# THE CHARACTERISATION OF SPECIFIC EHV-4 GENES

by

LESLEY NICOLSON, B.Sc. (Hons.)

Thesis submitted for the degree of Doctor of Philosophy,

Faculty of Veterinary Medicine, University of Glasgow.

February, 1991

Department of Veterinary Pathology,

University of Glasgow.

## CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## ABBREVIATIONS

| | |
|---|---|
| A | - adenosine |
| AVC | - acyclovir |
| APS | - ammonium persulphate |
| araT | - arabinosylthymidine |
| BVdU | - (E)-5-(2-bromovinyl)-2′-deoxyuridine |
| C | - cytosine |
| dCyd | - deoxycytidine |
| DMSO | - dimethyl suphoxide |
| DTT | - dithiothreitol |
| EDTA | - ethylenediaminetetraacetic acid |
| G | - guanosine |
| IVdU | - (E)-5-(2-iodovinyl)-2′-deoxuridine |
| SDS | - sodium dodecyl suphate |
| SLM | - special liquid medium |
| T | - thymidine |
| TEMED | - N,N,N′,N′- tetra-ethylenediamine |
| | |
| aa | - amino acids |
| BS | - Bluescript M13+ plasmid |
| cpm | - counts per minute |
| Ci | - Curies |
| ds | - double stranded |
| i.p. | - intraperitoneal |
| kbp | - kilobase pairs |
| kD | - kilodaltons |
| | - bacteriophage lambda |
| nm | - nanometres |
| OD | - optical density |
| RT | - room temperature |
| $TCID_{50}$ | - 50% tissue culture infective dose |
| ug | - micrograms |

| | |
|---|---|
| EBV | - Epstein-Barr virus |
| VZV | - varicella-zoster virus |
| HSV-1 | - herpes simplex virus type 1 |
| HSV-2 | - herpes simplex virus type 2 |
| BHV | - bovine herpesvirus |
| PRV | - pseudorabies virus |
| MDV | - Marek's disease virus |
| HVS | - herpesvirus saimiri |
| HVT | - herpesvirus of turkeys |
| EHV | - equine herpesvirus |
| HCMV | - human cytomegalovirus |
| | |
| ADCC | - antibody-dependent cellular cytotoxicity |
| CDL | - complement-mediated antibody lysis |
| DBP | - DNA binding protein |
| DIP | - defective interfering particles |
| E | - early |
| EFK | - equine foetal kidney |
| ELISA | - enzyme-linked immunoassay |
| IE | - immediate early |
| Ig | - immunoglobulin |
| ISCOM | - immune-stimulating complexes |
| LAT | - latency-associated transcript |
| MAb | - monoclonal antibody |
| MCS | - multicloning site |
| MCP | - major capsid protein |
| OBP | - origin binding protein |
| ORF | - open reading frame |
| PCR | - polymerase chain reaction |
| RE | - restriction enzyme |
| RR | - ribonucleotide reductase |
| SDS PAGE | - SDS polyacrylamide gel electrophoresis |
| TK | - thymidine kinase |
| TMD | - transmembrane domain |

## ACKNOWLEDGEMENTS

I wish to sincerely thank Professor David Onions for his help, advice, and encouragement throughout the course of this project.

I extend my gratitude to Professor W.F.H. Jarrett for the use of departmental facilities. Thanks to everyone in Vet Pathology who has helped me in my work and socialising - especially for all those nights in Lock 27 or Reids when everything seemed worthwhile!

My thanks to Dr. Ann Cullinane for the provision of an EHV-4 DNA library and to Dr. Anne Stokes for her participation in the peptide studies. I thank Mr. Alan May, Mr. John Fuller, and Dr. Helen Laird for their preparation of photographs and diagrams presented herein. I am grateful to the Horserace Betting Levy Board and the Equine Virology Research Fund for their generous financial support .

Finally, but most importantly, I'd like to thank my Mum and Dad for their support throughout all stages of this project.

With the exception of the collaborative work presented in Chapter 7, all work presented in this thesis was carried out by the author.

## ABSTRACT

Equine herpesvirus 4 (EHV-1) is an alphaherpesvirus primarily associated with respiratory disease. The genome of this virus has been characterised (strain 1942) and a BamHI library prepared (Cullinane et al., 1988).

This dissertation reports the localisation of five intact EHV-4 genes within the $U_L$ component of the genome . Homologues of HSV-1 genes UL22 (gH), UL23 (thymidine kinase), UL24, and UL44 (gC) were identified. An additional EHV-4 mapped to the left of the gC gene homologue has no apparent HSV-1 homologue. Partial nucleotide sequences were determined for EHV-4 homologues of HSV-1 genes UL21, and UL43. The map positions, nucleotide sequence, and predicited amino acid sequence of the products of these genes are detailed. The EHV-4 polypeptides are compared to their counterparts in other herpesviruses. Sequence analysis revealed a region of direct repeats and a putative origin of DNA replication proximal to the 3′ end of the gH gene.

Five peptides derived from the predicted amino acid sequences of the EHV-4 gp13 (gC) and gH polypeptides were tested for their immunogenicity in a hamster model.

The implications of these data are discussed in reference to the relatedness of EHV-4 to other herpesviruses in terms of sequence conservation (at the amino acid level) and in genome organisation. Finally, possible applications of the data to the development of novel EHV vaccines and diagnostic tests are outlined.

# CHAPTER 1

# GENERAL INTRODUCTION

# HERPESVIRUSES

## 1. The Family Herpesviridae

The family herpesviridae comprises approximately 90 viruses of similar biochemical and morphological characteristics isolated from members of all five animal families (Roizman, 1982; Wildy, 1985 ).

Herpesviruses are characterised by a linear double-stranded DNA genome of molecular weight 80-150 x $10^6$ which is encapsidated within an icosahedral capsid 100-200nm in diameter (Wildy et al., 1960). Electron micrography of a herpesvirus particle distinguishes four distinct morphological elements - an outer envelope, an electron-dense tegument, the capsid, and a core structure comprising an electron-dense outer area and an electron-translucent centre (Fig. 1). The capsid, an icosahedral structure composed of 162 capsomers, encapsidates the core which is composed of the DNA genome of the virus complexed with proteins. The tegument , an amorphous proteinaceous layer, separates the capsid from the virion envelope which is derived from modified regions of host cell internal membranes (Roizman and Furlong, 1974).

The most fully characterised herpesviruses are seven human herpesviruses - herpes simplex virus types 1 and 2 ( HSV-1 and HSV-2), varicella-zoster virus (VZV), human cytomegalovirus (HCMV) , Epstein-Barr virus (EBV), and human herpesviruses type 6 and type 7 (HHV-6 and HHV-7) - and viruses such as pseudorabies virus (PRV), bovine herpesvirus (BHV-1), and equine herpesviruses 1 and 4 (EHV-1 and EHV-4) which cause economically important disease in animals within the agricultural or leisure industries.

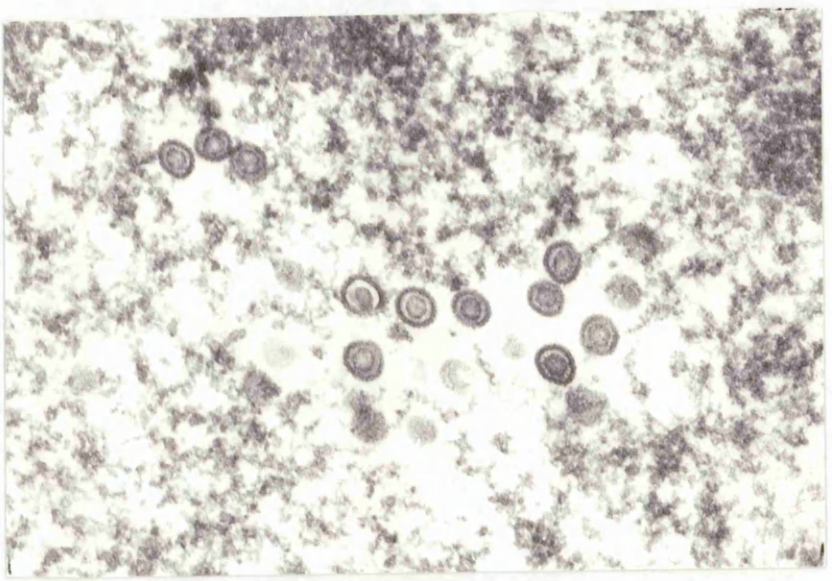
## 2. Classification of Herpesviruses

Herpesviruses can be differentiated by several criteria including antigenic structure, genome structure and biological properties. At

Figure 1.1

Herpesvirus Virion Structure

a)



b)

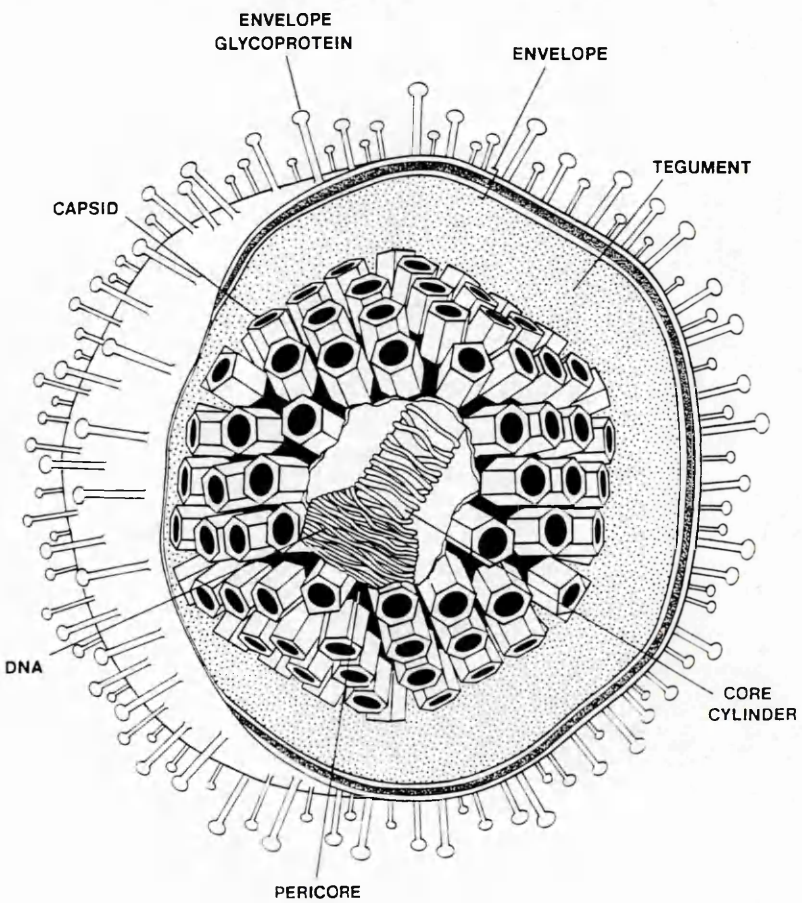



Figure 1.1 a) Electron micrograph of EHV-infected cells. Magnification x60000. b) Schematic diagram of a herpesvirus virion.

present only the latter is of use with regard to the subclassification of herpesviruses into defined groups since only a few herpesvirus genome structures have been well characterised and serological data gives rise to rather arbitrary groups.

**1) Classification on the basis of biological criteria**

Approximately two thirds of herpesviruses reported to date have been assigned to three subfamilies - alpha-, beta-, and gammaherpesviruses - on the basis of biological characteristics such as *in vivo* properties, *in vitro* host cell range, rate of propagation and features of latent infection and cytopathology (Roizman, 1982).

The alphaherpesvirus subfamily comprises herpesviruses of variable host range specificity *in vivo* and *in vitro*, a short reproductive cycle, and a rapid lytic infection *in vitro*. Further classification into alpha 1 and alpha 2 divisions separates viruses such as HSV-1, HSV-2, EHV-1, EHV-4 and PRV ( $\alpha_1$) which have a wide *in vitro* host range, a cytolytic capacity and which tend to establish latent infection within ganglionic tissue from viruses such as VZV ( $\alpha_2$) which has a narrow *in vitro* host range, a slower multiplication rate, a greater cell-associated character and an association with neurotropic as well as ganglionic latency.

The betaherpesvirus subfamily comprises the cytomegaloviruses, characterised by a restricted *in vivo* and *in vitro* host range, a slow replicative cycle, a cell-associated character and often a lymphotropic latent phase.

Gammaherpesviruses specifically infect B or T lymphocytes *in vivo* and establish latency within lymphoid and epithelial tissue. The length of the replicative cycle is variable. EBV and some other gammaherpesviruses are capable of immortalising lymphocytes *in vitro* and may initiate lymphoproliferative disease in susceptible hosts.

Classification of a gammaherpesvirus as $\gamma_1$ or $\gamma_2$ reflects the preference of the virus for B ( $\gamma_1$) or T ( $\gamma_2$) lymphocytes. Thus EBV ( $\gamma_1$) is typically associated with B lymphocytes whereas herpesvirus saimiri (HVS) ( $\gamma_2$) tends to infect and immortalise T-lymphocytes .

**2) Classification by genome biochemistry and size**

Differences in genome base composition, size and repeat element size and organisation also serve to distinguish herpesviruses. Classification according to genome size or base composition (which range from 80-150 x $10^6$ and 32-75% G+C moles% respectively) does not distribute viruses in a manner consistent with classification by biological criteria although there is a degree of correlation between %GC composition and biological sub-family in as much as the higher %GC (>60%) genomes tend to associated with alphaherpesviruses, the midrange (50-60%) with betaherpesviruses and the lower (34-50%) with gammaherpesviruses (Honess, 1984). The wide range of %GC compositions is regarded to be a consequence of the differential fixing of biased mutations presumably influenced by the host cell type and replicative characteristics of different herpesviruses. It was thought previously to be the result of different selective pressures on the protein coding regions of the genome but , as mentioned above, correlation between %GC and biological properties is limited. The dinucleotide rather than trinucleotide composition is now the more useful criteria for analysing herpesvirus genomes given that the CpG dinucleotide frequency is apparently reflective of the latent characteristics of a virus (Honess et al., 1989b).

**3) Classification by genome structure**

Five viral sub-groups - groups A to E - have been proposed on the basis of genome structure. Herpesviruses are classified according to the number, size and organisation of reiterated sequences exceeding 100

base pairs (Roizman, 1982) (Fig. 2). It has been hypothesised that the different arrangements of repeat regions within genomes may be a consequence of rearrangement events. Alternatively they may have arisen due to less selective pressure on different parts of the genome of the various herpesviruses.

Group A and B genomes possess direct repeat elements at their termini. Whereas in A genomes one copy of each element is present at the termini, in B genomes multiple reiterations of the element exist. These genome structures are exemplified by channel catfish herpesvirus (Chousterman et al., 1979) and by herpesvirus saimiri (Bornkamm et al., 1976).

Group C genomes such as EBV (Baer et al., 1984; Raab-Traub et al., 1980) possess internal tandem reiterations of sequences in addition to multiple direct repeats of distinct elements at the genome termini.

Group D genomes possess an internal inverted repeat of a terminal sequence element. Part of the repeat element is also reiterated, in the same orientation, at the other terminus. Group D genomes may exist in two isomeric forms due to the potential for intramolecular recombination events between the inverted repeat elements (Ecker and Hyman, 1982; Whalley et al., 1981). The two major genome components , the long unique component and the shorter component flanked by inverted repeats, are designated L and S respectively with the designations $U_L$ and $U_S$ serving to identify the unique portions of those components (i.e. the entire L component and the S component minus the repeat elements ). Genomes classified as group D include those of EHV-1, EHV-4, and PRV (Ben-Porat et al., 1979; O'Callaghan et al, 1983 ).

Group E genomes are characterised by two sets of inverted repeats, one at either terminus with its invert positioned internally. As in group D genomes, part of the sequence element is reiterated at each

Figure 1.2

Herpesvirus Genome Structures

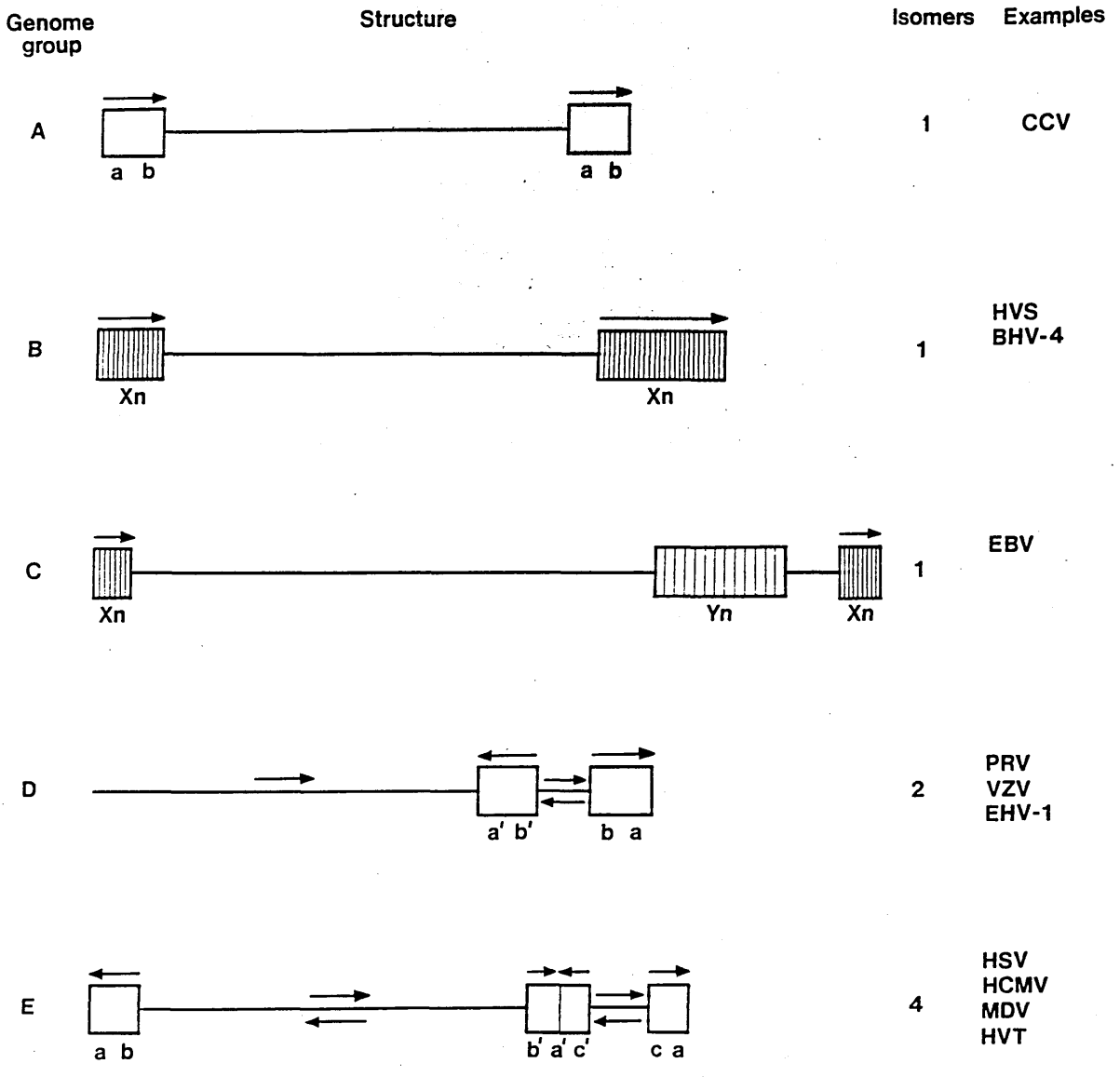



Figure 1.2 Schematic representation of the five types of herpesvirus genome structure as defined in Roizman, 1982.

termini to give rise to small direct repeat elements. Since two inverted repeats are present within E genomes, four isomers may exist differing only in the relative orientations of the unique sequences flanked by the repeats. The prototypic alphaherpesvirus HSV-1 possesses a type E genome (Sheldrick and Berthelot, 1975). The designations $U_L$ and $U_S$ are used as for D genomes while the repeat elements are denoted by their position at the termini or internally and the component in which they are positioned such that the major repeats are designated $TR_L$, $IR_L$, $IR_S$ and $TR_S$ from left to right.

**4) Classification by nucleotide sequence data**

The most sensitive determinant of herpesvirus classification is likely to be nucleotide sequence data which should allow the fine analysis of the evolutionary relatedness of different herpesviruses. As for genome structure though the usefulness of such a system is limited in taxonomic terms since such data is available for a minority of reported herpesviruses.

Nucleotide sequence data has revealed discrepancies in the current biological grouping of several herpesviruses including MDV, HVT, and HHV-6. Molecular analysis of MDV and HVT genes and genomes has indicated that although these viruses are classified as gammaherpesviruses on the basis of their lymphotropic nature (Roizman, 1982) they are more closely associated with alphaherpesviruses in molecular and genomic terms with gene sequence and CpG ratio more closely related to genes in other alphaherpesviruses than to those of gammaherpesviruses (Buckmaster et al., 1988; Cebrian et al., 1982; Honess et al., 1989b). Likewise, HHV-6 , classified as a gammaherpesvirus on account of its lymphotropic nature, is more closely related to cytomegalovirus in molecular terms and is now generally regarded as a $\beta_2$ herpesvirus (Lawrence et al., 1990). In these

instances the biological property of lymphotropism does not correlate with genome characteristics at the structural or molecular level and it seems likely that this tropism may be associated with different gene products in MDV, HVT, and HHV-6 as compared to those in gammaherpesviruses.

## 3. Herpesvirus Nucleotide Sequence Data and its Implications for Herpesvirus Evolution

The nucleotide sequence of the entire genome of four herpesviruses has been reported and their coding capacity analysed - HSV-1 ( McGeoch et al., 1985; McGeoch et al., 1986; McGeoch et al., 1988a; Perry and McGeoch, 1988), VZV (Davison and Scott, 1986), HCMV (Chee et al., 1990), and EBV (Baer et al., 1984). Comparative analyses of these genomes and of limited regions of other herpesvirus genomes have revealed the molecular relatedness of six biological subgroups of the herpesvirus family - alpha 1 and 2, beta 1 and 2, and gamma 1 and 2. On the basis of these comparisons albeit limited by the restricted sequence data available for certain herpesviruses, and of cross-hybridisation data ( Baumann et al., 1986; Cullinane et al., 1988; Davison and McGeoch, 1986; Davison and Taylor, 1987; Davison and Wilkie, 1983; Gompels et al., 1988a; Kouzarides et al., 1987; Lawrence et al., 1990 ) several conclusions can be drawn. The most diverse region of herpesvirus genomes is the short component such that viruses within the same biological subgroup possess significantly different coding arrangements in this region. For example, the short component of HSV-1 encodes 13 genes , 6 of which are unique to HSV-1, while that of VZV encodes 7 genes all of which have HSV-1 $U_S$ homologues. Neither genome possesses homologues of genes within the corresponding region of the HCMV and EBV genomes (Davison and McGeoch, 1986). There are also

significant differences in the arrangement of conserved S component genes in different alphaherpesviruses (Cullinane et al., 1988; Davison and McGeoch, 1986; Petrovskis and Post, 1987). Within the long component of the genome many genes are conserved throughout all herpesviruses sequenced to date with some arranged in conserved gene blocks . These include genes encoding DNA polymerase, the major DNA binding protein and two structural glycoproteins gB and gH. The number of genes which have homologues in other herpesviruses is obviously greater within a specific group such as the alphaherpesviruses than between groups and employing this criteria the beta- and gammaherpesviruses seem more closely related to each other than to the alphaherpesviruses.

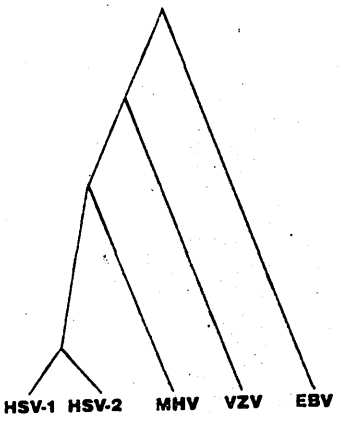
Given the extent of gene conservation throughout the herpesvirus family it is evident that all herpesviruses are descended from a common progenitor. Analysis of the position of these counterparts in the respective genomes indicates that in the course of evolution large scale rearrangements have occurred since blocks of conserved genes arbitrarily designated order ABC in VZV (and HSV-1) are positioned as CA'B (where A' represents inversion) in EBV and HVS (Gompels et al., 1988a). Whether either of these gene orders corresponds to that of the progenitor is unknown. Small scale rearrangements have occurred at later stages in evolution as evidenced by the the 0.1-0.4 map unit inversion in PRV relative to that in other alphaherpesviruses (Davison and Wilkie, 1987).

As increasing sequence data becomes available it is becoming more feasible to consider the fine evolutionary relationships within the herpesvirus family and herpesvirus evolutionary trees have been constructed on the basis of the nucleotide sequence of single genes such as thymidine kinase and gB genes (Gentry et al., 1988; Honess

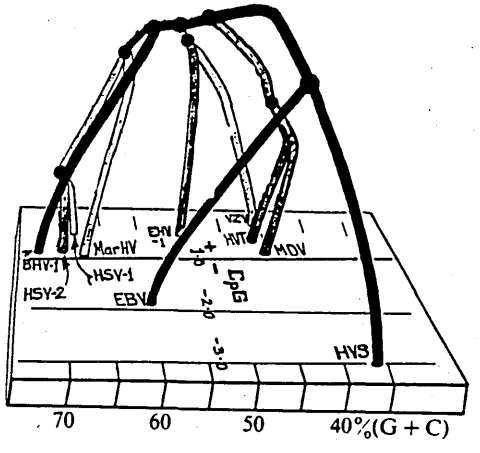
Figure 1.3

Derivation of Herpesvirus Evolutionary Trees from Nucleotide Sequence Data

a)




b)




c)

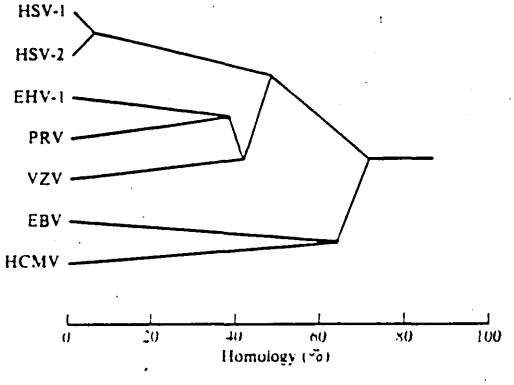



Figure 1.3 Herpesvirus evolutionary schemes derived from the analysis of the nucleotide sequence of genes encoding thymidine kinase, a) Gentry et al., 1988, b) Honess et al., 1989a, and glycoprotein gB, c) Whalley et al., 1989.

et al., 1989a; Whalley et al., 1989) (Fig. 3).

## 4. Herpesvirus Life Cycle

Herpesvirus lytic infection involves a series of events mediated by virus-specified and cell-specified factors which ultimately results in the release of infectious progeny from the infected cell (Fig. 4).

**1) Virus entry**

Two mechanisms have been proposed for the entry of herpesviruses into cells on the basis of electron micrography : i) entry within phagocytic vesicles (viropexis) (Dales and Silverberg, 1969) , and ii) fusion of virion envelope with cell plasma membrane releasing the nucleocapsid into the cytoplasm (Morgan et al., 1968). The generally regarded view is that the latter mechanism is the major mechanism. Constituents of the virion envelope are important in facilitating viral entry into permissive cells since naked nucleocapsids and DNA have a significantly lower specific infectivity than intact virions (Stein et al., 1970). In alphaherpesviruses initial adsorption seems to occur through the interaction of virion constituents , identified as glycoprotein gIII in BHV-1 and PRV (Mettenleiter et al., 1990 ; Okazaki et al., 1987), with heparan sulphate moieties on the cell surface (WuDunn and Spear, 1989). Although an HSV-1 glycoprotein with a direct role in this initial interaction has not been identified, glycoproteins have , however, been implicated in the stabilisation of adsorption through interaction with cellular receptors (Fuller and Spear, 1985; Kuhn et al., 1990).

The penetration process involves fusion of the envelope of an adsorbed virion with the plasma membrane of the host cell ( Morgan et al., 1968) and is mediated by viral glycoproteins such as gB, gD, and gH in HSV-1 (Spear, 1984) . Entry of intact herpesvirions into cells

Figure 1.4

Herpesvirus Life Cycle

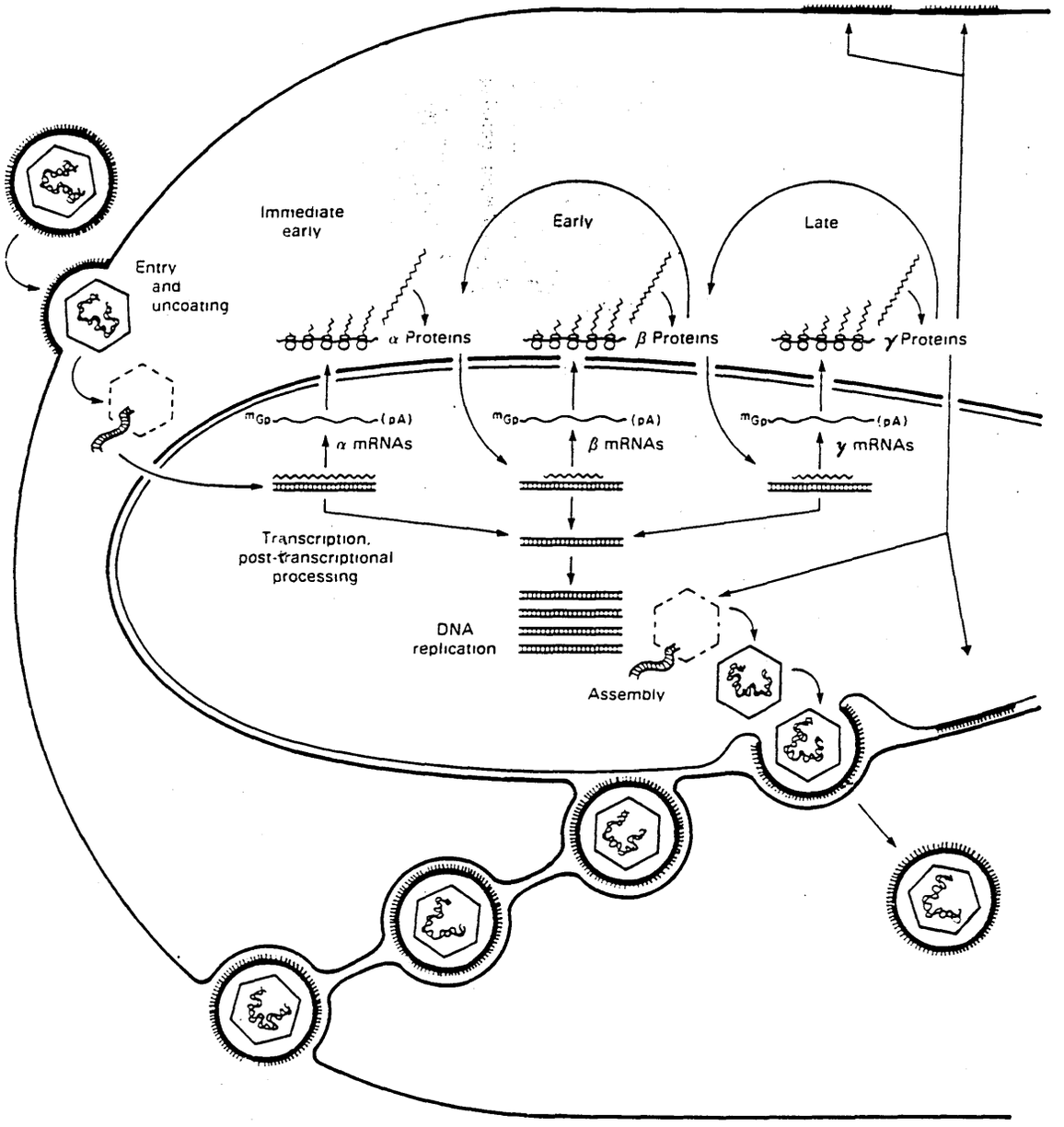



Figure 1.4 Schematic representation of the herpesvirus lytic cycle (Ginsberg, 1980).

thus apparently involves interaction between several virion envelope glycoproteins and cell surface components.

