-kb mRNA has been obtained, and signals involved in transcription initiation and termination have been identified. The putative coding region of the 1.2-kb mRNA and the C-terminal portion of the polypeptide specified by the 4.5-kb mRNA have been determined and show that the two coding regions do not overlap. Comparison of this HSV-2 region with the equivalent locus of HSV-1 (Draper *et al.*, 1982; McLauchlan and Clements, 1983) reveals high overall sequence conservation except for three discrete tracts of non-homology. Beyond the 3′ co-terminus in a region of low sequence homology, we have identified a conserved sequence which is present at a similar location in many eukaryotic genes.

Results and Discussion

A transcription unit located at the mtr-2 region comprises two early mRNAs

The HSV-1 locus between map coordinates 0.56 and 0.60 specifies two early, 3′ co-terminal mRNAs. To determine the approximate sizes and relative abundances of the mRNA species located at the equivalent HSV-2 locus, a nick-translated *Bgl*II/*Bam*HI fragment (Figure 1, map coordinates 0.580 – 0.583), from within mtr-2, was hybridised to a Northern blot containing HSV-2 early polyadenylated RNA. Two bands with approximate sizes of 5.4 kb and 1.5 kb were detected and the 1.5-kb mRNA species was ~5 times more abundant (Figure 2); an additional band of 6.6 kb was observed on prolonged exposure. Jenkins *et al.* (1982) have detected mRNA species with similar sizes using radiolabelled *Bgl*II fragment *n*. Nuclease S1 and exonuclease VII analyses have shown the arrangement of the mRNAs at the mtr-2 region to be identical to that observed at the corresponding HSV-1 locus.

The *Bgl*II/*Bam*HI fragment, used in the Northern blot analysis, was uniquely 5′-labelled at the *Bam*HI site and hybridised to infected and mock-infected cell RNAs. A band of 143 nucleotides was detected with the infected cell RNA only (Figure 5A, lane2), and thus the 5′ portion of a right-

*To whom reprint requests should be sent.

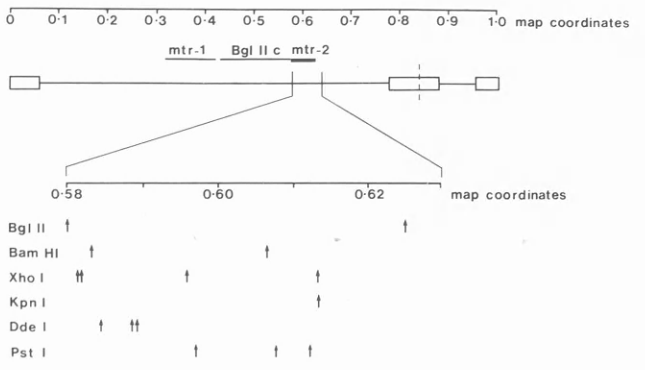


Fig. 1. Restriction endonuclease cleavage maps of HSV-2 strain HG52 at the genome region located between map coordinates 0.58 and 0.63. The cleavage sites for *Dde*I have been mapped between 0.58 and 0.61. The positions of the HSV-1 (mtr-1) and both HSV-2 (*Bgl*II fragment *c* and mtr-2) transforming regions are indicated.

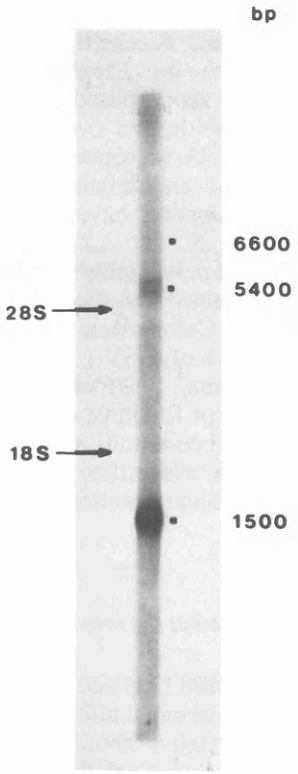


Fig. 2. Identification of the mRNAs located within a portion of *Bgl*II *n*. The Northern blot contains 5 μg of HSV-2 early polyadenylated cytoplasmic RNA and the DNA probe was a nick-translated *Bgl*II/*Bam*HI fragment (Figure 1, map coordinates 0.580 – 0.583). The sizes of the mRNA species detected are indicated.

wards-transcribed mRNA is located within this fragment. The 5′ portion of this mRNA was analysed further using *Bam*HI *e* (map coordinates 0.532 – 0.583), 5′-labelled at both ends. A nuclease S1 product of 143 nucleotides and an exonuclease VII-resistant product of 148 nucleotides were detected (Figure 3A, lanes 2 and 4). The small size difference between the protected fragments is a result of the processive nature of exonuclease VII activity as described in Materials and methods. The result indicates that the 5′ portion of the mRNA, transcribed across the *Bam*HI site, is unspliced.

The 3′ portion of the mRNA was analysed using a *Bam*HI/*Xho*I fragment (Figure 1, map coordinates 0.583 – 0.595), uniquely 3′-labelled at the *Bam*HI site. A band of 1.05 kb was detected in both nuclease S1- and exonuclease VII-digested samples (Figure 3B, lanes 5, 7 and 9); it follows that the 3′ portion is unspliced. Therefore, the total size of this mRNA is ~1.2 kb and it corresponds to the 1.5-kb species detected by Northern blot analysis.

A mRNA was shown to be 3′ co-terminal with and overlapping the 1.2-kb mRNA by using a *Bgl*II/*Kpn*I fragment (Figure 1, map coordinates 0.580 – 0.612) that had been uniquely 3′-labelled at the *Bgl*II site, which is upstream from the 5′ end of the 1.2-kb mRNA (Figure 3C). Nuclease S1- and exonuclease VII-resistant products of 1.6 kb were detected (Figure 3B, lanes 1 and 3) which places the 3′ terminus of this mRNA and of the 1.2-kb mRNA at the same location. The 5′ terminus of the larger overlapping mRNA lies outside *Bgl*II *n* and the mRNA is unspliced with a total size of 4.5 kb (J.McLauchlan and J.B.Clements, unpublished data). This mRNA corresponds to the 5.4-kb species detected on Northern blots.

In summary, the HSV-2 transcription unit comprises two early, 3′ co-terminal mRNAs of sizes 4.5 kb and 1.2 kb which are unspliced and the 5′ end of the smaller mRNA is located within the sequences specifying the larger mRNA (Figure 3C). The 6.6-kb mRNA species, detected by Northern blot analysis, is a late mRNA which also is 3′ co-terminal with the two early mRNAs (J.McLauchlan and J.B.Clements, unpublished data).

The 3′ co-terminus: an HSV-2/HSV-1 sequence comparison reveals conserved regions

The 3′ co-terminus was positioned within a *Dde*I/*Xho*I fragment (Figure 1, map coordinates 0.589 – 0.595) that had been uniquely 3′-labelled at the *Dde*I site. The major nuclease S1 product (Figure 4A, lane 2) positioned the 3′ co-terminus within the sequence TTGTT (Figure 4B, positions 1452 – 1456), 20 nucleotides downstream from the hexanucleotide sequence AATAAA (Figure 4B, positions 1432 – 1437). Minor nuclease S1 products were detected reproducibly which placed other 3′ termini at three preferred positions up to 70 nucleotides downstream from the major 3′ location.

Draper *et al.* (1982) have reported that the major 3′ co-terminus in HSV-1 strain KOS preceded the AATAAA poly(A) recognition signal (Proudfoot and Brownlee, 1976). Our data for both HSV-1 strain 17+ (McLauchlan and Clements, 1982) and HSV-2 strain HG52 unambiguously locate the 3′ co-termini downstream from the polyadenylation signal which is consistent with data for other eukaryotic genes. An AATAAA sequence (Figure 4B, positions 1605 – 1610), which could serve as a poly(A) signal for a leftwards-transcribed mRNA, is present in the strand complementary to that specifying the 3′ co-terminal mRNAs. Similarly, a poly(A) signal for a late, leftwards-transcribed mRNA is present at the equivalent HSV-1 locus (McLauchlan and Clements, 1982). The distance between the poly(A) signals on opposite strands is 168 nucleotides for HSV-2 and 125 nucleotides for HSV-1.

Upstream from the 3′ co-terminus, the HSV-2 DNA sequence was highly homologous with the corresponding HSV-1 region as far as the AATAAA signal (Figure 4B, positions 1361 – 1437); beyond this point, the homology decreased markedly. In the untranscribed region, beyond the 3′ co-terminus, three blocks of homology were observed; TTGGGTCTC (Figure 4B, positions 1455 – 1463); TGTGATT

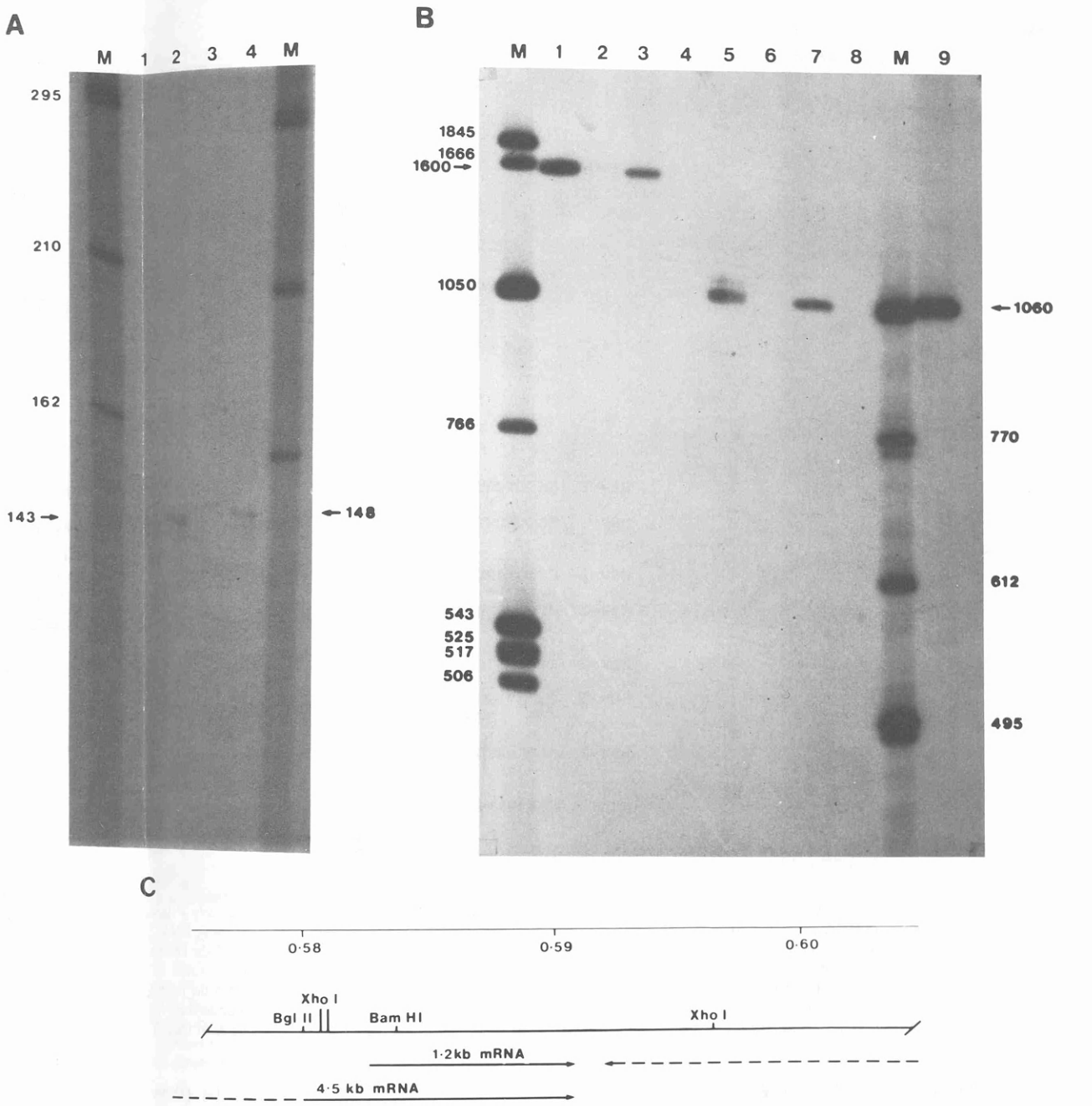


Fig. 3. Structural analysis of the mRNAs forming an early transcription unit. **(A)** Analysis of the 5′ portion of the 1.2-kb mRNA. The DNA probe was *Bam*HI *e* (Figure 1, map coordinates 0.532 – 0.583), 5′-labelled at both *Bam*HI sites. The RNA samples used were: **lanes 1** and **3**, mock-infected cytoplasmic RNA; **lanes 2** and **4**, early infected cell cytoplasmic RNA. **(B)** Analysis of the 3′ portions of the 1.2-kb and 4.5-kb mRNAs. The DNA probes used were: **lanes 1 – 4**, for the 4.5-kb mRNA, a *Bgl*II/*Kpn*I fragment (Figure 1, map coordinates 0.580 – 0.612) was uniquely 3′-labelled at the *Bgl*II site; **lanes 5 – 9**, for the 1.2-kb mRNA, a *Bam*HI/*Xho*I fragment (Figure 1, map coordinates 0.583 – 0.595) was uniquely 3′-labelled at the *Bam*HI site. The RNA samples were: **lanes 1, 3, 5** and **7**, early infected cell cytoplasmic RNA; **lanes 2, 4, 6** and **8**, mock-infected cytoplasmic RNA; **lane 9**, early polyadenylated infected cell cytoplasmic RNA. After hybridisation, samples 1 and 2 in **(A)** and samples 1, 2, 5, 6 and 9 in **(B)** were digested with nuclease S1. Exonuclease VII-treated samples were **3** and **4** in (A) and **3, 4, 7** and **8** in **(B)**. The nuclease-resistant material was electrophoresed on 8% denaturing polyacrylamide gels. Nuclease-resistant fragments are arrowed and their sizes indicated. The size standards (M) used were: ϕX174 DNA/*Hinc*II markers and pSV40/*Hinf*I markers in **(B)**. **(C)** Summary of the genome map locations and orientations of the HSV-2 mRNAs between map coordinates 0.58 and 0.60. The proposed location of a leftwards-transcribed mRNA also is shown.

(Figure 4B, positions 1467 – 1473), GGGGGTGG (Figure 4B, positions 1488 – 1495). The highly conserved sequence between positions 1467 and 1475 is related to a consensus sequence YGTGTTYY which we have found to be located ~15 – 35 nucleotides downstream from the poly(A) signal of many viral and eukaryotic cell mRNAs (J.McLauchlan and J.B.Clements, in preparation). Sequences similar to our consensus sequence have been noted by Taya *et al.* (1982). A portion of the sequence GGGGGTGG (Figure 4B, positions 1488 – 1495) is present at a similar location downstream from the 3′ end of the HSV-1 thymidine kinase gene (McKnight, 1980).

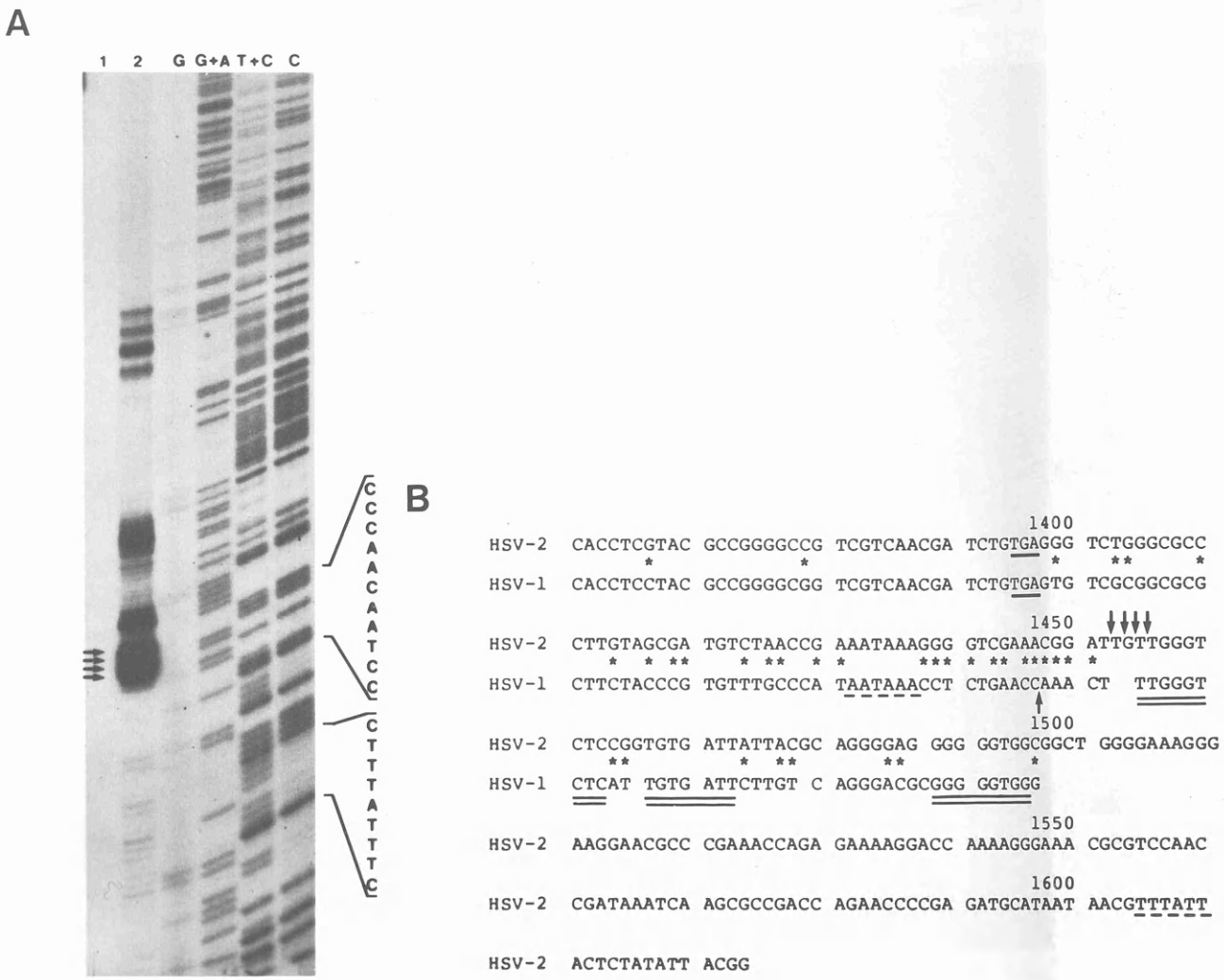


Fig. 4. (A) Precise location of the 3′ co-terminus. A *Dde*I/*Xho*I fragment (Figure 1, map coordinates 0.589 – 0.595), uniquely 3′-labelled at the *Dde*I site, was hybridised to the following RNA samples: **lane 1**, mock-infected cytoplasmic RNA; **lane 2**, early infected cell cytoplasmic RNA. Samples were digested with nuclease S1 and electrophoresed on an 8% denaturing polyacrylamide gel alongside the sequence reaction products of the DNA probe. Arrows indicate the major location of the 3′ co-terminus. **(B)** DNA sequence at the 3′ co-terminus; a comparison with a portion of the equivalent HSV-1 region. Locations of the 3′ co-termini are arrowed and the poly(A) signals for the 3′ co-terminal mRNAs are indicated by a dotted line as is the poly(A) signal for a leftwards-transcribed mRNA. Translational stop codons for the HSV-2 38 000 mol. wt. and HSV-1 40 000 mol. wt. proteins are underlined. The asterisks denote nucleotides which are not conserved and the homologous regions downstream for the 3′ co-termini are indicated by two solid lines. The sequences are numbered according to the HSV-2 sequence in Figure 6.

The 5′ terminus of the 1.2-kb mRNA: evidence for a stem and loops structure

We position the 5′ terminus (Figure 5A) on the guanosine residue within the sequence ATGTAC (Figure 6, positions 245 – 250). Upstream from the 5′ end, a TATA box sequence (Corden *et al.*, 1980) is located at positions −22 to −28 (Figure 6, positions 221 – 225) and a pentameric sequence GGTCC (Busslinger *et al.*, 1980) at positions −44 to −49 (Figure 6, positions 199 – 203). These signals, which are thought to be involved in the initiation of transcription, are completely conserved and are located at identical positions relative to the 5′ ends of the HSV-2 and HSV-1 1.2-kb mRNAs (Figure 7, positions 220 – 224 and 198 – 202).

In both HSV-2 (Figure 6, positions 164 – 233) and HSV-1 (Figure 7, positions 164 – 232), a region upstream from the 1.2-kb mRNA, within a transcribed portion of the 4.5-kb mRNA, contains inverted repeats which could form the stem and loops structures shown in Figure 5B. The structure comprises a stem of 42 nucleotides (shown by R1/R1′ and R2/R2′ in Figure 6) with a loop of 24 nucleotides and a loop of six nucleotides containing the TATA box. Under our DNA/RNA hybridisation conditions, a minor proportion of molecules were generated in which the stem and loops structure was formed by intra-molecular reassociation of the DNA probe. In these molecules, sequences within the stem and loops were unavailable for hybridisation to the 4.5-kb mRNA, and the intra-molecular loops were sensitive to nuclease S1. The remainder of the DNA probe was available for DNA/RNA hybridisation. The minor bands indicated in Figure 5A (lane 2) correspond to nuclease S1-sensitive sites located within both loops; no nuclease S1 cleavage sites were detected at the stem region. Similar nuclease S1-sensitive sites were observed for HSV-1 (McLauchlan and Clements, unpublished observations). This indicates that the stem and loops structure can be generated by intra-molecular reassociation but does not provide evidence for its natural occurrence in virus DNA.

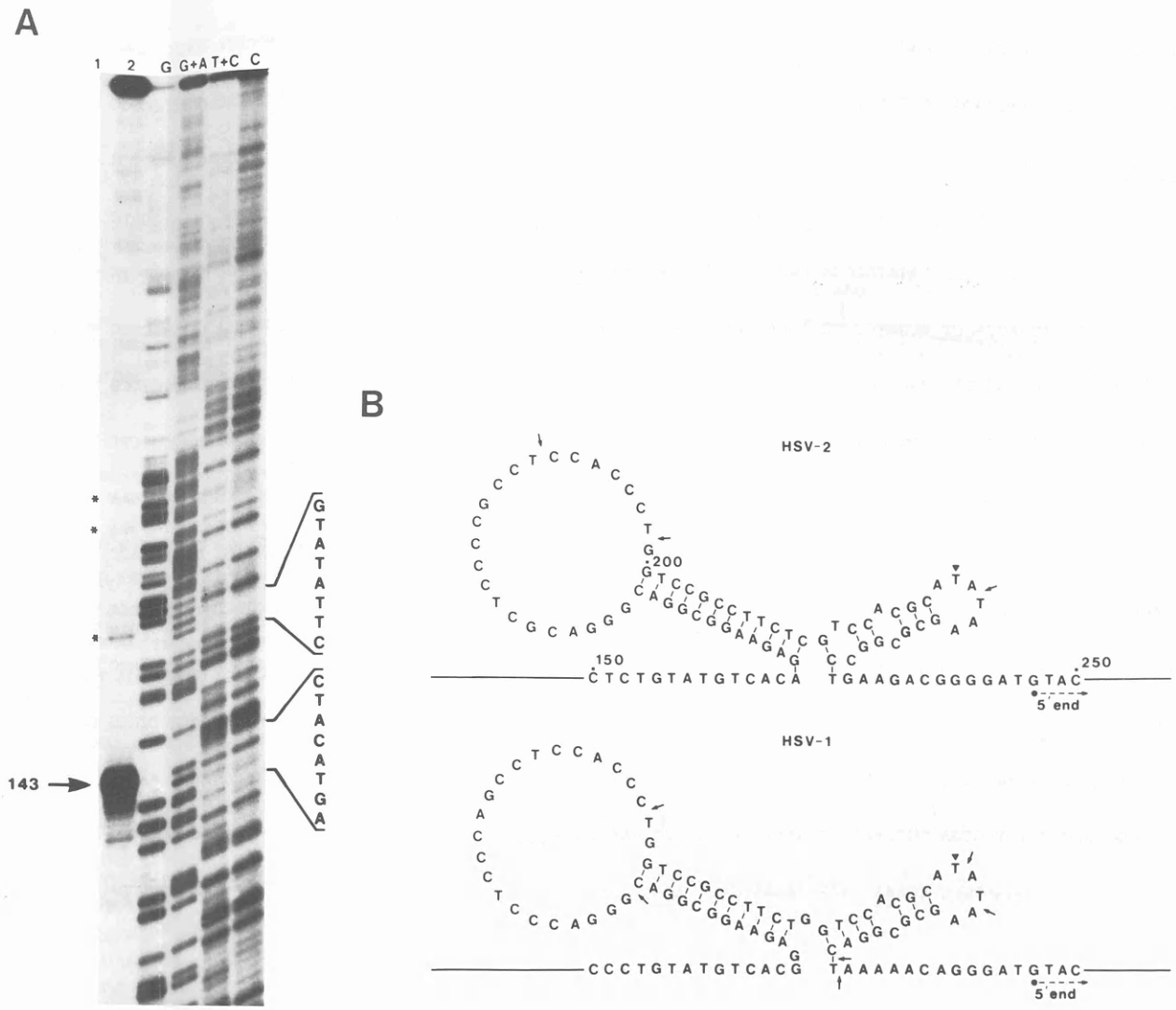


Fig. 5. (A) Precise location of the 5′ terminus of the 1.2-kb mRNA. A *Bgl*II/*Bam*HI fragment (Figure 1, map coordinates 0.580 − 0.583), uniquely 5′-labelled at the *Bam*HI site, was hybridised to the following RNA samples: **lane 1**, mock-infected cytoplasmic RNA; **lane 2**, early infected cell cytoplasmic RNA. Following hybridisation, the samples were digested with nuclease S1 and electrophoresed on an 8% denaturing polyacrylamide gel alongside the sequence reaction products of the DNA probe. The arrow indicates the location of the 5′ terminus and the size of the hybrid was determined relative to φX174/*Hin*cII size standards (not shown). Asterisks locate the minor nuclease S1 products described in the text. **(B)** Proposed stem and loops structures formed by inverted repeat sequences upstream from the HSV-2 and HSV-1 1.2-kb mRNAs. The first T in each of the TATA boxes is indicated (▼) and the 5′ termini are shown. Arrows denote minor nuclease S1 sites; the HSV-2 sequence is numbered as in Figure 6.

We propose that this structure may act as a promoter element in the synthesis of the 1.2-kb mRNA. The presence of the larger loop would tend to destabilise the structure in duplex DNA, however it would form much more readily in denatured regions of double-stranded DNA. Transcription of the overlapping 4.5-kb mRNA through this locus generates a single-stranded DNA region which may facilitate the intra-strand formation of the stem and loops and thereby potentiate initiation of the 1.2-kb mRNA. As a consequence of this feature, synthesis of the 1.2-kb mRNA would be dependent upon transcription of the 4.5-kb mRNA and the structure would function essentially as a transcription-activated promoter. Alternatively, if this structure is present in duplex DNA, it may function as a pause site for RNA polymerase molecules transcribing the 4.5-kb mRNA and facilitate initiation of the 1.2-kb mRNA.

Transcription regulatory signals of the 1.2-kb mRNA are located within the translated portion of the overlapping 4.5-kb mRNA

The DNA upstream from the 5′ end of the 1.2-kb mRNA contains signals involved in transcription initiation and also is transcribed as part of the 4.5-kb mRNA. A single open reading frame which extends for 255 nucleotides and terminates 82 nucleotides downstream from the 5′ end of the 1.2-kb mRNA is present also in HSV-1 and the sequences are highly conserved (Figure 7, positions 1 − 327). We have inferred previously that the C terminus of the HSV-1 140 000 mol. wt. protein lies within the 5′-untranslated leader of the HSV-1 1.2-kb mRNA (McLauchlan and Clements, 1983). The corresponding HSV-2 protein of 138 000 mol. wt. (Docherty *et al.*, 1981) appears to be encoded by the 4.5-kb

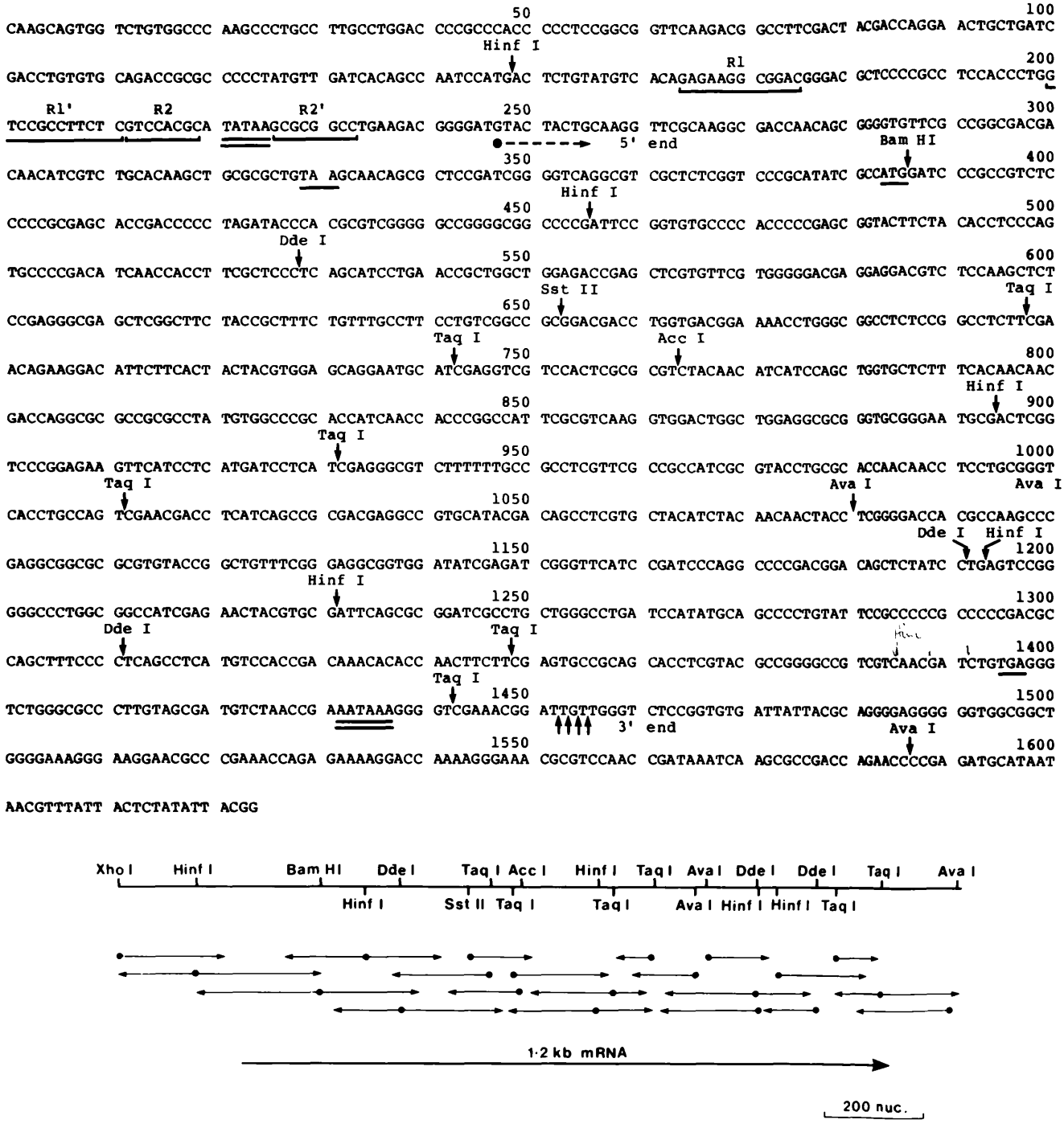


Fig. 6. Nucleotide sequence of the HSV-2 1.2-kb mRNA. The 5′ and 3′ termini are shown and the TATA box and AATAAA signals are indicated. The codons which represent the N and C termini of the 38 000 mol. wt. protein and the C terminus of the 138 000 mol. wt. protein are underlined as are the inverted repeats (R1, R1′, R2 and R2′) which form the stem and loops structure shown in Figure 5B. A sequencing strategy is shown below and the sites used to determine the nucleotide sequence are indicated on both diagrams.

mRNA. The same open reading frame within the equivalent genome regions of HSV-2 and HSV-1 strongly suggests that the HSV-2 coding region is organised in a similar fashion to HSV-1 and therefore, that the translational stop codon of the 138 000 mol. wt. protein is located at position 329 (Figure 6). Most of the HSV-2/HSV-1 differences in this region occur in the third base of the codons, and the predicted amino acid sequence is highly conserved (Figure 7, positions 1 – 327).