**2) Transcription**

HSV-1 mRNA, like cellular RNA, is transcribed by host cell RNA polymerase II and is capped, polyadenylated, and internally methylated (Wagner, 1982). Transcription is temporally regulated through mechanisms mediated by virus-specified proteins. mRNAs are classified into three major groups - alpha (immediate early), beta (early), or gamma (late) - according to their time of synthesis (Honess and Roizman, 1974, 1975) : alpha mRNAs are synthesised prior to virus-specified protein synthesis, beta mRNAs are synthesised in response to transactivation by products of the alpha mRNAs and prior to viral DNA replication, and gamma mRNAs are synthesised maximally ( $\gamma_1$ mRNAs) or exclusively ( $\gamma_2$ mRNAs) once viral DNA replication is underway (Wagner, 1982).

**3) Translation**

HSV-1-specified proteins are synthesised in a temporally regulated manner in a cascade system subject to controls on abundance. As for mRNA classification, three general groups of proteins are distinguished - alpha, beta , and gamma. Alpha, or immediate early (IE), proteins are expressed abundantly without de novo protein synthesis from 1hr postinfection to around 6hr postinfection with a peak at 3-4hr (Everett, 1987). Genes encoding these proteins map at or near the long and short major repeats and encode 5 mRNA species which are translated into proteins involved in the regulation of expression of later proteins. Beta, or early (E), proteins are expressed after alpha proteins and prior to viral DNA replication at 3-6hr postinfection. They are encoded by genes which map throughout the genome in noncontiguous regions and include proteins such as DNA polymerase which

are required for DNA synthesis. Gamma proteins are divided into two subsets , betagamma and true gamma, the latter expressed prior to DNA synthesis and the latter depending stringently on DNA synthesis prior to their expression. Genes encoding these proteins map throughout the genome and include structural virion protein genes (Wagner, 1982). Virus-specified proteins are subjected to co- and post-translational modification. Phosphorylation of several immediate early proteins and proteolytic cleavage and glycosylation of structural glycoproteins occur.

Functions have been assigned to approximately half of the proteins encoded by the 70 or so distinct genes of HSV-1 (McGeoch et al., 1988a; Roizman, 1978). Enzymes include DNA polymerase, thymidine kinase, ribonucleotide reductase, protein kinases, and DNAses, required directly, or indirectly , for viral DNA replication, host shut-off, and virion assembly. Regulatory proteins include immediate early proteins such as ICP0 and ICP4 which regulate transcription of later HSV genes and presumably proteins which might contribute to establishment or maintenance of a latent state. Structural proteins include constituents of the capsid, tegument and virion envelope including the glycoproteins.

**4) Viral DNA replication**

Replication of herpesviral DNA occurs in a semi-conservative manner and utilises both host-cell and virus-specified gene products (Challberg and Kelly, 1989). Replication of viral genomes is believed to occur by the rolling circle model since this would account for the presence of extensive head-to-tail concatamers of HSV-1 DNA in infected cells late in infection (Jacob et al., 1979; Jongeneel and Bachenheimer, 1981; Rabkin and Hanlon, 1990). The site of DNA replication initiation is dependent on the position of cis-acting

origin sequences of which there may be up to three depending on the type of herpesvirus (see chapter 5). HSV-1 encodes 7 functions essential for viral DNA replication ( McGeoch, et al., 1988b) including a major DNA binding protein (ICP8), DNA polymerase , an accessory DNA binding protein (65K DBP), an origin-binding protein (OBP), and a putative helicase/primase . There is evidence to suggest that some of these proteins are associated in a functional multienzyme complex (Vaughan et al., 1984). DNA isolated from virions is susceptible to fragmentation on denaturation with alkali indicating that gaps exist within the double stranded genome. It is still not clear whether these occur at unique or random sites and it is thought that at least some of these gaps may be a consequence of incorporation of ribonucleotides into the replicating DNA polymer possibly as constituents of the RNA primers of synthesis.

**5) Virion assembly**

The core and capsid are assembled within the nucleus. Envelopment of the capsid by a budding process through several internal membranes, most notably the inner nuclear membrane, has been observed by electron microscopy (Roizman and Furlong, 1974).

**6) Virus egress**

Several theories exist as to how the enveloped virion is transported to the extracellular space : the continuous theory proposes that the virion travels through the ER and is released at a point of connection between the ER and the plasma membrane; discontinuous theories propose i) that capsids bud into and out of any cellular membrane and that the envelope is derived from the last membrane encountered , or ii) that virions are transported in vacuoles and secreted from the cell (Roizman, 1978).

The effect of productive herpesvirus infection is cellular death

due to inhibition of cellular processes such as DNA, RNA, and protein synthesis, so-called host shut-off, alteration in organelle structure and alteration of the immunological characteristics of the plasma membrane.

## 5. Latency

Latency , the ability to remain within the host for a long period of time following primary infection with the potential to reinitiate productive infection, is a characteristic of herpesviruses (Hill, 1982; Roizman, 1982). In the course of primary lytic infection a proportion of the progeny virus enter specific target cells, primarily neurones in the case of HSV-1, lymphocytes in the case of EBV, in which a non-productive latent infection is maintained. On entry into these cells the virus is no longer directly accessible to host immune defence mechanisms. The host and/or viral factors which favour establishment and maintenance of latency as opposed to lytic infection have yet to be formally identified although lymphokines have been implicated in the maintenance of MDV latency (Buscaglia and Calneck, 1988). Virus-specified gene products involved in establishment and maintenance of latency would be expected to be immediate early or early functions since expression of beta or gamma genes would presumably lead to lytic cycle events.

In the latent state the viral genome seems to be maintained in various forms - HSV-1 as circular or concatameric molecules (Mellerick and Fraser, 1987), PRV as circular, concatameric or linear molecules (Rziha, 1986), and EBV as a super-coiled plasmid or integrated genome (Lindahl et al., 1976; Matsuo et al., 1984) - and is subject to limited expression. Transcripts expressed during HSV, PRV, and BHV-1 latent infection (latency-associated transcripts or LATs) have been characterised (Cheung, 1989; Rock et al., 1987; Wagner et

al., 1988 ) - these are partially complementary to immediate early (IE) gene products, ICP0 in the case of HSV-1 or the sole IE gene in PRV , and have thus been implicated in a role in the establishment and maintenance of latency by repressing expression of IE genes (Stevens et al., 1987). However, recently LAT$^-$ HSV-1 mutants which can establish latency have been isolated but which are defective in reactivation suggesting that LAT transcripts may be important in reactivation from latency rather than in establishment of latency (Javier et al., 1988; Steiner et al., 1989). Other gene products implicated in latency are thymidine kinase (TK) and ribonucleotide reductase (RR). The isolation of TK- and RR-deficient mutants capable of latent infection but impaired in reactivation has , as with LAT transcripts, implied that these functions are important for reactivation rather than for establishment or maintenance (Jacobson et al., 1989a; Tenser et al., 1989).

Reactivation of latent HSV results in productive infection with resultant death of the host cells and infection of peripheral tissues via axonal transport of infectious virus. This may or may not result in clinical disease. Mechanisms by which reactivation is triggered are unknown although they are known to be initiated by external or internal stimuli such as physical trauma or stress (Hill, 1982).

## HERPESVIRUS GLYCOPROTEINS

Herpesvirus glycoproteins are important in virus attachment and penetration into host cells and in the envelopment of progeny virions, membrane fusion events and egress. They thus play a role in the determination of host range, tissue tropism, and in the mediation of cell-virion interaction and cell-cell interaction. Furthermore, they are important targets of host humoral and cellular immune responses given their localisation on the surface of virions and of infected cells and their consequent exposure to the immune system (Norrild, 1985; Spear, 1984 ). Several herpesvirus glycoproteins form morphologically identifiable entities , so-called 'spikes,' which protrude from the surface of the envelope. Analysis of HSV-1 virions has revealed that gB and gD form distinct spikes (Stannard et al., 1987).

Herpesvirus glycoprotein complements have been determined by the radioactive labelling of virion surface components and their resolution by electrophoresis or isoelectric focussing. Such techniques are limited in that i) low abundance glycoproteins might not be detected, ii) co-migrating glycoproteins might not be distinguished, iii) different virus growth conditions and electrophoretic conditions result in different mobilities for proteins, making the comparison of glycoprotein profiles presented by different laboratories difficult, and iv) without further investigation the relationship between different proteins in terms of co- and post-translational processing is unknown. Perhaps the most straightforward analysis of herpesvirus glycoprotein complements is performed at the genome level through the determination of the number, map position, and coding characteristics of glycoprotein genes by nucleotide sequence analysis. Putative glycoprotein-encoding genes are identified by the potential of the

predicted polypeptide product to possess one or more of the following features - i) an N-terminal hydrophobic 'signal sequence' which is required for insertion of the primary translation product during synthesis into the endoplasmic reticulum , ii) a further hydrophobic sequence - the transmembrane domain (TMD) - towards the C-terminal end which is required for passage through the bilayer of cellular membranes or virion envelope, iii) a charged region of amino acids downstream of the predicted TMD, and iv) N-X-S/T amino acid sequences in the region of the polypeptide N-terminal to the TMD which might serve as targets for N-linked glycosylation. Such analysis is currently limited by the amount of complete genome sequence data available since this has been reported for only HSV-1, VZV, HCMV, and EBV. Antigenic, functional, and biochemical investigation of glycoprotein species will ultimately lead to the identification of the product(s) of each glycoprotein gene, their relatedness to each other in terms of post-translational processing, their similarity to glycoproteins specified by other herpesvirus, and their function(s). Identified homologues of HSV-1 glycoproteins in other herpesviruses are presented in Table 1.

HSV-1 specifies at least eight envelope glycoproteins - gB, gC, gH, gD, gE, gG, gI, and gJ - which are located on the virion envelope and on the surface of infected cells (Gao and Spear, 1990; McGeoch et al., 1988a; Spear, 1984). Five HSV-1 glycoproteins , gD, gE, gG, gI, and gJ, are encoded within the $U_S$ component of the genome. The full complement of functions of these glycoproteins has yet to be determined. However, functional and mutational studies have revealed several roles for glycoproteins in the virus life cycle. gB, gD, and gH are essential for viral replication in vitro (Cai et al., 1988; Desai et al., 1988; Little et al., 1981 ) and are involved in virus penetration (Fuller and Spear, 1987; Fuller et al., 1989;

Highlander et al., 1988). gH is involved in viral egress and cell-cell spread of virus (Buckmaster et al., 1984). gB and gC may be important in natural killer cell recognition (Bishop et al., 1986) and in cell-mediated immunity (Rosenthal et al., 1987). Glycoproteins gC, gE, gG, and gI are nonessential for viral replication in vitro ( Holland et al., 1984a; Longnecker and Roizman, 1987; Weber et al., 1987).

Three HSV-1 glycoproteins elicit the production of antibody which will neutralise viral infectivity in the presence or absence of complement , gB, gD, and gH , the major target being gD (Para et al., 1985; Spear, 1984 ). Other glycoproteins , including gC, specify complement-dependent neutralising epitopes. Three glycoproteins of HSV-1 have been implicated in the modulation of the host immune reponse: gE and gI in complexed form bind the Fc portion of IgG ( Johnson et al., 1988; Para et al., 1982) and gC binds complement component C3b fragment (Friedman et al., 1984; McNearney et al., 1987).

VZV specifies six to eight major glycoproteins which associate in three complexes - gpI, gpII, and gpIII - on the basis of reactivity of the glycoproteins with monoclonal and polyclonal antibody preparations (Davison et al., 1986). These glycoproteins elicit virus-neutralising antibody and apparently possess all the virus-neutralising epitopes of the VZV virion reported to date with gpI the most immunoreactive and abundant followed by gpII and gpIII (Keller et al., 1984, 1986). The nucleotide sequences of genes encoding gpI, gpII, and gpIII and two minor glycoproteins , gpIV and gpV have been determined (Davison, 1983; Davison and Scott, 1986; Davison et al., 1985; Keller et al., 1987 ). All five VZV glycoprotein genes have counterparts in HSV-1 (Table 1).

PRV specifies at least 7 glycoprotein species , the genes for 6 of

Table 1.1

Analogues of HSV-1 glycoproteins identified within specific herpesviruses

| HSV-1 | VZV | PRV | BHV-1 | EHV-1 | EHV-4 | HCMV | EBV |
|---|---|---|---|---|---|---|---|
| gB | gpII | gII | gI | gp14 | | gp58 | gp110 |
| gC | gpV | gIII | gIII | gp13 | | - | - |
| gD | - | gp50 | gIV | gD | | - | - |
| gE | gpI | gI | gII | gE | gE | - | - |
| gG | - | gX | | | | - | - |
| gH | gpIII | gH | | gH | | gp86 | gp85 |
| gI | gpIV | gp63 | | gI | | - | - |
| gJ | - | | | | | - | - |

Table 1.1 Analogues of HSV-1 glycoproteins as identified by analyses of viral proteins or nucleotide sequence data. Dashes indicate no analogue exists; blanks indicate published data is insufficient to determine whether an analogue is or is not encoded.

which have been mapped and sequenced - gI, gII, gIII, gp50, gp63, and gX (Hampl et al., 1984; Mettenleiter et al., 1986; Petrovskis et al., 1986a,b; Rea et al., 1985; Robbins et al., 1984, 1986a, 1987; Wathen and Wathen, 1984, 1986). Two further glycoproteins, gIV and gV, have been identified in cell extracts. The relationship of these gene products to the mapped glycoprotein genes is unknown although it has been speculated that the gp50 gene may encode either gIV or gV. A gH-type glycoprotein has yet to be identified in infected cell extracts but a homologue of the HSV-1 gH gene has been mapped to the central region of $U_L$ (Petrovskis et al., 1988). Four glycoproteins are nonessential for viral replication in vitro - gI, gIII, gp63, and gX. gX, the homologue of HSV-1 gG is secreted into the medium and is not a virion component (Rea et al., 1985). One of the essential glycoproteins, gp50, is unusual in that it has no sites for potential N-linked glycosylation (Petrovskis et al., 1986a). Glycoproteins gI and gp63 are functional as a complex and , with gIII, affect virus virulence (Zuckermann et al., 1988). gIII is also important for the adsorption of virions to cells (Schreurs et al., 1988).

The glycoprotein complement of beta and gamma herpesviruses is significantly different from that of the alphaherpesviruses and is likely to be a major factor contributing towards the difference in biological properties manifested by herpesviruses of the three subgroups. Only two virion glycoprotein species are conserved throughout members of all three subgroups of the herpesvirus family - glycoproteins gB and gH (Baer et al., 1984; Cranage et al., 1986, 1988; Davison and Scott, 1986; Gompels and Minson, 1986; McGeoch and Davison, 1986; McGeoch et al., 1988a; Mettenleiter et al., 1986; Petrovskis et al., 1988) (Table 1) . This implies these glycoproteins have an essential role in the herpesvirus life cycle.

## EHV-1 AND EHV-4

### 1. Equine Herpesviruses

The equine herpesviruses comprise a group of serologically distinct herpesviruses of related biological and biochemical features (Plummer et al., 1973; O'Callaghan et al., 1983). Four types of equine herpesviruses are officially recognised by the International Committee on nomenclature of viruses - EHV-1, or equine abortion (Randall et al., 1953), EHV-2, or equine cytomegalovirus (Plummer and Waterson, 1963), EHV-3, or equine coital exanthema virus (Ludwig et al., 1971), and EHV-4, or rhinopneumonitis virus (Studdert et al., 1981). One further candidate EHV type has been proposed, EHV-5, a novel equine betaherpesvirus (Browning and Studdert, 1987a).

The equine betaherpesviruses are ubiquitous viral agents present in approximately 90% of the equine population (Kemeny and Pearson, 1970). These viruses are associated with upper respiratory tract disease (Blakeslee et al., 1975) and immunosuppression in foals ( Browning and Studdert, 1987b; Sugiura et al., 1983) although they have yet to be identified as a causative agent of these disorders. Viral isolates differ in their growth rate with a cytopathic effect being manifested from 24 hours to one month post infection although the majority of isolates tend to exhibit the characteristically slow growth of cytomegaloviruses in general and their cell-associated nature (O'Callaghan et al., 1983). They also exhibit a narrow host cell range (Wharton et al., 1981) and have the capacity to establish persistent infection (O'Callaghan et al., 1983). Two serologically distinct types of cytomegalovirus have been identified - EHV-2, and EHV-5 - which differ in their genome size, EHV-2 being 192 kbp and EHV-5 ,148 kbp, and which have limited sequence homology (Browning and Studdert, 1987a). The genome of EHV-2 is unusual in structure in that it

comprises a unique region (m.u. 0.09-0.83) flanked by repeat elements (m.u. 0.00-0.90 and 0.83-1.00). The 0.00-0.09 repeat consists of two 2.4kbp repeats flanking a 6kbp unique sequence. Apparently only one genomic isomer exists (Browning and Studdert, 1989; Colacino et al., 1989). Such a genome structure does not strictly fit into any of the five genome classes specified by Roizman (1982). The EHV-2 genome shares less than 3% homology with EHV-3 and EHV-1 genomes (Staczek et al., 1983; Wharton et al., 1981 ).

Three equine alphaherpesviruses have been identified - EHV-1, EHV-3, and EHV-4. EHV-3, or equine coital exanthema virus, induces lesions on the external genitalia and, less commonly, on facial features (O'Callaghan et al., 1983). Mild systemic disease may occur. EHV-3 does not seem to induce abortion in nature or in experimental situations . The 96MDa EHV-3 genome possesses a group D type structure akin to that of EHV-1 and EHV-4 and exists in two isomeric forms as a consequence of recombination at the inverted repeats (Sullivan et al., 1984). Cross neutralisation does not occur between EHV-3 and EHV-1 implying the viruses are antigenically distinct. Sequence homology does , however, exist between the EHV-1, EHV-4, and EHV-3 genomes at limited regions throughout the genome in a colinear manner ( Baumann et al., 1986 ). No homology was detected between the unique short components of the EHV-1 and EHV-3 genomes suggesting that the diverse gene products in $U_S$ of the two viruses may be responsible in part for the different biological characteristics of these viruses.

EHV-1 and EHV-4 are the most important of the equine herpesviruses in economic terms. Considerable financial losses are incurred by the horse racing and stud industries as a consequence of the illness and death of racing and stud animals and their offspring through the respiratory, abortigenic , or paralytic sequelae of infection by these

viruses. Virus-induced abortion in mares was initially reported in 1933 (Dimmock and Edwards, 1933). By 1954 it became apparent that the viral agent involved, named equine abortion virus (EAV), was also associated with respiratory disease in young horses (Doll and Wallace, 1954). Two viral subtypes were subsequently distinguished on the basis of serological characteristics and pathological characteristics, one , subtype 1, being frequently isolated from aborted foetal tissue and the other, subtype 2, from respiratory tract tissue (Burrows and Goodridge, 1973; Shimizu et al., 1959 ). It became clear that the two subtypes differed not only in their antigenic and pathogenic characteristics but also in their in vitro host range and in the electrophoretic mobility of their structural proteins ( Allen and Bryans , 1986; Studdert and Blackney, 1979; Turtinen et al., 1981). Analysis of the genomes by restriction endonuclease analysis and by liquid hybridisation indicated that the viruses are genetically distinct (Allen and Turtinen, 1982; Allen et al., 1983b; Sabine et al., 1981; Studdert, 1983; Studdert et al., 1981) bringing about the proposal that the two subtypes be redesignated as separate viral types, EHV-1 subtype 1 becoming EHV-1, and EHV-1 subtype 2, or equine rhinopneumonitis virus, becoming EHV-4. This revised nomenclature was ratified by the International Committee on Virus Taxonomy in 1987.

All equine herpesviruses studied to date are capable of inducing oncogenic transformation in primary hamster embryo cells and of establishing persistent infection (O'Callaghan et al., 1983). However, there is no evidence that equine herpesviruses are associated with cancer in the horse.

## 2. EHV-1 and EHV-4 Pathogenesis

EHV-1 and EHV-4 are a major cause of acute viral respiratory disease in horses. Although primarily respiratory pathogens the viruses can also induce abortion in susceptible hosts , neurological disease and, in rare cases, mild coital exanthema (Bryans and Allen, 1973; O'Callaghan et al., 1983) (Fig. 5). The viruses are endemic in horse populations worldwide and may also infect and cause disease in ruminants (Chowdhury et al., 1988; Crandell et al., 1988). In Great Britain horses are exposed to EHV-1 or EHV-4 for the first time at a young age, usually within their first year of life . Sources of infectious virus include virus shed from infected cohorts, aborted foetal tissue and infected bedding or food (O'Callaghan et al., 1983). Following primary infection the animal is immune to reinfection for only 3-6 months (Bryans, 1969). Horses can thus be infected periodically throughout their life.

The frequency with which EHV-1 and EHV-4 are identified as the causative agents of respiratory disease and abortion varies from country to country . A study of 246 American and Canadian isolates between 1960 and 1984 revealed that 99% of abortions were caused by EHV-1 and 86% of respiratory disease incidences by EHV-4 ( Allen and Bryans, 1986; Allen et al., 1983b). Similar analysis of 43 EHV-1/EHV-4 isolates from Australian horses revealed EHV-1 as the cause of 86% of abortions and EHV-4 as the cause of all cases of respiratory disease (Studdert, 1983). In Australia EHV-1, apparently introduced into the continent in 1977, is primarily associated with abortigenic sequelae with only one published incidence of EHV-1 induced paresis and none of EHV-1 induced respiratory disease (Sabine and Whalley, 1989) . There is evidence that the dominance of the two types of virus, and even of specific virus strains (Allen and Bryans, 1986) with respect to

Figure 1.5

Clinical Manifestations of EHV-1 and EHV-4 Infections

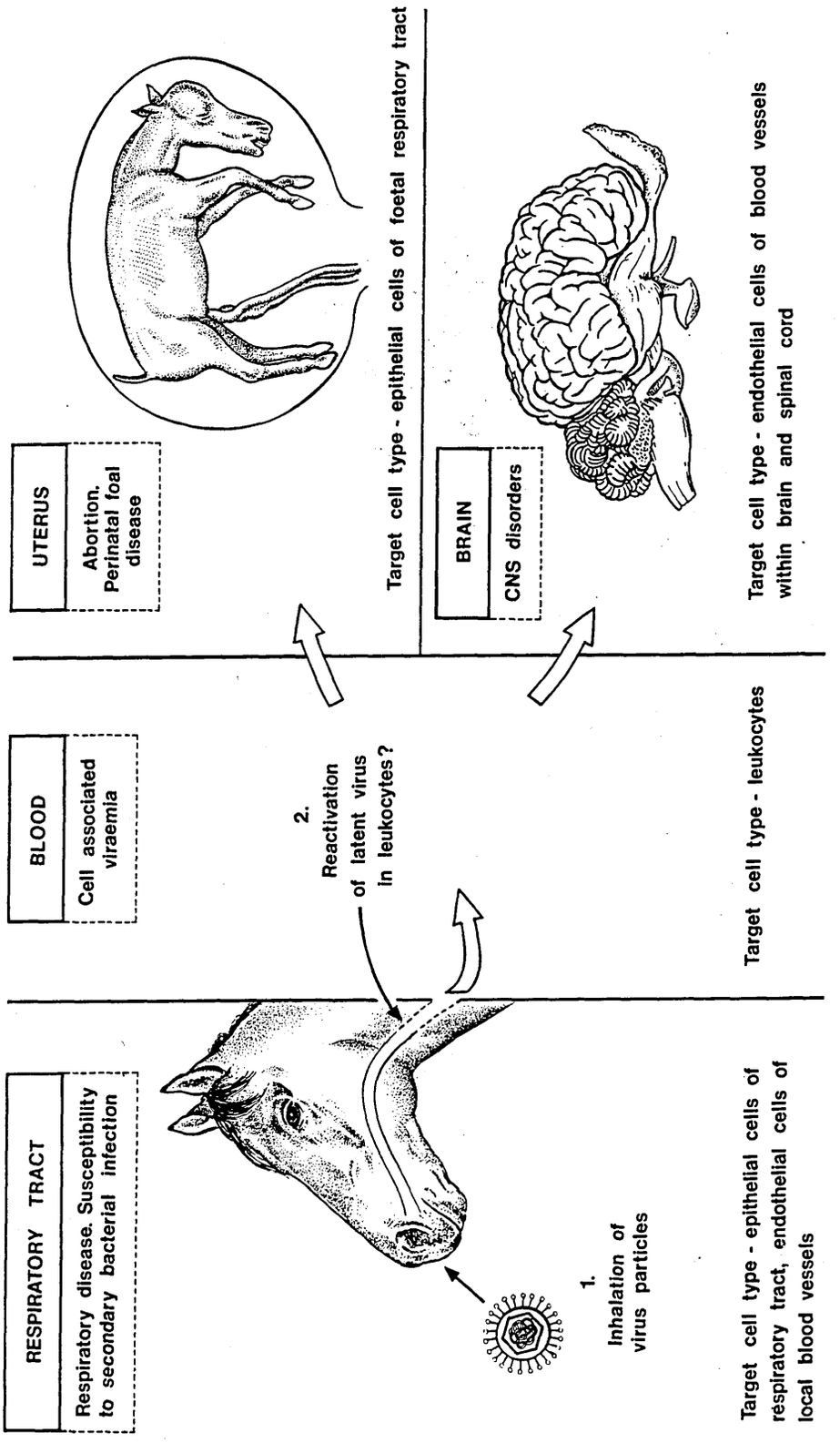



Figure 1.5 Pathogenesis of EHV-1 and EHV-4. Adapted from Allen and Bryans, 1986.

severity and nature of disease (i.e. whether respiratory or abortigenic) may vary temporally. For example, prior to 1979 in the U.K. EHV-4 was the major cause of respiratory disease. However, in recent years EHV-1 has been isolated from a significant number of 'respiratory isolates' and has been the major respiratory pathogen in certain years (Anon, 1989).

**1) EHV-1- and EHV-4-induced respiratory disease**

The usual entry route of EHV-1 and EHV-4 into the host is via the upper respiratory tract and as a consequence of this, and of the tissue tropism of the viruses, initial viral replication occurs in the respiratory epithelium. First exposure to EHV-1 or EHV-4 in foals often results in an acute febrile respiratory infection which may last up to one week (O'Callaghan, et al., 1983). EHV-1 induced respiratory disease tends to be more severe than EHV-4 induced disease since more virus tends to be produced in an EHV-1 infection. However, fatalities from EHV-induced respiratory disease are rare and often result from secondary bacterial infection rather than from respiratory sequelae . Virus is disseminated from the respiratory tract to draining lymph nodes by macrophages. EHV-1 can also infect endothelial cells of blood vessels lining the respiratory tract and of the brain ( Patel and Edington, 1983; Patel et al., 1982). An important difference between the two viruses is that although both viruses can be recovered from the respiratory tract up to 12 days post infection, only EHV-1 is routinely recovered, from 2-12 days post-infection, from the white cell population due to leukocyte-associated viraemia (Bryans, 1969). However, transient EHV-4 viraemias do seem to occur (A.A. Cullinane, personal communication). With increased exposures to virus, resistance develops against clinical respiratory disease such that an infected horse may suffer only mild respiratory symptoms or incur loss of form

or fitness with no definitive clinical symptoms. In such cases it may not be readily apparent that the causative agent is EHV-1 or EHV-4 and diagnosis may require identification of the pathogen by virus isolation or serological analyses .

Resistance to reinfection is limited to a 3-6 month period following infection. Subsequent infections result in productive infection of the upper respiratory tract ( with shedding of virus into the environment), a cell-associated viraemia , a secondary immune response and in some animals abortion or neurological disease (Bryans, 1969). It is the latter sequelae of abortion and neurological disease , rather than respiratory sequelae, which pose a serious potential threat to the life or well-being of the infected animal or its unborn young.

**2) EHV-1 and EHV-4-induced abortion**

As stated above both EHV-1 and EHV-4 induce abortion in susceptible mares. Infection by EHV-1 may lead to sporadic abortion from a single mare or may result in a multiple-abortion storm. In contrast, EHV-4 is associated with sporadic abortions and reports of EHV-4 -induced multiple abortion storms are exceedingly rare. Herpesvirus induced abortions tend to occur from months 8 to 12 (Doll, 1952; Doll and Bryans, 1963). The cause of foetal death is hypoxia due to separation of the placenta from the endometrium (Bryans and Prickett, 1972). Foals infected with EHV-1 or EHV-4 <u>in utero</u> and born alive tend to die soon after birth as a consequence of EHV-induced tissue damage and their increased susceptibility to secondary bacterial infection ( Hartley and Dixon , 1979). It is thought that EHV-1 reaches the foetus by transplacental transport of EHV-1 infected leukocytes. Since EHV-4 viraemia is occasionally detected it seems possible that EHV-4 might reach unborn animals by a similar route. Mares infected with EHV-1 may show no outward manifestation of disease due to a

resistance to respiratory symptoms and thus multiple abortions may occur within a cohort without prior indication of EHV-1 infection.

**3) EHV-1-induced neurological disease**

EHV-4-induced neurological disease is extremely rare with few instances reported (Meyer et al., 1987). Thus, it is not generally regarded as a neurological pathogen. EHV-1, however, causes neurological damage with greater frequency and paralytic outbreaks result in serious economic losses in stud and racing stables. EHV-1-induced neurological disease is a consequence of the infection of endothelial cells of central nervous system blood vessels by virus transported in the blood in association with leukocytes (Allen and Bryans, 1986). Nervous tissue damage is incurred due to hypoxia and inflammatory responses which result from endothelial infection (Patel et al., 1982). Clinical symptoms in affected horses range from mild lack of co-ordination and weakness to paresis, ataxia, and paralysis, usually of the hindlimbs. Horses suffering mild disease may recover fully or may be left with mild neurological disabilities. Horses with more severe symptoms have a poorer prognosis and may die as a consequence of secondary bacterial infection, pulmonary congestion or paralytic sequelae. As yet no marker has been discovered which distinguishes neurological strains of EHV-1 from abortigenic strains. The general view is that a combination of host suceptibility and virus neurotropicity lead to the development of neurological sequelae (Allen and Bryans, 1986).

## 3. EHV-1 and EHV-4 Genome Structure

**1) Standard genomes**

The EHV-1 and EHV-4 genomes are type D genomes in that they are composed of a long component ( L or $U_L$) of 111kbp and 109kbp and a

short component ( S) of 39kbp and 35kbp respectively, the latter of which comprises a short unique sequence ($U_S$) bounded by identical inverted repeat elements ($IR_S$ and $TR_S$) (Baumann et al., 1986 ; Cullinane et al., 1988; Henry et al., 1981; Ruyechan et al., 1982; Whalley et al., 1981 ). The %GC content of the EHV-1 genome is 57% (Soehner et al., 1965) while that of the EHV-4 genome has not been determined. Isomerisation of the EHV-1 genome is known to occur such that two isomers exist differing with respect to the orientation of the S component relative to the L component . This was deduced by analysis of the molarity of terminal restriction fragments relative to that of restriction fragments mapping internally ( Henry et al., 1981; Whalley et al., 1981). Predicted BamHI and EcoRI terminal restriction fragments of the two hypothetical isomers of EHV-4 are of similar size thus isomerisation of the EHV-4 genome has not yet been experimentally proved (Cullinane et al., 1988 ). However, it seems likely that as for EHV-1 , the EHV-4 genome exists as two isomers. Restriction maps of EHV-1 strain HVS-25 and of EHV-4 strain 1942 have been reported and are detailed in Figure 6 (Cullinane et al., 1988; Robertson and Whalley, 1988; Whalley et al., 1981).

Comparison of the EHV-1 and EHV-4 genomes to those of other alphaherpesviruses by Southern hybridisation indicates that the $U_L$ components of EHV-1 and EHV-4 are apparently colinear with each other (Cullinane et al., 1988) and with the corresponding regions of EHV-3 (Baumann et al., 1986), HSV-1 ($I_L$ isomer), HSV-2, and VZV ( Davison and Wilkie, 1983) . The $U_L$ component of PRV is colinear apart from map units 0.1-0.4 which are inverted relative to this region in the other genomes mentioned (Ben-Porat et al., 1983; Davison and Wilkie, 1983).

Figure 1.6

EHV-1 and EHV-4 Genome Restriction Maps

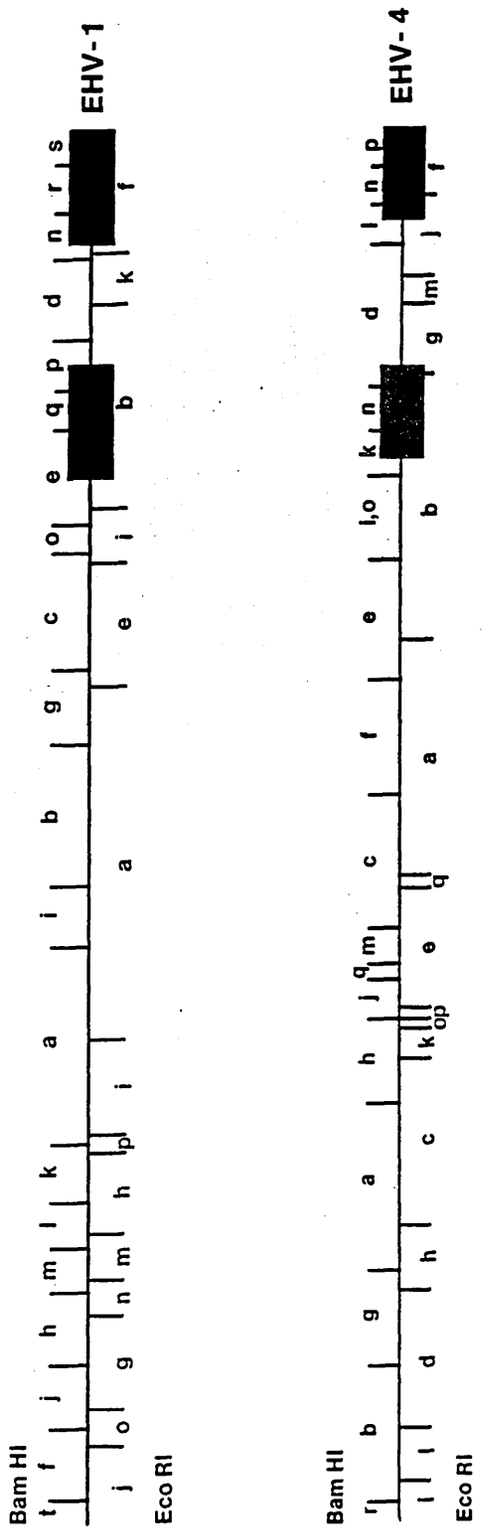



Figure 1.6 BamHI and EcoRI restriction maps of EHV-1 (strain HVS25, Whalley et al., 1981; Robertson and Whalley, 1988) and EHV-4 (strain 1942, Cullinane et al., 1988).

**2) Defective genomes**

Passage of EHV-1 and EHV-4 at high multiplicity results in the generation of defective interfering particles (DIP) (O'Callaghan et al., 1983) which inhibit the replication of standard virus. EHV-1 defective interfering particle genomes are composed of DNA derived from map units 0-0.4 (left end of $U_L$), 0.78-0.79 and 0.83-0.87 ($IR_S$), and 0.91-0.95 and 0.99-1.00 ($TR_S$) of the standard EHV-1 genome (Baumann et al., 1987). Studies of herpesvirus DI particle replication have indicated that DI DNA specifies cis-acting sequences such as cleavage-encapsidation signals and origins of replication (Frenkel et al., 1980) but that replication is dependent on functions supplied in trans by standard viral DNA sequences. Analysis of DI genomes has led to the mapping of an origin of replication within the central region of the inverted repeats of EHV-1 (Baumann et al, 1989).

**3) Genetic diversity of EHV-1 and EHV-4 genomes**

An indication of the genetic diversity of EHV-1 and EHV-4 can be obtained through the analysis of the banding pattern of viral DNA fragments produced by digestion of genomic DNA with restriction enzymes. Viral isolates with different arrangements of restriction endonuclease cleavage sites throughout the genome possess distinct electropherotypes. As mentioned above, this technique distinguishes between EHV-1 and EHV-4 field isolates since the viruses have no common restriction sites (Allen et al., 1983b). Extension of such analyses to isolates of a given virus type enables the determination of the variabilty in number and position of restriction enzyme sites and thus of the genetic diversity of the viruses. Comparison of the electropherotypes of individual U.S.A. field isolates of both viruses by Allen and Bryans (1986) indicated that EHV-1 is less diverse than EHV-4 . Of 235 USA field isolates of EHV-1, 90% possessed one of two

dominant electropherotypes out of a total of 16 identified electropherotypes. In contrast, 13 electropherotypes were identified within 21 EHV-4 isolates and of these, 50% had unique electropherotypes. Similar studies have yet to be performed on British EHV-1 and EHV-4 isolates.

Liquid hybridisation of EHV-1 and EHV-4 by Allen and Turtinen (1982) suggested the genomes share only 17% homology at the nucleotide level. A more accurate analysis of the homology of the two genomes would be achieved by comparison of their nucleotide sequences and on the basis of sequence data available for restricted regions of the $U_L$ component of EHV-1 and EHV-4, 17% would appear to be an underestimate (see Chapter 8).

## 4. Virus Propagation

### 1) Transcription

Investigation of the transcription of EHV has been restricted mainly to EHV-1. In EHV-1 transcription is coordinately regulated such that immediate early, early and late phases can be distinguished as in HSV transcription (Gray et al., 1987a). EHV-1 apparently specifies only one immediate early (IE) transcript of 6kb derived from part of the inverted repeats (m.u. 0.79-0.83 and 0.96-1.00) . At least four antigenically related IE proteins are derived from the transcript by differential 5′ splicing events . A similar IE transcriptional and translational arrangement is found in PRV (Fenwick and McMenamin, 1984). It has been suggested that different splicing patterns of the IE transcript may exist within persistently infected and oncogenically transformed cells. A gene homologue of HSV-1 US1 which encodes immediate early protein Vmw68 has been identified in EHV-4 $U_S$. However, immediate early transcripts have yet to be associated with this region

of the EHV-4 or EHV-1 genome (Cullinane et al., 1988). Around 45 early and 20 late transcripts have been identified (Gray et al., 1987b).

**2) Translation**

EHV-1 protein synthesis is coordinately regulated in a manner similar to that in HSV with immediate early, early, and late phases (Caughman et al., 1985, 1988). Four antigenically related immediate-early proteins are synthesised, IE1, IE2, IE3, and IE4, which are presumably important trans-acting regulatory proteins serving to activate early gene expression (Caughman et al., 1985). 18 early (beta) and 12 late proteins (beta/gamma rather than true late proteins) have been identified. Around thirty virion polypeptides have been identified which are distributed between the core of the virion, the nucleocapsid, the tegument , and the envelope. Although most of the proteins isolated from EHV-1 virions have electrophoretic counterparts in EHV-4, at least 8 have distinct electrophoretic mobilities in EHV-1 and EHV-4 (Allen and Bryans, 1986). Studies have indicated that the nucleocapsid contains six polypeptide species (O'Callaghan et al., 1983 ) including VP9 (148 kDa), VP22 (46kDa), and VP24. Scanning electron microscopy studies have indicated that the 150 hexavalent capsomers of the 100nm capsid comprise hexamers of p9 (Baker et al., 1990; Newcomb et al., 1989) which accounts for 55-75% of nucleocapsid protein (O'Callaghan and Randall, 1976). The constituents of the pentameric capsomers have yet to be determined. Eight putative tegument proteins have been identified and twelve envelope glycoproteins (Allen and Bryans, 1986; Turtinen and Allen, 1982).

**3) Virus DNA replication**

Replication of EHV DNA is believed to occur, as for other herpesvirus genomes, by a rolling-circle mechanism. Consistent with this is the finding that the genomic termini of EHV-1 seem to comprise

complementary 3′ single-base extensions which could presumably base pair within the infected cell to circularise the genome. Sequence analysis of the terminal region of EHV-1 has also revealed potential signals for cleavage of concatameric replicated DNA (Chowdhury et al., 1990).

## 5. EHV-1 and EHV-4 Glycoproteins

EHV-1 and -4 specify at least twelve glycoproteins, eight of which are highly abundant within the virion envelope as identified by electrophoretic analysis of virion proteins (Allen and Bryans, 1986; Turtinen and Allen, 1982). These have been designated gp 2, 10, 13, 14, 17, 18, 21, 22a and assigned molecular masses of 200, 125, 95, 90, 68, 63, 45, and 41 kd respectively by Turtinen and Allen (1982) and by Allen and Bryans (1986) . Ten glycosylated proteins of EHV-1 and nine of EHV-1 have been identified by similar techniques by Meredith and coworkers (Meredith et al., 1989) and seven by Bridges et al., 1988 (Table 2). The relatedness of glycoproteins separated by different groups cannot be conclusively deduced from comparisons of mobilities since the mobility of the proteins depends not only on the electrophoretic system and running conditions but on the virus strain and cells in which the virus is propagated. Identification of glycoproteins will also depend on the labelling system used and the limits of sensitivity of the system. Glycoprotein homologues of HSV-1-specified glycoproteins have been identified by antigenic cross-reactivity and, more recently, by sequence analysis (Table 1). Thus gC and gB-type EHV-1 glycoproteins and their genes have been identified by antigenic analyses, gene expression studies, and sequence data ( Allen and Coogle, 1988; Allen and Yeargan, 1987; Bell et al., 1990; Guo et al., 1989; Whalley et al., 1989). Currently two nomenclature systems

Table 1.2

Glycoproteins of EHV-1 and EHV-4

| EHV-1 | | | EHV-4 | | |
|---|---|---|---|---|---|
| a) | b) | c) | a) | b) | c) |
| 250 (1) | >250 (1/2) | 300 | 250 (1) | >250 (1/2) | 300 |
| 190-240 (2) | - | - | 190-240 (2) | - | 260 |
| - | - | - | - | - | 220 |
| 140 (9a) | 143 (9/9a) | - | - | 147 (9/9a) | - |
| 128 (10a) | - | - | - | - | - |
| 124 (10) | 134 (10) | 118 | 124 (10) | 134 (10) | 138 |
| - | - | 108 | - | - | 112 |
| 96 (13) | 99 (13) | 88 | 110 (13) | 113 (13) | 92 |
| 90 (14) | 87 (14) | 76 | 87 (14b) | 94 (14a) | 74 |
| 74 (16) | - | 66 | 74 (16) | - | - |
| 68 (17) | - | 60 | 68 (17) | - | 61 |
| 63 (18) | 64 (18) | 58 | 61 (18) | 62 (18) | 59 |
| 45 (21) | 46 (21) | 45 | 45 (21) | 46 (21) | 45 |
| 41 (22a) | - | - | 41 (22a) | - | - |
| - | - | - | 33 (23a) | - | - |
| 24 (25) | - | - | 25 (25) | - | - |

Table 1.2 Molecular weights of the envelope glycoproteins of EHV-1 and EHV-4 identified by three research groups, a) Turtinen and Allen (1982) and Allen and Bryans (1986), b) Bridges et al. (1988), and c) Meredith et al. (1989) are presented. Authors' proposed designations of glycoproteins of specific molecular weights under the alternative nomenclature system, gp13, gp14 etc. are indicated in brackets ( a) and b) only).

are in use - glycoproteins are identified either by their mobility relative to that of all proteins, as in gp13 or gp14 (Allen and Bryans, 1986), or by their molecular weight, as in gp88 or gp76 (Meredith et al., 1989). As increasing data has become available it has become clear that although up to 12 glycoprotein species may be specified by EHV-1 and EHV-4 some of these glycoproteins are specified by the same gene and differ only in the extent of their post-translational processing. For example, three species have been identified which react with gB polyclonal antisera. One of these, gp108/gp112, (EHV-1/EHV-4) is a precursor and the others , gp76/gp74 and gp58/59, mature glycoprotein subunits which are disulphide linked in active dimeric 'gB' but separated by electrophoresis under reducing conditions (Meredith et al., 1989; Sullivan et al., 1989) .

Little is known of the precise functions of equine herpesvirus glycoproteins although more is becoming known of the immunogenicity of individual glycoproteins.

## 6. Sequence Data and Mapping Studies

Sequencing and mapping of EHV genes has been restricted to glycoprotein genes and the immediate early gene to date although several other genes have been sequenced as a consequence of their proximity to the primary genes of interest. Allen and Yeargan (1987) were the first to report genomic positions of EHV-1 glycoprotein genes. They constructed a lambda gt11 library of EHV-1 and identified recombinant bacteriophage expressing epitopes of six major glycoproteins using anti-glycoprotein monoclonal antibodies. The reported map positions of EHV-1 gp13, gp14, gp10, gp21/22a, gp2, and gp17/18 are detailed in Figure 7. Genes gp13 and gp14 map at positions colinear with HSV-1 genes encoding gC and gB and sequence data has

confirmed that gp13 and gp14 encode gC and gB homologues (Allen and Coogle, 1988; Hudson et al., 1988; Whalley et al., 1989). That only one gene was mapped to the $U_S$ component which is usually the major site of glycoprotein genes in alphaherpesviruses and that no gH gene was mapped to the $U_L$ component suggests the technique used is limited to detection of highly abundant and/or immunogenic glycoproteins and that further glycoprotein genes will be mapped as increasing sequence data becomes available. To that end, the map positions of EHV-1 gH, gD, gI, and gE gene homologues have recently been elucidated on the basis of sequence analysis (Audonnet et al., 1990; Robertson and Whalley, 1988; Robertson et al., 1991). Several late proteins have been mapped to the $U_S$ component of EHV-1 (Robertson et al., 1988). These might include products of gD, gI, gE, and as yet unreported $U_S$ glycoprotein genes. To date no glycoprotein genes have been mapped in other herpesvirus genomes at positions colinear with those reported for gp10 , gp2, or gp21/22a. Other genes sequenced in part or in entirety include the thymidine kinase gene (UL23), the major capsid protein gene (UL19) and gene homologues of HSV-1 genes UL15, UL18, UL24, UL25 , and UL28 (Robertson and Whalley, 1988; Sabine and Whalley, 1989; Whalley et al., 1989). In $U_S$, an IE175 immediate early gene homologue and US9 gene homologue have been characterised (Audonnet et al., 1990; Grundy et al., 1989).

Published EHV-4 sequence data is limited to gB (UL27) and UL28 gene homologues mapped to $U_L$ (Riggio et al., 1989) and to gE (US8), US9, US10, and IE68 (US1) gene homologues mapped to $U_S$ (Cullinane et al, 1988).

## 7. EHV-1 and EHV-4 Latency

EHV-1 and EHV-4 are capable of persisting in the host in a latent state until reactivation is triggered and productive infection , with its potential for inducing pathological sequelae and virus spread, is reinitiated. Strategies for demonstration of latent infection are based on detection or experimental reactivation of latent genomes in an animal which has recovered from productive infection and include i) attempted isolation of virus from tissue explants, ii) demonstration of corticosteroid-induced reinitiated lytic infection in animals , iii) detection of viral genomes *in situ* or in whole tissue extract by DNA hybridisation, and , more recently, iv) detection of latent viral DNA using the polymerase chain reaction technique. The major evidence for EHV-1 and -4 latency until recently was circumstantial in nature and comprised a report by Burrows and Goodridge (1984) that EHV-1/4 (untyped) had been isolated from ponies in a herd isolated from other horses for ten years. More recently the induction of reactivation of EHV-1 and EHV-4 by corticosteriod administration has been reported (Browning *et al.*, 1988; Edington *et al.*, 1985). The site of latency of EHV-1 and EHV-4 has not been firmly established. A common site of herpesvirus latency is ganglionic tissue yet attempts to consistently re-isolate EHV-1 or -4 from explanted ganglia of latently infected horses have been unsuccessful with one exception - the isolation of EHV-4 from a trigeminal ganglion (Allen and Bryans, 1986). Another common target tissue for herpesvirus latency is leukocytes (Roizman, 1982) . On current evidence this seems the most probable site of EHV-1 and -4 latency since EHV-1 is known to induce a cell associated viraemia infecting T and B lymphocytes and monocytes (Gleeson and Coggins, 1980; Scott *et al.*, 1983) and although EHV-4 viraemia is comparatively rare, cases have been reported. Due to the differences

in pathology of the two viruses it cannot be assumed, however, that the viruses have identical sites of latency (Mumford, 1985).

## 8. The Equine Immune Response to EHV-1 and EHV-4 Infection

The importance of EHV-1 envelope and nucleocapsid antigens in inducing a virus-neutralising response in the host was established by Papp-Vid and Derbyshire (1978, 1979). Specific key antigens, glycoproteins 2, 10, 13, 14, and , to a lesser extent, 22a and the major capsid protein, VP9, were later identified by testing the reactivity of equine serum with whole viral proteins by an immunoblot procedure (Allen and Bryans, 1986). Glycoproteins 2, 10, 13, and 14 were identified as type-common in that sera from convalescent horses infected with one viral type was reactive with these glycoproteins in immunoblots of both EHV-1 and EHV-4 protein preparations (Allen and Bryans, 1986). Glycoproteins 18, and 22a are thus apparently type-specific. Bridges and Edington (1987a) reported reactivity of equine sera with additional non-glycoprotein polypeptides, VP11, 15, 16, 20, 21, and 23a.

Monoclonal antibody studies have indicated that four types of EHV-1 and EHV-4 epitopes exist - those which are present throughout all EHV-1 and EHV-4 isolates (type-common, conserved), those present in all isolates of one or other type (type-specific, conserved), those present in some isolates of both types (type common, variable), and those present only in some isolates of one type (type-specific, variable) (Allen and Bryans, 1986).

Cellular immune mechanisms may make a greater contribution to the protective immune response in horses than humoral immune mechanisms since virus neutralising antibody responses do not seem to correlate with resistance to infection ( Bryans, 1969; Mumford and Bates, 1984).

Information on the key EHV-1 and -4 antigens stimulating a cellular immune response in equines is less extensive than the equivalent knowledge on humoral immune system targets. Bridges and coworkers (1988) found that fractions of EHV-1 proteins which contained glycoprotein were capable of inducing lymphocyte transformation. Responses were directed against fractions containing glycoproteins of molecular mass 140k (gp9a), 128k (gp10a), 124k(gp10), 96k (gp13), 90k (gp14),and 45k (gp21) of EHV-1 and 134k (gp10), 113k (gp13), 94k (gp14), 62 (gp18), and 46k (gp21) of EHV-4.

In general, the gross virus-neutralising and complement-fixing (CF) antibody response and T-cell mediated cytotoxic immune response are type-specific after primary infection but increase in their crossreactivity (to heterologous virus targets) following subsequent exposures to EHV (Allen and Bryans, 1986; Bridges and Edington, 1987b; Fitzpatrick and Studdert, 1984).

Antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent lysis (CDL) mediation by sera of EHV-infected horses has been demonstrated . These activities , which do not appear to be involved in protection against reinfection, are crossreactive with targets infected with the heterologous virus (Bridges and Edington, 1987a; Stokes and Wardley, 1988).

Experimental primary infection of horses with EHV-1 or EHV-4 induces short-lived immunity to challenge by the homologous virus type but not to the heterologous virus type. However, following multiple administration of one viral type at suitable intervals, the horse is protected against challenge by both the administered virus type and the heterologous type (Allen and Bryans, 1986; Burrows et al., 1984; Edington and Bridges, 1990; Fitzpatrick and Studdert, 1984; Mumford and Bates, 1984). Conflicting viewpoints exist on which virus type would be

the more effective in a single virus vaccine preparation administered to protect against both EHV-1 and EHV-4-induced disease (Edington and Bridges, 1990; Fitzpatrick and Studdert, 1984). It has been suggested that a bivalent EHV-1/4 vaccine might be the most effective protective agent.

## 9. Immunoprophylaxis and Chemotherapy

Two vaccines are at present marketed in the U.K. for protection against EHV-1 and EHV-4 induced disease : Pneumabort-K , marketed by ScanVet, a formaldehyde-inactivated oil adjuvant vaccine administered to prevent abortion and Rhinomune, marketed by Smith Kline, a live attenuated virus administered to protect against respiratory disease (Anon, 1987). Multiple, appropriately spaced administrations of both viruses are recommended by the manufacturers. These vaccines are of limited efficacy and , due to the financial losses incurred as a result of EHV-1 and EHV-4 infections, development of a more effective vaccine would be of some benefit to the equine industry.

The ultimate aim of a novel EHV vaccine would be to induce an immune response in the horse superior to that induced by current vaccines or indeed wild type EHV in terms of longevity and quality particularly with respect to protection against reinfection. Obviously a considered approach to the development of such a vaccine requires i) more information on the pathogenicity and latency of the viruses, ii) identification of the proteins, and ultimately epitopes, which are important in eliciting a protective immune response, iii) identification of virus-specified immunosuppressive functions or epitopes, and iv) further characterisation of the mechanisms of the equine immune response involved in protection against reinfection.

A novel vaccine should possess the immunodominant epitopes with respect to a protective immune response. Since some of these may be type-specific or even strain-specific, identification of these epitopes would be beneficial to the selection of an optimal vaccine strain. However, such dissection of the antigenicity of the key glycoproteins is time-consuming and more empirical vaccine selection strategies have greater feasibility in the short term.

Approaches to the development of novel EHV vaccines are outlined in Figure 7. Four types of live vaccine could be developed - i) a monovalent attenuated EHV-1 or EHV-4, ii) a bivalent EHV-1/4 recombinant comprising immunodominant glycoprotein genes of one virus inserted into the attenuated genome of the heterologous virus , iii) a recombinant in which genes of non-EHV origin are inserted into an attenuated EHV-1 or EHV-4 genome, or iv) a recombinant comprising genes of EHV-1 or -4 inserted into a live vector other than EHV which can be safely administered to horses. The basis for three of these vaccines is an EHV-1 or EHV-4 vector from which virulence functions have been deleted. Functions linked to virulence in other herpesviruses include thymidine kinase, ribonucleotide reductase, gC, and in $U_S$ several glycoproteins and a protein kinase (Cameron et al., 1988; Kit et al., 1987c;. Mettenleiter et al., 1988). Cornick and coworkers (1990) produced an attenuated TK- EHV-1 virus by nucleoside analogue-induced mutagenesis. The contribution of TK deletion to overall attenuation was not, however, evaluated.

A major consideration in the development of live vaccines is their safety : the vaccine strain must be incapable of inducing severe disease within the horse. It must also be borne in mind that vaccine virus could potentially recombine with wild type virus, either at the time of inoculation or at a later date, on reactivation, to produce a

Figure 1.7

Potential vaccine design strategies for the control of EHV-1 and EHV-4

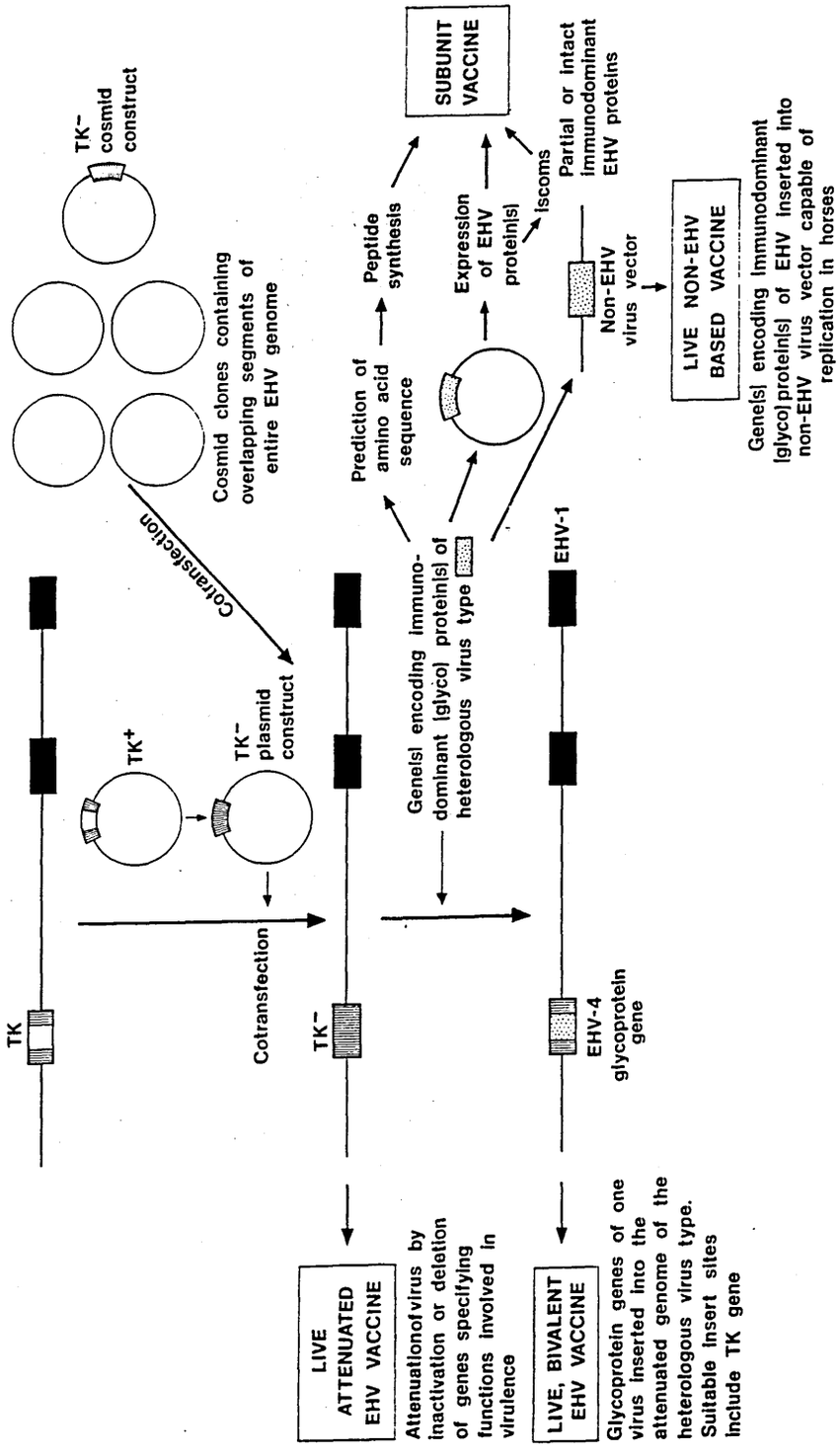



Figure 1.7 Schematic diagram outlining several approaches to the development of a novel vaccine against EHV-1 and/or EHV-4.

recombinant of unknown pathogenicity (Henderson et al., 1990).

Subunit vaccines could be developed comprising mixtures of several immunodominant EHV-1 or -4 proteins or of peptides derived from these proteins . Candidates for inclusion in subunit vaccines are obviously envelope glycoproteins since these have been associated with the humoral and cellular immune response in horses (Allen and Bryans, 1986; Bridges et al., 1988) . However, it must be borne in mind that other virion components such as the major capsid protein and nonstructural polypeptides may be important particularly in eliciting cellular immunity (Martin et al., 1988). The combination of gp13 and gp14 expressed in vaccinia virus recombinants elicited a protective immune response in hamsters (Guo et al., 1989, 1990) suggesting these might be potential subunit vaccine candidates. gH, gD, gp17/18, and indeed as yet uncharacterised EHV glycoproteins , may specify epitopes which might be important constituents of an EHV vaccine of any type i.e. peptide, subunit, or live virus (Stokes et al., 1989) . An ISCOM vaccine containing the major glycoproteins of EHV-1 protects hamsters against lethal challenge with EHV-1 (Cook et al., 1990).

An alternative approach to the control of EHV-1 and EHV-4 is to develop antiviral therapies. Targets for antiviral drugs include ribonucleotide reductase and thymidine kinase enzymes (Field and Awan, 1990; Telford et al., 1990). The development and characterisation of hamster and mouse models of EHV infection should provide an opportunity to evaluate potential chemotherapeutic and immunoprophylactic strategies in the laboratory prior to expensive clinical trials in horses (Awan et al., 1990; Stokes et al., 1989).

## OBJECTIVE OF THE PROJECT

The objective of this project was to determine the localisation of specific EHV-4 genes with the aim of using this information in the future development of recombinant EHV-1 or EHV-4 based vaccines. From the point of view of live EHV vaccine development three types of gene are of particular interest : genes specifying proteins which elicit a protective immune response; genes specifying functions involved in pathogenicity; genes encoding functions which are involved in immunosuppression and/or mis-direction of the host protective immune response. On the basis of the background of information available on other herpesviruses, HSV-1 in particular, we selected three genes for investigation. The thymidine kinase (TK) gene was selected due to the apparent role of TK in herpesvirus virulence. Its localisation would ultimately enable construction of TK- EHV-4 recombinants which could be tested for reduced pathogenicity. The other genes selected were those encoding homologues of HSV-1 genes encoding gH and gC . Herpesvirus gC-type glycoproteins tend to be a major glycoprotein component of the virion envelope, specify neutralising antibodies , may stimulate cellular immunity, and, in the case of HSV-1 and HSV-2 gC at least, may interact with the host complement system. Herpesvirus gH-type glycoproteins tend to be minor components of the envelope but nevertheless may be essential for viral replication and specify one or more epitopes which elicit a complement-independent neutralising antibody response.

# CHAPTER 2

# MATERIALS AND METHODS

# MATERIALS

## 1. Bacteria and DNAs

*E. coli* strains JM101 and JM109 were obtained from Northumbria Biotechnology Limited (NBL).

pUC8 and lambda HindIII and HindIII/EcoRI DNAs were obtained from Bethesda Research Laboratories. ØX174 , pBR322 HaeIII and HpaII size marker DNAs and salmon sperm DNA were obtained from Boehringer-Mannheim. Bluescript M13+ , manufactured by Stratagene, was obtained from NBL.

BamHI libraries of EHV-1 strain HVS25 in pBR322 and of EHV-4 strain 1942 in pUC9 were kindly provided by J.M. Whalley (Macquarie University, Australia) and A.A. Cullinane (Irish Equine Centre, Eire) respectively.

## 2. Chemicals

Unless stated to the contrary all chemicals were obtained from B.D.H. Chemicals Ltd., Poole, Dorset.

Agarose (type II), low melting point agarose, ethidium bromide, lysozyme, triton X100, dextran sulphate, ampicillin, spermidine - Sigma.

Caesium chloride, IPTG, X-Gal - Boehringer-Mannheim, UK.

TEMED , APS , Sephadex G50, Bind-Silane , Repelcote and nucleotides - Pharmacia/LKB

Formamide - Fluka

Bacterial agar and tryptone - Oxoid

Yeast extract and bactopeptone - Difco

Tris-equilibrated phenol - Rathburn Chemicals Ltd.

Radiochemicals, $\alpha$-$^{35}$S-ATP and $^{32}$P-dCTP - Amersham Radiochemicals Ltd.

Gamma $^{32}$P-ATP - ICN.