Two highly conserved polypeptides differing at the N-terminal region

Previous reports have mapped a 38 000 mol. wt. polypeptide to the region examined here (Docherty *et al.*, 1981; Galloway *et al.*, 1982). The first ATG triplet at position 384 (Figure 6), within the sequences specifying the 1.2-kb mRNA, precedes a single open reading frame which extends for 337

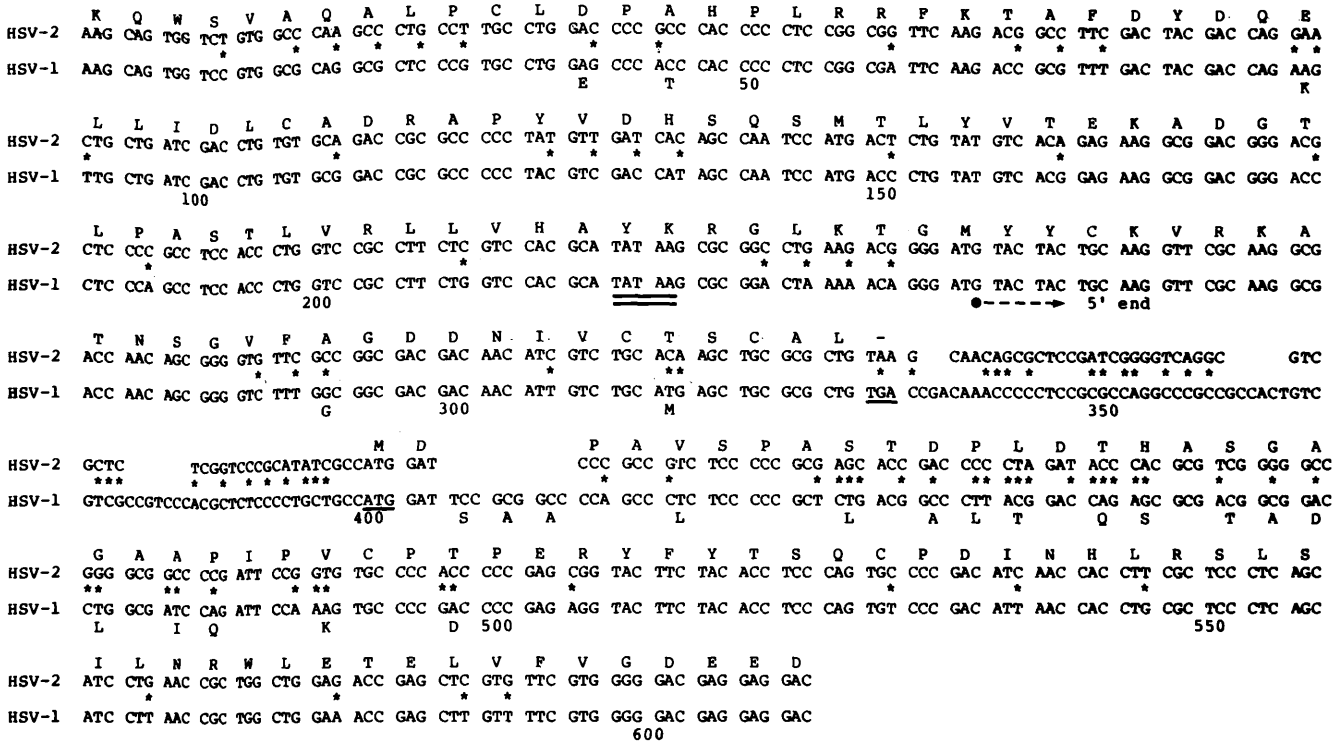


Fig. 7. A comparison of the HSV-2 and HSV-1 sequences in a region encoding the C terminus of the 138 000 mol. wt. protein and the N terminus of the 38 000 mol. wt. protein. The coding regions are arranged in triplets and the amino acid sequences of the HSV-2 polypeptides are shown above the nucleotide sequence. HSV-1 amino acids which are not conserved are shown below the nucleotide sequence and the asterisks indicate nucleotides which are non-homologous. The C termini of the 138 000/140 000 mol. wt. and the N termini of the 38 000/40 000 mol. wt. proteins are underlined. The 5′ termini of the 1.2-kb mRNAs are shown and the TATA boxes are indicated. The HSV-1 sequence (McLauchlan and Clements, 1983) is numbered.

codons and terminates with a TGA stop signal at position 1395 (Figure 6). This sequence would specify a 37 500 mol. wt. polypeptide, which is close to the 37 800 mol. wt. obtained by *in vitro* translation of mRNA selected with the HSV-2 *Bgl*II fragment *n* (Docherty *et al.*, 1981).

For the HSV-1 1.2-kb mRNA, we inferred that the first ATG initiation codon (Figure 7, position 313), which was followed by a stop codon in the same reading frame (Figure 7, position 328), did not encode the N terminus of the 40 000 mol. wt. protein (McLauchlan and Clements, 1983). Interestingly, this ATG triplet is absent in the HSV-2 sequence (Figure 7). The sequences flanking the first ATG codon of the HSV-2 1.2-kb mRNA are identical to those at the second ATG triplet of the HSV-1 1.2-kb mRNA (Figure 7, positions 396 – 404) which reinforces our view that the second ATG codon is used in the HSV-1 1.2-kb mRNA.

If the HSV-1 and HSV-2 sequences are aligned to maximise DNA homology, the sequence predicts that the HSV-2 protein is three amino acids shorter than the HSV-1 polypeptide at the N terminus (Figure 7, positions 405 – 413). Further downstream, there is a discrete region of low homology which results in 14 amino acid differences within a stretch of 22 amino acid residues (Figure 7, positions 432 – 497). This may represent a different functional domain between the HSV-2 and HSV-1 polypeptides. Following this non-homologous region, from positions 498 – 611 (Figure 7), which is the limit of our HSV-1 sequence data, the amino acid residues are identical in both serotypes. Recently, Draper *et al.* (1983) realised that their previously published sequence of the HSV-1 strain KOS 1.2-kb mRNA (Draper *et al.*, 1982) was incorrect. A comparison of the HSV-2 amino acid sequence with their amended HSV-1 DNA sequence (obtained from E.K.Wagner) reveals that the 38 000 mol. wt. and 40 000 mol. wt. polypeptides are highly conserved apart from the already mentioned region of ~90 nucleotides at the N terminus. Most nucleotide differences from position 474 (Figure 6) occur in the third base of the codons and the HSV-2 and HSV-1 reading frames terminate at the same stop codon (Figure 6, position 1395).

The polypeptide coding regions of the 4.5-kb and 1.2-kb mRNAs do not overlap

From their immunoprecipitation studies with monoclonal antibodies, Galloway *et al.* (1982) concluded that the coding regions of the 138 000 and 38 000 mol. wt. polypeptides were at least partially co-linear. This is incompatible with our findings. Galloway's results were obtained in the absence of any structural analysis from this region and we have shown that the mRNAs encoding these proteins are unspliced. Our locations for the C terminus of the 138 000 mol. wt. polypeptide and the N terminus of the 38 000 mol. wt. polypeptide (Figure 6, positions 329 and 384, respectively) indicate that the two coding regions do not overlap, as also is the case for the corresponding HSV-1 proteins (Figure 7, positions 328 and 399); the DNA sequences between these locations do not encode protein and this untranslated portion shows low homology between HSV-2 and HSV-1. The sequences in this region are non-co-linear as the HSV-1 sequence contains an additional 16 nucleotides (Figure 7, positions 331 – 398).

The immunoprecipitation data could be the result of a common antigenic site within the 138 000 and 38 000 mol. wt. polypeptides, however, a computer-assisted comparison

of the amino acid sequence of the 38 000 mol. wt. polypeptide with the available amino acid sequence of the 138 000 mol. wt. protein revealed no homology. An alternative explanation for the immunoprecipitation data might be that the two polypeptides are associated; hence, an antibody directed against one protein could precipitate both polypeptides.

Significance of non-homologous regions for transforming ability

Overall homology between the HSV-2 and HSV-1 DNA sequences in this genome region is high, however there are tracts of low conservation. Two such tracts are located within the 2.1-kb *Bam*HI/*Pst*I sub-fragment (Figure 1, map coordinates 0.583–0.596) of HSV-2 *Bgl*II *n* which causes the morphological transformation of cells *in vitro* (Galloway and McDougall, 1983). One of these tracts is downstream from the 3′ co-terminus, in the intergenic region, where there are additional nucleotides in the HSV-2 sequence. The other tract is at the N terminus of the HSV-2 38 000 mol. wt. protein which contains 14 out of 22 amino acid changes compared with the equivalent HSV-1 40 000 mol. wt. polypeptide. The 2.1-kb transforming fragment contains all the coding sequences of the 38 000 mol. wt. polypeptide except for the ATG initiation codon. Galloway *et al.* (1982) were unable to detect a viral 38 000 mol. wt. antigen in transformed cells. However, a 35 000 mol. wt. protein has been isolated from HSV-2-transformed hamster cells which has an identical proteolytic cleavage pattern to that of a 35 000 mol. wt. product detected in HSV-2-infected cells (Suh, 1982). This 35 000 mol. wt. polypeptide may be the 38 000 mol. wt. protein described here. Ribonucleotide reductase activity is likely to be a function of the 38 000 and/or the 138 000 mol. wt. polypeptides, and mechanisms have been proposed whereby alterations in the regulatory control of this enzyme may cause mutations in cellular DNA (Weinberg *et al.*, 1981).

The mechanism of HSV-induced cell transformation is unclear; Galloway and McDougall (1983) favour a 'hit and run' mechanism whereby transformation could result either from transient expression of a viral gene or by insertion of an element which may stimulate a cellular function. In the latter case, the activating element could be a small segment of the *Bgl*II *n* fragment which was undetectable by the hybridisation procedures used. A candidate for activating cellular gene expression would be the putative stem and loops structure present in both HSV-2 and HSV-1. This explanation would not account for the different transforming abilities of this region between HSV-2 and HSV-1 unless, for example, the efficiency of integration differed. Alternatively, the region responsible for the different transforming abilities could lie outside the segment which we have sequenced. Whatever the mechanism of HSV-mediated cell transformation, it is clear that maintenance of the transformed phenotype is not necessarily dependent upon either the stable acquisition of viral DNA sequences or the prolonged expression of virus gene products.

Materials and methods

Cells and virus

Baby hamster kidney 21 (C13) cells were grown as monolayers in 800 ml plastic tissue culture flasks (Clements *et al.*, 1977). For the production of early RNA, cell monolayers were infected with HSV-2 (Glasgow strain HG52) at a multiplicity of infection of 10 p.f.u./cell and incubated at 31°C for 3 h.

Cell fractionation and isolation of RNA

Cytoplasmic cell fractions were prepared and RNA was isolated as described previously (Easton and Clements, 1980).

Selection of polyadenylated mRNA

Polyadenylated mRNA from HSV-2-infected cells was prepared by elution from oligo(dT)-cellulose as described previously (Watson *et al.*, 1979).

Enzymes

All enzymes were obtained from Bethesda Research Laboratories or New England Biolabs, with the exception of T4 polynucleotide kinase (P-L Biochemicals) and nuclease S1 (Boehringer Corporation Limited). DNA was digested with restriction endonucleases at 37°C in 50–200 μl 6 mM Tris HCl, pH 7.5, 6 mM $MgCl_2$ and 6 mM 2-mercaptoethanol.

Cloning procedures

Fragments of HSV-2 DNA, generated using restriction endonucleases, were cloned within the Institute of Virology under Category I Containment (UK Genetic Manipulation Advisory Group). The host bacterium was *Escherichia coli* K12 HB101 and the cloning vector was pAT 153 (Twigg and Sherratt, 1980). Isolation of cloned virus DNA was as described by Davison and Wilkie (1981).

Purification and radiolabelling of DNA fragments

Purification of DNA fragments from agarose or polyacrylamide gels, and labelling of the 5′ and 3′ ends was carried out as described by McLauchlan and Clements (1982).

In order to generate fragments with uniquely-labelled ends, the DNA fragments, either 5′- or 3′-labelled at both ends, were redigested with a second restriction endonuclease.

DNA was labelled by nick-translation with DNA polymerase I and [α-^{32}P]-deoxyribonucleoside triphosphates (3000 Ci/mmol) as described by Rigby *et al.* (1977).

Northern blot analysis

Cytoplasmic polyadenylated RNA was electrophoresed on denaturing formaldehyde gels and transferred to nitrocellulose as described by Spandidos and Paul (1982). Prehybridisations and hybridisation of the DNA probe to the nitrocellulose were performed in 50% formamide, 1 x Denhardt's (1966) solution (0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin), 2 x SSC (1 x SSC = 0.15 M NaCl, 0.015 M sodium citrate) and 50 μg/ml calf thymus DNA. Prior to hybridisation, the DNA probe was denatured in 100% formamide at 100°C for 10 min and then rapidly chilled in ice. Hybridisations were carried out at 42°C for 16 h. After hybridisation, the nitrocellulose was washed extensively in 2 x SSC at 60°C and the bands visualised by autoradiography at −70°C using Kodak X-Omat-S film.

Structural analysis of mRNAs

Structural analysis of mRNAs was performed as described previously (McLauchlan and Clements, 1983), using the nuclease S1 and exonuclease VII digestion procedures of Berk and Sharp (1978), which were modified by using either 5′- or 3′-end-labelled DNA probes instead of uniformly-labelled DNA (Weaver and Weissmann, 1979).

Due to the processive nature of the exonuclease VII activity (Chase and Richardson, 1974), this enzyme will leave several undigested nucleotides at a hybrid end. This accounts for the slightly larger size of the exonuclease VII-resistant bands as compared with the equivalent nuclease S1-resistant bands.