## 3. Enzymes

Restriction enzymes and DNA ligase and buffers were obtained from BRL or NBL.

T4 polymerase, Exo III - BRL.

Klenow enzyme (sequencing grade), S1 nuclease, alkaline phosphatase - Boehringer-Mannheim, UK.

T7 DNA polymerase was supplied as part of a sequencing kit by Pharmacia/LKB.

# METHODS

Methods used throughout the thesis are described in this chapter while methods specific to one section of the work are described in later chapters.

## 1. Recombinant DNA Techniques

### 1) Storage and growth of bacterial strains

All plasmids were maintained in *E. coli* strains DH1, JM101, or JM109. To facilitate long term storage of these bacterial stocks and of transformants obtained in the course of the project, glycerol stocks were prepared. A small aliquot of a given bacterial culture was streaked out on a 1.5% agar plate (1.5% bacterial agar in L-broth (Table 1)) supplemented with 100ug/ml ampicillin since all the vector plasmids used confer ampicillin resistance to their host bacterium. The plate was incubated overnight at 37°C and single colonies picked into 10ml L-broth /100ug/ml ampicillin the next day. The culture was incubated for 24 hours in an orbital incubator at 37°C. Confirmation that the overnight culture was derived from a bacterium containing the desired recombinant plasmid was achieved by DNA extraction and restriction digestion (see below). 5ml of the culture was then centrifuged at 2K for 10 minutes and the bacterial pellet resuspended in 1ml 80% glycerol, 2% bactopeptone in L-broth. Glycerol stocks were stored at -20°C.

Bacteria were routinely cultured in 2 litre sterile flasks to which 0.5ml of an overnight culture of the desired bacterial clone and 500 ml L-Broth/100ug/ml ampicillin had been added. Cultures were incubated for 16-24 hours at 37°C in an orbital incubator.

### 2) Extraction of plasmid DNA

Bacterial pellets were obtained from 500ml cultures by

centrifugation of the culture at 7K for 15 minutes and plasmid DNA isolated by a modification of the technique of Birnboim and Doly, 1979). Pellets were resuspended in 50ml solution 1 (25mM Tris, 50mM glucose, 10mM EDTA) and 200mg of lysozyme on ice for 30 minutes. 80ml solution 2 ( 1% SDS, 200mM NaOH) was added and the lysate left on ice for a further 5 minutes. 40ml 3M KOAc was added and , following a further 15 minutes on ice , the preparation was spun at 7K for 15 minutes. The supernatant was filtered through sterile gauze, 100ml isopropanol added and DNA precipitated at -20$^{o}$C for 5-30 minutes. DNA was pelleted by centrifugation at 7K for 30 minutes at 0$^{o}$C and was then resuspended in 17ml 1xTE (Table 1). To facilitate formation of a density gradient during centrifugation and to induce differential density of plasmid and chromosomal DNAs 18g CsCl and 1.8 ml ethidium bromide were added to the DNA preparation. The optical density (OD) of the solution was adjusted to 1.388-1.390 by adding more 1xTE or CsCl and determining the OD of the solution in a refractometer. The preparation was transferred to quick seal tubes which were heat sealed. Preparations were spun in a Beckman ultracentrifuge for 36-48 hours at 45K in a Ti50 rotor or alternatively overnight at 55k in a VTi65 rotor . The plasmid DNA band was removed from the gradient using a sterile needle and syringe and ethidium bromide extracted using three to four equal volume extractions with isopropanol. The aqueous phase was then transferred to a pre-washed collodian tube and dialysed for a minimum of 2 hours against 0.5xTE. DNA was precipitated from the dialysed preparation by the addition of 2 volumes of ethanol and 1/10th volume of 5M NaCl at -20$^{o}$C. DNA pellets recovered by centrifugation at 18K for 30 minutes were washed with 70% ethanol, dried and resuspended in 2ml $dH_2O$. DNA was reprecipitated, washed, dried and resuspended in water or 1xTE at a concentration of 0.5 ug/ul.

**3) Determination of nucleic acid concentration**

The concentration of DNA within a given preparation was determined by the analysis of the optical density of the sample at 260 nm. 5 ul of the DNA preparation was diluted 200-fold with water and its $OD_{260}$ determined with reference to a water standard. An $OD_{260}$ of 1.00 represents a dsDNA concentration of approximately 50 ug/ml or an RNA or oligonucleotide concentration of approximately 20ug/ml .

**4) Restriction endonuclease digestion and electrophoresis of plasmid DNA**

Restriction digests of plasmids were carried out in the appropriate buffer supplemented with spermidine to a concentration of 1mM. Routinely, 0.5-1.0 ug DNA were digested for a minimum of 1 hour at 37°C in an excess of enzyme. For isolation of specific restriction fragments larger quantities of plasmid DNA , 10-20 ug, were digested.

Fractionation of fragments generated by restriction endonuclease digestion was achieved by electrophoresis of digests through agarose or polyacrylamide gels. 0.7% -1.0% agarose gels were prepared by adding the appropriate amount of agarose to 1xTBE buffer and heating to melting point in a microwave oven. Gel mixes were cooled to approximately 50°C before pouring into the gel former, with appropriate well comb, of a horizontal gel apparatus. The set gel was placed in the electophoretic cell and submersed in 1xTBE. The comb was removed under buffer.

Analysis of restriction fragments of under 500 bp in size was facilitated by electrophoresis of restriction digest products through vertical polyacrylamide gels. 5%-8% gels were used according to the size of fragments to be distinguished. 5% acrylamide gels were prepared by adding 10ml of a 20% acrylamide/1% bisacrylamide solution to 4ml 10xTBE and 25.3 ml $dH_2O$. 480 ul 10% APS and 240 ul 10% TEMED were

added, the solution mixed and poured between two glass plates separated by 0.8mm spacers and clamped firmly with metal clips to prevent leakage of gel. A gel comb was placed in position and the gel allowed to set. On setting the gel was placed in a vertical electrophoresis kit, clamped in position and 1xTBE buffer added to the upper and lower chambers. The comb was removed and the wells were flushed out with 1xTBE to remove any unpolymerised acrylamide. Any air bubbles at the gel/buffer interface were removed.

An appropriate amount of 5 x bromophenol blue loading buffer (Table 1) was added to the restriction digests prior to electrophoresis. Size markers used were lambda HindIII or HindIII/EcoRI digests for agarose gels or pUC8 HaeIII or HpaII digests for polyacrylamide gels. Samples were then loaded into the gel wells and the gel run for an appropriate time - 2-3 hours at 70V/cm for 0.8% agarose gels, 2-3 hours at 35 mA for 5% polyacrylamide gels. Gels were stained with ethidium bromide for 30 minutes, destained for 30 minutes and viewed and photographed under UV illumination using instant black and white film.

**5) Isolation of purified restriction fragments**

Purified restriction fragments required for utilisation as hybridisation probes or for recombinant plasmid construction were isolated from low melting point agarose. DNA was electrophoresed in 0.75% low melting point agarose, the gel stained and the desired fragment excised from the gel under UV illumination using a sterile scalpel blade. Two volumes of TE were added to the gel slices and the agarose melted by heating to $65^{0}$ C for 10 minutes. The preparation was phenol extracted and the aqueous phase transferred to a fresh tube. The phenol phase was back-extracted with 1 volume of TE and the aqueous phase pooled with the previously collected aqueous extract. A further two phenol extractions were performed followed by two chloroform

extractions. DNA released from the gel slices was precipitated at -20° C by the addition of 1/10th volume of 5M NaCl and 2.5 volumes of ethanol. The DNA pellet obtained by centrifugation of the preparation at 13K for 20 minutes was washed with 70% ethanol, dried and resuspended in 1xTE.

An alternative method of restriction fragment purification utilised involved spinning gel slices through a 0.2 um filter according to the protocol of Vogelstein (1987). The restriction fragment of interest was excised from a 0.75% normal melting point agarose gel and the slices suspended on the 0.2 um filter of a Costar Spin-X tube. Gel slices were subjected to freeze-thawing at -20°C and 37°C and the tubes then spun for 40 minutes at 13K. Fluid which had passed through the filter was phenol extracted twice and chloroform extracted. DNA was precipitated and treated as above.

**6) Phosphatase treatment of vector DNA**

Subcloning of fragments terminating in blunt ends or in common restriction endonuclease sites required plasmid vector DNA digested with a single restriction enzyme : an enzyme identical to that used to generate the target fragment to be cloned or SmaI in the case of cloning blunt-ended fragments into Bluescript M13+ (BS). Digested vectors were phosphatase-treated to prevent ligation of vector-self or vector-vector at the cleavage site. Since dephosphorylated termini can only ligate to compatible phosphorylated termini ligation events were restricted to those of a vector-fragment or fragment-fragment nature. Dephosphorylation of vector termini was achieved by the addition of 1 unit of alkaline phosphatase to the restriction enzyme digest in the final 30 minutes of the incubation period.

**7) Ligation of vector and target DNA**

Vector and target DNAs were mixed at a concentration of 1:3 in the

presence of ligase buffer and 2 units T4 DNA ligase. Blunt-ended ligations were performed overnight at room temperature. Standard ligations were carried out at 15°C overnight.

**8) Transformation of competent bacteria**

Bacterial cultures were prepared for transformation with recombinant plasmids by innoculating 100ml L-broth with 1ml of an overnight culture of JM101 or JM109 E. Coli and shaking the culture in an orbital incubator for 2 - $2^1/_2$ hours at 37°C. Cells were pelleted by spinning the culture at 1k for 10 minutes. The pellet was resuspended in 40ml 100mM $MgSO_4$ on ice for 10 minutes. Cells were spun down at 1K for 10 minutes and resuspended in 25ml 50mM $CaCl_2$ on ice for 15 minutes. Cells were pelleted once more and resuspended in 2.5ml 50mM $CaCl_2$. Transformed cells were kept for up to 48 hours at 4°C. For longer term storage 1/2 volume of 50% glycerol was added and aliquots stored in the vapour phase of a liquid nitrogen tank or alternatively at -80°C.

The products of a ligation reaction mix were added to 200 ul competant bacteria and placed on ice for a minimum of 15 minutes. The bacteria were then heat shocked by transferring the tube to 37°C for 5 minutes and then returning it to ice for 30 minutes. 1ml of L-broth was added to each tube and the contents incubated for 90 minutes at 37°C to permit expression of the plasmid ampicillin-resistance function. Bacteria were then pelleted by centrifugation at 13K for approximately 20 seconds and resuspended in 100 ul L-broth. 75ul of the suspension were spread on to ampicillin-supplemented agar plates which were subsequently incubated overnight at 37°C.

Cloning of inserts into the multicloning site (MCS) of Bluescript M13+ disrupts beta-galactosidase expression thus when using BS, transformed bacteria were spread on plates supplemented with X-gal (40ug/ml) and IPTG (25ug/ml) in addition to ampicillin to assist in

selection procedures : a blue colony phenotype is associated with transformants containing a religated vector expressing beta-galactosidase; a white colony phenotype is associated with transformants containing vectors with a disrupted lac gene due to insertion of DNA.

**9) Screening of transformants for desired recombinant plasmids**

Several colonies were picked from each overnight plate of transformed bacteria. Only white colonies were picked on occasions when colour selection was utilised. Minipreps were performed to establish whether the selected colonies contained plasmid DNA and, if so, whether the plasmid constituted a desired recombinant. A 1ml aliquot of each single colony-derived overnight was spun at 12K for 3 minutes and the cell pellet resuspended in 70ul 1xSTET (Table 1). Following addition of 25ul of a 10mg/ml lysozyme solution, each preparation was placed in a boiling water bath for 45 seconds. Supernatant obtained by spinning at 13K for 5 minutes was phenol extracted and chloroform extracted . Plasmid DNA was recovered by precipitation in two volumes of ethanol and 1/10th volume 5M NaCl at -70$^{o}$C. DNA was pelleted by spinning at 13K for 10 minutes and was resuspended in 30 ul 1xTE. Restriction digests were then performed in order to establish the nature of the plasmid construct within selected colonies. Cultures derived from colonies with the desired recombinants were then used to prepare glycerol stocks and large scale plasmid preparations.

## 2. Southern Hybridisation

**1) Preparation of radiolabelled probe**

Radioactive labelling of a purified restriction fragment was achieved by introducing nicks into the DNA at which a fill-in reaction

could be initiated, in the presence of a radioactively labelled deoxynucleotide. An Amersham nick translation kit was used. To 1 ug DNA on ice the following was added in sequence - 10 ul of nucleotide/buffer solution ( 100 uM dATP, dGTP and dTTP in Tris-HCl pH7.8, $MgCl_2$ and 2-mercaptoethanol ), 10 ul of $\alpha$ - $^{32}$P-dCTP (10mCi/ml), $dH_2O$ to bring the final volume to 50 ul and 5 ul of enzyme solution (2.5U DNA polymerase 1, 50pg DNase 1 in tris-HCl, pH 7.5, $MgCl_2$ , glycerol and bovine serum albumin ). The reaction mix was incubated at 15$^{o}$C for 2 hours and the labelled DNA separated from unincorporated radionucleotides by column chromatography.

**2) Preparation of radiolabelled lambda DNA**

HindIII-digested lambda DNA was radioactively labelled for use as size markers. 1ug lambda HindIII DNA was incubated with 10ul $^{32}$P-dCTP (10mCi/ml), 5ul Amersham nick translation buffer, and 1ul T4 polymerase in a reaction mix of total volume 50ul for 5 minutes at 37$^{o}$C . A further 25 minute incubation at 37$^{o}$C was performed on addition of 1ul 100mM dCTP. The reaction mix was phenol extracted, dialysed and radioactivity of the markers determined. An aliquot of radiolabelled lamda DNA corresponding to an activity of $5x10^3$ cpm was added to the marker lane of a gel intended for Southern blotting.

**3) Gel filtration of radiolabelled DNA or oligonucleotides**

Nick-translated restriction fragments or $\gamma$ -$^{32}$P-ATP labelled oligonucleotides were purified by passage through a Sephadex G50 column or a commercially available nick-column. Restriction fragments and oligonucleotides of a length greater than 20 bases were purified using pre-packed Pharmacia 'nick columns'. The buffer supplied with the column was discarded and three 1ml aliquots of 1xTE applied. The radio-labelled sample was added once the 3ml of buffer had completely entered the gel bed and 400 ul of 1xTE added. The leading peak was

eluted by adding a further 400 ul of 1xTE to the column and collecting the elutant. Alternatively, a polypropylene column packed with pre-equilibrated Sephadex G50 was equilibrated with STE buffer (Table 1) . Immediately prior to loading the labelling reaction product mix, buffer was allowed to completely enter the gel phase. The reaction products mixed with 1/20th volume of 1% Orange G were layered on to the top of the column. Once the sample solution had entered the gel bed continous addition of STE buffer to the column was resumed. As transport of products down the column occurred two peaks of radioactivity could generally be distinguished. The leading peak corresponded to radionucleotide incorporated into either DNA or oligonucleotide while the lagging, Orange G tagged, peak corresponded to unincorporated radionucleotides. Purified labelled product was collected as the leading peak was eluted from the column. Care was taken to avoid collecting eluant at the trailing end of the leading peak since it was likely to overlap with the elution profile of the unincorporated radionucleotides.

**4) Transfer of DNA to membrane**

Transfer of DNA from an agarose gel to a nitrocellulose membrane was performed according to the technique of Southern (1975). Restriction digests of recombinant constructs containing EHV-4 DNA of interest were electrophoresed through 0.7-1.0% agarose gels . Radiolabelled lambda HindIII DNA was added to a lane as size markers. Electrophoresis was carried out until the bromophenol blue dye had run 75% of the length of the gel. The gel was then stained with ethidium bromide, destained, and photographed. The gel was trimmed to remove agarose at the top, bottom, and edges of the gel outwith the region of DNA separation. The gel was then prepared for DNA transfer by soaking in denaturation buffer (Table 1) for two 15 minute periods, in

neutralisation buffer (Table 1) for two 15 minute periods, and in 1xGeneScreen buffer (GS) (Table 1) for three 20 minute periods. Meanwhile, GeneScreen membranes were prepared by cutting a piece of membrane to the same size as the gel and by soaking this for 10 minutes in 1xGS. Transfer of DNA to GeneScreen was achieved by placing the gel on top of wet 3MM paper suspended in 1xGS, arranging the GeneScreen membrane on top of the gel and positioning successive layers of dry 3MM paper and tissue above the membrane. This created a capillary effect drawing 1xGS from the buffer reservoir through the layers of tissues, transferring gel DNA to the membrane in the process. On completion of transfer (8-12 hours) the membrane was rinsed in 1xGS and baked at 80$^{o}$C for 2 hours.

**5) Prehybridisation**

Prior to prehybridisation the baked membrane was soaked in 1% triton to ensure even distribution of the prehybridisation mix. The membrane was placed in a polythene bag sealed on three sides and prehybridisation mix added. The constitution of this mix was dependent on the stringency of hybridisation required, a parameter affected by formamide concentration. Most commonly a 45% prehybridisation mix was utilised composed of 45% formamide, 4.2xSSC, 4.2xDenhardts solution , 8% dextran sulphate, 0.04M NaPPi, 80ug/ml salmon sperm DNA, and 0.08% SDS (Table 1). The bag was sealed and placed in a shaking water bath at 42$^{o}$C overnight.

**6) Hybridisation**

Hybridisation was initiated by the addition of $2x10^{7}$ cpm of radiolabelled probe to the prehybridisation mix. The bag was resealed and returned to the shaking water bath for a further overnight incubation at 42$^{o}$C.

**7) Washing of blots**

On completion of the hybridisation procedure the membrane was rinsed, at room temperature (RT), in 2xSSC (Table 1) and washed for three 30 minute periods at 65$^{o}$C in either 1xSSC;0.1% SDS or 2xSSC;0.1% SDS depending on the level of stringency required. The membrane was rinsed in 0.1xSSC at room temperature, dried, sealed in a polythene bag, and exposed to autoradiograph film overnight at -70$^{o}$C in the presence of an intensifying screen.

**8) Autoradiography**

Membranes were autoradiographed at -70$^{o}$C with Amersham MP hyperfilm in the presence of a Dupont lightning fast intensifying screen. Films were developed in an automatic processor.

## 3. DNA Sequence Analysis

**1) Preparation of oligonucleotide sequencing primers**

17 base oligonucleotides for use as sequencing primers were manufactured in an Applied Biosystems oligonucleotide synthesiser on a 0.2 um column with trityl off. Oligonucleotides were detached from the column by flushing the column with 2ml $NH_4OH$ over a 1-2 hour period. The $NH_4OH$ and the eluted oligonucleotide were then heated at 55$^{o}$ C for a minimum of 5 hours to deprotect the oligo. The oligo was then precipitated by the addition of 2 volumes of ethanol and 1/10th volume 5M NaCl at -70$^{o}$C for 30 minutes. The pellet obtained after centrifugation at 13K for ten minutes was dried and the oligonucleotide resuspended at a concentration of 10ng/ul.

**2) Sequencing reactions**

Sequencing reactions were carried out by the Sanger dideoxy technique (Sanger et al., 1977) using Klenow fragment in the earlier part of the project and later with a modified T7 polymerase which was

found to give superior results in terms of quality of reaction and number of base pairs elucidated per reaction. Template DNA in both instances was denatured plasmid DNA. Denaturation of plasmid DNA was achieved by the addition of 4ul 1M NaOH to DNA resuspended in 16ul $dH_2O$. In the case of a 6kbp plasmid, 3ug DNA were denatured; the DNA amount was adjusted for larger and smaller constructs. After 5 minutes at RT the mix was neutralised by the addition of 2ul $2MNH_4OAc$ (pH5.3) and DNA precipitated at -70°C on addition of 55ul ethanol. DNA was pelleted by spinning at 13K for 10 minutes, dried, and resuspended in 6.5ul for Sequenase-catalysed reactions or in 9.5ul for Klenow-catalysed reactions.

Klenow-catalysed sequencing was performed as follows. Annealing of primer to denatured template was achieved by incubating resuspended DNA with 2ul primer (10ng/ul), 1.5ul annealing buffer (100mM Tris;50mM $MgCl_2$), and 2ul $\alpha$-$^{35}$S-dATP for 15 minutes at 37°C. 3.5ul aliquots of the annealing mix were transferred to eppendorfs containing 2ul of A, C, G, or T nucleotide mix (Table 2) and incubated as before. A final incubation , again at 37°C for 15 minutes, was performed following the addition of 1ul chase solution (Table 2) to each reaction mix. Enzyme activity was terminated by the addition of 5ul stop mix to each tube . Reactions were stored at -20°C prior to electrophoresis.

T7 polymerase catalysed sequencing was performed using solutions supplied by Pharmacia. 1.5ul of an appropriate sequencing primer (10ng/ul) and 2ul of kit buffer (final concentration :40mM Tris-HCl pH7.5; 20mM $MgCl_2$; 50mM NaCl ) were added to the resuspended template and the mix incubated for 15 minutes at 37°C to permit annealing of primer to template DNA. The extension and labelling reactions were carried out for 5 minutes at RT on addition of 1ul 0.1M DTT, 2ul dGTP mix (1,5uM dGTP,dCTP,dTTP), 0.5ul $^{35}$SdATP, and 2ul of T7 polymerase

diluted 1:8 to 10ul of the annealed primer/template mix . The reaction was terminated by the addition of 3.5ul of labelling reaction mix to each of four termination mixes (80uM dGTP, dATP, dCTP, dTTP plus 8uM ddATP,50mM NaCl (mix A only), 8uM ddCTP, 50mM NaCl (mix C only), 8uM ddGTP, 50mM NaCl (mix G only), or 8uM ddTTP, 50mM NaCl (mix T only)) prewarmed for 5 minutes at 37$^{o}$C. Reaction mixes were incubated at 37$^{o}$C for 5 minutes. 4ul of stop solution was added to terminate all enzyme activity and mixes stored at -20$^{o}$C until required.

**3) Electrophoresis**

Sequencing reaction products were electophoresed through a 6% acrylamide wedge gel prepared and run on an LKB-Macrophor apparatus. To ensure that the gel bound only to the glass plate and not to the thermostatic back plate plates were treated with binding or repelling agents as described in LKB literature.

The acrylamide gel solution was prepared by the addition of 0.4ml APS and 40ul TEMED to a 6% acrylamide gel preparation. The gel was poured according to LKB literature and left for one hour to set. The prepared gel was positioned in the electrophoretic chamber with the thermostatic plate connected to a pump circulating water at a temperature of 62$^{o}$C, and the buffer compartments filled with 1xTBE. The gel comb was removed under buffer and the wells immediately flushed out to remove any unpolymerised acrylamide solution. The gel was electrophoresed for 30-60 minutes at 2000V prior to addition of samples.

Sequencing reaction products were placed in a boiling water bath for 3 minutes. The gel wells were flushed out again and the samples loaded using capillary tubes. Electrophoresis was performed at 2000V for approximately 2 hours for a short run or 5 hours for an extended run. On completion of electrophoresis the plate assembly was removed from

the electrophoretic chamber and the notched plate prised apart from the thermostatic plate. The notched plate plus gel were then transferred to a 10% acetic acid/10% methanol solution and soaked for 30 minutes . The gel was dried and exposed to Amersham MP film overnight in a light tight box.

## Table 2.1
## General Stock Solutions

L-Broth
20g Tryptone
20g NaCl
10g Yeast extract
to 2l with $dH_2O$
Adjusted to approx. pH 7.0 with NaOH. Autoclaved.

100xTE
121g Tris
37g EDTA
to 1l with $dH_2O$

10xTBE
216g Tris
110g Boric acid
200ml 0.25M EDTA
to 2l with $dH_2O$

STE
5ml 1M Tris
10ml 5M NaCl
2.5ml 0.2M EDTA
to 500ml with $dH_2O$

1xSTET
5ml Triton X100
25ml 0.25M EDTA pH8
5ml 1M Tris pH8
8g Sucrose
to 100ml with $dH_2O$

5x bromophenol blue loading buffer
0.125g Bromophenol blue
12.5g Ficoll 400
to 100ml with $dH_2O$

IPTG

50mg/ml in $dH_2O$

stored at -20°C

X-Gal

80mg/ml in dimethylformamide

stored at -20°C

Denaturation buffer

60g NaOH

263g NaCl

to 3l with $dH_2O$

Neutralisation buffer

500ml 1M Tris pH8

175g NaCl

to 1l with $dH_2O$

40xGS

356g $Na_2HPO_4$

312g $NaH_2PO_4$

20xSSC

438g NaCl

220g Tri-sodium citrate

to 2.5l with $dH_2O$

45% Prehybridisation mix

13.5ml Formamide

6.25ml 20xSSC

2.5ml 50xDenhardts

5.0ml 50% Dextran sulphate

1.25ml 40xGS

1.2ml $dH_2O$

0.25ml Salmon sperm DNA (10mg/ml)

0.125ml SDS

Denhardts Solution

5g Ficoll

5g Polyvinylpyrrolidone

5g BSA (fraction V)

to 500ml with $dH_2O$

stored at $-20^{o}C$

Table 2.1 Stock solutions used in molecular biological techniques applied throughout the project.

Table 2.2

DNA Sequencing Stock Solutions

Klenow Sequencing Mixes

A mix - 20ul 0.5mM dCTP, dGTP, dTTP; 20ul 1xTE; 80ul 0.1mM ddATP

C mix - 20ul 0.5mM dGTP, dTTP; 1ul 0.5mM dCTP; 20ul 1xTE; 61ul 0.2mM ddCTP

G mix - 20ul 0.5mM dCTP, dTTP; 1ul 0.5mM dGTP; 20ul 1xTE; 61ul 0.1mM ddGTP

T mix - 20ul 0.5mM dCTP, dGTP; 1ul 0.5mM dTTP; 20ul 1xTE; 61ul 0.5mM ddTTP

Chase Solution

0.5mM dATP, dCTP, dGTP, dTTP

Stop Solution

95% Formamide

20mM EDTA

0.05% Bromopenol blue

0.05% Xylene cyanol FF

6% Acrylamide Gel Mix

25,2g Urea

8.8ml 38% Acrylamide:2% Bisacrylamide

6ml 10xTBE

24.4 ml $dH_2O$

Table 2.2 Stock solutions used in DNA sequencing procedures

# CHAPTER 3

# LOCALISATION AND SEQUENCING OF THE EHV-4 THYMIDINE KINASE GENE AND UL24 GENE HOMOLOGUE

## INTRODUCTION

The cellular enzyme thymidine kinase (ATP:thymidine 5'-phosphotransferase; EC 2.7.1.21; TK) operates within the salvage pathway of pyrimidine biosynthesis catalysing the conversion of thymidine to thymidylate (dTMP) using ATP as phosphate donor (Kit, 1985a) (Fig. 1). During lytic infection many herpesviruses induce a virus-specific thymidine , or more accurately deoxypyrimidine, kinase which acts in conjunction with the cellular TK to increase the efficiency of deoxythymidine salvage thereby increasing the dTTP pool available for viral DNA synthesis. Herpesviruses with TK-coding capacity have been identified in all three subgroups, alpha, beta, and gamma, through the detection and characterisation of i) virus-specified TK activity in infected cells and/or ii) potential TK-coding genes in the viral genome (reviewed in Kit, 1985a)(Table 1). The betaherpesvirus HCMV is the only herpesvirus sequenced in entirety which does not seem to encode a TK function although it does stimulate the host cell TK during productive infection (Estes and Huang, 1977). Until recently the only evidence that TK was specified in other betaherpesviruses was provided by the existence of a bovine cytomegalovirus (BHV-4)- specific TK (Kit , 1985a ) . The current view is that this herpesvirus should be reclassified as a gammaherpesvirus on the basis of genome structure and composition and its gammaherpesvirus-type TK activity (Bublot et al., 1990; Kit et al., 1986). However, the betaherpesvirus EHV-2 does seem to encode a TK function so presumably TK-coding capacity was deleted from the HCMV genome relatively late in the cytomegalovirus evolutionary scheme (Colacino et al., 1988; Staschke et al., 1990).

Herpesvirus TKs differ from cellular TKs and from each other in their substrate and phosphate donor specificities. Whereas cellular TKs use only ATP and dATP as phosphate donor, some herpesvirus TKs can

utilise CTP and GTP in addition. Similarly, while cellular TKs efficiently utilise only thymidine as a substrate herpesvirus TKs are less fastidious and can use deoxycytidine (Kit, 1985a). Herpesvirus TKs can be divided into three groups on the basis of their substrate specificity - a broad specificity group including HSV-1, HSV-2, MarHV, EHV-1, and VZV, an intermediate group including PRV, BHV-1, and EHV-3, and a narrow specificity group including BHV-4 and HVS which inefficiently phosphorylate deoxycytidine (Kit , 1985a) (Table 1).

These ranges of specificity extend to the affinity of different TKs for nucleoside analogues such as acyclovir (ACV), which, in a phosphorylated form, inhibit the DNA replicative cycle of specific herpesviruses (Collins, 1983; Elion, 1983; Kit et al., 1987a,b)(Table 1). This has led to the development of antiviral therapies based on the application of nucleoside analogues to infected hosts. Preferential uptake of the drug by herpesvirus-infected cells and the inability of cellular TK to phosphorylate such analogues protects uninfected cells from the harmful effects of the activated analogue. The effectiveness of the phosphorylated analogue in inhibiting viral replication and killing infected cells is subject to one or more of the following occurring - i) competition with dTTP for incorporation into DNA (Coen and Schaffer, 1980) , ii) inhibition of the virus-specified DNA polymerase (Furman et al., 1984), iii) incorporation into replicating DNA such that its stability and/or function are affected (Mancini et al., 1983), iv) induction of unbalanced nucleotide pools through allosteric effects on ribonucleotide reductase and/or dCMP aminase or, v) inhibition of viral replication or infectivity through interference with envelope glycoprotein synthesis (Olofsson et al., 1985).

The ultimate target of most antiherpetic drugs is the virus specified DNA polymerase - the role of TK is limited to the activation

of the analogue. Thus the effectiveness of antivirals depends not only on the substrate specificity of the herpesvirus TK but also on the substrate specificity of the DNA polymerase, the specifity of cellular enzymes required to convert the mono-phosphorylated analogue to its di- and tri-phosphorylated form, and on the requirement that the analogue should be effective at a dose level which is not toxic to uninfected cells *in vivo*. An experimentally useful outcome of the broader specificity of herpesvirus TKs is that it enables *in vitro* selection of $TK^+$ or $TK^-$ viruses from a mixed background in either $TK^-$ or $TK^+$ host cells depending on the virus type and the analogue used.

The HSV-1, HSV-2 and VZV TKs are multifunctional in that they have a thymidylate kinase activity in addition to TK activity phosphorylating thymidine monophosphate to thymidine diphosphate ( Chen and Prusoff, 1978; Fyfe, 1982). In HSV-1 these two catalytic activities are thought to share a common or at least overlapping site within the enzyme (Chen *et al.*, 1979).

A functional TK gene is not essential for viral growth ***in vitro*** with the exception of cells cultured in serum-starved conditions (Jamieson *et al.*, 1974). This implies that herpesvirus TK activity may play a particularly important role in the infection of resting or terminally differentiated cells *in vivo* such as neural cells - the site of latency for neurotropic herpesviruses. However, TK gene function also seems to be important in the infection of dividing cells *in vivo* in that it has been linked to herpesvirus virulence since $TK^-$ mutants are often attenuated *in vivo* (Kit *et al.*, 1983b, 1985a,b). A link between TK activity and neurovirulence has been established for HSV-1 TK in a mouse model (Field and Wildy, 1978). Although TK gene expression has been regarded as a requirement for viral latency (Tenser and Dunstan, 1979), TK-deficient mutants capable of

establishing a latent infection in animal models have recently been isolated (Coen et al., 1989; Efstathiou et al., 1989; Meigner et al., 1988; Tenser et al, 1989 ). It seems that the extent to which decreased TK expression affects latency may depend not only on the genetic background of the mutant virus but also on the animal model used (Coen et al., 1989).

The mapping of genes encoding herpesvirus TKs has been approached from several experimental angles including i) amplification of TK DNA by polymerase chain reaction using primers derived from conserved sequence regions of characterised TK genes (Nunberg et al., 1989) , ii) analysis of intertypic HSV-1/HSV-2 recombinants (Halliburton et al., 1980), and iii) marker rescue of the TK function (Bello et al., 1987; Stow et al., 1978). The most common approach has been the transfection of fragments of herpesvirus DNA into $TK^-$ cells followed by the analysis of the ability of these fragments to biochemically transform transfected cells from a $TK^-$ phenotype to a $TK^+$ phenotype (McDougall et al., 1980; Otsuka et al., 1981; Robertson and Whalley, 1988; Sawyer et al., 1986; Wigler et al., 1977 ). Despite TK expression in lytic infection being dependent on expression of immediate early functions, expression of TK in transfected fragments tends to be constitutive. In HSV-1 for example, the TK gene is regulated as an early gene during productive viral infection, its expression dependent on expression of immediate early gene products, in particular ICP4 and ICP0 (Gelman and Silverstein, 1985). Nonetheless, fragments of HSV DNA containing an intact TK gene are capable of transforming $TK^-$ cells to a $TK^+$ phenotype on transfection with the integrated TK gene expressed in the absence of immediate early HSV gene products. Superinfection with HSV-1 does, however, increase the level of TK transcription and translation 5-20 fold suggesting that

certain regulatory controls are maintained (Kit and Dubbs, 1977 ).

The nucleotide sequences of the TK genes of a variety of herpesviruses have been published including EHV-1 (Robertson and Whalley, 1988), PRV (Kit, 1985b), VZV (Davison and Scott, 1986), BHV-1 (Kit, 1985a; Mittal and Field, 1989), BHV-2 (Sheppard and May, 1989), HSV-1 (McKnight, 1980; Wagner et al., 1981), HSV-2 ( Kit et al., 1983a; Swain and Galloway, 1983), FHV (Nunberg et al., 1989), MarHV (Otsuka and Kit, 1984), MDV (Scott et al., 1989), HVT (Martin et al., 1989; Scott, et al., 1989 ), EBV (Baer et al., 1984), and HVS (Honess et al., 1989a). We utilised transfection analysis in conjunction with Southern analysis to map the EHV-4 TK gene . The nucleotide sequence of the EHV-4 TK gene and flanking genes was then determined and the predicted protein products compared to those encoded by the gene homologues in other herpesvirus (Nicolson et al., 1990b).

## MATERIALS AND METHODS

### 1. Materials

**1) Cells**

A BHK TK$^-$ cell line was kindly provided by the Beatson Institute for Cancer Research, Glasgow.

**2) Tissue culture**

Special liquid medium was obtained from Flow Laboratories and bromodeoxyuridine and DMSO from Sigma. All other media supplements, including HAT supplement, and trypsin-versene were supplied by Gibco-BRL. A CellPhect kit used in transfections was supplied by Pharmacia.

### 2. Methods

**1) Tissue Culture**

BHK TK$^-$ cells were propagated in Special Liquid Medium supplemented with L- glutamine (2mM), penicillin (2000IU/ml), streptomycin (2mg/ml), and 10% foetal calf serum. Bromodeoxyuridine (BrdU), a nucleotide analogue phosphorylated by cellular TK and thus toxic to TK$^+$ cells, was added at 25ug/ml to eliminate spontaneous TK+ revertants in cell culture. This media is referred to as 10% SLM/BrdU. Cells were propagated in 175cm$^2$ flasks in 15ml 10% SLM/BrdU at 37$^o$C. Confluent monolayers were detached from the flask by the addition of trypsin-versene. Cells were pelleted, resuspended in 10% SLM and transferred to fresh 175 cm$^2$ flasks at a density of 5x10$^5$ to 1x10$^6$ cells per flask.

**2) Cell Stocks**

Cells recovered from a confluent 125cm$^2$ flask were pelleted and resuspended in 4ml 10% special liquid medium, 0.5ml foetal calf serum, and 0.5ml DMSO. Cell suspensions were transferred to 2cm plastic vials and stored in the vapour phase of a liquid nitrogen tank for 12-24 hours. Vials were then immersed in liquid nitrogen and stored until

required. Cell lines were recovered by retrieval of appropriate vials from the liquid nitrogen store and rapid thawing of frozen cells at 37$^{o}$C. Thawed cells were added slowly to 8ml 10% SLM, spun, and resuspended in fresh medium.

**3) Southern hybridisation**

0.5 ug aliquots of recombinant plasmid DNAs were digested to completion with appropriate restriction enzymes to release fragments of interest. Reaction products were electrophoresed through 0.75% agarose and transferred to a Gene Screen membrane (New England Nuclear, Dupont) . Membranes were baked for 2 hours at 80$^{o}$C, prehybridised overnight at 42$^{o}$C in 45% formamide, 4.2x SSC, 4.2x Denhardts solution, 8% dextran sulphate, 0.04M NaPPi , 80ug/ml salmon sperm DNA, 0.08% SDS and hybridised for a similar period following the addition of 2x10$^{7}$ cpm of denatured, $^{32}$P-labelled EHV-1 DNA. The blot was washed in 2xSSC,0.1% SDS at 65$^{o}$C, dried and exposed overnight.

**4) Construction of recombinant plasmids**

Recombinant plasmids were constructed by subcloning EHV-4 fragments into Bluescript M13+ by standard techniques ( Maniatis et al., 1982; Chapter 2 ). 'First generation' plasmids pBSBC10 and pBSCS6 were constructed by isolating restriction fragments of C and subcloning these in a Bluescript vector cleaved at appropriate sites within the multicloning site (MCS). A series of second and third generation plasmids pBSSB4, pBSSS4, pBSRX3, pBSRS3, and pBSMS2 were derived by digestion of parental plasmids with an appropriate enzyme followed by religation as outlined in Figure 2.

TK-deleted plasmids were constructed by deletion of a central portion of cloned EHV-4 DNA by restriction endonuclease digestion. Cleavage overhangs were filled in using T4 pol ( for 3' overhangs ) or Klenow enzyme (for 5' overhangs) and plasmids religated at the site of

deletion (Fig. 3).

**5) Transfections**

Recombinant plasmids were transfected into monolayer BHK $TK^-$ cells by a modification of the technique of Graham and van der Eb (1973). An appropriate amount of plasmid DNA (0.2-20 ug) was prepared in 120 ul of sterile water and mixed with Buffer A (0.5M $CaCl_2$, 0.1M HEPES ) of a CellPhect transfection kit . After a 10 minute incubation at room temperature, 120ul of Buffer B (0.28M NaCl, 0.05M HEPES, 0.75mM $NaH_2PO_4$, 0.75mM $Na_2HPO_4$) was added while vortexing the solution. The precipitate was left at room temperature for 15 minutes prior to its addition to 3ml 10%SLM covering a semi-confluent monolayer culture of BHK $TK^-$ cells (seeded at $10^4$ cells/$cm^2$ 24 hours earlier). Following a 20-24 hour incubation at $37^oC$, the medium was decanted from the flasks and the cell layer washed twice with unsupplemented special liquid medium. Cells were incubated in 10% SLM for a further 24 hours and thereafter in HAT-supplemented medium ( Hypoxanthine $10^{-4}M$, Aminopterin 4 x $10^{-5}M$, Thymidine 1.6 x $10^{-5}M$ in SLM). Flasks were screened for colonies derived from cells surviving selection within 2 weeks of transfection.

**6) Exo III Deletions**

Exonuclease III deletion of plasmids pBSRX2 and pBSRS3 was performed in order to generate deletions in EHV-4 DNA proximal to one of the BSM13+ sequencing primer sites. A commercially available sequencing primer could then be utilised to read the sequence of several deleted inserts thus generating more sequence data per primer than the alternative approach involving synthesis of custom primers every 200-300bp into a sequence. The protocol used was a modified version of that described by Henikoff (1984).

25ug of plasmid DNA was digested with two carefully selected enzymes

cleaving within the BS multicloning site but not within insert EHV-4 DNA, one enzyme generating a 5′ overhang proximal to target DNA and the other a 3′ overhang at a site close to the first enzyme but nearer the BS T7 or T3 promoter. Plasmid pBSRS3 was digested with SstI and BamHI to produce deletions from the EcoRV terminus of the EHV-4 insert and plasmid pBSRX2 with KpnI and XhoI to produce deletions from the XhoI terminus of the insert (Fig. 4). In theory, Exo III should delete from the 5′ overhang into insert DNA but should be prevented from deleting into plasmid DNA, thus destroying essential sequencing primer sites, by the 3′ overhang. Linearised plasmid DNA was phenol extracted, chloroform extracted, precipitated, and dried. The pellet was resuspended in 120ul ExoIII buffer (66mM Tris-HCl (pH8); 0.66mM $MgCl_2$). A 30ul aliquot was incubated with 200 units of ExoIII at 32$^{o}$C and 2.5ul aliquots removed into 7.5ul S1 mix (0.04M KOAc; 0.07M NaCl; 7% glycerol; 0.02M $ZnSO_4$; 60 units S1 nuclease ) at 2 minute intervals. The mixes were incubated at room temperature for 30 minutes, mixed with 1ul S1 stop (0.3M Tris-OH; 0.05M EDTA), and incubated 10 minutes at 70$^{o}$C. Mixes were cooled to 37$^{o}$C and 1ul Klenow mix (0.02M Tris-HCl; 0.09M $MgCl_2$; 3 units Klenow fragment ) added. Following a 5 minute incubation 1ul of dNTPs (0.125mM dATP, dCTP, dGTP, dTTP) was added to promote a filling-in reaction at each of the plasmid termini. 40ul ligation mix (1/8th volume 10xligase buffer ; 25U T4 DNA ligase) was added to each aliquot and the reaction mix incubated overnight at room temperature. Competent E. Coli JM101 or JM109 were transformed with the products of selected ExoIII incubation time points and screened for inserts with suitable size deletions. Selected recombinant inserts were then sequenced using a primer proximal to the restriction sites used in the deletion process.

**7) Sequencing**

Sequencing of RX2 and flanking DNA was performed using single stranded pBSRX2 DNA and pBSRX2-derived Exo III-deleted plasmids as template and Bluescript-derived and custom-made oligonucleotides as primers in a Sanger dideoxy sequencing strategy ( Sanger et al., 1977)(Fig. 7b).

# RESULTS

## 1. Localisation of the EHV-4 TK gene

### 1) Southern hybridisation

The EHV-4 genome is a double stranded DNA molecule of 144 kbp, composed of a long (L) and short component (S), the latter bounded by a set of inverted repeats. Given the HSV-1 TK gene map position of approximately 0.3 (P isomer)(Holland et al., 1984b) and evidence for colinearity of the $I_L$ isomer with EHV-1 (Davison and Wilkie, 1983), and of EHV-1 with EHV-4 (Cullinane et al., 1988), the EHV-4 TK gene was expected to map at around 0.5 map units.

In a series of Southern hybridisations EHV-4 BamHl fragments mapping between 0.39 and 0.62 map units, BamHI Q, M, C, and F, were hybridised initially to an HSV-1 TK DNA probe - a 2kbp PvuII fragment of BamHI P (strain CL101). No hybridisation was detected at the stringency used (2xSSC) presumably due to a lack of sufficient DNA homology between HSV-1 and EHV-4 across this region of the genome (data unshown). At that time the EHV-1 TK gene had been mapped to a 5.35 kbp BamHI/ClaI fragment mapping at the left end of EHV-1 BamHI B (Fig. 5a). This fragment, later found to include the entire TK gene and flanking sequence (Robertson and Whalley, 1988), was used as a hybridisation probe. Hybridisation to EHV-4 DNA was confined to the BamHI C fragment mapping between 0.43 and 0.53 . Southern analysis of subfragments of C indicated the region of crosshybridisation mapped to the right hand end of C: the probe hybridised to a 10kbp Cla1/BamHl fragment (BC10) and to three Sma1 subfragments mapping between 0.48 and 0.53 (Fig. 6). Since the precise position of the EHV-1 TK gene within the probe fragment was unknown at this time, more accurate mapping of the EHV-4 TK gene by Southern analysis was not possible.

**2) Transfection Analyses**

In order to confirm correct mapping of the EHV-4 TK gene and to determine more sensitively its position within BamHI C prior to sequence analysis a crude transfection assay was employed. Fragments of EHV-4 DNA encoding a functional TK gene should, on transfection, be capable of inducing biochemical transformation of $TK^-$ cells to a $TK^+$ phenotype. Transfection of 8ug of recombinant plasmid pUC 9:C of the EHV-4 BamHI library resulted in the transformation of BHK TK- cells to a TK+ phenotype confirming the presence of a functional TK gene within EHV-4 C (Table 2). As positive controls plasmids containing the HSV-1 and EHV-1 TK genes, pTK (HSV-1 CL101 BamHI P in pAT153) and pBSBC5 (EHV-1 HVS25 5.35kbp BamHI/ClaI fragment of BamHI B in BSM13+), were transfected into cells under identical conditions. Both plasmids exhibited a significantly greater transformation efficiency than pUC9:C. The negative control BamHI A of EHV-1 in pBR322 induced no $TK^-$ to $TK^+$ transformation. In order to more accurately map the gene the ability of subfragments of EHV-4 BamHI C to biochemically transform $TK^-$ cells was investigated.

A partial restriction map of BamHI C was determined by restriction analysis of C (Table 3). Suitable restriction sites for subcloning of the hybridising region of C were identified and a family of recombinant Bluescript plasmids constructed containing EHV-4 fragments of 2kbp to 10kbp in size (Table 3; Fig. 2).

Transfection of plasmids pBSBC10, pBSCS6, pBSSS4, pBSRS3, and pBSRX2 into $TK^-$ cells resulted in biochemical transformation of a small proportion of the transfected cell population (Tables 3,4,5). This mapped the TK gene to a 2kbp EcoRV/XhoI fragment, RX2, of approximate genome map position 0.48 (Table 3).

Plasmid pBSSB4, which contains EHV-4 DNA mapping outwith the RX2

region did not induce biochemical transformation of cells. One $TK^+$ colony was detected in pBSMS2-transfected cultures. However, the transformed $TK^+$ colony was not tested for the presence of integrated MS2 DNA in cytoplasmic DNA and it is possible the colony was a spontaneous revertant.

Having localised the EHV-4 TK gene to a 2kbp fragment sequence analysis was performed to establish the precise position and coding potential of the gene.

## 2. Sequence Analysis of EHV-4 C Fragment RX2 and Flanking DNA

The nucleotide sequence of RX2 and adjoining DNA was determined according to the strategy outlined in Fig. 7 and is detailed in Fig. 8. A typical autoradiograph of a sequence gel is shown in Figure 9. The position of three major open reading frames (ORF), two complete and one partial, are indicated in a VAX analysis of potential start and stop codons within all six frames (Fig. 7c). One ORF was mapped to the lower strand (ORF1: bp 934-119) and two to the upper strand (ORF2:bp 949-2004 and ORF3: 2225-2400) (Figs. 7,8).

ORF1 encodes a 272 amino acid polypeptide with 36% identity to the predicted products of genes UL24 of HSV-1 (McGeoch et al., 1988a) and 35 of VZV (Davison and Scott, 1986). A search for potential regulatory elements immediately upstream of ORF1 revealed no elements of identical sequence to the consensus TTATA (Corden et al., 1980 ) element. Potential TTATA boxes map at bp 1034-1029 (TAAATT) and bp 1166-1171 (TAATTA) . Two potential polyadenylation signals are located at the 3′ end of the gene - AATAAA which spans the termination codon TAA, and ATTAAA located 10bp downstream of the end of the coding region of the gene (Birnstiel et al., 1985; Proudfoot and Brownlee, 1976). Since the initiation Met codon of ORF1 lies only 15 bp upstream of the

first Met codon of ORF2, the 5′ regulatory and N-terminal coding regions of the ORF2 gene and gene UL24 homologue overlap.

ORF2 is the only intact open reading frame within RX2, the smallest tested fragment with TK-transforming activity. It encodes a 352 amino acid product which shares 36% identity with the HSV-1 TK gene product. TATA box homologues (Corden et al., 1980), TTATA and TATTAA, are located 85 and 305 bp upstream of the initial methionine codon of ORF1 (at 865bp and 645bp in Fig. 8). A potential RNA Polymerase initiation site (Corden et al., 1980) is located 22 bp downstream of the TTATA motif, within the sequence TAC (bp 886-888). The initial Met codon of most viral and eukaryotic genes is the first 'in frame' ATG from the RNA initiation site and resides within the consensus sequence (A/G)CCATGG (Kozak, 1984, 1986), the most strongly conserved feature of which is a purine at position -3. The predicted EHV-4 TK initiation codon is positioned at bp 949-951 within the sequence GTAATGG. The coding region of the gene commences at bp 949 and extends to a TGA termination codon at bp 2005 .

The G/C content of the coding region of the EHV-4 TK gene is 52% as compared to 65% in HSV-1 (Wagner et al., 1981 ). This difference in G/C contents is reflected in the distribution of codon usage within the two genes: G/C frequency in the third codon position of amino acids specified by any third position base is significantly higher in the HSV-1 TK gene (80%) than in the EHV-4 TK (45%) or UL24 (44%) genes and may represent differential adaptation to the overall G/C contents of the two viral genomes (Tables 6,7). A Poly A signal, AATAAA (Proudfoot and Brownlee, 1976), is positioned 42 base pairs downstream of the termination codon.

ORF3 extends for 2565 bp and encodes the EHV-4 gH gene (Chapter 4). A 5 amino acid sequence encoded upstream of UL24 with 100% identity to the opening 5 amino acids of the EHV-1 UL25 homologue may correspond to the N-terminal coding region of an EHV-4 UL25 homologue (Fig.7).

## 3. Transfection of ExoIII-deleted RX2 and RS3 DNA into TK- cells and Generation of TK- plasmid constructs.

A series of plasmids derived by deletion of pBSRX2 or pBSRS3 were analysed for their TK-transformation activity by DNA mediated transfer into $TK^-$ cells. Deletion mutants investigated included pBSD1, pBSD2, pBSD3, pBSD4, pBSD5, and pBSMS2 the EHV-4 inserts of which extend from bp 257 ,515, 703, 942, 983, and 966 respectively through to the SalI site, and pBSD6 which comprises bp 181-1714 (Fig. 10) .

A further two deleted constructs were created by the production of SmaI-BstEII and SmaI-BstXI internal deletions within fragments RX2 or RS3 of EHV-4 cloned into Bluescript M13+ or pUC9 (Fig. 3).

The results of transfections with these constructs are outlined in Table 5. Constructs pBSRX2, pBSD1, pBSD2, pBSD3,and pBSD4 efficiently transformed cells in comparison to pBSMS2 , pBSD5, pBSME, and pPRSMB . Several points can be inferred from these data with the caveat that a limited number of transfections were performed . The first is that the internal SmaI-BstEII and SmaI-BstXI deletions apparently render the DNA incapable of TK-transformation. Presumably a TK transcript is inititiated at the intact Met codon but is truncated . Since the deletion spans the sites essential for nucleotide binding and nucleoside binding a functional TK would not be expressed as evidenced by the transfection results. Other points can be made with regard to the possible location of the promoter of the EHV-4 TK gene. Comparison of the number of HAT-resistant colonies produced on transfection with plasmids pBSD2, pBSD3, pBSMS2, and pBSD5 to that of undeleted pBSRX2 indicated that pBSD2 and pBSD3 efficiently expressed TK in transfected cells suggesting that the TK promoter was intact in these constructs and thus maps downstream of bp 703 - the 5′ limit of D3 (Fig. 10) . Of the two potential TATA boxes with a good fit to the consensus TATA

sequence, only T2 at bp 861-865 lies downstream of bp703 (Fig. 8). I therefore tentatively suggest that the EHV-4 TK promoter TATA box spans bp 861-865. Two GC rich boxes located approximately 50 and 105 bp upstream of this TATA box may be the EHV-4 equivalent of the GC rich boxes which are intrinsic components of the distal promoter elements of the HSV-1 TK gene (Jones et al., 1985). GC boxes have also been located within the upstream region of the HSV-2 (Swain and Galloway, 1983) and marmoset herpesvirus (Otsuka and Kit, 1984) TK genes . The CAAT box of the HSV-1 promoter is located asymetrically between the two GC boxes in an inverted orientation (McKnight and Kingsbury, 1982). Putative EHV-4 CAAT boxes exist in both orientations within the DNA sequence between the GC boxes : CCAATT at 781-786 and TGAATT at 788-783.

## 4. Analysis of the Predicted Protein Products of ORF1 and ORF2

### 1) EHV-4 UL24 gene homologue protein product

Gene homologues of HSV-1 gene UL24 (McGeoch et al., 1988a) have been identified in HSV-2 (Kit et al., 1983a; Swain and Galloway, 1983), EHV-1 (Robertson and Whalley, 1988), MDV (Scott, 1989), HVT (Jacobson et al., 1989b), HCMV (Jacobson et al., 1989b), EBV (Baer et al., 1984), and HVS (Jacobson et al., 1989b). The degree of identity of the products of these genes to the predicted EHV-4 UL24 gene homologue product is detailed in Table 8. Identities were of the order of 20% with beta- and gammaherpesvirus UL24-encoded proteins and of 30 to 40% with non-equine alphaherpesvirus UL24-encoded proteins. The EHV-4 and EHV-1 proteins exhibited 83% amino acid identity.

Alignment of these UL24 gene sequences indicates five regions of conserved sequence within the N-terminal and central portion of the protein (Fig. 11). 100% conservation of cysteine residues at two positions (II and III) suggests these might be involved in disulphide

bond formation. The functional significance of these domains has yet to be determined. As yet all that is known of the UL24 gene product function is its importance in HSV-1 for viral growth in vitro (Jacobson et al., 1989b) : amino acid changes , particularly within the more strongly conserved domains of the protein, often result in a virus mutant with a small plaque phenotype. The predicted EHV-4 UL24 product exhibits a particularly high degree of conservation with the EHV-1 UL24 product across the N-terminal two thirds of the polypeptide, to the level of 94% . In contrast, identity within the C-terminal third is only 53%. This suggests that the N-terminal domain may be constrained in its sequence by functional requirements. The extreme N-terminal sequence may incur additional constraints due to the overlap of its coding sequence with the TK gene promoter region or, as in HSV-1, TK coding sequence.

**2) EHV-4 TK gene protein product**

The translated product of the EHV-4 TK gene is a 352 amino acid protein of predicted molecular weight 38800. The TK of EHV-1 has been identified as a dimer of approximate molecular weight 80 000 (Allen et al., 1979). Comparisons of its predicted amino acid sequence to that of other herpesvirus TKs indicated identities of 16 and 18% with gammaherpesvirus TKs and 29-47% with non-equine alphaherpesviruses. The degree of identity of the EHV-4 and EHV-1 TKs was 89% (Table 8).

A multiple alignment of all herpesvirus TK sequences reported to date and of the EHV-4 TK is presented in Figure 12. A particularly divergent region within herpesvirus TKs corresponds to the extreme N-terminal portion of the protein as defined as the part of the protein N-terminal to the first region of strong conservation between all herpesvirus TKs. In the case of the EBV and HVS TKs which, at 607 and 527 amino acids in length are considerably larger than the other

enzymes , the bulk of the extra amino acids are located N-terminal to this conserved site. The corresponding region of the HSV-1 TK polypeptide (amino acids 1-45) is non-essential for thymidine phosphorylating activity (Halpern and Smiley, 1984; Haarr and Flatmark, 1987) although it affects enzyme stability.

Multiple alignments (reviewed in Honess et al. (1989a), Mittal and Field (1989), Robertson and Whalley (1988), Scott et al. (1989)) have highlighted several regions of conserved sequence. Analysis of the degree of identity of the EHV-4 TK with other herpesvirus TKs using limits defined by the position of the most N-terminal and C-terminal conserved regions indicated higher identities. This was particularly noticeable in the case of the two gammaherpesvirus TKs since these limits eliminate inclusion of the additional N-terminal portion of these polypeptides in the comparison (Table 8).

The limits of the conserved regions of thymidine kinases quoted in the literature varies since different groups use different criteria for conservation. For the purposes of this thesis six conserved regions are designated A-F in Figure 12. The most N-terminal of these conserved regions , A, shares identity with a sequence domain within many nucleotide binding proteins and has been proposed to form part of a nucleotide binding site (Gentry, 1985; Otsuka and Kit, 1984) (Fig. 13a). Residues which are critical to HSV-1 nucleotide binding activity and which seem to be particularly well conserved between herpesvirus TKs correspond to $Gly_{56}$, $Gly_{59}$, $Gly_{61}$, $Lys_{62}$ and $Thr_{63}$ (Liu and Summers, 1988) . Substitution of valine for glycine at positions 56, 59, and 61 and substitution of isoleucine for lysine 62 abolishes enzyme activity in mutant HSV-1 TKs. Substitution of $Thr_{63}$ with alanine dramatically reduces enzyme activity. However, substitution with a serine residue results in a functional enzyme although its

kinetic properties differ from those of the wild type enzyme (Liu and Summers, 1988) (Table 9). This suggests that a hydroxyl group might be involved in the transfer of the phosphoryl group of the phosphate donor. The presence of only GXXGXGKT and GXXGXGKS motifs within the nucleotide binding domain of the alphaherpesviruses is consistent with the apparent restriction on residues which can be tolerated at positions equivalent to HSV-1 amino acid 63 within a functional herpesvirus TK. The predicted EHV-4 TK nucleotide binding domain is of the form GXXGXGKS in common with those of the EHV-1, PRV, FHV and MarHV TKs. The glycine residues lie within a flexible part of the polypeptide which changes conformation on binding ATP to allow a stretch of alpha helix following the GXXGXGK(S/T) motif to enfold the ATP molecule within a cleft in the enzyme (Fry et al., 1986; Pai et al., 1977). The lysine residue following GXXGXG is conserved and may interact with the alpha phosphate of the phosphate donor in order to neutralise its charge (Pai et al., 1977). In adenylate kinase other basic amino acid residues could not substitute for lysine (Kamps and Sefton, 1986) and it is interesting that lysine is 100% conserved throughout the virus TKs outlined in Fig. 13a . The stretch of hydrophobic residues N-terminal to the GXXGXGK(S/T) motif of region A may, as in adenylate kinase, form a hydrophobic pocket which binds the aromatic ring of the phosphate donor ( Fry et al., 1986).

Region C corresponds to the putative nucleoside binding site (Fig. 13b). This region of the polypeptide is of particular interest in that its amino acid sequence may determine the substrate specificity of the enzyme which may in turn affect the sensitivity of the viral reproductive cycle to nucleotide analogues. Sequence analysis of HSV-1 mutants with altered substrate specificity has indicated that residues $Ala_{168}$ and $Arg_{176}$ play a direct role in substrate binding since

viruses specifying $Thr_{168}$ and $Gln_{176}$ are BVdU and ACV resistant respectively (Darby et al., 1986; Larder et al., 1983) (Table 9). The TKs of EHV-1, EHV-4, BHV-2, MarHV, FHV, HSV-2 and VZV possess a serine at a position corresponding to HSV-1 $Ala_{168}$ and in those viral TKs tested , including HSV-2 and EHV-1, resistance to BVdU- and IVdU-induced inhibition of viral replication tends to be greater than that of wild type HSV-1 which is consistent with the analysis of the HSV-1 BVdU resistant mutant (De Clerq, 1982; Kit et al., 1987a,b). Resistance to the effect of BVdU in these viruses is thought to be a reduced efficiency of binding of monophosphorylated analogue rather than a reduced binding affinity for the nonphosphorylated analogue since the mutant Ala-Thr TK of HSV-1 and the wild type HSV-2 and EHV-1 TKs efficiently catalyse the first phosphorylation step (Darby et al., 1986; De Clerq, 1982; Fyfe et al., 1983; Kit et al., 1987a). This suggests that a change of residue may confer different binding activity for the substrate of the thymidylate kinase-catalysed step without changing affinity of the enzyme for the substrate of the thymidine kinase function.

Other mutations may, however , affect both TK and thymidylate kinase activity, indirectly through tertiary structure alteration, or directly since in HSV-1 at least, the thymidine kinase and thymidylate functions are thought to overlap and may even be specified by the same site. Thus, the DRH motif of the nucleoside binding site may be involved in both nucleoside and nucleotide binding (Robertson and Whalley, 1988) . An involvement in nucleotide binding has been implied due to the observation that the aspartate residue within DRH is the only one conserved throughout the herpesvirus TKs and evidence for a role for aspartate residues in binding nucleotide cosubstrates in a range of enzymes (Dever et al., 1987; Fry et al., 1986) . The DRH Asp residue

may be critical for the binding of GTP as a donor in addition to ATP (Folkers and Trumpp, 1987). The Arg residue of the DRH motif within VZV has been implicated in substrate binding since VZV TK mutants with changes at this residue no longer bind acyclovir (Sawyer et al., 1988).

The only information on the functional activity of region D is that deletion of the valine residue within sequence NLV of BHV-1 TK results in BVdU resistance (Mittal and Field, 1989).

Region E may be involved in movement towards region A in the presence of substrate (Pai et al., 1977) and Arg residues might be important in this interaction since ACV resistance is conferred to HSV-2 by the substitution of $Arg_{223}$ with His (Kit et al., 1987d) (Table 9).

Region F possesses weak similarity to a region towards the C-terminus of adenylate kinase, human thymidine kinase and pox virus TK across the region corresponding to KRL through to CRD in HSV-1 TK. Apart from the nucleotide binding site this is the only part of the non-herpesvirus enzymes to show similarity in primary amino acid sequence to herpesvirus TKs although overall similarities in secondary structure throughout the proteins have been predicted (Folkers and Trumpp, 1987).

The precise role of the other conserved regions of the herpesvirus thymidine kinases with respect to the structural and functional integrity of the protein remains to be determined.

Figure 3.1

Thymidine Kinase-Catalysed Reaction

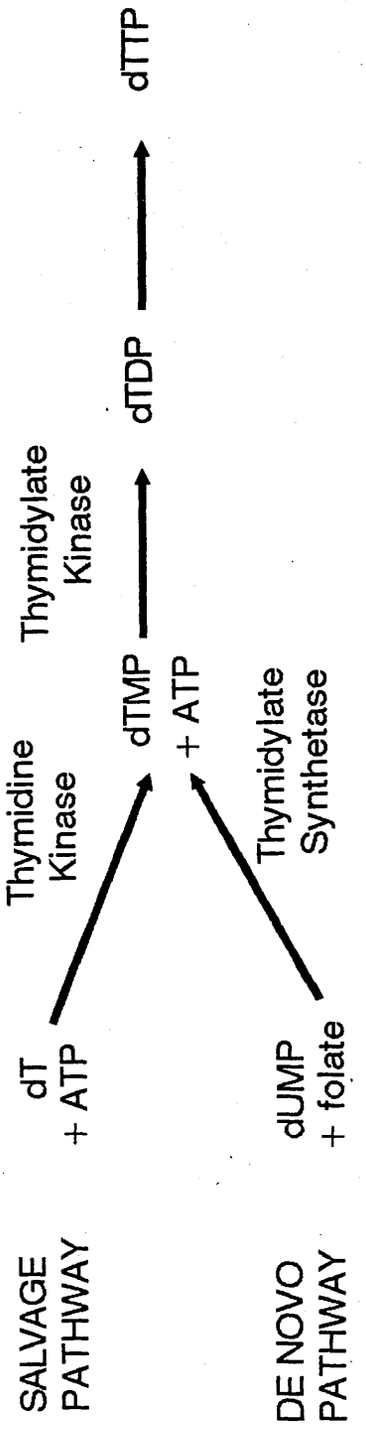



Figure 3.1 Schematic outline of the thymidine kinase and thymidylate kinase activities of multifunctional herpesvirus thymidine kinases (adapted from Robertson and Whalley (1988)).