Gel electrophoresis

Denaturing polyacrylamide gels, essentially as described by Maxam and Gilbert (1980), were run in 90 mM Tris, 90 mM boric acid, pH 8.3, 1 mM EDTA and the gels contained 9 M urea. Samples were dissolved in deionised formamide and denatured at 90°C for 2 min before loading. Electrophoresis was carried out at room temperature for 3–6 h at 40 watts. The radiolabelled bands were detected by autoradiography at −70°C.

DNA sequencing

DNA sequences were determined by chemical degradation (Maxam and Gilbert, 1980) of 5′- and 3′-labelled DNA fragments. The restriction endonuclease sites which were used to determine the nucleotide sequence are shown in Figure 6 and all the sequences shown were obtained for both strands of DNA.

Acknowledgements

We wish to thank Professor J.H.Subak-Sharpe for helpful discussion and for his critical reading of the manuscript.

References

Anderson,K.P., Frink,R.J., Devi,G.B., Gaylord,B.H., Costa,R.H. and Wagner,E.K. (1981) *J. Virol.*, **37**, 1011-1027.

Berk,A.J. and Sharp,P.A. (1978) *Cell*, **14**, 695-711.
Busslinger,M., Portmann,R., Irminger,J.C. and Birnstiel,M.L. (1980) *Nucleic Acids Res.*, **8**, 957-977.
Camacho,A. and Spear,P.G. (1978) *Cell*, **15**, 993-1002.
Chase,J.W. and Richardson,C.C. (1974) *J. Biol. Chem.*, **249**, 4533-4561.
Clements,J.B., Watson,R.J. and Wilkie,N.M. (1977) *Cell*, **12**, 275-285.
Corden,J., Wasylyk,B., Buchwalder,A., Sassone-Corsi,P., Kedinger,C. and Chambon,P. (1980) *Science (Wash.)*, **209**, 1406-1414.
Davison,A.J. and Wilkie,N.M. (1981) *J. Gen. Virol.*, **55**, 315-331.
Denhardt,D.T. (1966) *Biochem. Biophys. Res. Commun.*, **23**, 641-646.
Docherty,J.L., Subak-Sharpe,J.H. and Preston,C.M. (1981) *J. Virol.*, **40**, 126-132.
Draper,K.G., Frink,R.J. and Wagner,E.K. (1982) *J. Virol.*, **43**, 1123-1128.
Draper,K.G., Frink,R.J. and Wagner,E.K. (1983) *J. Virol.*, **45**, 899 (Erratum).
Dutia,B.M. (1983) *J. Gen. Virol.*, **64**, 513-521.
Easton,A.J. and Clements,J.B. (1980) *Nucleic Acids Res.*, **8**, 2627-2645.
Galloway,D.A., Goldstein,L.C. and Lewis,J.B. (1982) *J. Virol.*, **42**, 530-537.
Galloway,D.A. and McDougall,J.K. (1983) *Nature*, **302**, 21-24.
Huszar,D. and Bacchetti,S. (1983) *Nature*, **302**, 76-79.
Jarriwalla,R.J., Aurelian,L. and Ts'o,P.O.P. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 2279-2283.
Jenkins,F.J., Howett,M.K., Spector,D.J. and Rapp,F. (1982) *J. Virol.*, **44**, 1092-1096.
Kudler,L., Jones,T.R., Russell,R.J. and Hyman,R.W. (1983) *Virology*, **124**, 86-99.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
McKnight,S.L. (1980) *Nucleic Acids Res.*, **8**, 5949-5964.
McLauchlan,J. and Clements,J.B. (1982) *Nucleic Acids Res.*, **10**, 501-512.
McLauchlan,J. and Clements,J.B. (1983) *J. Gen. Virol.*, **64**, 997-1006.
Proudfoot,N.T. and Brownlee,G.G. (1976) *Nature*, **263**, 211-214.
Reyes,G.R., LaFemina,R., Hayward,S.D. and Hayward,G.S. (1979) *Cold Spring Harbor Symp. Quant. Biol.*, **44**, 629-641.
Rigby,W.J.B., Dieckermann,M., Rhodes,C. and Berg,P. (1977) *J. Mol. Biol.*, **113**, 237-251.
Spandidos,D.A. and Paul,J. (1982) *EMBO J.*, **1**, 15-20.
Suh,M. (1982) *J. Virol.*, **41**, 1095-1098.
Taya,Y., Devos,R., Tavernier,J., Cheroutre,H., Engler,G. and Fiers,W. (1982) *EMBO J.*, **1**, 953-958.
Twigg,A.J. and Sherratt,D. (1980) *Nature*, **283**, 216-218.
Watson,R.J., Preston,C.M. and Clements,J.B. (1979) *J. Virol.*, **31**, 42-52.
Weaver,R.F. and Weissmann,C. (1979) *Nucleic Acids Res.*, **7**, 1175-1193.
Weinberg,G., Ullman,B. and Martin,D.W.,Jr. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2447-2451.

Volume 13 Number 4 1985 Nucleic Acids Research

The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini

John McLauchlan, Dairena Gaffney, J.Lindsay Whitton and J.Barklie Clements*

MRC Virology Unit, Institute of Virology, Church Street, Glasgow G11 5JR, UK

Received 10 December 1984; Revised and Accepted 22 January 1985

ABSTRACT

Our previous DNA sequence comparisons of 3' terminal portions from equivalent herpes simplex virus type 1 (HSV-1) and HSV-2 genes identified a conserved sequence (consensus YGTGTTYY; Y=pyrimidine) located approximately 30bp downstream from the AATAAA signal. We report here that this signal is located downstream from 67% of the mammalian mRNA 3' termini examined. Using constructions with the bacterial chloramphenicol acetyl transferase (CAT) gene linked to an HSV 'terminator' fragment, we show that deletions in the 'terminator' reduce CAT activities and the levels of CAT mRNA 3' termini. Specifically: (1) deletions of downstream sequences which extend up to the consensus YGTGTTYY signal reduce CAT levels to values 35% of those obtained with undeleted plasmids, (2) a deletion of a further 14bp, which removes the YGTGTTYY consensus but not the poly A site, reduces CAT activities to 1%-4%. The levels of CAT mRNA 3' termini reflect the reductions in CAT activities howevever, levels of mRNA 5' termini are unaffected by these deletions. The RNA produced in the absence of the YGTGTTYY signal is present in the cytoplasm although no CAT activity is detectable.

INTRODUCTION

Sequence comparison of genes transcribed by RNA polymerase II has identified a number of consensus sequences certain of which have been shown by functional assays to be important in the control of transcription. For example, the TATA homology, initially identified by comparison of 5' flanking regions (1), was shown to be required for transcription initiation (2). We have used HSV-1/HSV-2 DNA sequence comparison of equivalent genomic regions to identify discrete blocks of homology and have shown that these correspond to transcription regulatory signals which are located upstream from mRNA 5' termini (3, 4).

In contrast to 5' end formation, analysis of DNA sequences required for the formation of mRNA 3' termini has received less

attention. The highly conserved hexanucleotide sequence AATAAA (5) is required for polyadenylation of mRNAs transcribed by RNA polymerase II (6); a point mutation within this sequence abolishes the ability to form correctly terminated mRNA (7), and a ∝-thalassaemic defect is due to an altered AATAAA signal (8). However, it is clear that sequences apart from the AATAAA signal are required for the polyadenylation process since this sequence is present at sites other than at mRNA 3' termini.

Our HSV DNA sequence comparisons also identified a conserved sequence (consensus YGTGTTYY) located downstream from the poly A site at about 30 nucleotides from the AATAAA signal (3, 9, 10, 11) which is similar to a sequence noted by Taya et al (12). Our study of DNA sequences at the 3' termini of mammalian poly A^+ mRNAs reveals that the YGTGTTYY signal is present downstream from many genes. We also report, using plasmid constructions with the bacterial CAT gene, that the YGTGTTYY signal, together with other further downstream signals, is required for efficient formation of an mRNA 3' terminus.

MATERIALS AND METHODS

Growth of Cells and Calcium Phosphate Transfection

HeLa cells were grown as monolayers on 90mm Petri dishes in Dulbecco's modified Eagles medium supplemented with 2.5% foetal calf serum (Flow Laboratories) and 2.5% calf serum.

Subconfluent HeLa cells were transfected (13) with calcium phosphate precipitates containing 15µg of pTER plasmid, 10µg of the pseudorabies (PRV) plasmid, pPRVKpnh (14, 15), and 10µg of pβ(244+)β, a β-globin plasmid (16). In experiments where CAT activities alone were assayed, the β-globin plasmid was omitted. At 24h after addition of the calcium phosphate precipitates, the medium was removed and 15ml of fresh medium was added to each plate.

Cloning Procedures

The host bacterium was E.coli K12 strain JM 83 (17) and the cloning vector was pUC9 (18). Cloning procedures for the construction of the pTER plasmid series and for pBTER1 are given in Results, and the construction of plasmid pLW1 is described in Gaffney et al (19). The various deletions in plasmid pTER3 were

generated by linearisation with Bam HI followed by treatment with exonuclease III; blunt-ended DNA was produced using Mung Bean nuclease and then ligated to give the plasmid series pTER3 Δ. Deletion end-points were determined by DNA sequencing (20).

Radiolabelling of DNA Fragments

DNA fragments were either 5'- or 3'-labelled as described by McLauchlan and Clements (21). In order to generate single-stranded DNA probes, the end-labelled fragments were denatured and electrophoresed on a 5% polyacrylamide gel at 4°C (22).

RNA Isolation and Structural Analysis of mRNAs

Cytoplasmic cellular fractions were prepared and RNA was isolated as described by Easton and Clements (23).

Structural analysis of mRNAs was performed using the nuclease S1 procedure (24, 25). Either 5'- or 3'-labelled DNA probes and 15-20μg of cytoplasmic RNA were mixed in a buffer containing 50% (v/v) formamide (deionised with Amberlite monobed resin MB-2), 400mM NaCl, 40mM PIPES pH6.8 and 1mM EDTA to a final volume of 30μl. This mixture was heated to 90°C for 3min and then incubated at 42°C for 16h. Prior to nuclease S1 treatment, the hybridisation mixtures were quenched in ice.

Nuclease S1 digestion was performed at 30°C for 2-3h in 300μl of 400mM NaCl, 40mM sodium acetate pH 4.5, 1mM $ZnSO_4$ with 5000 units of nuclease S1. Following phenol/chloroform extraction and ethanol precipitation, the nuclease S1-resistant hybrids were electrophoresed on a 6% denaturing polyacrylamide gel.

CAT Assays

Cell extracts were prepared as described by Gorman et al (26). The levels of CAT activity were measured by incubating aliquots from cell extracts in a mixture containing 1mM acetyl CoA, 250mM TRIS pH7.8, and 0.125μCi ^{14}C chloramphenicol (68nmoles/μCi) at 37°C. Aliquots were removed at various times, extracted with 200μl of ethyl acetate and the organic phase was dried down. The ^{14}C chloramphenicol was resuspended in 30μl ethyl acetate and spotted onto 0.25mm silica gel thin layer chromatography plates. Ascending chromatography was performed in 95% chloroform: 5% methanol, after which the chromatograms were

autoradiographed. The percentage conversion to acetylated ^{14}C chloramphenicol was calculated by removing the spots from the chromatograms and counting in a scintillation counter. CAT activities were normalised by measuring the protein content of each extract (27) which allowed calculation of CAT specific activities.

RESULTS

To test the effect of removing sequences downstream from the poly A site, plasmid constructions were made using the CAT gene. An HSV-2 IE promoter fragment was inserted upstream from the CAT gene giving plasmid pLW1 (19), and a 'terminator' fragment from the same IE gene was inserted downstream to give plasmid pTER5.

The HSV-2 IE promoter was a 210bp SmaI/Sau3A fragment from IE gene-4/5 which comprised 91bp of 5' flanking sequence and 119bp of leader sequence (10). This promoter fragment contains all the sequences required for the initiation of transcription and the location of the mRNA 5' terminus is identical to that determined for the HSV-2 IE mRNA-4/5 synthesised *in vivo* (10, 19). Plasmid pLW1 has no ATG in the HSV-2 leader sequence; the first initiation codon in the mRNA is the CAT gene ATG.

The 'terminator' was a 100bp Sma I/Xba I fragment from the 3' terminus of HSV-2 IE gene-5. This fragment contains an AATAAA signal, and located 8 nucleotides downstream from the poly A site is the sequence TGTGTTGC which fits the YGTGTTYY consensus (10) except in one position. A comparison of the HSV-2 and HSV-1 DNA sequences from the 3' terminal portions of IE gene-5 (Fig. 1) reveals conservation of the AATAAA signal together with two blocks of homology downstream from the poly A site. The conserved block between positions 62-65 (Fig. 1) contains the sequence corresponding to the YGTGTTYY consensus.

Construction of 'terminator' plasmids

Fig. 2 details the construction of plasmids used for short-term expression assays. Briefly, the 'terminator' fragment was cloned into pUC9 to give plasmid pTER3. The 'terminator' was then inserted into the pLW1 promoter/CAT plasmid at a site just downstream from the CAT gene to give plasmid pTER5. In pTER5,

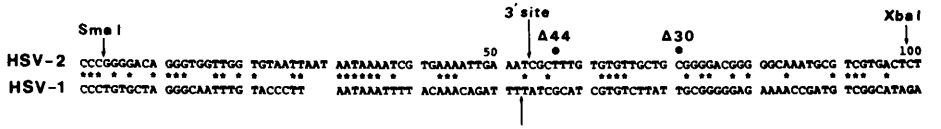


FIG. 1. DNA sequence comparison at the mRNA 3' termini of the HSV-2 and HSV-1 IE genes-5 (10). The asterisks denote nucleotides which are conserved. Locations of the 3' termini and the deletion end-points in pTER5 plasmids are indicated. The AATAAA signal and the sequences homologous to the YGTGTTYY motif are at positions 31-36 and 61-68 respectively.

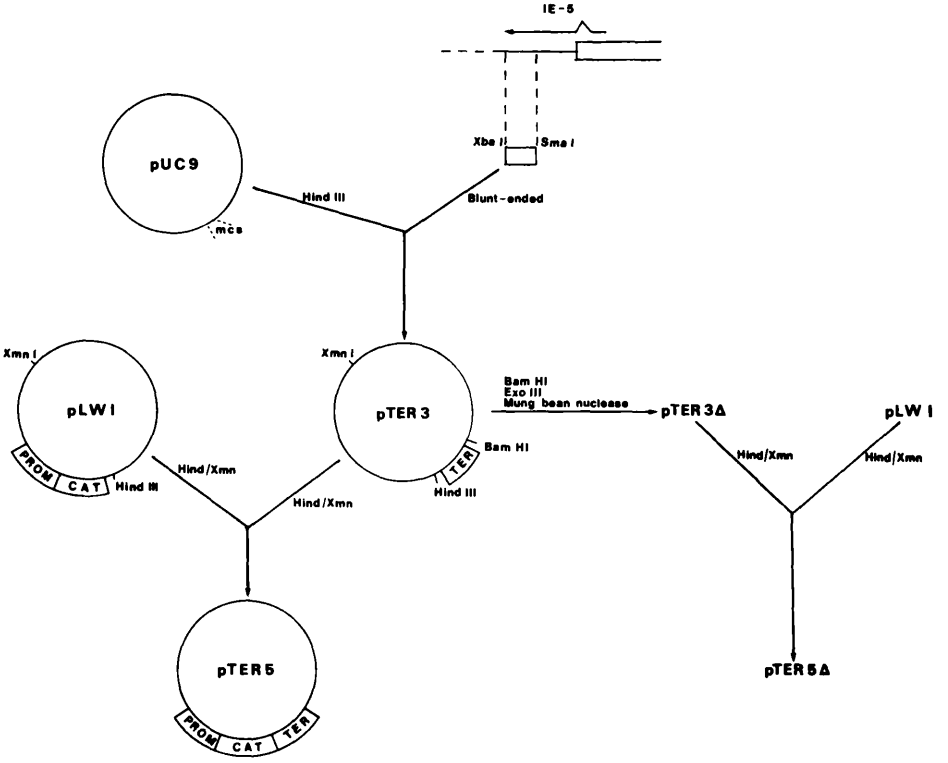


FIG. 2. Construction of pTER plasmids. The Sma I/Xba I fragment (Fig. 1) was derived from the HSV-2 plasmid pBam b' (10) and blunt-ended using Klenow fragment polymerase with all four dNTP's. The multiple cloning site of pUC9 is designated mcs.