Figure 3.2 Production of recombinant plasmids used in transfections and as sequencing templates.

a) 'First generation' plasmids pBSBC10 and pBSCS6 were constructed by isolating restriction fragments of C and subcloning these in a Bluescript vector cleaved at appropriate sites within the multicloning site.

b) 'Second and third generation' plasmids pBSSB4, pBSSS4, pBSRX2, pBSRS3,and pBSMS2 were derived from plasmids pBSBC10, pBSCS6, and pBSRS3 by digestion of parental plasmids with an appropriate enzyme followed by religation of the vector at the site of the deletion in EHV-4 DNA.

## Figure 3.2

## Production of Recombinant Constructs

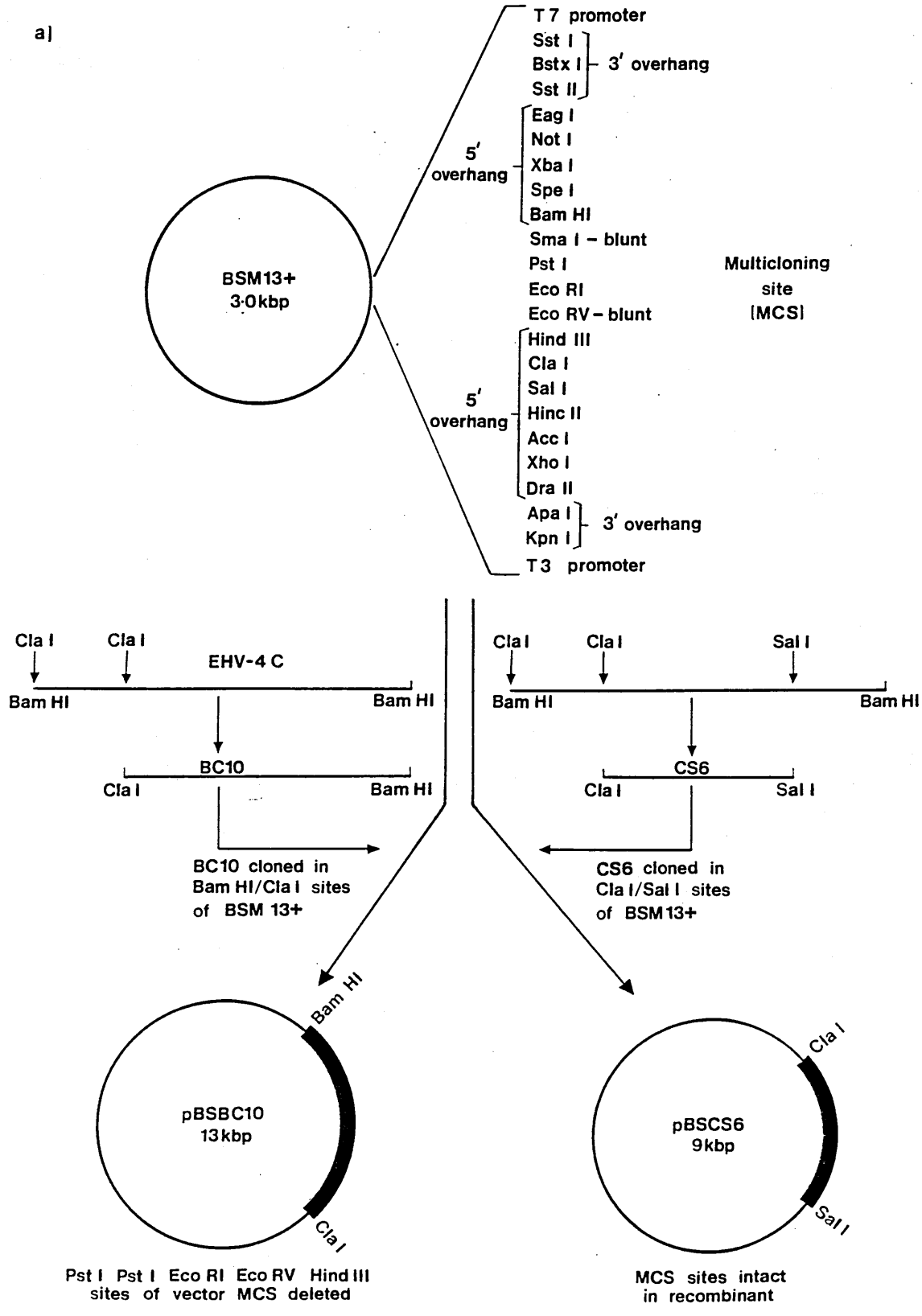


Figure 3.2 (Cont.)

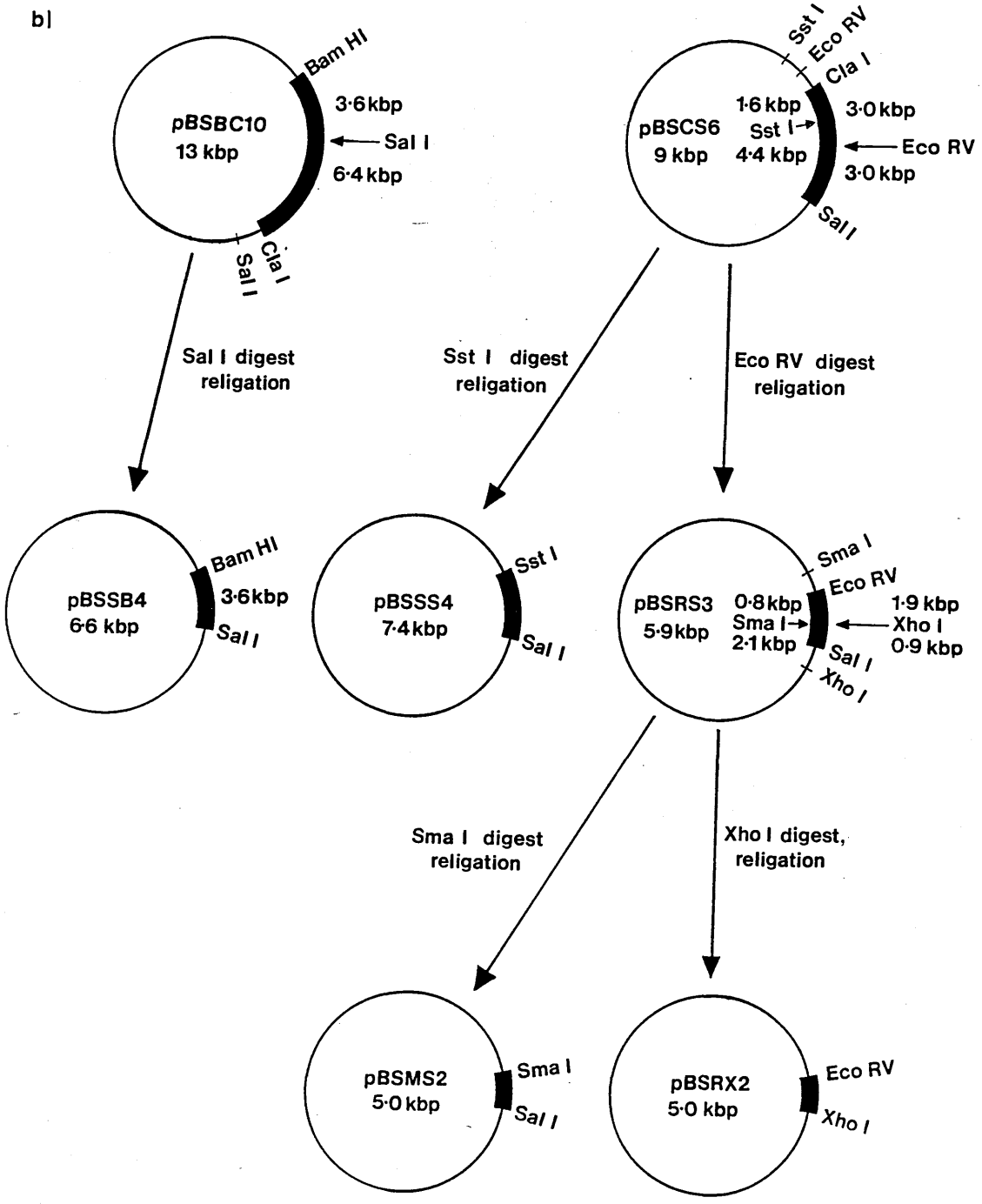


Figure 3.3

TK-Deleted Constructs

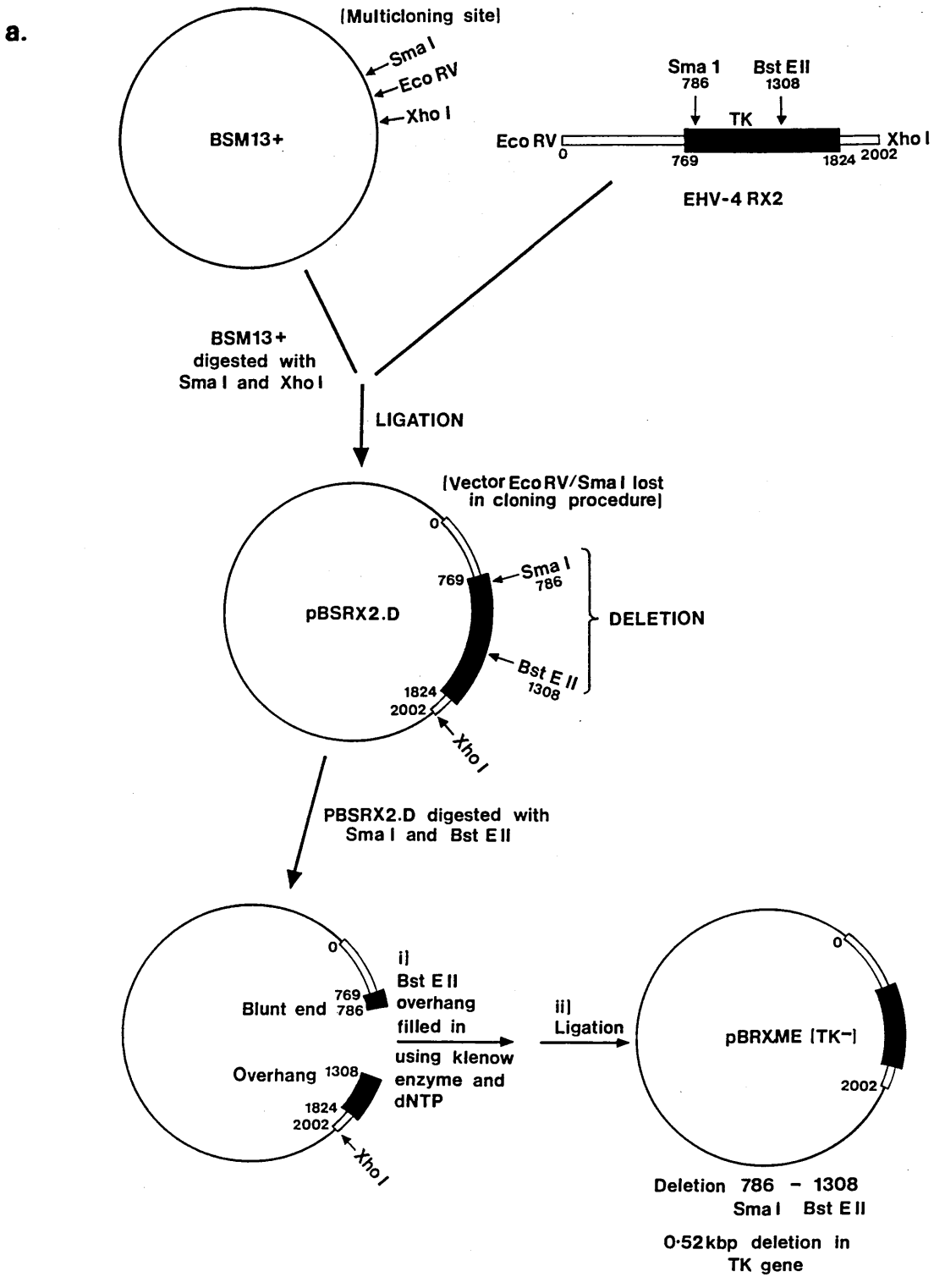


Figure 3.3 (Cont.)

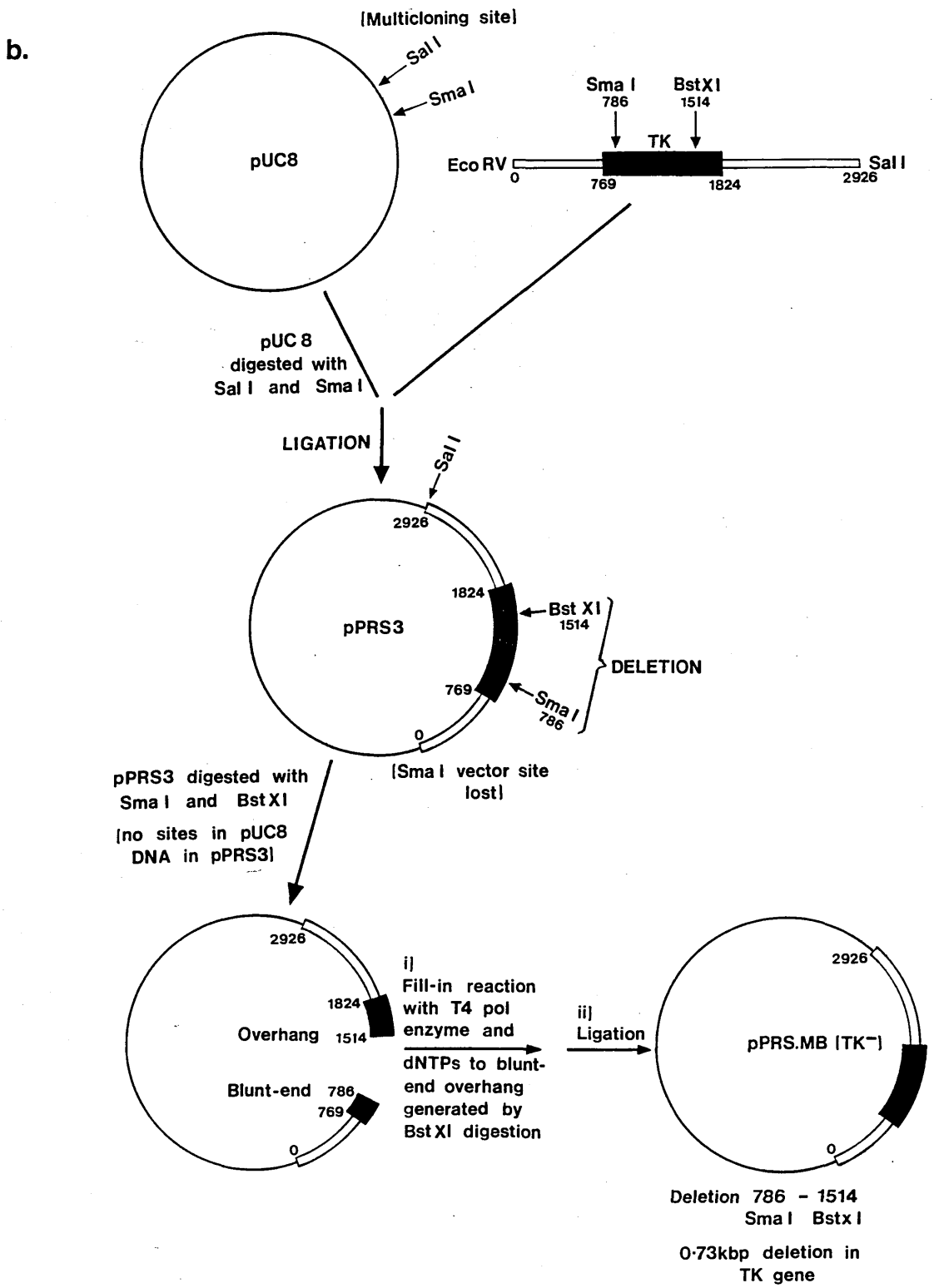



Figure 3.3 Strategies utilised in the production of two TK-deleted constructs, pBRX.ME (a) and pPRS.MB (b). See text for details.

Figure 3.4

ExoIII Deletion of pBSRX2 and pBSRS3

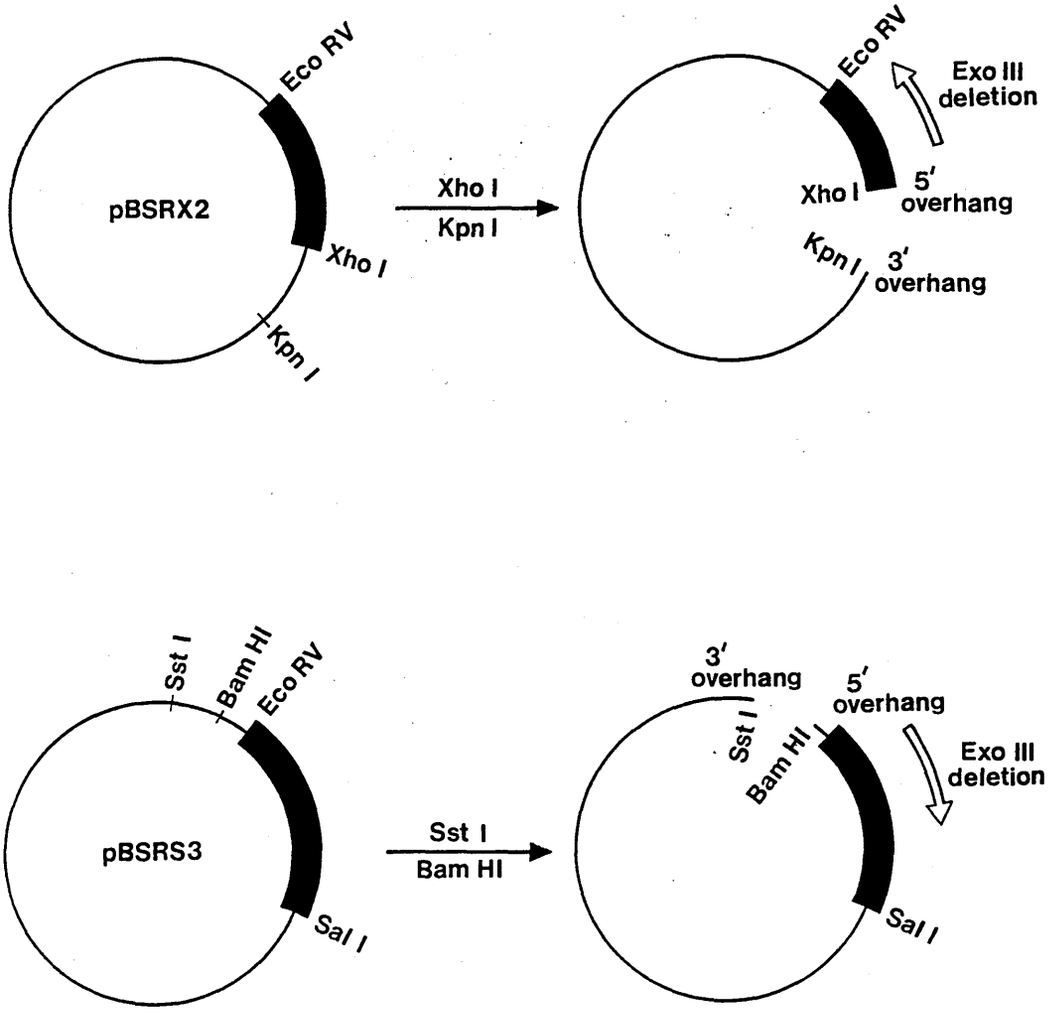



Figure 3.4 Schematic diagram indicating the BSM13+ MCS restriction enzyme sites utilised in the production of sets of ExoIII-deleted constructs derived from plasmids pBSRX2 and pBSRS3.

Figure 3.5 a) BamHI restriction map of EHV-1 (strain HVS25; Whalley et al., 1981; Robertson and Whalley, 1988) with expansion of BamHI B to show the position of the 5.35kbp TK DNA probe utilised in Southern hybridisations. b) BamHI restriction map of EHV-4 (strain 1942; Cullinane et al., 1988) presented in such a way that the EHV-1 and EHV-4 genomes align in terms of map units. On the basis of colinearity of the EHV-1 and EHV-4 $U_L$ components the EHV-4 TK gene would be expected to map within BamHI C .

Figure 3.5

Map Position of EHV-1 TK DNA Probe

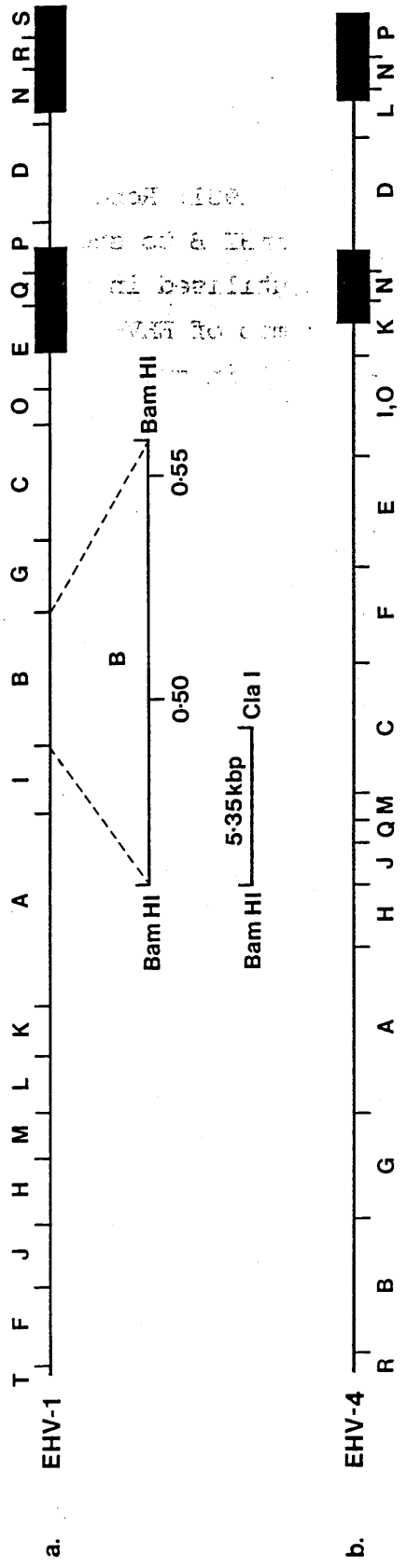


Figure 3.6 Southern hybridisation of an EHV-1 TK probe, a 5.35kbp BamHI/ClaI fragment of EHV-1 BamHI B, to EHV-4 BamHI library constructs comprising EHV-4 BamHI M, Q, C, and F fragments cloned in pUC 9.

a) 0.8% agarose gel : (1) HindIII size markers ($^{32}$P-labelled) (2) BamHI digest of pUC9:M (3) BamHI digest of pUC9:Q (4) BamHI digest of pUC9:C (5) BamHI digest of pUC9:F (6) BamHI/ClaI digest of pUC9:C (7) BamHI/SmaI digest of pUC9:C. Approx. 0.5ug DNA loaded per lane with the exception of lane (4) to which 0.1ug DNA was added.

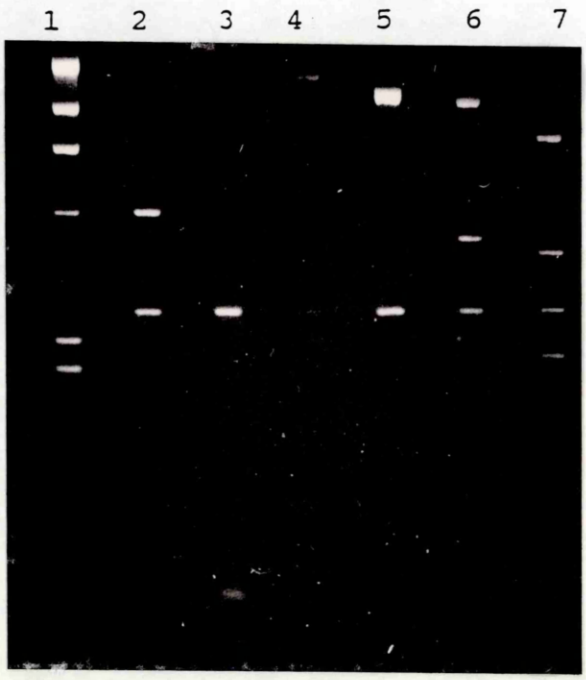
b) Autoradiograph showing hybridisation of the EHV-1 TK DNA probe to a Southern blot of the above gel : (1) HindIII size markers ($^{32}$P-labelled) (2) BamHI digest of pUC9:M (3) BamHI digest of pUC9:Q (4) BamHI digest of pUC9:C (5) BamHI digest of pUC9:F (6) HindIII size markers ($^{32}$P-labelled) (7) BamHI/SmaI digest of pUC9:C (8) BamHI/ClaI digest of pUC9:C.

c) ClaI and SmaI restriction map of EHV-4 BamHI C. The EHV-1 TK probe hybridised to the rightward 10kbp ClaI/BamHI fragment and to three rightward SmaI/SmaI or SmaI/BamHI fragments.

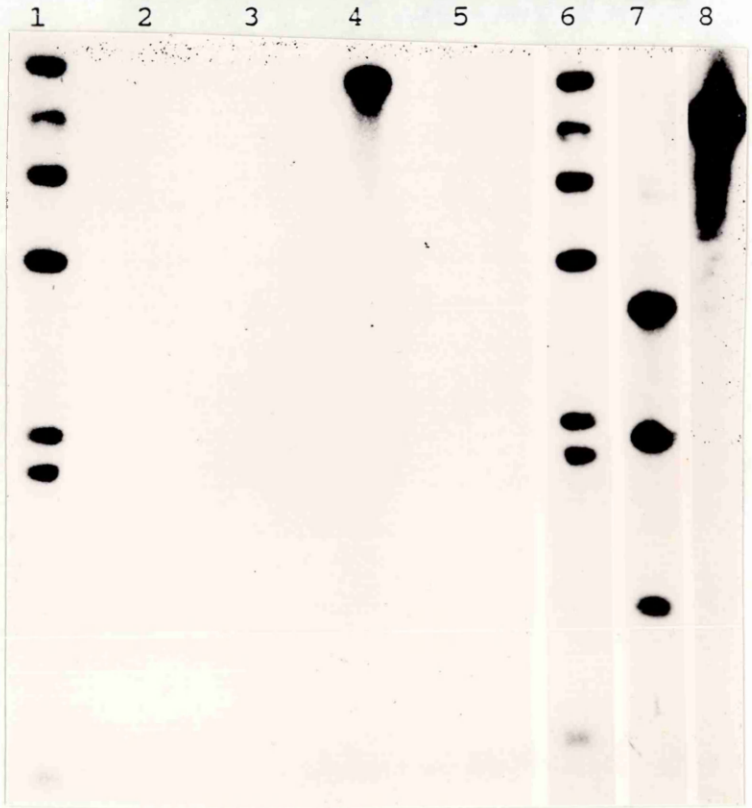
Figure 3.6

Southern Hybridisation

a)




b)




c)

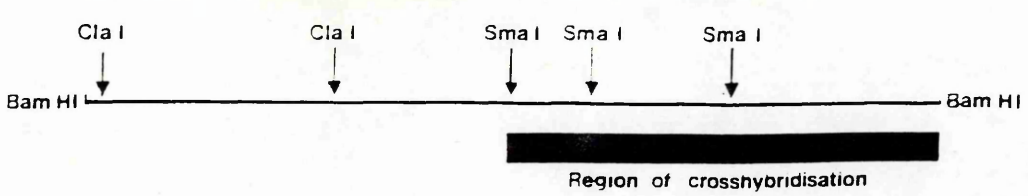


Figure 3.7 Sequence analysis of RX2 and flanking DNA.
a) Localisation of two major ORF , a UL24 gene homologue and the TK gene, within RX2. A further partial ORF was identified as the N-terminal coding region of the gH gene of EHV-4. Transcriptional direction is indicated by arrows. Numbering corresponds to that used in Figure 3.8.
b) Sequencing strategy utilised. Arrows indicate regions sequenced using individual custom-made primers or BS primers (with ExoIII-deleted EHV-4 DNA). The arrow direction indicates whether data was obtained for the upper or lower strands.
c) VAX analysis of the coding potential of the sequence. Potential ORFs are indicated by boxes. As shown in a) two major and one minor ORF were detected.

Figure 3.7

Open Reading Frames Within RX2

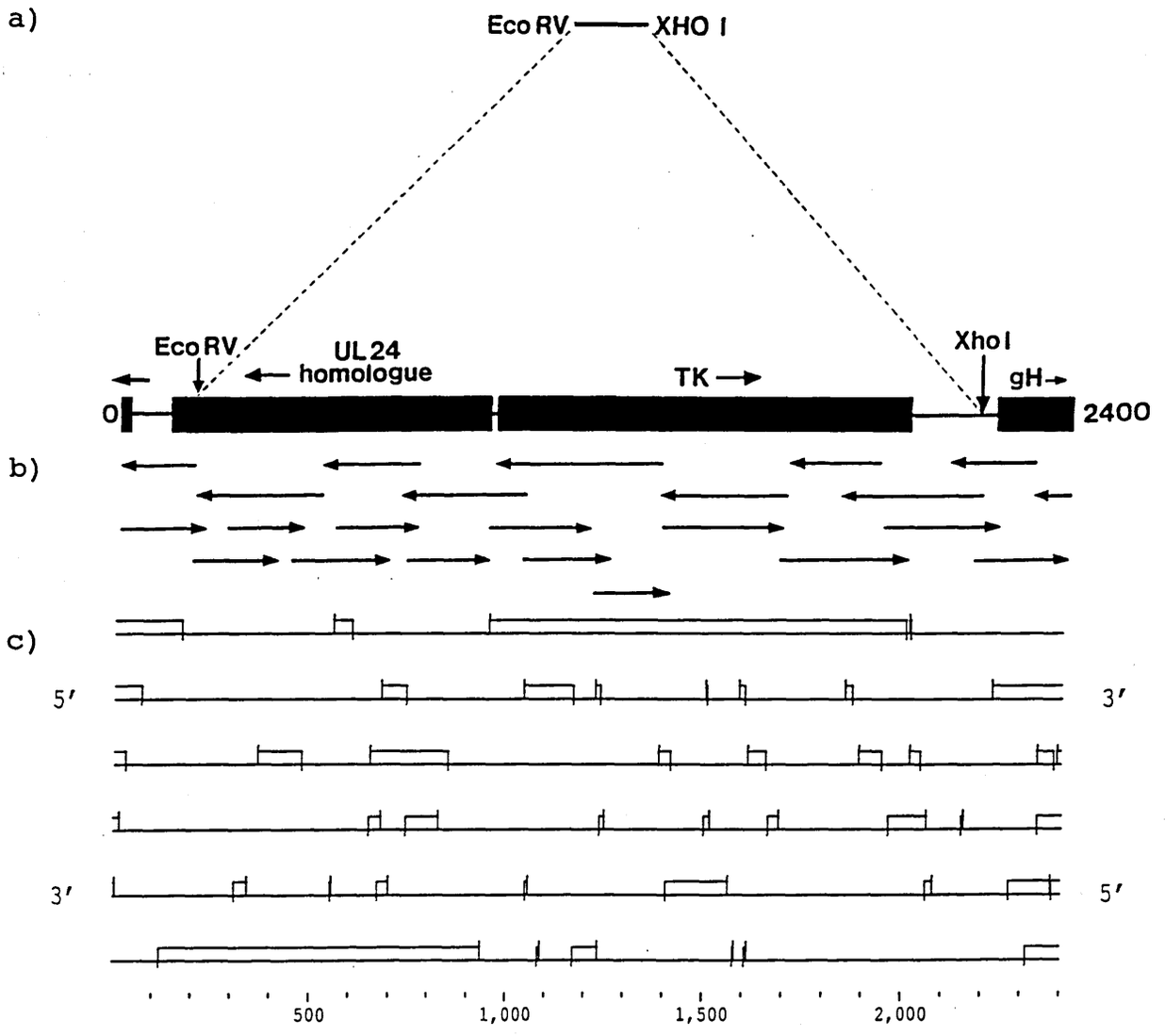


Figure 3.8 Nucleotide sequence analysis of RX2 and flanking DNA and predicted amino acid sequences of the products of the UL24 gene homologue and of the thymidine kinase gene. The predicted N-terminal amino acid sequence of the gH gene homologue is also indicated. Putative TATA boxes (T), polyadenylation signals (PS), and transcriptional start sites (R) are underlined as are the EcoRV and XhoI restriction site termini of RX2.

# Figure 3.8

## Nucleotide Sequence Analysis of 2400bp of EHV-4 BamHI C

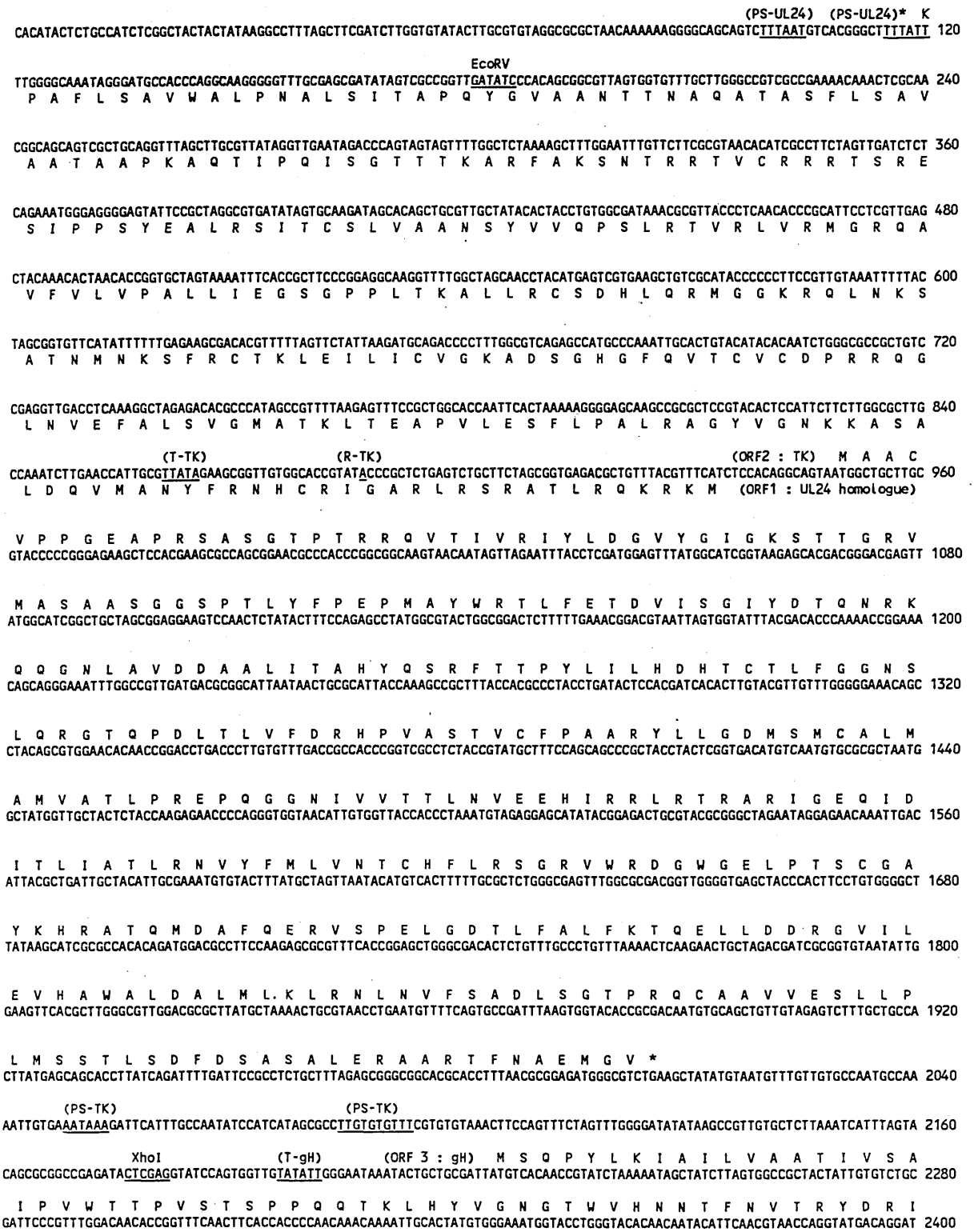


Figure 3.9 Autoradiograph of sequence analysis of EHV-4 C using custom-made primers. Reaction mixes were loaded in order A,C,G,T from left to right. The gel was run at 2000V/cm for 3.5 hours to elucidate sequence unreadable in a previous 2.0 hour run. Thus the sequences shown below commence approximately 70bp 3′ of the primer sites.

Lanes 1-4: Primer UPSBS bp 390-374 (lower strand); Template pBSCS6.

Lanes 5-8: Primer LN5 bp 1329-1345 (upper strand); Template pBSRX2.

Lanes 9-12: Primer LN8 bp 1391-1375 (lower strand); Template pBSRX2.

Lanes 13-16: Primer LNC bp 1064-1047 (lower strand); Template pBSRX2.

Figure 3.9

Autoradiograph of TK-Sequencing Reactions

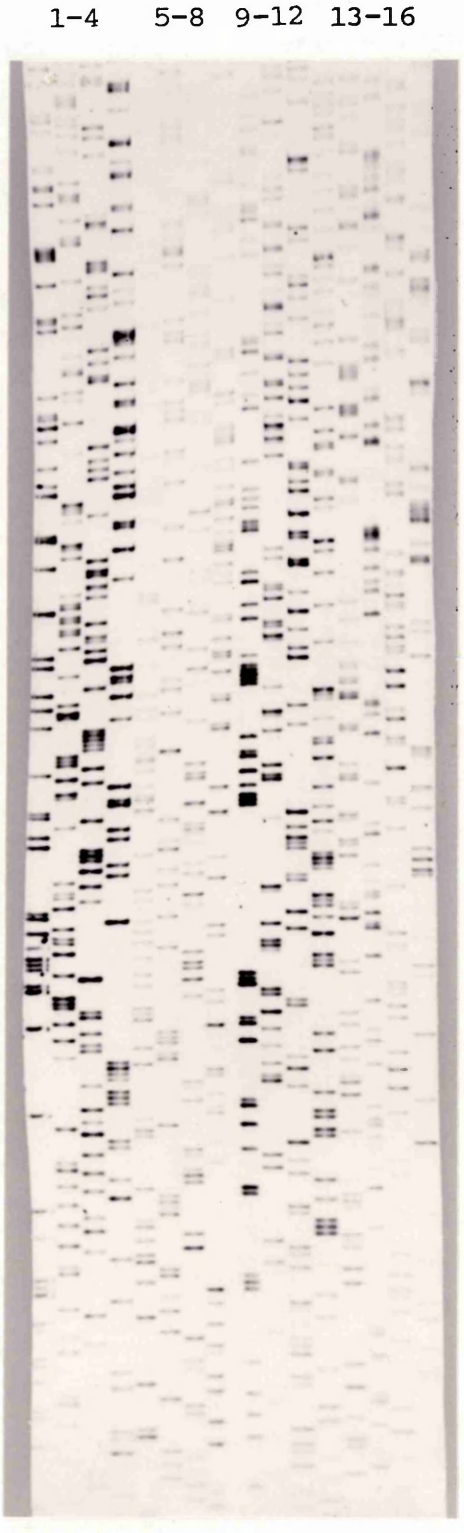
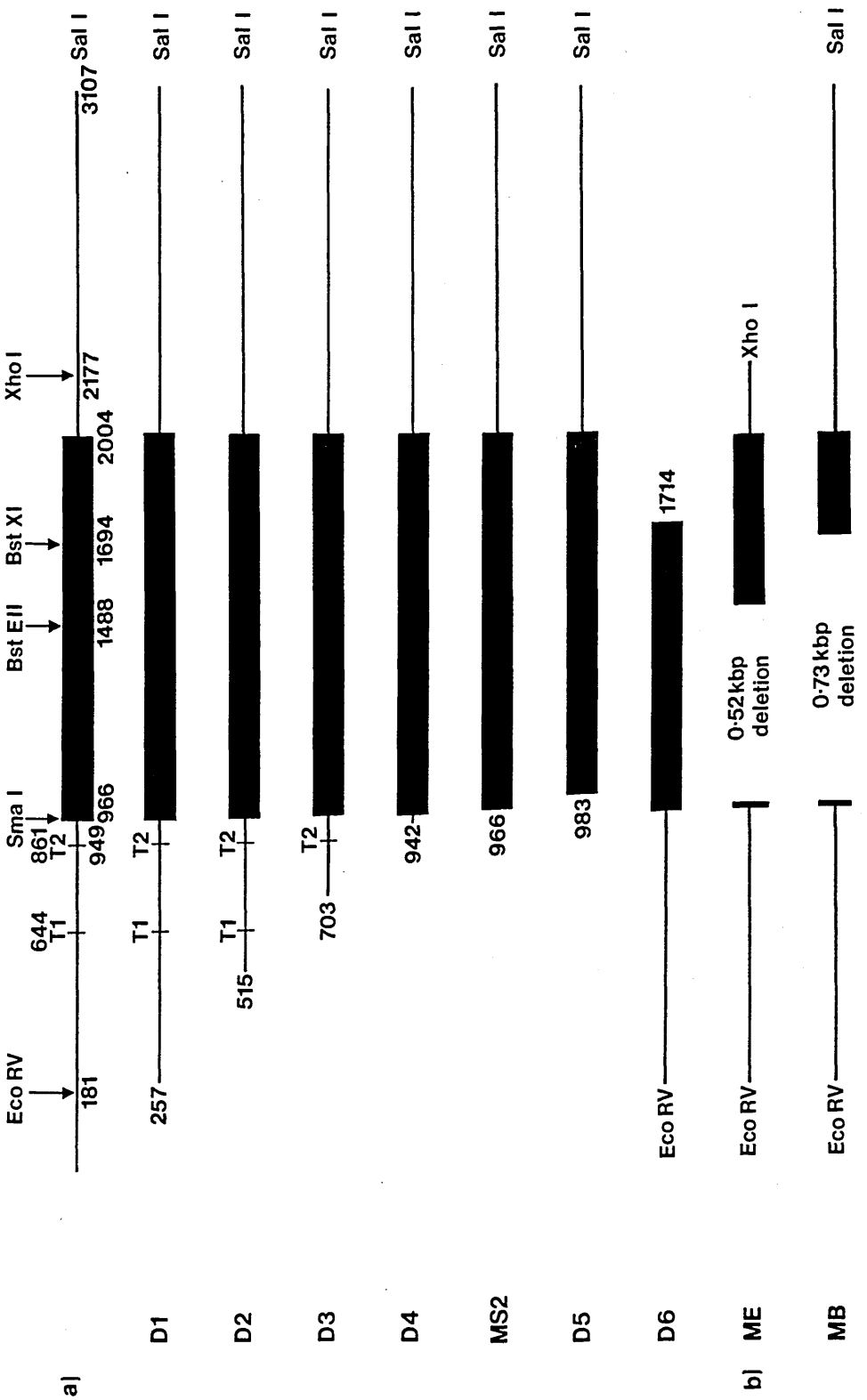


Figure 3.10 EHV-4 DNAs cloned in BSM13+ or pUC vectors which were analysed for their capacity to induce TK- to TK+ transformation within BHK TK- cells (see Table 3.5).
a) Schematic representation of EHV-4 RS3 DNA indicating restriction sites used in deletion procedures, the position of the TK gene (black box), and the position of putative TK TATA boxes (T1 and T2). Numbering corresponds to that utilised in Figure 3.8. Deletions produced within RX2 and RS3 by ExoIII digestion are represented below.
b) Deletions produced within RX2 and RS3 by restriction endonuclease digestion (see Figure 3.3).

Figure 3.10

RX2- and RS3-Deleted DNAs



## Figure 3.11

## Alignment of the Predicted Products of Herpesvirus UL24 Gene Homologues

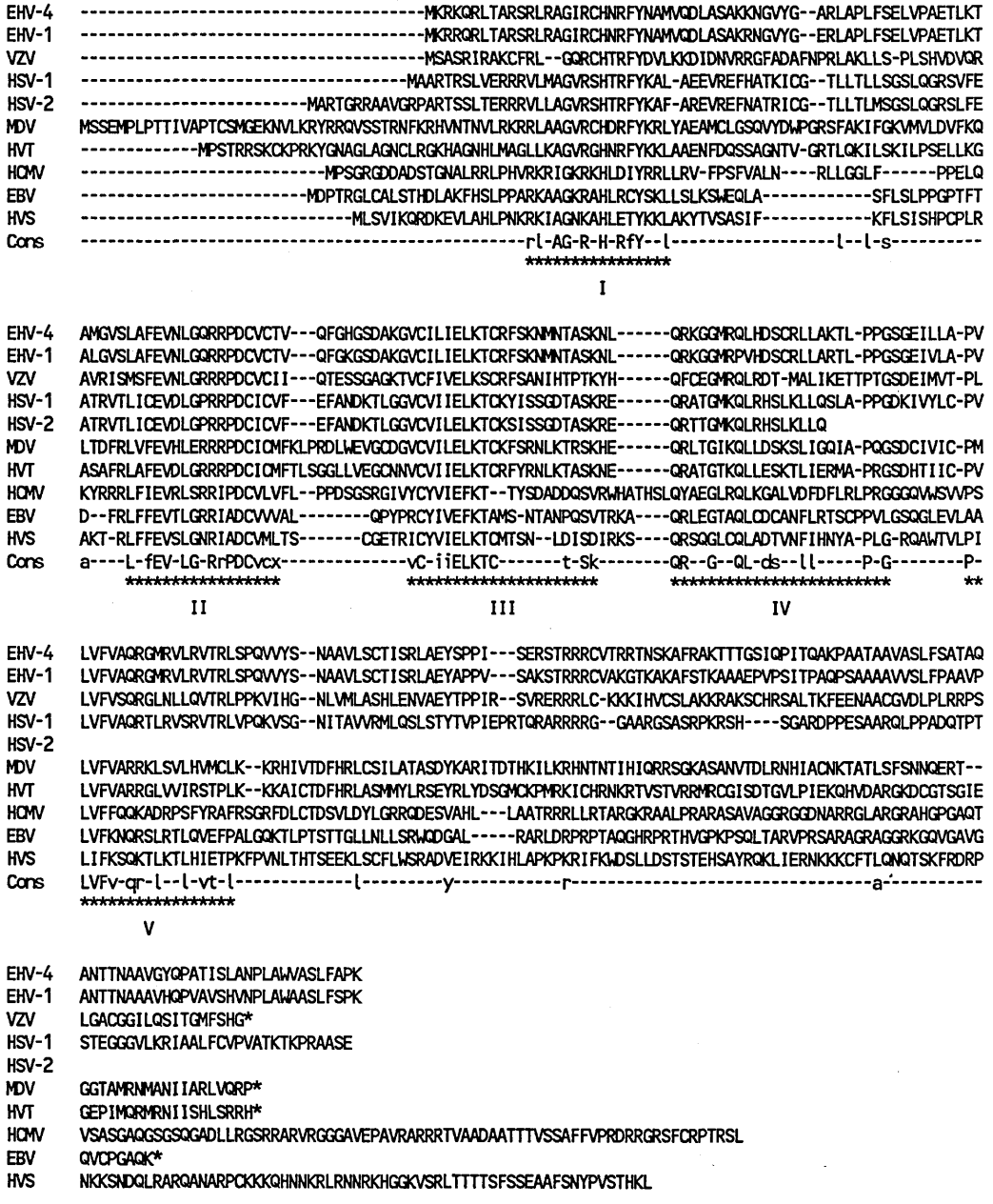
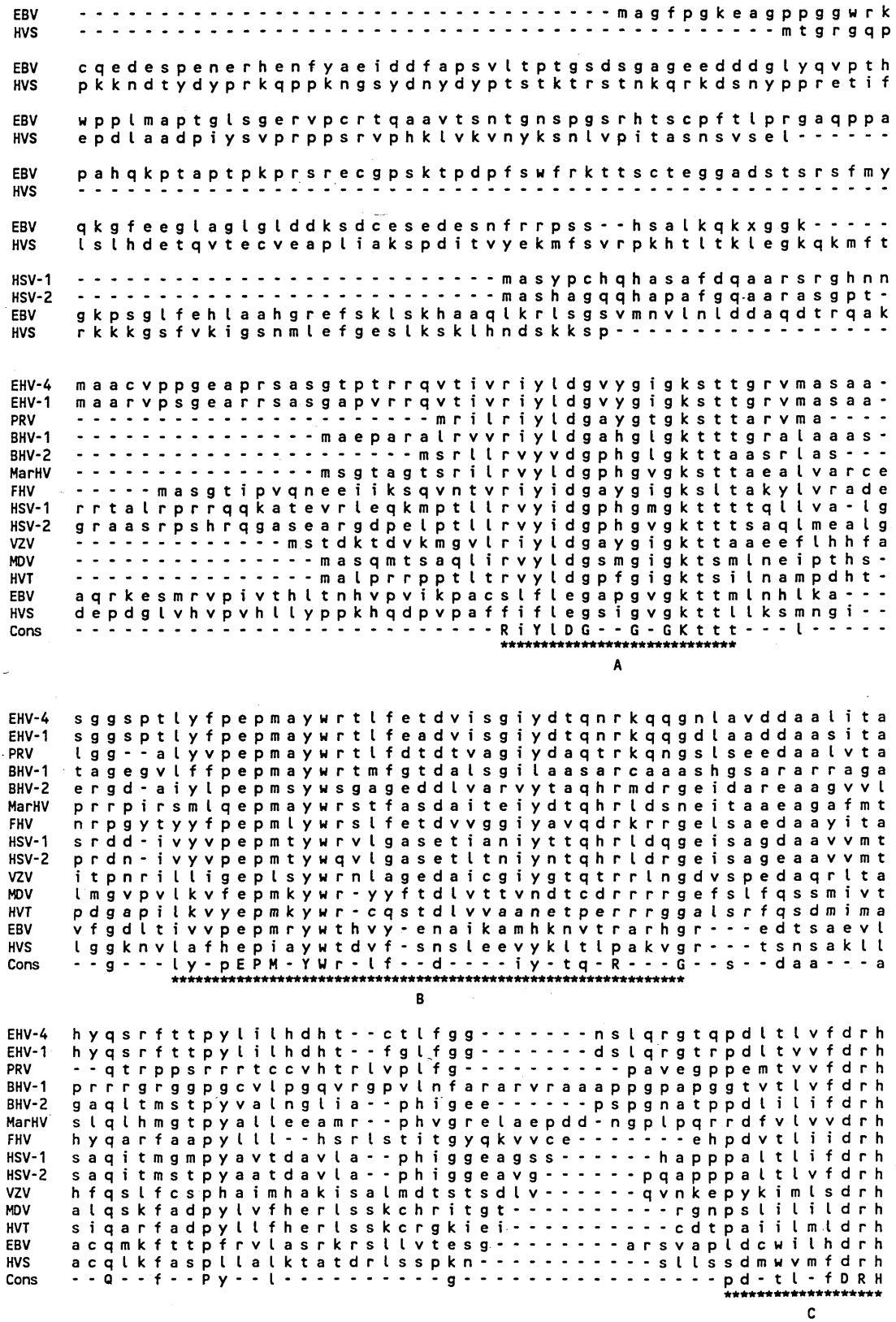



Figure 3.11 Alignment of the predicted amino acid sequences of the HSV-1 UL24-specified polypeptide and its counterpart in other herpesviruses. Data obtained from Robertson and Whalley, 1988 (EHV-1), Davison and Scott, 1986 (VZV), McGeoch et al., 1988a (HSV-1), Scott, 1989 (MDV and HVT), and Jacobson et al., 1989b. A consensus sequence line is detailed in which residues conserved in 6 to 8 polypeptides are presented in lower case, those conserved in 8-10 in upper case.

## Figure 3.12

## Alignment of the Predicted Amino Acid Sequences of Herpesvirus Thymidine Kinases



# Figure 3.12 (Cont.)

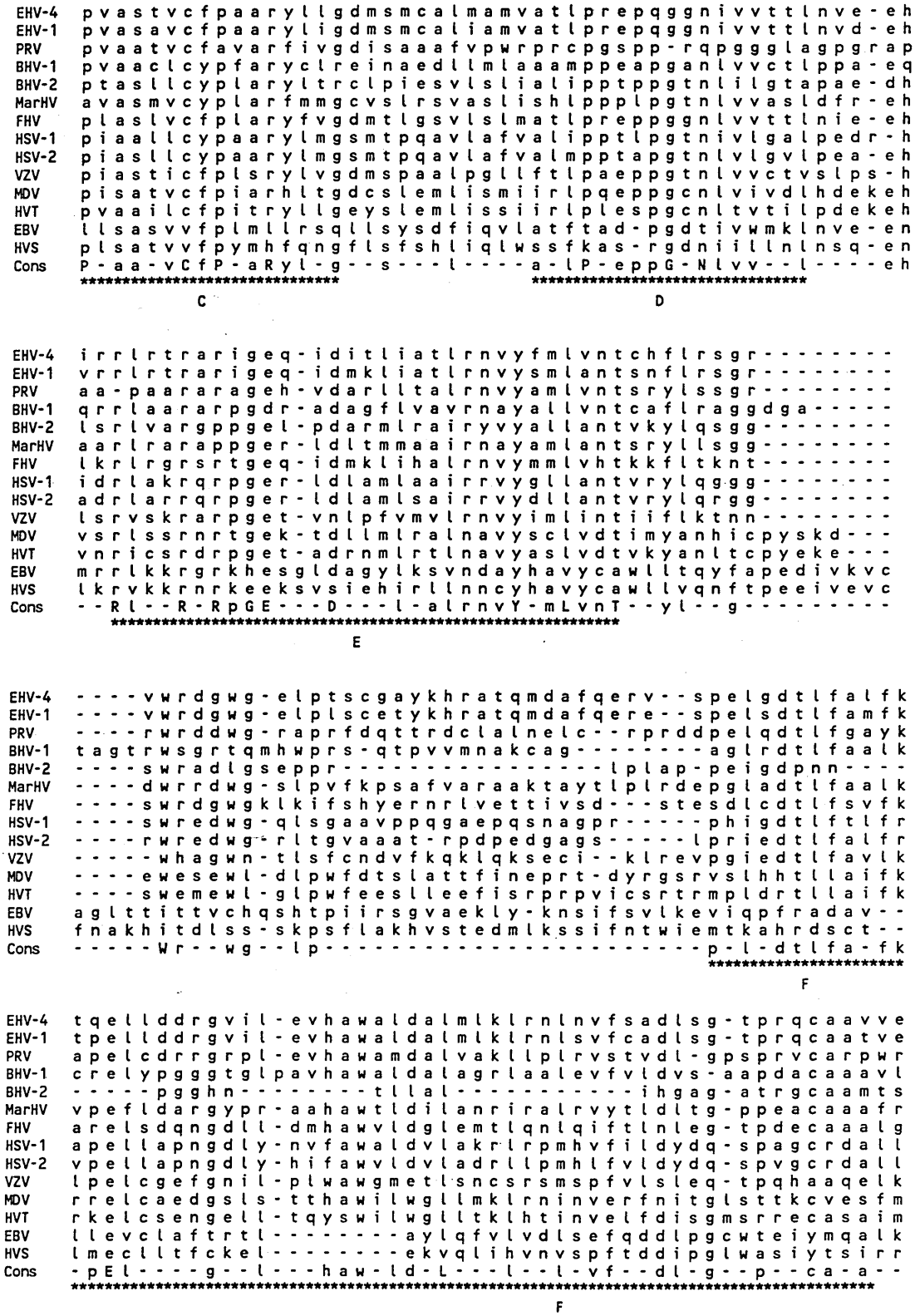


Figure 3.12 (Cont.)

```
EHV-4   s l l p l m - s s t l s d f d s a s a l e r a a r t f n a e m g v *
EHV-1   s l i p l m - s s t l s d s e s a s s l e r a a r t f n a e m g v *
PRV     r g a r r r *
BHV-1   d m r p a m - q a a c a d g a a g a t l a t l a r q f a l e m a g e a t a g p r g l *
BHV-2   w t l d l l a d r l r s m n m f t w t t a r p r s l c g s d e e l a p l d a v h r d p k t r i s s r *
MarHV   r l c a g l v l t e g s h p g a l c e l k r a a a a y a r e m s v v g s r e p t t a e v e s a *
FHV     a l r q d m d m t f i a a c d m h r i s e a l t i y h *
HSV-1   q l t s g m v q t h v t t p g s i p t i c d l a r t f a r e m g e a n *
HSV-2   r l t a g m i p t r v t t a g s i a e i r d l a r t f a r e v g g v *
VZV     t l l p q m t p a n m s s g a w n i l k e l v n a v q d n t s *
MDV     d t m s e r l v t h m s w n d a f e i e a d v l - a y n k e m a m *
HVT     h t m p e r l s t l a s w n d l c e l e d d v i - s y n k g m c n e v g a s r *
EBV     n p a i r s q f f - - d w a g l - - - - s k v i s d f e r g n r d *
HVS     n s a i k p n r v - - n w l a l - - - - e d l a r t f n s q *
Cons    - l - - - m - - - - - - - - - - - - - - - - - - - - - - - - - - -

EHV-4       (352)
EHV-1       (352)
PRV         (295)
BHV-1       (357)
BHV-2       (306)
MarHV       (360)
FHV         (344)
HSV-1       (376)
HSV-2       (375)
VZV         (341)
MDV         (343)
HVT         (350)
EBV         (607)
HVS         (527)
```

Figure 3.12 Alignment of the predicted amino acid sequences of the thymidine kinases of 14 herpesviruses. Data obtained from Robertson and Whalley, 1988 (EHV-1), Kit, 1985b (PRV), Mittal and Field, 1989 and Kit, 1985a (BHV-1), Sheppard and May, 1989 (BHV-2), Otsuka and Kit, 1984 and Scott et al., 1989 (MarHV), Nunberg et al., 1989 (FHV), McGeoch et al., 1988a (HSV-1), Kit et al., 1983a (HSV-2), Davison and Scott, 1986 (VZV), Scott et al., 1989 (MDV and HVT), Baer et al., 1984 (EBV), and Honess et al., 1989a (HVS). A consensus sequence line is detailed in which residues conserved in 7 to 11 polypeptides are presented in lower case, those conserved in in 12 to 14 herpesvirus TKs in upper case. Six regions of conservation of sequence referred to in the text are denoted by asterisks and the letters A-F.

Figure 3.13 Alignment of the putative nucleotide and nucleoside binding domains of herpesvirus thymidine kinases .

a) Conservation of sequence within the putative nucleotide binding domain of thymidine kinases of herpesviruses, vaccinia virus (Hruby et al., 1983), and humans (Bradshaw and Deininger, 1984) and of two ATP-binding enzymes, adenylate kinase (Heil et al., 1974) and bovine ATPase (Walker et al., 1982). Consensus lines are detailed for herpesvirus TK sequences and for all enzyme sequences presented.

b) Conservation of sequence within the putative nucleoside binding domain of herpesvirus thymidine kinases.

In the consensus lines of a) and b) capital letters indicate 100% conserved residues, small letters indicate 70-99% conservation.

## Figure 3.13

## Conservation of Sequence within the Putative Nucleotide and Nucleoside binding Sites of Thymidine Kinases

| | | | |
|---|---|---|---|
| a) | EHV-4 | R I Y L D G V Y G I G K S T T | (27-41) |
| | EHV-1 | R I Y L D G V Y G I G K S T T | (27-41) |
| | PRV | R I Y L D G A Y G T G K S T T | ( 5-19) |
| | BHV-1 | R I Y L D G A H G L G K T T T | (12-26) |
| | BHV-2 | R V Y V D G P H G L G K T T A | ( 6-20) |
| | FHV | R I Y I D G A Y G I G K S L T | (22-36) |
| | MarHV | R V Y L D G P H G V G K S T T | (12-26) |
| | HSV-1 | R V Y I D G P H G M G K T T T | (51-65) |
| | HSV-2 | R V Y I D G P H G V G K T T T | (51-65) |
| | VZV | R I Y L D G A Y G I G K T T A | (14-28) |
| | EBV | S L F L E G A P G V G K T T M | (286-300) |
| | HVS | F I F L E G S I G V G K T T L | (211-225) |
| | MDV | R V Y L D G S M G I G K T S M | (12-26) |
| | HVT | R V Y L D G P F G I G K T S I | (12-26) |
| | Cons (HV) | r - y - d G - - G - G K - t - | |
| | VV | - I - I - G P M F S G K S T E | ( 9-20) |
| | Human TK | - I - L - G P M F S G K S T E | (24-35) |
| | Bov.ATPase | K I G L F G G A G V G K T V F | (151-165) |
| | Ad.Kinase | K IIF V V G G P G S G K G T Q | (10-25) |
| | Cons (all) | - - - - - G - - g - G K - t - | |

| | | | |
|---|---|---|---|
| b) | EHV-4 | D R H P V A S T V C F P A A R Y L L G D M | (138-158) |
| | EHV-1 | D R H P V A S A V C F P A A R Y L I G D M | (138-158) |
| | PRV | D R H P V A A T V C F A V A R F I V G D I | (108-128) |
| | BHV-1 | D R H P V A A C L C Y P F A R Y C L R E I | (132-152) |
| | BHV-2 | D R H P T A S L L C Y P L A R Y L T R C L | (115-135) |
| | FHV | D R H P L A S L V C F P L A R Y F V G D M | (134-154) |
| | MarHV | D R H A V A S M V C Y P L A R F M M G C V | (130-150) |
| | HSV-1 | D R H P I A A L L C Y P A A R Y L M G S M | (162-182) |
| | HSV-2 | D R H P I A S L L C Y P A A R Y L M G S M | (163-183) |
| | VZV | D R H P I A S T I C F P L S R Y L V G D M | (129-149) |
| | EBV | D R H L L S A S V V F P L M L L R S Q L L | (392-412) |
| | HVS | D R H P L S A T V V F P Y M H F Q N G F L | (315-335) |
| | MDV | D R H P I S A T V C F P I A R H L T G D C | (121-141) |
| | HVT | D R H P V A A I L C F P I T R Y L L G E Y | (121-141) |
| | Cons | D R H p - a - - - c - p - a r - - - g - - | |

Table 3.1

Substrate Specificities of Herpesvirus Thymidine Kinases

| Herpesvirus | dCyd kinase activity | donor | | Nucleoside analogues | | | |
|---|---|---|---|---|---|---|---|
| | | ATP | CTP | BVDU | araT | DHPG | ACV |
| alpha | | | | | | | |
| HSV-1 | + | + | + | + | + | + | + |
| HSV-2 | + | + | + | + | + | + | + |
| EHV-1 | + | + | + | + | + | + | + |
| VZV | + | + | + | + | + | na | + |
| MarHV | + | + | + | + | + | + | + |
| PRV | - | + | + | + | + | - | - |
| BHV-1 | - | + | + | + | + | - | - |
| EHV-3 | - | + | + | na | + | na | na |
| HVT | - | + | + | na | na | na | + |
| gamma | | | | | | | |
| BHV-4 | - | + | + | + | -+ | - | - |
| HVS | - | + | na | + | na | na | - |
| EBV | + | + | + | na | + | na | na |

Table 3.1 Specificity of herpesvirus thymidine kinases for nucleoside substrates and phosphate donors. Data compiled from Kit, 1985a, and Scott et al., 1989. Unavailable data are represented by 'na'.