TABLE 1 CAT activites obtained with various plasmid constructions are shown for 4 separate transfection experiments.

	1		2	3	4	poly A site	YGTGTTYY
pTER5	100	(4.01)	100	100	100	+	+
pTER5 Δ 30	35	(1.41)	34	32	39	+	+
pTER5 Δ 44	3	(0.12)	4	1	1	+	-
pTER6	1.5	(0.02)	2	0.6	1	inverted	
pLW1	0.4	(0.01)	U	U	U	-	-
pBTER1	110	(4.45)				+ β-globin	+

Activities are expressed as a percentage of the pTER5 value; U indicates undetectable CAT activity. The figures in brackets from experiment 1 are typical CAT values expressed as nmoles converted chloramphenicol/ ug protein/ h. Plasmids pTER Δ 30 and pTER5 Δ 44 contain deletions in the HSV 'terminator' fragment; pTER6 has an 'inverted terminator' ; pLW1 has no 'terminator' fragment; PBTER1 contains a 1250bp rabbit β-globin fragment. The presence (+) or absence (-) of the poly A site and YGTGTTYY signal in each plasmid is indicated.

the 'terminator' is orientated with the AATAAA signal on the mRNA sense strand.

A series of deletions was made in plasmid pTER3 from a BamHI site located in the multiple cloning site downstream from the 'terminator'. These deletions removed sequences from the 3' portion of the 'terminator' and extended for different lengths through the YGTGTTYY homology towards the AATAAA sequence to give the plasmid series pTER3Δ. These 'terminator' deletions were then inserted into the pLW1 promoter/CAT plasmid to give the expression plasmid series pTER5 Δ (Fig. 2).

Removal of 'terminator' sequences reduces CAT activity

Plasmid pLW1 lacks the pTER5 'terminator' and Table 1 shows that Hela cells transfected with pLW1 produced little detectable CAT activity whereas high levels of activity were obtained with pTER5. The 'terminator' fragment therefore was required for detectable CAT gene expression. To determine the effect on CAT activity of an alternative 'terminator', a 1250bp Bam HI/Xho I fragment of the rabbit β-globin gene containing all the IVS 2 sequences and some 600bp of flanking DNA (28) was inserted into the Hind III site of pLW1. Plasmid pBTER1 contained this rabbit β-globin fragment with the AATAAA signal on the mRNA sense strand. Table 1 shows that CAT activity obtained with pBTER1 was

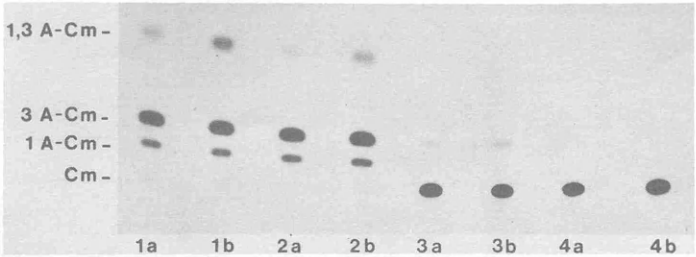


FIG. 3. CAT activities produced by the pTER plasmids as determined by thin layer chromatography. Aliquots were removed from each reaction after (a) 20min and (b) 40min. The positions of the acetylated and non-acetylated forms of ^{14}C chloramphenicol are indicated. Lanes 1a, 1b pTER5; lanes 2a, 2b pTER5△30; lanes 3a, 3b pTER5△44; lanes 4a, 4b pTER6.

comparable to that obtained with pTER5.

The CAT levels obtained with the pTER5 △ series were compared with pTER5 CAT activity. On this basis, two deleted plasmids with reduced CAT activities were selected, and the deletion end-points were determined by DNA sequencing. These two plasmids pTER5 △ 30 and pTER5 △ 44 had 30bp and 44bp respectively deleted from the 3' portion of the 'terminator'. Locations of the deletion end-points are shown in Fig. 1; pTER5 △ 30 has a deletion extending to 3bp from the YGTGTTYY signal while in pTER5 △ 44 this signal is removed but the poly A site is retained.

Fig. 3 shows an autoradiograph of the CAT assays obtained with the pTER5 plasmid series and the relative CAT activities from a number of separate experiments are listed in Table 1. The CAT levels produced by pTER5 △ 30 and pTER5 △ 44 were low and transactivation by co-transfection with the PRV plasmid, pPRVKpnh, was required to boost activities to levels which could be accurately estimated. Plasmid pPRVKpnh expresses the IE regulatory protein (29) which increases transcription from other plasmid-borne eukaryotic promoters and from certain cellular promoters (15, 30, 31).

Deletion of 30bp from the 3' portion of the 'terminator' fragment reduced CAT levels to approximately 35% of the pTER5 value while removal of a further 14bp markedly reduced activity to a level of 1%-4%; pTER5 △ 44 produced CAT activities only just above the background levels obtained with plasmids pLW1 and pTER6

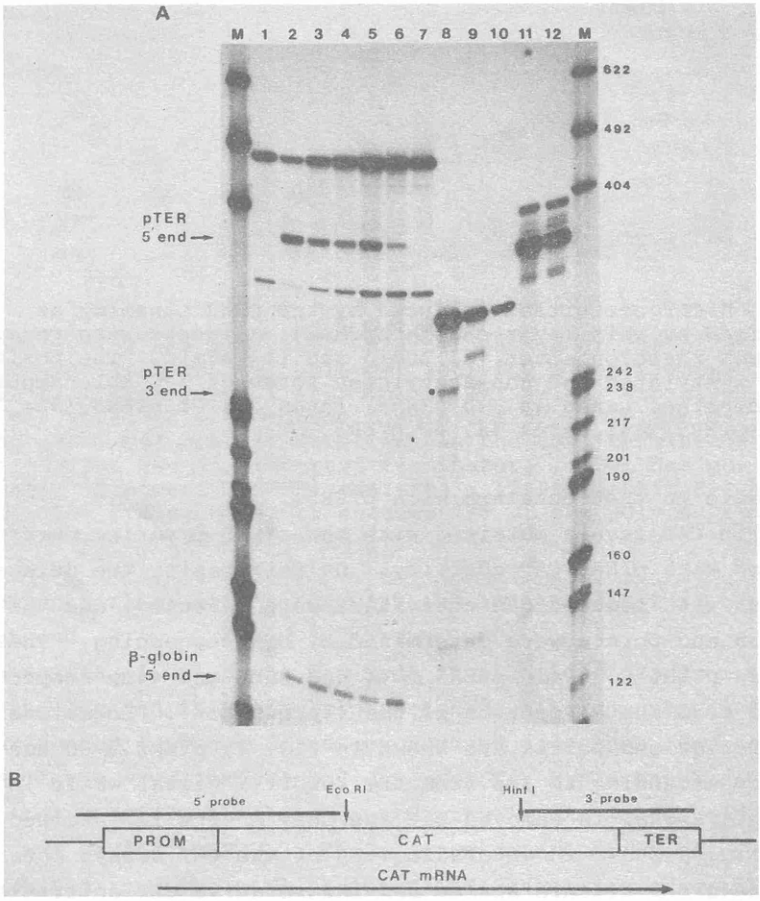


FIG. 4. (A) The abundance of mRNA 5' and 3' termini produced by the pTER plasmids. The DNA probes used in lanes 1-7 were (1) an Eco RI/Sma I fragment 5'-labelled at the Eco RI site to determine the pTER CAT mRNA 5' termini and (2) a 5'-labelled Bst NI fragment to detect the β-globin mRNA 5' termini. The single-stranded DNA probes used to detect the pTER CAT mRNA 3' termini were a Hinf I fragment from pTER5 3'-labelled at the Hinf I site in the CAT gene coding sequences (lanes 8-10) and a Hinf I/Bgl I fragment from pTER5△44 again 3'-labelled at the Hinf I site (lanes 11 and 12). RNA samples shown; lanes 1, 7, 10 and 12, mock-infected RNA; lanes 2 and 8, pTER5 RNA; lanes 3 and 9, pTER5 △30 RNA; lanes 4 and 11, pTER5△44 RNA; lane 5, pTER6 RNA; lane 6, pLW1 RNA. Molecular weight markers (M) were pAT 153 DNA cleaved with Hpa II.

(B) Approximate locations for the 5' and 3' termini of pTER CAT mRNA are shown. Restriction endonuclease cleavage sites within the CAT coding sequences which were used for making end-labelled DNA probes (see Fig. 4A) are indicated.

(Table 1). Plasmid pLW1 has no 'terminator' fragment and pTER6 has the 'terminator' inserted in the opposite orientation to that of pTER5.

Levels of mRNA 3' and 5' termini produced by the pTER5 plasmid series

Cytoplasmic RNA isolated from cells transfected with the pTER5 plasmid series was hybridised to strand-separated DNA probes which were 3'-labelled at the Hinf I site within the CAT coding sequences (Fig. 4B). RNAs produced by pTER5 and pTER5 △ 30 were hybridised to a probe produced from pTER5 DNA. As the deletion end-point of pTER5 △ 44 is only 4bp from the poly A site (Fig. 1), hybrids produced by RNA transcribed across the poly A site could not be resolved from correctly terminated RNAs, therefore RNA from pTER5 △ 44 was hybridised to a probe from pTER5 △ 44 DNA. The DNA probes were end-labelled to similar specific activities. After treatment with nuclease S1, the protected DNA fragments were electrophoresed on 6% denaturing polyacrylamide gels (Fig. 4A).

RNA produced from pTER5 protected a fragment of 247 nucleotides (Fig. 4A, track 8) and this positioned the 3' end at nucleotide 54 (Fig. 1) which is the same location as the mRNA 3' terminus for IE mRNA-5 from HSV-2 infected cells (10). The 3' end of mRNA produced by pTER5 △ 30 was located at the same position as that obtained with pTER5 but the amount of protected DNA fragment was reduced to approximately 30% of the pTER5 value (Fig. 4A, track 9). No discrete 3' end was detected with pTER5 △44 RNA (Fig. 4A, track 11). Hence, the levels of mRNA 3' termini obtained with the pTER5 series reflected the CAT activities obtained with these plasmids.

Plasmids pTER5 and pTER5△30 have identical sequences up to 18bp beyond the poly A site then the DNA sequences are dissimilar due to the terminator deletion (30bp) in plasmid pTER5△30. Therefore, the band below the fully protected probe (Fig. 4A, track 9) represents hybridisation of the pTER5 DNA probe to pTER5 △30 RNA transcribed through the poly A site. Hence the size of this band reflects the sequence homology between pTER5△30 RNA and the pTER5 DNA probe.

To assess the validity of the comparative 3' terminal data,

plasmid, pβ(244+)β, containing the β-globin promoter and the polyoma virus enhancer (16), was co-transfected with the various pTER5 plasmids. A strand-separated DNA probe, 5'-labelled at a Bst NI site approximately 130bp from the β-globin mRNA 5' terminus was used to estimate β-globin mRNA levels (32). As Fig. 4A tracks 2-6 indicate, the β-globin mRNA levels were similar in all the RNA samples.

Levels of mRNA 5' termini were estimated using a strand-separated DNA probe 5'-labelled at an EcoRI site within the CAT coding sequences of pTER5 (Fig. 4B). Using this probe, a protected DNA fragment of 369 nucleotides was obtained with RNA produced by all the pTER5 series plasmids and by pLW1 (Fig. 4A, tracks 2-6). A similar amount of labelled probe was protected with the different RNA samples. Therefore, unlike the situation with the 3'-labelled probe, the levels of protected 5' probe obtained with pTER5 △ 30 and pTER5 △ 44 were similar to that obtained with pTER5 although the CAT activities were lower than the pTER5 values. This suggests that cytoplasmic RNA with a normal 5' terminus but an incorrect 3' terminus does not produce functional protein.

DISCUSSION

Our results have demonstrated that removal of sequences downstream from the poly A site of HSV-2 IE gene-5 reduces the level of correctly processed mRNA 3' termini, and this effect is reflected in lower levels of CAT activity. Deletion plasmid pTER5 △ 30 gave reduced mRNA 3' termini and the CAT levels were 35% of the values obtained with pTER5 which contains all the 'terminator' sequences. Moreover, the removal of a further 14bp, which includes the YGTGTTYY signal but not the poly A site, markedly reduced CAT activity to 1%-4% of the pTER5 value, and no correct 3' termini were detected.

When the DNA sequences at the 3' terminal region of HSV-2 IE gene-5 were compared with those of HSV-1 IE gene-5, a portion of the 14bp segment required for efficient 3' end formation was conserved (10). Furthermore, a similar downstream sequence was conserved in other equivalent HSV-1 and HSV-2 genes at approximately the same position relative to the AATAAA signal (3,

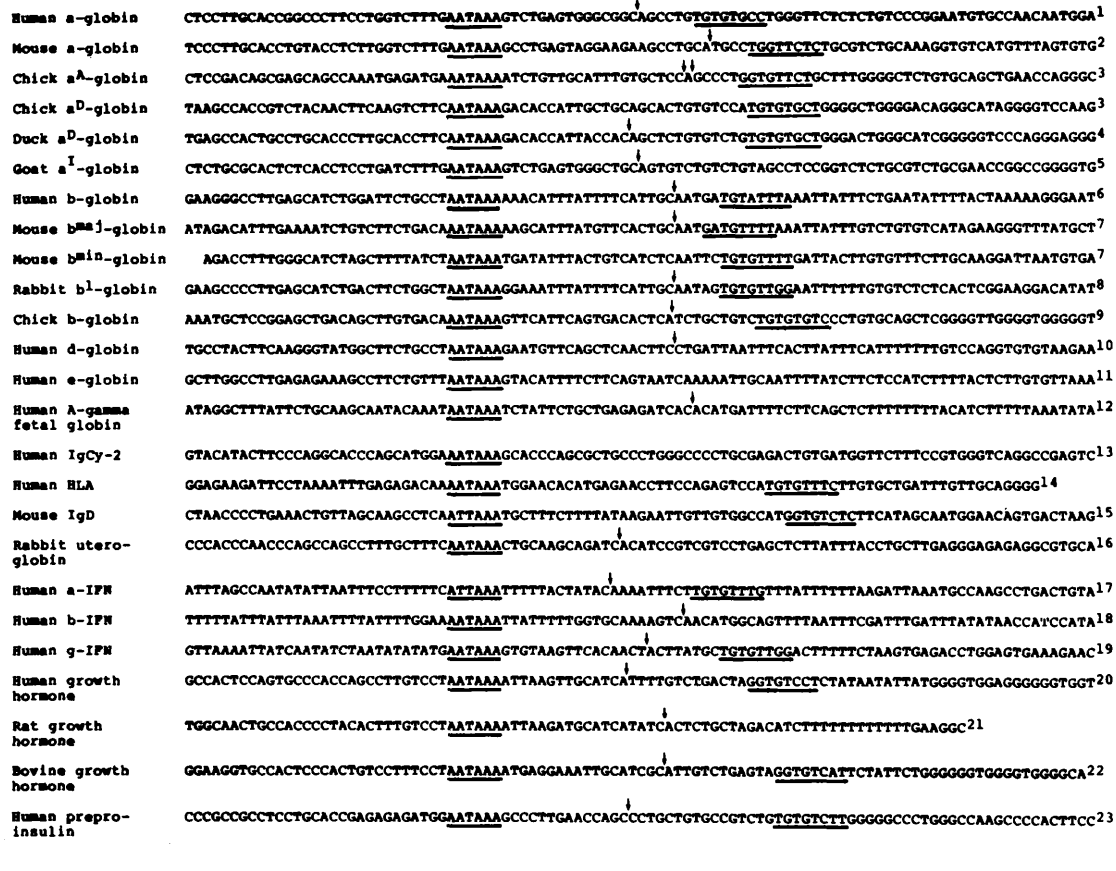
9). The 3' terminal DNA sequences from a number of mammalian and eukaryotic viral genes which specify poly A^+ mRNAs were examined (Fig. 5) and sequences similar to those conserved downstream from the poly A sites of HSV genes were identified in 67% of the genes examined. From a study of 3' terminal sequences from 100 genes, we have derived the consensus sequence YGTGTTYY, and the preferred location for this sequence is 24-30 nucleotides downstream from the AATAAA signal. We conclude from our analysis that this conserved element is likely to play an important role in the formation of mRNA 3' termini for a wide variety of mammalian genes.

Nevertheless, it appears that other sequences further downstream are required for efficient polyadenylation. McDevitt et al (33) have shown that sequences between 20 and 35 nucleotides downstream from the poly A site (ie downstream from YGTGTTYY) are required for the production of functional adenovirus 2 (Ad2) E2a mRNA however correctly terminated mRNA was produced at detectable levels. Our data also shows that removal of sequences downstream from the YGTGTTYY signal reduces the level of 3' end formation but comparison of the HSV-2 and Ad2 sequences in this region revealed no significant homology. A deletion which removes the poly A site of Ad2 E2a mRNA and the sequence AGTGTCTC (this sequence is 26 nucleotides from the AATAAA signal and fits our consensus) resulted in no detectable mRNA. Similarly, a 52bp deletion which removes the poly A site together with sequences containing TGTGTTGG also abolishes rabbit β-globin 3' end formation (34).

Indirect but compelling evidence for the importance of sequences downstream from the poly A site may be derived from an analysis of gene duplications and from the locations of small tandem repeat sequences which are present downstream from the 3' termini of HSV mRNAs. The human α 1 and α 2 globin genes have arisen from a gene duplication event and have areas of high homology (35). Sequence conservation at the 3' portions of α 1 and α 2 genes is markedly reduced beyond 15bp downstream from the poly A site, and it has been proposed that this decrease in homology delimits those sequences involved in gene duplication. We consider it significant that the sequence TGTGTGCC, which fits

COMPARISON OF MAMMALIAN AND EUKARYOTIC VIRAL mRNA 3' TERMINI

Mammalian 3' termini



Chick prepro-insulin	CTTACTCTATCGCACCTTCAAAGCATTTGAATAAACCTTGTTGGTCTACTGGAAGACTTGTGCCATTTTATGTGTCCATGGTT[24]
Dog prepro-insulin	CCTTCGGCTCTCTGCACCCCAAGTGATTCAATAAACCCTCTGAATGAGCCCTAGTGGTGTTGTCTGTGCGGCGCAGGGGTTGAGGTGTGGGCCAGGGGTC[25]
Rat prepro-insulin	CACCACTCCCCGCCCACCCCTCTGCAATGAATAAAGCCTTTGAATGAGCACCAAAATGAGAGAGTTTTTATGAATACAAAGGGATTGTGTGA[26]
Rat prepro-insulin II	ACCACTACCCTGTCCACCCCTCTGCAATGAATAAAACCTTTGAAAGAGCACTACAAGTTGTGTGTACATGCGTGCATGTGCATATGTGGTGC[27]
Human chorionic gonadotropin	GAATACAGCATTAAGCTTGTTCCACTGCAAATAAAGCCTTTTAAATCATCATTCAATCACTGAATTATCATTTTTCTTCAAAGTAAG[28]
Human prepro-enkephalin	TCAATAAATGCTTACTTGTATATAGAGATAATAAACCTATTACCCCAAGTGCATAATATCCTTGTAAGTCTCTTTTTCTCCAAGGCTC[29]
Human proopio-melanocortin	AGCCTCTTAAAGCTTGCCTGTAGTTAGGAAATAAAACCTTTCAAATTTCACATCCACCTCTGACTTTGAATGTAAACCGTGTGAATAAAGTAAAAAATAC[30]
Human para-thyroid	AAAACAAAAATGTAATGATCATAAGTCTAAATAAATGAAGTATTTCTCACTCATTGCAAGTATATCTTTTTGGTTATCACTGATACCCACATGTTTACAT[31]
Human prepro-corticotropin-releasing factor	GTAAGAGAATGTTGAATACTTATATTTTTAATAAAAGTTGCAAAGGTAATCATGACTTTACTGTTTCTCTTCAGTGCTCATACACTGGAAATACTCTCTT[32]
Rat a-tubulin	CGAGCCGTCCACTTCACTGATGGGTTTTAAATAAAATACTCCCTGTCTTAAACGTTGTCTTGCGTTCTGTATTTGGGTGCGTCCCAGCTTTCATGCTTTG[33]
Rat prostatic steroid binding protein	GAAACCTCACATTTATGTCTTTCCATTTTATTAAATTATCTGCATACAATTCTCTCTCGTGTTGTTTCATAAAAAATAACCTAAGTGGGCTGCTCTGTCT[34]
Rat b-actin	TGGCCTGTACACTGACGTGAGACCGTTTTAATAAAAGTGCACACCTTACAAACAAGTTTGTGGCTCTGTGGCTTCTACTGGGTGTGGGGAGCAGGCTGGG[35]
Chick skeletal muscle a-actin	GTGGAAGGAAACATAAAACTTCAAGAAGCATTAAATCATCAGTCATTCTGTCACACCCCCCAAGCAGTTGTTTCTGTCATCA[36]
Chick pro a2(I) collagen	AACCAAAAAAAAAAAAAACCAAAAAAATAAAATAAAAAATGATGAAAGTTTGCACCACTTGTGGCCTTTGACTATCTTCCACAGAGGAAGTTTAAAACCAC[37]
Chick pro a2(I) collagen (alternative 3' terminus)	AAATGTATTTTGTATATGTGAGATGTTTAAATAAATTGTGAAAAGAAAAAAAAATAAAGCATGTCCAGTGTTCCAAAAGAAGGTATCGAATGATTAAAGT[37]
Chick ovalbumin Y	TGTTAATTAATGGTCTTTCTATTTATTCTAATAAAAGCTTCGCAAACAAAACATGTCATTACCTATTCTGGGTTACTGTACTACACAACCTGAAAAATAT[38]
Mouse kallikrein 1	AACCCCTGAGTGACACACTGTCTGTTCTCAATAAAATCCACCATGCAACAAATGAGTTCAAGGTCTGACATTCTCTGTCCTACTAGAGTGTGACACCCAA[39]
Mouse kallikrein 2	GCCTAAGTGTCACATTATCCTCTGTTCTCAATAAAATCCACTATGCAACAAATGAGTCAAGGTCTGACAGTCTCTATCCTACTGGAATACAGTACACAAC[39]
Human metallo-thionein-II	TCCTTTTTCTATAAAATAATGTGAATGATAATAAAACAGCTTTGACTTGATTCTGTCTCTGGTTTTCTTTATATGCCTTAAAA[40]
Mouse metallo-thionein	AACCCCCGTTTTCTACCGAGTACGTTAATAATAAAAGCCTGTTTGAGTCTAACTCTGGTTTCTTGGTGTGGTTTGGCAATAAGAAACTGGGGTGACTTGA[41]
Adenovirus 3' termini	
Ad2 Ela	TTGTTTGCTGAATGAGTTGATGTAAGTTTAATAAAGGGTGAGATAATGTTTAACTTGCATGGCGTGTTAAATGGGGCGGGGCTTAAAGGGTATATAATGC[42]
Ad7 Ela	AGCTGTGTTTATTTAATGTGACGTCATGTAATAAAATTATGTCAGCTGCTGAGTGTTTTATTACTTCTTGGGTGGGGTCTTGGATATATAAGTAGGAGCA[43]

Ad12 E1a	GGGTAATGTGGTTTTTGTGAGTCATGTATAATAAAACTGGTTTCGGTTGAAGTGTCTTGTTAATGTTTGTTTGGGCGTGGTTAAACAGGGATATAAAGCT[44]
Ad2 E1b	CCTCCCCTCCCAATGCGGTTTAAAACATAAATAAAAACCAGACTCTGTTTGGATTTTGATCAAGCAAGTGTCTTGCTGTCTTTATTTAGGGGTTTTGCGC[42]
Ad7 E1b	AAAGTCTAAATAAAGATCTCAAATCAATAAATAAAGAAATACTTGTTATAAAAACAAATGAATGTTTATTTGATTTTTCGCGCGCGGTATGCCCTGGACC[43]
Ad12 E1b	TACCCATACCCCAGCCCAACCTGTAACCCAATAAAGAAAAAACTTAAATTGAGATGGTGTTATGAATCTTTATTGATACTTGTTTTTTCTGACATGGTAA[45]
Ad2 E2a	GGGGGTAAATAATCACCCGAGAGTGTACAAATAAAAGCATTTGCCTTTATTGAAAGTGTCTCTAGTACATTATTTTTACATGTTTTTCAAGTGACAAAAA[46]
Ad2 E2b	CCGGGCCTACCGCGCGCGCAAAACCCCTAAATAAAGACAGCAAGACACTTGCTTGATCAAAATCCAAACAGAGTCTGGTTTTTATTTATGTTTTAAACCG[42]
Ad7 E2b	GTCCAGGGCATACCGCGCGCGAAAAATCAAATAAACATTCATTTGTTTTTATAACAAGTATTTCTTTATTTATTGATTTGAGATCTTTATTTAGACTTTG[43]
Ad12 E2b	AGCTTACCATGTCAGAAAAAACAAGTATCAATAAAGATTCATAACACCATCTCAATTTAAGTTTTTTCTTTATTGGGTTACAGGTTGGGCTGGGGTATGG[45]
Ad2 E3	CTTATTCCATTCAACTAACATAAACACACAATAAATTACTTACTTAAAATCAGTCAGCAAATCTTTGTCCAGCTTATTCAGCATCACCTCCTTTCCTTCC[47]
Ad2 E4	AAATGACTTGAAATTTTCTGCAATTGAAAAATAAACACGTTGAAACATAACATGCAACAGGTTCACGATTCTTTATTCCTGGGCAATGTAGGAGAAGGTG[42]
Ad2 L2	TAAAAACAAGTTGCATGTGGAAAAATCAAAATAAAAAGTCTGGAGTCTCACGCTCGCTTGGTCCTGTAACTATTTTGTAGAATGGAAGACATCAACTTTG[48]
Ad2 L3	TGTAAAAATAATGTACTAGAGACACTTTCAATAAAGGCAAATGCTTTTATTTGTACACTCTCGGGTGATTATTTACCCCCACCCTTGCCGTCTGCGCCGT[46]
Ad2 L4	ATCTTTGTTGTCATCTCTGTGCTGAGTATAATAAATACAGAAATTAGAATCTACTGGGGCTCCTGTCGCCATCCTGTGAACGCCACCGTTTTTACCCACC[49]
Ad2 L5	CTCTTACACCTTCTCCTACATTGCCCAGGAATAAAGAATCGTGAACCTGTTGCATGTTATGTTTCAACGTGTTTATTTTTCAATTGCAGAAAATTTCAAG[42]
Papovavirus 3' termini	
SV40 late	TTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTT[50]
Polyoma late	CTGTATTTCCTGGAAATTAATGTTTATTCAATAAACTGTGTATTCAGCTATATCACATATGTCCTTCAAGGGGTCGTCGCCTTCTTGCACATTACAAAGT[51]
BK late	ACAGGTGCTTTTATTGTACATATACATTTAATAAATGCTGCTTTTGTATAAGCCACTTTTAAGCTTGTGTTATTTTGGGGGTGGTGTTTTAGGCCTTTTA[52]
SV40 early	AATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTG[50]
Polyoma early	TGTGATATAGCTGAATACACAGTTTATTGAATAAACATTAATTTCCAGGAAATACAGTCTTTGTTTTTCCAAAGCGGTCAACATAGCGCGTCATATCAGG[51]
BK early	AAAGCAGCATTTATTAAATGTATATGTACAATAAAAGCACCTGTTTAAAGCATTTTGGTTTGCAATTGTCCCTGTTTGTCAATATATCTTATCATATCTG[52]
BPV-1 (E mRNAs)	CTCAGATTTTATATGGTTTAAGCTGCAGCAATAAAAATGAGTGCACGAAAAAGAGTAAAACGTGCCAGTGCCTATGACCTGTACAGGACATGCAAGCAAG[53]
HPV-1a (E mRNAs)	TATCAAACACCTGCCCACCTGACATTCAAAATAAAATTGAGCATACAACAATTGCTGATAAAATATTGCAATATGGCAGTCTGGGAGTTTTTTTGGGAGG[54]
BPV-1 (L mRNAs)	GTAAAACAGAAGGTAAGTCAACTGCACCTAATAAAAATCACTTAATAGCAATGTGCTGTGTCAGTTGTTTATTGGAACCACACCCGGTACACATCCTGTC[53]
HPV-1a (L mRNAs)	CTTTCTTTTTGTTAATCCCTCCCTACTCCAATAAAAAATCCCTACCCCTAAAATCTGTTTGTGCTGGTTTTATTAATAATTGCGCTCTTTTATATAATAA[54]
Herpes simplex virus 3' termini	
HSV-1 IE mRNA 1	AATAAAAAAAACCGGGTTTGCCGTCCGTGGGAATAAAGGTTTACTTTGTAACTTTCCCGTCGTGTTGGTGTACGC(GGGGTGCGTGGGAGTC)$_{x9}$[55]
HSV-1 IE mRNA 2	TTTTTATTGCAACTTCCTGTTTTAGGTACAATAAAAACAAAACATTTCAAACAAATCGCCCCTCGTGTTGTCCTTCTTTGCTCATGGCCGGCGGGGCGTG[56]
HSV-2 IE mRNA 2	CTTCTACTGCAACTCCCTATTTTGGCAGGAATAAACATATTGCGTCAACCCAGTGGTTTCGTGTGTGTTCTGCGCG[56]
HSV-1 IE mRNA 3	GGCGGGGCGCTTGTGAGACCCGAAGACGCAATAAACGGCAACAACCTGATTAAGTTTTGCAGTAGCGTTGTTTATT(CGAGGGGCGGGAGGGGG)$_{x19}$[57, 58]

HSV-1 TK mRNA	ATACCGGAGGAGGAACCCGCGCTATGACGGCAATAAAAGACAGAATAAAACGCACGGGTGTGGGTCGTTTGTTCATAAACGCGGGGTTCGGTCCCAGGGC[59]
HSV-2 TK mRNA	CGAACGGAAGGCTCCCGGCGATGACGGCAATAAAGAACAGAATAAAAGCATTGTTGTCGTGGTGTCCATAAGCGCGGGGGTTCGGGGCCAGGG[60]
HSV-1 3' co-terminal mRNAs (map coords.0.54-0.60)	GCGGCGCGCTTCTACCCGTGTTTGCCCATAATAAACCTCTGAACCAAACTTTGGGTCTCATTGTGATTCTGTGCAGGGACGCGGGGGTGGGAGAGGATAA[61]
HSV-2 3'co-terminal mRNAs (map coords. 0.54-0.60)	TGGGCGCCCTTGTAGCGATGTCTAACCGAAATAAAGGGGTCGAAACGGATTGTTGGGTCTCCCGGTGTGATTATTACGCAGGGAGGGGGGTGGCGGCTGG[62]
HSV-1 3' terminus at map coord. 0.60	GAGCGCGCTGCCCGAACAAACGACCCACCCAATAAATTATGGTATCCTATGCCCGCAGAATCTGAACGAACCTGGTTACTGCTTTTTGCGCCGCCTTTTA[61]
HSV-2 3' terminus at map coord. 0.60	CCGTAATATAGAGTAATAAACGTTATATGCATCTCGGGGTTCTGGTCGGCGCTTGATTTATCGGTTGGACGCGTTTCCCTTTTG[62]
HSV-1 gC 3' terminus	GTTGGGGGGGTCCGGCACGGACCCAAAATATAAACACACAATCACGTGCGATAAAAAGAACACGCGGTCCCCTGTGGTGTTTTTGGTTATTTTTATTAA[63]
HSV-1 3' terminus at map coord. 0.646	GTCGACGAGATTTAATAAAATAACCAAAAACACCACAGGGGACCGCGTGTCTTTTTATCGCACGTGATTGTGTGTTTATTAF[63]
RNA Tumour Viruses	
HTLV-1	CCTAGACTCAGCCGGCTCTCCACGCTTTGCCTGACCCTGCTTGCTCAACTCTACGTCTTTGTTTCGTTTTCTGTTCGCGCGTTACAGATCGAAAGTTC[64]
A-MuLV	GACTGAGTCCGCCCGGGTACCCGTGTATCCAATAAACCCTCTTGCAGTTGCATCCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGCGT[65]
AMV (MC29)	ATGTTGTATTTAAGTGCCTAGCTCGTAACAATAAACGCCATTTTACCATCCACCACATTGTCTGCACCTGGGTAGATGGACAGACCGTTGAGTCCCTAA[66]
SSV	GACTGAGTCGCCCGGGTACCTGTGTGTTCCAATAAACCTCTTGCTATTTGCATCCGAAGCCGTGGTCTCGTTGTTCCTTGGGAGGGTCTCTCCTACTGA[67]
GA-FeSV	GACTTGACCGCCCCCGGGTACCGTGTACGAATAAACCTCTTGCTGTTTGCATCTGACTCGGTGTCTCCGGTGTTCCGTGGGTACGGGGTCTCATCGCCGAG[68]
MSV (strain 124)	GACTGCGTCGCCCGGGTACCCGTATTCCCAATAAAGCCTCTTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGAT[69]
Avian endogenous provirus ev-2	TACGATTATATAAGCTGTTGCCACCATCAAATAAACGCCATTTTACCATCACCACATTGGTGCACCTGGGTAGATGGACAGACCGTTGAGTCCCTAA[70]
BaEV	ATTCCTTAGAGGGAGGGCCCTGGTGCACCAGTAAACGACTTTCTGCCGAAATTTGTGTTTGGGTGGTTCCTCGCGCCGACTCTCAAACCCTAAGGAACA[71]
SNV	TGTGACGTGCGGCCCAGATTCGAAATCTGTAATAAACTTTTTCTTCTGAATCCTCAGATTGGCAGTGAGAGGAGATTTTGTTCGTGGTGTTGGCTGGCCT[72]
GaLV	ACTGAGTCCGCCCGGGTACCCGTGTGTACGAATAAACCTCTTGCTATTTGCATCCGAAGTCGTGGACTCGTTGTTCCTTGGGAGGGTCTTTCCCAACTGA[73]
ASV	ATATTGTATTTAAGTGCCTAGCTCGATACAATAAACGCCATTTGACCATTCACCACATTGGTGTGCACCTGGGTTCATGGCTGGACCGTCGATTCCCT[74]
MTV	CCATAATATAAAAGAGTGCTGATTTTTTGAGTAAACTTGCAACAGTCCTAACATTCTTCTCTCGTGTGTTTGTCTGTTCGCCATCCCGTCTCCGCTCG[75]

FIG. 5. Compilation of DNA sequences at the 3' termini of mammalian and eukaryotic viral mRNAs. Sequences are aligned using the AATAAA signals which are underlined. Locations of poly A sites are indicated by arrows and the sequences which resemble the YGTGTTYY consensus are underlined. The references numbered at the right of each sequence are listed in the Compilation References section.

TABLE 2 Derivation of the YGTGTTYY consensus from sequences located 24bp to 38bp downstream from the AATAAA signal.

A	9	6	2	2	3	0	8	9
G	13	57	6	66	10	7	8	13
C	12	1	0	0	2	20	15	15
T	36	6	62	2	55	43	39	33
	Y	G	T	G	T	T	Y	Y

This represents an analysis of those 70 sequences listed in Fig. 5 which contain the signal; numbers indicate the frequency of occurrence for each nucleotide.

our consensus, is contained within this 15bp stretch and suggest that retention of this sequence reflects the ability to correctly process α 1 and α 2 mRNAs. In HSV, small tandem repeats occur downstream from the 3' termini of at least three mRNAs and these repetitive sequences are believed to have no specific function (11). The HSV reiterations begin a few nucleotides downstream from sequences which fit our consensus and hence could mark the boundary of regions important in transcription.

Approximately 33% of the 3' terminal DNA sequences examined do not contain a YGTGTTYY sequence motif. Hence, absence of this sequence may be overcome by other downstream signals, and we note that many of these 3' terminal regions contain T-rich residues. The downstream YGTGTTYY sequence could therefore be analagous to the transcription regulatory signals present upstream of the TATA box, eg the GC-rich segments present in many but not all promoters. Removal of GC-rich sequences reduces the level of initiation (32, 36, 37, 38), however, as these sequences are not present in the upstream regions of every promoter presumably they are not obligatory, or other sequences mimic their function.

While the function of sequences downstream from 3' termini is unclear it is likely that they are involved in the cleavage/processing of nuclear RNAs which extend beyond the poly A site. Cleavage/processing of primary transcripts is thought to be the mechanism whereby 3' termini are generated for both poly A^+ mRNAs (39), and histone messages (40, 41, 42). Histone genes do not contain AATAAA signals and a palindromic sequence frequently is present close to mRNA 3' termini (43, 44). Point mutations within this palindromic sequence abolish the ability to

form a correctly terminated mRNA, and sequences at least 24 nucleotides downstream from the 3' terminus are required for efficient processing (45). We have not observed any sequences similar to our consensus downstream from histone mRNA 3' termini. This is not surprising since, although the mechanism for generating 3' termini in poly A^+ and histone mRNAs may be similar, the processing events could be different and therefore alternative signals may be used.

We have no evidence yet as to whether the YGTGTTYY signal either in DNA or in pre-mRNA is involved in mRNA termination or processing. Its presence downstream from most HSV genes analysed so far may indicate a role in termination as the different transcription units are closely located on the virus genome (46). By contrast, genes in mammalian cells are more scattered so that transcription beyond the poly A site may less frequently disturb transcription of an adjacent gene. Examples of alternative processing of mRNA 3' termini include the pro $\alpha 2$(I) collagen (47), β^2-microglobulin (48) and mouse α amylase (49) genes, and there are situations where alternative poly A sites are used with differing efficiencies. Our analysis of these situations indicates that the YGTGTTYY signal is not associated with 3' end selection.

Deletion of the YGTGTTYY signal reduced 3' end formation but had no detectable effect on 5' end initiation; the RNA synthesised runs through into the pUC vector sequences and does not appear to be polyadenylated (data not shown). However, these extended RNAs are transported to the cell cytoplasm but, as reflected in CAT levels, are not translated. The inability to translate these RNAs may imply that their mode of transport to the cytoplasm differs from that of poly A^+ mRNAs thus suggesting a coupling of mRNA polyadenylation, transport and translation.

The lack of transcript polyadenylation is not surprising as *in vitro* only pre-mRNAs that contain a 3' end located at or just downstream from the *in vivo* poly A site are efficiently polyadenylated (50). *In vitro* studies with pre-mRNA suggest that the specificity of polyadenylation is linked to RNA synthesis and indicate that the poly A polymerase interacts with a specific signal located at the 3' end of pre-mRNA (51).

ACKNOWLEDGEMENT

We thank Professor J.H. Subak-Sharpe for his continued interest in this work.

*To whom correspondence should be addressed

LITERATURE CITED

1. Gannon, F., O'Hare, K., Perrin, F., LePennec, J.P., Benoist, C., Cochet, M., Brethnach, R., Royal, A., Garapin, A., Cami, B. and Chambon, P. (1979). Nature 278; 428-434.
2. Wasylyk, B., Derbyshire, R., Guy, A., Molko, D., Roget, A., Teolue, R. and Chambon, P. (1980). Proc. Natl. Acad. Sci. U.S.A. 77; 7024-7028.
3. Whitton, J.L., Rixon, F.J., Easton, A.J. and Clements, J.B. (1983). Nucl. Acids Res. 11; 6271-6287.
4. Whitton, J.L. and Clements, J.B. (1984). Nucl. Acids Res. 12; 2061-2079.
5. Proudfoot, N.J. and Brownlee, G.G. (1976). Nature 263; 211-214.
6. Fitzgerald, M. and Shenk, T. (1981). Cell 24; 251-260.
7. Montell, C., Fisher, E.F., Caruthers, M.H. and Berk, A.J. (1983). Nature 305; 600-605.
8. Higgs, D.R., Goodbourn, S.E.Y., Lamb, J., Clegg, J.B., Weatherall, D.J. and Proudfoot, N.J. (1983). Nature 306; 398-400.
9. McLauchlan, J. and Clements, J.B. (1983). EMBO J. 2; 1953-1961.
10. Whitton, J.L. and Clements, J.B. (1984). J. Gen. Virol. 65; 451-466.
11. Rixon, F.J., Campbell, M.E. and Clements, J.B. (1984). J. Virol. 52; 715-718.
12. Taya, Y., Devos, R., Tavernier, J., Cheroutre, H., Engler, G. and Fiers, W. (1982). EMBO J. 1; 953-958.
13. Corsalo, C.M. and Pearson, M.L. (1981). Genetics 7; 603-616.
14. Davison, A.J. and Wilkie, N.M. (1983). J. Gen. Virol. 64; 1927-1942.
15. Everett, R.D. (1984). EMBO J. in press.
16. de Villiers, J. and Schaffner, W. (1981). Nucl. Acids Res. 9; 6251-6264.
17. Messing, J. (1979). Recombinant DNA Bulletin, NIH Publication No 79-99, 2; No 2 43-48.
18. Vieira, J. and Messing, J. (1982). Gene 19; 259-268.
19. Gaffney, D., Whitton, J.L., Lynch, M., McLauchlan, J. and Clements, J.B. manuscript in preparation.
20. Maxam, A.M. and Gilbert, W. (1980). Methods Enzymol. 65; .499-560.
21. McLauchlan, J. and Clements, J.B. (1982). Nucl. Acids Res. 10; 501-512.
22. Favaloro, J.M., Treisman, R.H. and Kamen, R.I. (1980). Methods Enzymol. 65; 718-749.
23. Easton, A.J. and Clements, J.B. (1980). Nucl. Acids Res. 8; 2627-2645.
24. Berk, A.J. and Sharp, P.A. (1978). Cell 14; 695-711.
25. Weaver, R.F. and Weissmann, C. (1979). Nucl. Acids Res. 7; 1175-1193.

26. Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982). Mol. Cell. Biol. 2; 1044-1051.
27. Lowry, O.H., Rosebrough, N.J., Farr, A.L. and Randall, R.J. (1951). J. Biol. Chem. 193; 265-275.
28. Banerji, J., Rusconi, S. and Schaffner, W. (1981). Cell 27; 299-308.
29. Ihara, S., Feldman, L., Watanabe, S. and Ben-Porat, T. (1983). Virology 131; 437-454.
30. Green, M.R., Treisman, R. and Maniatis, T. (1983). Cell 35; 137-148.
31. Imperiale, M.J., Feldman, L.T. and Nevins, J.R. (1983). Cell 35; 127-136.
32. Everett, R.D. (1983). Nucl. Acids Res. 9; 6647-6666.
33. McDevitt, M.A., Imperiale, M.J., Ali, H. and Nevins, J.R. (1984). Cell 37; 993-999.
34. Gil, A. and Proudfoot, N.J. (1984). Nature 312; 473-474.
35. Proudfoot, N.J. and Maniatis, T. (1980). Cell 21; 537-544.
36. McKnight, S.L. and Kingsbury, R. (1982). Science 217; 316-324.
37. Everett, R.D., Baty, D. and Chambon, P. (1983). Nucl. Acids Res. 11; 2447-2464.
38. Dierks, P., van Ooyen, A., Cochran, M., Dobkin, C., Reisser, J. and Weissman, C. (1983). Cell 32; 695-706.
39. Salditt-Georgieff, M. and Darnell, J.E. (1983). Proc. Natl. Acad. Sci. U.S.A. 80; 4694-4698.
40. Krieg, P.A. and Melton, D.A. (1984). Nature 308; 203-206.
41. Price, D.H. and Parker, C.S. (1984). Cell 38; 423-429.
42. Birchmeier, C., Schumperli, D., Sconzo, G. and Birnstiel, M.L. (1984). Proc. Natl. Acad. Sci. U.S.A. 81; 1057-1061.
43. Hentschel, C.C. and Birnstiel, M.L. (1981). Cell 25; 301-313.
44. Kedes, L.H. (1979). Ann. Rev. Biochem. 48; 837-870.
45. Birchmeier, C., Grosschedl, R. and Birnstiel, M.L. (1982). Cell 28; 739-745.
46. Rixon, F.J. and McGeoch, D.J. manuscript submitted.
47. Aho, S., Tate, U. and Boedtker, H. (1983). Nucl. Acids Res. 11; 5443-5450.
48. Parnes, J.R., Robinson, R.R. and Seidman, J.G. (1983). Nature 302; 449-452.
49. Tosi, M., Young, R.A., Hagenbuchle, O. and Schibler, U. (1981). Nucl. Acids Res. 9; 2313-2323.
50. Manley, J.L. (1983). Cell 33; 595-605.
51. Moore, C.L. and Sharp, P.A. (1984). Cell 36; 581-591.

COMPILATION REFERENCES FOR FIG. 5

1. Liebhaber, S.A., Goossens, M.J. and Wai Kan, Y. (1980). Proc. Natl. Acad. Sci. U.S.A. 77; 7054-7058.
2. Nishioka, Y. and Leder, P. (1979). Cell 18; 875-882.
3. Dodgson, J.B. and Engel, J.D. (1983). J. Biol. Chem. 258; 4623-4629.
4. Erbil, C. and Niessing, J. (1983). EMBO J. 2; 1339-1343.
5. Schon, E.A., Wernke, S.M. and Lingrel, J.B. (1982). J. Biol. Chem. 257; 6825-6835.
6. Lawn, R.M., Efstratiadis, A., O'Connell, C. and Maniatis, T. (1980). Cell 21; 647-651.
7. Konkel, D.A., Maizel, J.V. and Leder, P. (1979). Cell 18; 865-873.

8. van Ooyen, A., van den Berg, J., Mantei, N. and Weissman, C. (1979). Science 206; 337-344.
9. Dolan, M., Dodgson, J.B. and Engel, J.D. (1983). J. Biol. Chem. 258; 3983-3990.
10. Spritz, R.A., Deriel, J.K., Forget, B.G. and Weissman, C. (1980). Cell 21; 639-646.
11. Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980). Cell 21; 621-626.
12. Shen, S., Slightom, J.L. and Smithies, O. (1981). Cell 26; 191-203.
13. Ellison, J., and Hood, L. (1982). Proc. Natl. Acad. Sci. U.S.A. 79; 1984-1988.
14. Malissen, M., Malissen, B., and Jordan, B.R., (1982). Proc. Natl. Acad. Sci. U.S.A. 79; 893-897.
15. Cheng, H.L., Blattner, F.R., Fitzmaurice, L., Mushinski, J.F., and Tucker, P.W. (1982). Nature 296; 410-415.
16. Suske, G., Wenz, M., Cato, A.C.B. and Beato, M. (1983). Nucl. Acids Res. 11; 2257-2271.
17. Ullrich, A., Gray, A., Goeddel, D.V. and Dull, T.J. (1982). J. Mol. Biol. 156; 467-486.
18. Ohno, S. and Taniguchi, T. (1981). Proc. Natl. Acad. Sci. U.S.A. 78; 5305-5309.
19. Taya, Y., Devos, R., Tavernier, J., Cheroutre, H., Engler, G. and Fiers, W. (1982). EMBO J. 1; 953-958.
20. Denoto, F.M., Moore, D.D. and Goodman, H.M. (1981). Nucl. Acids Res. 9; 3719-3730.
21. Page, G.S., Smith, S. and Goodman, H.M. (1981). Nucl. Acids Res. 9; 2087-2104.
22. Woychik, R.P., Camper, S.A., Lyons, R.H., Horowitz, S., Goodwin, E.C. and Rottman, F.M. (1982). Nucl. Acids Res. 10; 7197-7210.
23. Bell, G.I., Pictet, R.L., Rutter, W.J., Cordell, B., Tischer, E. and Goodman, H.M. (1980). Nature 284; 26-32.
24. Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R. and Dodgson, J. (1980). Cell 20; 555-566.
25. Kwok, S.C.M., Chan, S.J. and Steiner, D.F. (1983). J. Biol. Chem. 258; 2357-2363.
26. Cordell, B., Bell, G., Tischer, E., DeNoto, F.M., Ullrich, A., Pictet, R., Rutter, W.J. and Goodman, H.M. (1979). Cell 18; 533-543.
27. Lomedico, P., Rosenthal, N., Efstratiadis, A., Gilbert, W., Kolodner, R. and Tizard, R. (1979). Cell 18; 545-548.
28. Fiddes, J.C. and Goodman, H.M. (1981). J. Mol. Appl. Genet. 1; 3-18.
29. Noda, M., Teranishi, Y., Takahashi, H., Toyosato, M., Notake, M., Nakanishi, S. and Numa, S. (1982). Nature 297; 431-434.
30. Chang, A.C.Y., Cochet, M. and Cohen, S.N. (1980). Proc. Natl. Acad. Sci. U.S.A. 77; 4890-4894.
31. Vasicek, T.J., McDevitt, B.E., Freeman, M.W., Fennick, B.J., Hendy, G.N., Potts, J.T., Rich, A. and Kronenberg, H.M. (1983). Proc. Natl. Acad. Sci. U.S.A. 80; 2127-2131.
32. Shibahara, S., Morimoto, Y., Furutani, Y., Notake, M., Takahashi, H., Shimizu, S., Horikawa, S. and Numa, S. (1983). EMBO J. 2; 775-779.
33. Lemischka, I. and Sharp, P.A. (1982). Nature 300; 330-335.
34. Hurst, H.C. and Parker, M.G. (1983). EMBO J. 2; 769-774.

35. Nudel, U., Zakut, R., Shani, M., Neuman, S., Levy, Z. and Yaffe, E. (1983). Nucl. Acids Res. 11; 1759-1771.
36. Fornwald, J.A., Kuncio, G., Peng, I. and Ordahl, C.P. (1982). Nucl. Acids Res. 10; 3861-3876.
37. Aho, S., Tate, V. and Boedtker, H. (1983). Nucl. Acids Res. 11; 5443-5450.
38. Heilig, R., Muraskowsky, R., Kloepfer, C. and Mandel, J.L. (1982). Nucl. Acids Res. 10; 4363-4382.
39. Mason, A.J., Evans, B.A., Cox, D.R., Shine, J. and Richards, R.I. (1983). Nature 303; 300-307.
40. Karin, M. and Richards, R.I. (1982). Nature 299; 797-802.
41. Glanville, N., Durnam, D.M. and Palmiter, R.D. (1981). Nature 292; 267-269.
42. Gingeras, T.R., Sciaky, D., Gelinas, R.E., Bing-Dong, J., Yen, C.E., Kelly, M.M., Bullock, P.A., Parsons, B.L., O'Neill, K.E. and Roberts, R.J. (1982). J. Biol. Chem. 257; 13475-13491.
43. Dijkema, R., Dekker, B.M.M. and van Ormondt, H. (1982). Gene 18; 143-156.
44. Sugisaki, H., Sugimoto, K., Takanami, M., Shiroki, K., Saito, I., Shimojo, H., Sawada, Y., Uemizu, Y., Uesugi, S.-I. and Fujinaga, K. (1980). Cell 20; 777-786.
45. in Tooze, J. (ed.); DNA Tumour Viruses 2nd ed.; 1023-1036.
46. Akusjarvi, G., Zabrielski, J., Perricaudet, M. and Pettersson, U. (1981). Nucl. Acids Res. 9; 1-17.
47. Herisse, J. and Galibert, F. (1981). Nucl. Acids Res. 9; 1229-1240.
48. Akusjarvi, G. and Persson, H. (1981). J. Virol. 38; 469-482.
49. Herisse, J., Courtois, G. and Galibert, F. (1980). Nucl. Acids Res. 8; 2173-2192.
50. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., van der Voorde, A., van Heuverswyn, H., van Herreweghe, J., Volckaert, G. and Ysebaert, M. (1978). Nature 273; 113-120.
51. Soeda, E., Arrand, J.R., Smolar, N., Walsh, J.E. and Griffin, B.E. (1980). Nature 283; 445-453.
52. Seif, I., Khoury, G. and Dhar, R. (1979). Cell 18; 963-977.
53. Chen, E.Y., Howley, P.M., Levinson, A.D. and Seeburg, P.H. (1982). Nature 299; 529-534.
54. Danos, O., Katinka, M. and Yaniv, M. (1982). EMBO J. 1; 231-236.
55. Rixon, F.J., Campbell, M.E.M. and Clements, J.B. (1984). J. Virol. 52; 715-718.
56. Whitton, J.L., Rixon, F.J., Easton, A.J. and Clements, J.B. (1983). Nucl. Acids Res. 11; 6271-6287.
57. Rixon, F.J., Campbell, M.E.M. and Clements, J.B. (1982). EMBO J. 1; 1273-1277.
58. Davison, A.J. and Wilkie, N.M. (1981). J. Gen. Virol. 55; 315-331.
59. McKnight, S.L. (1980). Nucl. Acids Res. 8; 5949-5964.
60. Swain, M.A. and Galloway, D.A. (1983). J. Virol. 46; 1045-1050.
61. McLauchlan, J. and Clements, J.B. (1982). Nucl. Acids Res. 10; 501-512.
62. McLauchlan, J. and Clements, J.B. (1983). EMBO J. 2; 1953-1961.
63. Frink, R.J., Eisenberg, R., Cohen, G. and Wagner, E.K. (1983). J. Virol. 45; 634-647.

64. Seiki, M., Hattori, S., Hirayama, Y. and Yoshida, M. (1983). Proc. Natl. Acad. Sci. U.S.A. 80; 3618-3622.
65. Reddy, E.P., Smith, M.J. and Srinivasan, A. (1983). Proc. Natl. Acad. Sci. U.S.A. 80; 3623-3627.
66. Reddy, E.P., Reynolds, R.K., Watson, D.K., Schultz, R.A., Lautenberger, J. and Papas, T.S. (1983). Proc. Natl. Acad. Sci. U.S.A. 80; 2500-2504
67. Devare, S.G., Reddy, E.P., Law, J.D., Robbins, K.C. and Aaronson, S.A. (1983). Proc. Natl. Acad. Sci. U.S.A. 80; 731-735.
68. Hampe, A., Gobet, M., Even, J., Sherr, C.J. and Galibert, F. (1983). J. Virol. 45; 466-472.
69. Van Beveren, C., van Straaten, F., Galleshaw, J.A. and Verma, I.M. (1981). Cell 27; 97-108.
70. Scholl, D.R., Kahn, S., Malavarca, R., Astrin, S. and Skalka, A.M. (1983). J. Virol. 45; 868-871.
71. Tamura, T., Noda, M. and Takano, T. (1981). Nucl. Acids Res. 9; 6615-6626.
72. Shimotohno, K., Mizutani, S. and Temin, H.M. (1980). Nature 285; 550-554.
73. Lovinger, G.G. and Schochetman, G. (1979). J. Virol. 32; 803-811.
74. Czernilofsky, A.P., DeLorbe, W., Swanstrom, R., Varmus, H.E., Bishop, J.M., Tischer, E. and Goodman, H.M. (1980). Nucl. Acids Res. 8; 2967-2984.
75. Ucker, D.S., Firestone, G.L. and Yamamoto, K.R. (1983). Mol. Cell. Biol. 3; 551-561.

SUMMARY

This study was undertaken to examine the organisation and structures of mRNAs mapping at the herpes simplex virus type 2 (HSV-2) ribonucleotide reductase locus. From comparisons between equivalent HSV-2 and HSV-1 nucleotide sequences, putative transcription control signals were identified, and functional analysis of certain control elements was performed. The positions of translated regions within transcripts encoding ribonucleotide reductase were identified as were conserved amino acid domains between viral and cellular reductases. In addition, a conserved DNA sequence, YGTGTTYY (Y = pyrimidine), located 30nuc downstream from the polyadenylation (AATAAA) signal of both viral and cellular genes, was identified. Functional tests showed that the YGTGTTYY signal was required for efficient mRNA 3' end formation.

Arrangement of mRNAs and DNA Sequence Comparisons at the HSV-2 Ribonucleotide Reductase Locus.

Four overlapping mRNAs have been identified at the ribonucleotide reductase locus. These mRNAs are unspliced and share common 5' and 3' termini. Two early, 3' co-terminal mRNAs with sizes of 4.5kb and 1.2kb encode HSV-2 polypeptides Vmw138 and Vmw38 respectively which are almost certainly components of the viral ribonucleotide reductase. The other two mRNAs are late, 5' co-terminal species with sizes of 6.4kb and 1.7kb which appear to encode an identical 54,000 mol. wt. protein; the 6.4kb

transcript is 3' co-terminal with the early mRNAs. As a consequence of this mRNA arrangement, the 4.5kb and 1.2kb mRNA 5' termini and the 1.7kb mRNA 3' terminus are located within larger, overlapping transcripts. The genome region adjacent to the 6.4kb and 1.7kb mRNAs specifies three late transcripts synthesised in the opposite orientation; the 5' terminal region of one of these species overlaps with the 6.4kb and 1.7kb mRNAs.

The 1.7kb mRNA 3' terminus is located within the transcribed region of the 6.4kb species. Previous proposals suggested that read-through of HSV poly A sites is due to inefficient processing at certain 3' termini. Bacterial chloramphenicol acetyltransferase (CAT) activities produced in HeLa cells by a plasmid carrying the 1.7kb mRNA 3' terminal sequences were comparable to activities obtained with a plasmid containing the 3' processing signals from a non-internal poly A site. Read-through of HSV poly A sites may therefore reflect a general reduction in mRNA 3' processing efficiency rather than differences in the 3' processing signals of individual genes.

The HSV-1 transcripts equivalent to the 4.5kb and 1.2kb mRNAs have sizes of 5.0kb and 1.2kb and encode HSV-1 polypeptides Vmw136 and Vmw38 respectively. Nucleotide sequence comparisons between the 5' flanking regions of equivalent HSV-2 and HSV-1 mRNAs revealed the following homologies:

1) upstream from the 4.5kb and 5.0kb mRNA 5' termini, blocks of conserved sequences are present which resemble

transcriptional signals in the promoters of other genes, namely TATA box homologues, C-rich tracts and A+C-rich elements. The C-rich tracts resemble the binding site consensus sequence of Sp1, a transcription control factor, and, from other studies, the A+C-rich sequences could be a component of other HSV early promoter regions.
2) the 5' terminal region of the HSV-2 1.2kb mRNA is located within the polypeptide coding region of Vmw138; an identical situation exists for the equivalent HSV-1 transcript. The HSV-2 and HSV-1 5' flanking sequences contain inverted repeats which could form stem and loop structures; possible effects of these structures on transcription of the HSV-2 and HSV-1 1.2kb mRNAs are discussed.

Amino Acid Conservation Between Herpesvirus and Cellular Ribonucleotide Reductases.

Amino acid sequences for the Vmw138 N- and C-terminal regions and all of the Vmw38 polypeptide have been predicted from nucleotide sequences. Comparisons with coding regions of the corresponding HSV-1 proteins, Vmw136 and Vmw38, reveal high homology apart from areas at the N-termini. Poorly conserved N-terminal regions have been identified in other equivalent HSV-2 and HSV-1 proteins; these may be non-essential for function.

In common with ribonucleotide reductases with similar enzymatic properties, HSV ribonucleotide reductase appears to contain two non-identical polypeptides, and

HSV-2 polypeptides Vmw138 and Vmw38 almost certainly represent components of the virus enzyme. Both Vmw138 and Vmw38 amino acid sequences share areas of homology with coding regions on Epstein-Barr virus (EBV) and varicella-zoster virus genomes; this allows identification of putative proteins which may be ribonucleotide reductase components specified by these viruses. These blocks of homology extend to cellular ribonucleotide reductases of eukaryotic and prokaryotic origin and may represent domains involved in enzyme function.

Vmw138 and Vmw136 contain large N-terminal domains which are absent in equivalent herpesvirus and in cellular ribonucleotide reductase proteins. These N-terminal domains could have arisen by gene fusion and may not be essential for ribonucleotide reductase activity.

The EBV polypeptide which shares homology with HSV-2 Vmw138 has a size of 93,000 mol. wt. The C-terminal portions of both proteins contain the 5' terminal regions of mRNAs specifying the proposed small subunit of ribonucleotide reductases. However, these 5' termini are not located at equivalent positions in the Vmw138 and 93,000 mol. wt. polypeptides and separate evolutionary routes are proposed in the formation of these internal promoters.

The transforming ability of HSV may be linked to its action as a mutagen. This mutagenic effect may be due to the expression of HSV ribonucleotide reductase and this is discussed.

Sequences Downstream from the AATAAA Signal Required for mRNA 3' End Formation.

Comparisons of the regions flanking HSV mRNA 3' termini identified conserved sequences downstream from poly A sites. From more extensive comparisons, 67% of mammalian and viral genes examined contained 3' flanking sequences similar to those conserved in HSV. A consensus sequence, YGTGTTYY (Y = pyrimidine), was derived for a conserved downstream element, preferentially located 24 to 35nuc beyond the AATAAA signal. Downstream from YGTGTTYY, other G-rich and T-rich segments also were conserved but to a lesser extent. To test the functional significance of YGTGTTYY in mRNA 3' end formation, deletions produced in the 3' flanking sequences of HSV-2 IE gene-5 were inserted into a CAT plasmid. Deletion of downstream sequences which extended up to YGTGTTYY reduced CAT levels to values 35% of those obtained with undeleted plasmids. Removal of a further 14bp, which deleted YGTGTTYY but not the poly A site, reduced CAT activities to levels of 1% to 4%. The levels of CAT mRNA 3' termini reflected the reductions in CAT activities, but, in contrast, levels of mRNA 5' termini were unaffected by these deletions.

These results demonstrate a requirement of 3' flanking sequences, in particular the YGTGTTYY signal, for efficient 3' end formation of poly A mRNAs. Alternative functions for the YGTGTTYY sequence in either mRNA termination or mRNA processing are discussed.