Table 3.2

Transfection of HSV-1, EHV-1, and EHV-4 TK DNAs into TK- Cells

| Construct | ug DNA | HAT-resistant colony count |
|---|---|---|
| HSV-1 pTK | 4 | >1000 |
| EHV-1 pBSBC5 | 9 | >100 |
| EHV-1 pBR:BamHI A | 10 | 0 |
| EHV-4 pUC9:BamHI C | 8 | 8 |

Table 3.2 Transfection of HSV-1 and EHV-1 TK DNAs , EHV-1 BamHI A, and EHV-4 BamHI C into BHK TK- cells. The final column details the number of HAT-resistant colonies induced by transfection of the constructs into $5x10^5$ TK- cells. These results suggested that EHV-4 BamHI C might encode a functional TK activity.

Table 3.3 The TK-transforming capacity of several regions of BamHI C was assessed by monitoring the ability of recombinant constructs to induce a TK+ phenotype on transfection into BHK TK- cells. Derivation of the constructs is detailed in Figure 3.2 and in the text. Fragments MS2 cloned in construct pBSMS2 is entered as (-) since on one occasion a TK+ colony was detected in the post-transfection screening process (see text).

Table 3.3

TK-Transforming Activity of Recombinant Plasmids

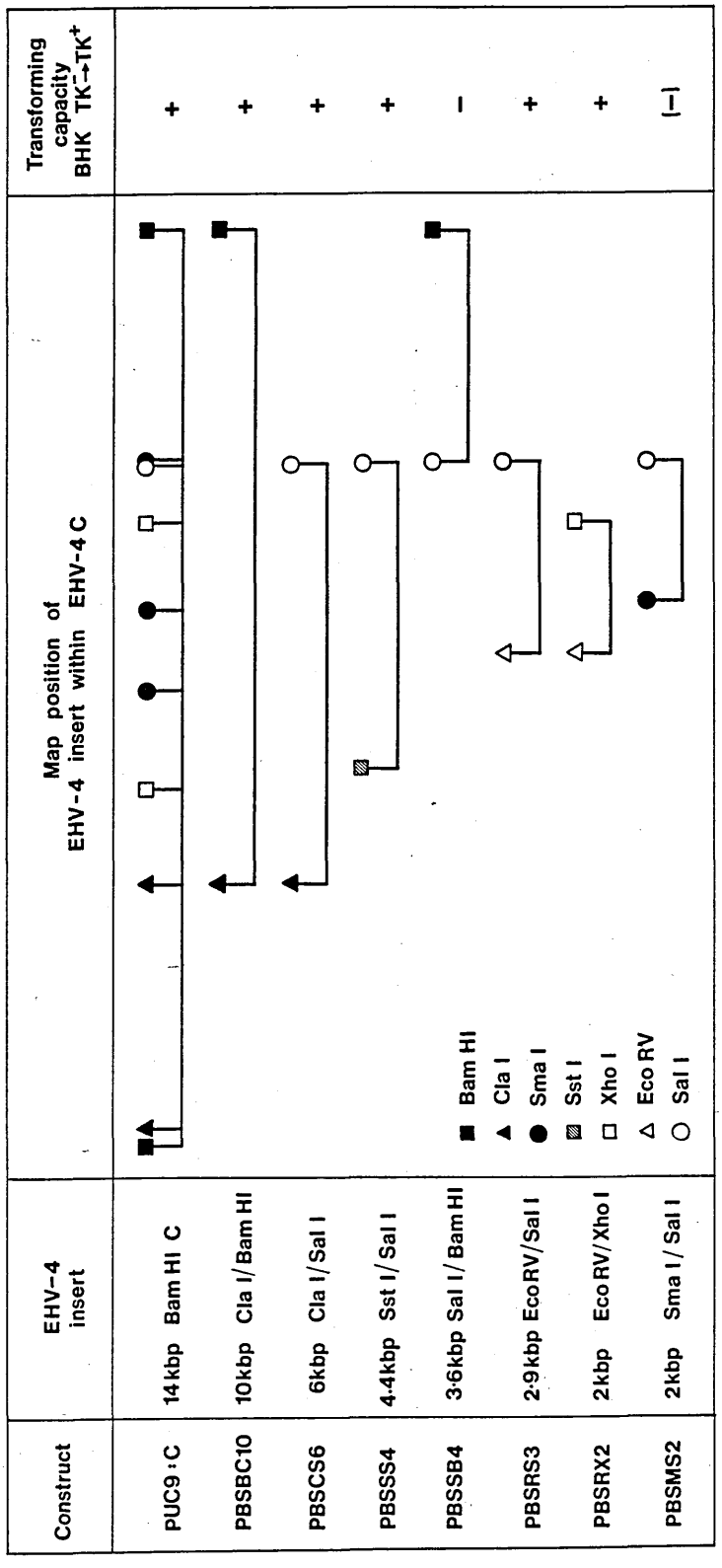


Table 3.4

Transfection of Subfragments of EHV-4 BamHI C into TK- Cells

| Construct | ug DNA | HAT-resistant colony count |
|---|---|---|
| pBSBC10 | 20 | 8 |
| pBSCS6 | 20 | 8 |
| pBSSS4 | 20 | 40 |
| pBSSB4 | 20 | 0 |
| pBSRS3 | 20 | 90 |

Table 3.4 Transfection of restricted regions of EHV-4 BamHI C into BHK TK- cells. A TK+ phenotype was evident within a proportion of TK- cells transfected with constructs containing fragments BC10, CS6, SS4, and RS3 of EHV-4 C. These data mapped the EHV-4 TK gene to the 3kbp EcoRV/SalI fragment of C.

Table 3.5

Transfection of RX2 and Deleted DNAs into TK- Cells

| Construct | limits (bp) | ug DNA | HAT-resistant colony count |
|---|---|---|---|
| pBSRX2 | 181-3107 | 2 | 14, 5, 15 |
| pBSD1 | 257-3107 | 2 | 11, 6, 11 |
| pBSD2 | 533-3107 | 2 | 18, 4, - |
| pBSD3 | 703-3107 | 2 | 41, 25, 29 |
| pBSD4 | 942-3107 | 2 | 6, 4, 3 |
| pBSMS2 | 966-3107 | 2 | 0, 1, 0 |
| pBSD5 | 981-3107 | 2 | 0, 0, 0 |
| pBSD6 | 181-1714 | 2 | 0, 0 |
| pBSME | 181-966:1488-2181 | 2 | 0, 0 |
| pBSMB | 181-966:1694-3107 | 2 | 0, 0 |

Table 3.5 Transfection of recombinant constructs containing deletions within RX2 or RS3 of BamHI C as detailed in Figure 3.10. TK+ colonies were evident in cultures transfected with pBSRX2 and pBSRS3 constructs with deletions spanning the 5′ 761 bp (181 to 942, Figure 3.10). With the exception of 1 HAT-resistant colony in the pBSMS2-transfected cultures (see text) no HAT-resistant colonies were observed in transfections with other deleted DNAs including pBSME and pBSMB.

Table 3.6

Codon Usage of EHV-4 UL24 gene Homologue

| 1st \ 2nd | A (22%) | C (31%) | G (21%) | T (26%) | 3rd | Residue | f | (%) |
|---|---|---|---|---|---|---|---|---|
| A | Lys 12 | Thr 2 | Arg 7 | Ile 5 | A (30%) | Ala | 36 | (13.2) |
| | Asn 8 | Thr 4 | Ser 4 | Ile 1 | C (20%) | Arg | 26 | ( 9.5) |
| (29%) | Lys 3 | Thr 6 | Arg 3 | Met 6 | G (24%) | Asn | 12 | ( 4.4) |
| | Asn 4 | Thr 9 | Ser 3 | Ile 3 | T (26%) | Asp | 4 | ( 1.5) |
| | | | | | | Cys | 8 | ( 2.9) |
| C | Gln 8 | Pro 4 | Arg 4 | Leu 7 | A | Gln | 12 | ( 4.4) |
| | His 2 | Pro 3 | Arg 4 | Leu 4 | C | Glu | 7 | ( 2.6) |
| (23%) | Gln 4 | Pro 3 | Arg 5 | Leu 0 | G | Gly | 15 | ( 5.5) |
| | His 1 | Pro 4 | Arg 3 | Leu 5 | T | His | 3 | ( 1.1) |
| | | | | | | Ile | 9 | ( 3.3) |
| G | Glu 5 | Ala 8 | Gly 6 | Val 4 | A | Leu | 27 | ( 9.9) |
| | Asp 2 | Ala 10 | Gly 2 | Val 2 | C | Lys | 15 | ( 5.5) |
| (30%) | Glu 2 | Ala 8 | Gly 4 | Val 11 | G | Met | 6 | ( 2.2) |
| | Asp 2 | Ala 10 | Gly 3 | Val 3 | T | Phe | 9 | ( 3.3) |
| | | | | | | Pro | 14 | ( 5.1) |
| T | End 1 | Ser 4 | End 0 | Leu 5 | A | Ser | 22 | ( 8.1) |
| | Tyr 2 | Ser 3 | Cys 3 | Phe 2 | C | Thr | 21 | ( 7.7) |
| (18%) | End 0 | Ser 3 | Trp 1 | Leu 6 | G | Trp | 1 | ( 0.4) |
| | Tyr 3 | Ser 5 | Cys 5 | Phe 7 | T | Tyr | 5 | ( 1.8) |
| | | | | | | Val | 20 | ( 7.3) |

Table 3.6 Codon usage of the EHV-4 UL24 gene homologue and amino acid frequencies within its predicted product. The frequencies of each base within codon positions 1,2, and 3 are presented as percentages.

## Table 3.7

## Codon Usage of EHV-4 TK gene

| 1st \ 2nd | A (23%) | C (26%) | G (22%) | T (29%) | 3rd | Residue | f | (%) |
|---|---|---|---|---|---|---|---|---|
| A | Lys 3 | Thr 6 | Arg 4 | Ile 6 | A (24%) | Ala | 35 | ( 9.9) |
| | Asn 5 | Thr 9 | Ser 8 | Ile 1 | C (24%) | Arg | 28 | ( 7.9) |
| (25%) | Lys 2 | Thr 8 | Arg 0 | Met 12 | G (24%) | Asn | 10 | ( 2.8) |
| | Asn 5 | Thr 8 | Ser 4 | Ile 7 | T (28%) | Asp | 19 | ( 5.4) |
| | | | | | | Cys | 7 | ( 2.0) |
| C | Gln 8 | Pro 6 | Arg 5 | Leu 10 | A | Gln | 13 | ( 3.7) |
| | His 5 | Pro 5 | Arg 9 | Leu 3 | C | Glu | 15 | ( 4.2) |
| (26%) | Gln 5 | Pro 5 | Arg 7 | Leu 11 | G | Gly | 25 | ( 7.1) |
| | His 3 | Pro 1 | Arg 3 | Leu 4 | T | His | 8 | ( 2.3) |
| | | | | | | Ile | 14 | ( 4.0) |
| G | Glu 6 | Ala 5 | Gly 10 | Val 7 | A | Leu | 39 | (11.0) |
| | Asp 12 | Ala 9 | Gly 3 | Val 2 | C | Lys | 5 | ( 1.4) |
| (33%) | Glu 9 | Ala 8 | Gly 3 | Val 3 | G | Met | 12 | ( 3.4) |
| | Asp 7 | Ala 13 | Gly 9 | Val 12 | T | Phe | 14 | ( 4.0) |
| | | | | | | Pro | 17 | ( 4.8) |
| T | End 0 | Ser 3 | End 1 | Leu 4 | A | Ser | 22 | ( 6.2) |
| | Tyr 8 | Ser 2 | Cys 3 | Phe 2 | C | Thr | 31 | ( 8.8) |
| (16%) | End 0 | Ser 1 | Trp 4 | Leu 7 | G | Trp | 4 | ( 1.1) |
| | Tyr 2 | Ser 4 | Cys 4 | Phe 12 | T | Tyr | 10 | ( 2.8) |
| | | | | | | Val | 24 | ( 6.8) |

Table 3.7 Codon usage of the EHV-4 TK and amino acid frequencies within its predicted product. The frequencies of each base within codon positions 1,2, and 3 are presented as percentages.

Table 3.8

Identity of the EHV-4 UL24 and TK Gene Products with Counterparts in Other Herpesviruses

| | UL24 (Entire) | TK (Entire) | TK (Restricted) | |
|---|---|---|---|---|
| EHV-1 | 83 | 89 | 90 | (27-326) |
| PRV | | 42 | 47 | ( 5-295) |
| BHV-1 | | 37 | 40 | (12-321) |
| BHV-2 | | 29 | 31 | (6-262) |
| MarHV | | 36 | 40 | (12-320) |
| FHV | | 47 | 51 | (22-323) |
| HSV-1 | 36 | 36 | 38 | (51-346) |
| HSV-2 | 18 | 35 | 37 | (50-345) |
| VZV | 37 | 35 | 40 | (14-316) |
| MDV | 29 | 34 | 37 | (12-317) |
| HVT | 31 | 33 | 35 | (12-318) |
| HCMV | 23 | | | |
| EBV | 23 | 18 | 25 | (286-586) |
| HVS | 21 | 16 | 25 | (211-509) |

Table 3.8 Percentage identities exhibited by counterparts of the EHV-4 UL24 and TK gene products in a variety of herpesviruses. The entire UL24 and TK gene products are compared in columns 1 and 2. The final column details comparison of amino acids 27-326 of the EHV-4 TK polypeptide with the equivalent region within other TKs (as indicated by residue numbers in brackets). This part of the TK polypeptide corresponds to the most conserved or 'core' region of the polypeptide commencing at residue 1 of the nucleotide binding domain and terminating at a position equivalent to the final residue of the PRV TK polypeptide.

Table 3.9 The effect on TK activity of substitution of specific amino acids within the HSV-1 and HSV-2 TK polypeptides. Abbreviations used are as follows - ns (nucleoside), nt (nucleotide), and wt (wild-type).

Table 3.9

Analysis of HSV-1 and HSV-2 TK mutants

| Herpesvirus | Residue | Substituted Residue | Effect on Enzyme Activity | Reference |
|---|---|---|---|---|
| HSV-1 | $Gly_{56}$ | Val | inactivated | Liu and Summers, 1988 |
| HSV-1 | $Gly_{59}$ | Val | inactivated | Liu and Summers, 1988 |
| HSV-1 | $Gly_{61}$ | Val | inactivated | Liu and Summers, 1988 |
| HSV-1 | $Lys_{62}$ | Ile | inactivated | Liu and Summers, 1988 |
| HSV-1 | $Thr_{63}$ | Ala | inactivated | Liu and Summers, 1988 |
| HSV-1 | $Thr_{63}$ | Ser | altered | Liu and Summers, 1988 |
| HSV-1 | $Ala_{168}$ | Thr | altered ns binding<br>BVdU resistance | Darby et al., 1986 |
| HSV-1 | $Arg_{176}$ | Gln | altered ns binding<br>ACV resistance | Darby et al., 1986 |
| HSV-1 | $Cys_{336}$ | Tyr | altered nt,ns binding | Darby et al., 1986 |
| HSV-1 | $Cys_{171}$ | Ser | wt-activity | Inglis and Darby, 1987 |
| HSV-1 | $Cys_{171}$ | Gly | wt-activity | Inglis and Darby, 1987 |
| HSV-2 | $Arg_{223}$ | His | altered ns binding<br>ACV resistance | Kit et al., 1987d |

## DISCUSSION

The TK gene of EHV-4 maps at around 0.48 map units within the long component of the genome. Similar map positions have been reported for the TK genes of other alphaherpesviruses including HSV-1 in its $I_L$ isomeric form. The TK gene lies within a region of the herpesvirus genome conserved in gene order such that the genes UL24 homologue-TK-gH are arranged in conserved relative transcriptional sense. However, the position of this block of genes within the genome varies between alpha-, beta-, and gamma herpesviruses and a TK gene is notably absent from this gene block in the betaherpesvirus HCMV genome. Thus in alphaherpesviruses ($I_L$ isomer HSV), with the exception of MDV which lacks a UL24 homologue, the gene block is positioned in the central region of the long component of the genomes, and in EBV it maps around 30kbp from the left genome terminus. Within the alphaherpesvirus group the fine gene arrangement differs slightly thus in HSV-1 the N-terminal coding regions and 5′ regulatory regions of the UL24 and TK genes overlap whereas in EHV-4 the overlap is between the N-terminal coding region of one gene and the 5′ regulatory region of the other gene.

Translation of the HSV-1 TK gene is unusual in that in addition to the 5′ terminal Met codon, internal Met codons are used by the ribosomal translational apparatus as initiation codons giving rise to variant sized transcripts and intact and truncated TK proteins (Marsden et al., 1983). A similar translational pattern has not been reported for any other herpesvirus TKs.

It is evident from the comparative analyses of the TKs encoded by different herpesviruses that the gene has evolved from a common progenitor gene. Robertson and Whalley (1988) have proposed that the gene may have been acquired from a cellular thymidylate kinase gene and that thymidine kinase activity evolved at a later stage. Honess and

coworkers (1989a) have proposed a tentative evolutionary divergence scheme on the basis of the comparative analyses of all herpesvirus TK sequences. Their tree clearly separates the EBV and HVS sequences from that of the alphaherpesviruses. In this scheme EHV-1 TK is most closely related to that of VZV (Chapter 1, Fig. 3).

The thymidylate kinase activity of EHV-1 , at least with respect to BVdU and IVdU monophosphate substrates, is significantly lower than that of HSV-1 (Kit et al., 1987a,b). The poor affinity of TKs for these analogues has been attributed to the presence of a serine residue at a position corresponding to $Ala_{168}$ in HSV-1. Since the EHV-4 TK possesses a serine and not an alanine at this position it might be expected to exhibit similar monophosphate substrate binding characteristics to those of EHV-1. Thus it would be predicted that EHV-4 would exhibit resistance to BVdU and IVdU.

On the basis of the study of the affinity of HSV-1 TK mutants for thymidine and other nucleoside substrates Darby et al (1986) have proposed that the HSV-1 TK active site is composed of three distinct regions of polypeptide - region A, region C, and Cys 336 - arranged to form overlapping nucleoside and nucleotide binding sites . The cysteine residue at 336 in HSV-1 is critical for HSV-1 TK nucleoside and nucleotide binding since substitution with tyrosine alters affinity for both substrate and cosubstrate and is thus proposed to interact with both nucleotide and nucleoside binding sites (Darby et al., 1986). Although it is conserved in most of the herpesvirus TKs including EHV-4 TK, at position 315, it is not present in HVS or VZV . Despite the 100% conservation of a cysteine residue within the putative nucleoside binding site (Fig. 13b) of alpha and beta herpesviruses, the residue can, in HSV-1 at least, be substituted with other side chains with little effect on enzyme activity (Table 9). This ruled out

a role for HSV-1 $Cys_{171}$ in disulphide bond formation with $Cys_{336}$ (Inglis and Darby, 1987).

The interest in herpesvirus TK genes is directed, in the main part, towards the following - i) use of the gene in genetic manipulation either as a target insertion site in herpesviral genomes or as a cotransfectable phenotypic marker (Shih et al., 1985), ii) investigation of the role of TK in virulence and latency, and iii) analysis of TK nucleoside specificity and its exploitation in antiviral drug development. Our interest in the EHV-4 TK gene is directed to its potential as a target site for insertion of foreign genes into the EHV-4 genome. It is possible that, as in other herpesviruses, inactivation of the TK gene might lead to viral attenuation. An EHV-1 $TK^-$ mutant produced by Cornick and coworkers (1990) was found to be attenuated in equines. However, the mutant was selected under pressure in the presence of nucleoside analogues, a strategy which induces multiple mutations within the genome. Thus, it cannot be concluded that attenuation was a direct consequence of mutations within the TK gene. Definitive analysis of the role of the TK gene in EHV pathogenicity will require the production of mutants with defined mutations solely within the TK gene .

The determination of the precise genomic position of the EHV-4 TK gene is a prerequisite to the production, by recombinant DNA techniques, of such mutants. As a preliminary step towards the construction of recombinant $TK^-$ viruses, we have constructed two EHV-4 TK-deletion plasmids. These could be co-transfected into cells with wild type EHV-4 and progeny recombinant $TK^-$ virus isolated. Studies on the growth characteristics and latency capacity of such mutants should lead to a greater understanding of the role of thymidine kinase, if any, in equine herpesvirus pathogenesis.

# CHAPTER 4

## NUCLEOTIDE SEQUENCE OF THE EHV-4 gH GENE

## INTRODUCTION

As detailed in the general introduction, map positions have been reported for the six major glycoproteins of EHV-1 (Allen and Yeargan , 1987). A candidate for a glycoprotein gene not mapped in that study is a homologue of HSV-1 gH. Glycoprotein H is one of only two glycoproteins , the other being gB, encoded throughout members of all three subgroups, alpha , beta, and gamma, of the herpesvirus family (Baer et al., 1984; Cranage et al., 1988; Gompels and Minson, 1986; Gompels et al., 1988b; Heineman et al., 1988; McGeoch and Davison, 1986; Oba and Hutt-Fletcher, 1988). A gH gene homologue has been mapped to the EHV-1 genome by sequence analysis but its protein product has yet to be identified (Robertson and Whalley, 1988; Robertson et al., 1991).

The gH-type proteins of VZV (gpIII), HSV-1 (gH), EBV (gp85), and HCMV (gp86) are important in mediating virus infectivity since antibodies against gH can neutralise infectivity independent of complement (Buckmaster et al., 1984; Cranage et al., 1988; Forghani et al., 1984; Gompels and Minson, 1986; Gretch et al., 1988; Grose et al., 1983; Keller et al., 1984, 1987; Montalvo and Grose, 1986; Oba and Hutt-Fletcher, 1988; Rasmussen et al., 1984; Strnad et al., 1982). Unlike gB, gC, and gD, gH of HSV-1 is apparently not involved in adsorption of the virion to cellular receptors (Kuhn et al., 1990). Accordingly, the neutralising effect of HSV-1, and also of EBV, anti-gH antibody appears to occur at some stage after adsorption of virions to the host cell membrane indicating a role for gH in penetration events (Fuller et al., 1989; Miller and Hutt-Fletcher, 1988) . HSV-1 and VZV anti-gH antibody inhibits plaque formation in infected monolayer cultures (Buckmaster et al., 1984; Gompels and Minson, 1986; Grose et al., 1983; Keller et al., 1987) and a monoclonal antibody to HSV-1 gH

inhibits syncitium formation in HSV-1 syncitial strain infected cells (Gompels and Minson, 1986). Thus, it seems the gH-type glycoprotein may play a role in membrane fusion events of both the virion-host cell and intercellular type thus being important for initial penetration into target cells and for cell-to-cell spread of the virus.

Analysis of the target glycoproteins of antibodies raised against whole herpesvirus virions has revealed that the gH-type protein of alpha-, beta- and gammaherpesviruses is low in abundance, immunoreactivity, or both since few antibodies are directed to this glycoprotein (reviewed in Gompels et al., 1988b). However, as stated above, monoclonal antibodies with a gH target tend to neutralise viral infectivity in the absence of complement and , in HSV-1 at least, gH is essential for viral replication (Desai et. al, 1988; McGeoch and Davison, 1986). Taken with the observation that gH is specified within all herpesviruses investigated to date, this indicates an important, if not critical, role for this glycoprotein in the herpesvirus life cycle.

In HSV-1, VZV, and EBV , the gH gene is located within part of the genome of conserved genetic organisation with gene order UL24-TK-gH-UL21-UL20-major capsid protein (MCP) gene . In HCMV, the gene arrangement differs : the gH gene maps proximal to the UL24 gene but is separated by 10 genes from the MCP gene and no TK function is encoded in its vicinity (Chee et al., 1990).

Southern hybridisation and sequencing studies have indicated that the UL24-gH-TK-x-x-MCP gene block is located at a colinear position within a wide range of alpha- and gammaherpesviruses (Davison and Wilkie, 1983; Davison and Taylor, 1987; Gompels et al., 1988a ; Cullinane et al., 1988 ). Thus in HVS (Gompels et al., 1988b), EHV-1 (Robertson and Whalley, 1988), and PRV and BHV-1 (Petrovskis et al., 1988), the gH gene homologue is located within the central region of the $U_L$

component and is positioned immediately downstream of the thymidine kinase gene. Having previously located the thymidine kinase gene of EHV-4 by virtue of colinearity of the EHV-4, EHV-1 and HSV-1 genomes (Nicolson et al., 1990b) we sequenced downstream of this gene in order to search for an open reading frame with the capacity to encode a glycoprotein H homologue. In this chapter the nucleotide sequence of the EHV-4 gH gene homologue is reported and its predicted amino acid sequence compared to that of other herpesvirus gH-type proteins as presented in Nicolson et al., 1990a.

## RESULTS

### 1. Nucleotide Sequence of EHV-4 genome - map units 0.49-0.51

The EHV-4 strain 1942 thymidine kinase (TK) gene maps within a 2kbp EcoRV/XhoI subfragment of the Bam HI C fragment in the long unique component of the genome at approximate map position 0.48 as reported in the previous chapter (Nicolson et al., 1990b). This corresponds to a position just upstream of the rightward XhoI site of BamHI C as represented in Figure 1a. The nucleotide sequence of 2.88kbp of DNA immediately downstream of the TK gene , determined according to the strategy outlined in Figure 1c, is detailed in Figure 2 . A typical sequence gel autoradiograph is presented in Figure 3.9 . Analysis of potential open reading frames (ORF) within all six frames (Fig. 1b) revealed a major open reading frame of 2565 base pairs, in transcriptional sense left to right, positioned 170bp downstream of the putative polyadenylation signal of the TK gene (AATAAA bp 9-14, Fig. 2). This ORF is predicted to encode a polypeptide possessing features characteristic of a membrane glycoprotein . Comparison of the predicted protein sequence of the gene to that of published herpesvirus glycoprotein sequences indicated similarity to the glycoprotein H of HSV-1 and to its homologue within other alpha, beta and gamma herpesviruses. The gene is thus hereafter identified as the EHV-4 gH gene pending identification of its protein product and review of equine herpesvirus glycoprotein gene and gene product nomenclature.

The DNA sequence upstream and downstream of this open reading frame was searched for putative gene control elements. Several TATA box homologues - consensus sequence TATA($^{A}_{T}$)A($^{T}_{A}$) (Corden et al., 1980) - are located upstream of the initiation codon. Two sequences with a particularly good fit to the consensus sequence are located at positions 87-93 (TATATAA) and 157-162 (TATATT). The polyadenylation

signal AAUAAA (Proudfoot and Brownlee, 1976) is not specified within 3′ non-coding DNA proximal to the termination codon, TAA, at bp 2750. However, the minor polyadenylation signals AUUAAA and AGUAAA (Birnstiel et al., 1985) are specified within the extreme 3′ terminal coding region of the gene. The sequence, TGTGTTGT (2770-2777) shares homology with the consensus element of McLauchlan et al. (1985), YGTGTTYY, often present 24-30 bp downstream of the polyadenylation signal element of HSV and mammalian mRNAs, and is located at an appropriate distance from the ATTAAA element at 2737-2742. However, since consensus sequence elements may occur fortuitously within DNA, we cannot rule out the possibility that the EHV-4 gH transcript is terminated considerably downstream of the gene, and that it may encode a contiguous gene.

The modified scanning hypothesis of translation (Kozak, 1984, 1986) proposes that the initiation codon of a transcript is the first ATG downstream of the mRNA start site which is accessible to the ribosomal translational apparatus. The ATG at bp 185-187, within the sequence CGATTATGT, is the first ATG downstream of the in-frame stop codon at bp116-118. Although the local sequence of this codon is only moderately homologous to Kozak's consensus sequence, CC($^{A}_{G}$)CCATGG, it does retain the purine at position -3, the most highly conserved feature of the 5′ flanking sequence of the initiation codon of eukaryotic mRNAs. Assignment of this codon as the initiation codon of the gH gene is supported by the potential of the DNA immediately downstream to encode a functional signal peptide domain.

The gene has a GC content of 45% which contrasts with the 67% GC content of the HSV-1 gH (McGeoch and Davison, 1986). Each of the 61 amino acid-specifying codons is used to encode the EHV-4 gH (Table 1) with no preference for codons with a third position G or C. In the

HSV-1 gH gene there is a clear bias towards codons possessing a 3′ G or C. This difference in the codon usage profiles of the two genes reflects the difference in the overall %GC content of the genes and, presumably, of the two viral genomes.

## 2. The EHV-4 gH Polypeptide

The EHV-4 gH gene encodes a protein 855 amino acids in length with a predicted molecular weight of 94 100. The predicted polypeptide has the potential to exhibit the following structural features characteristic of membrane glycoproteins as depicted in Figure 3 -

**i) N-terminal signal sequence peptide** - an N-terminal signal sequence is required for the targetting of membrane proteins to the endoplasmic reticulum, where the signal peptide is cleaved and the nascent protein undergoes processing (Walter and Lingappa, 1986). Signal sequences are characterised by an N-terminal basic region (n), a hydrophobic core (h) and a polar C-terminal region (c) followed by a signal peptidase cleavage site, usually near a predicted beta turn (Perlman and Halvorson, 1983; von Heijne, 1983). An N-terminal domain of hydrophobic character is evident from a hydropathic plot of the EHV-4 gH in which the local hydrophilicity of each residue is plotted (Fig. 4) (Hopp and Woods, 1981). Residues Tyr(5) to Ala(14) are predominantly hydrophobic in character and are likely to constitute the h region of the gH signal sequence. Application of the eukaryotic weight matrix developed by von Heijne (1986) for the prediction of signal sequence cleavage sites indicates that Ala(19) is the most probable terminal residue of the signal peptide . The predicted molecular weight of gH after cleavage of the signal peptide is 92130.

**ii) Hydrophilic external domain** - in membrane glycoproteins this major domain projects from the virion envelope or infected cell plasma

membrane and comprises the glycosylated portion of the polypeptide. It is therefore usual to find potential N-linked glycosylation sites (N-X-S/T) within this domain. Residues 20-816 may constitute the external domain of EHV-4 gH. Within this region there are 11 putative N-linked glycosylation sites, seven of which are located within the N-terminal half of the polypeptide and four in the C-terminal half (Fig. 3). Due to its exposure on the virion or cell surface this part of the protein is likely to contain the major antigenic determinants of gH with respect to the host humoral immune response and, in this case, possibly one or more complement-independent neutralisation epitopes as in the HSV-1, VZV, EBV, and HCMV gH-type polypeptides. Prediction of parts of the polypeptide which may be exposed on the surface of the native protein is aided by the analysis of the local hydropathic character of each residue and by prediction of the secondary structure of the protein using algorithms such as those of Hopp and Woods (1981), Chou and Fasman ( 1978) and Kyte and Doolittle (1982). Such analyses were applied to the EHV-4 gH using VAX and Microgenie software (Devereux et al., 1984; Queen and Korn, 1984) (Figures 4-6).

The region of maximum hydrophilicity in a hydropathic plot often corresponds to an antigenic determinant (Hopp and Woods, 1981 ). The hydrophilic peak of the EHV-4 gH plot occurs at amino acids 759-763 (Fig. 4), within a region of the polypeptide predicted to adopt an alpha helical conformation. We might therefore expect an epitope to map to this portion of gH given that it is likely to be oriented on the surface of the protein in an aqueous environment. Other algorithms result in a different position for the hydrophilic peak (Fig. 5) although, in gross terms the plots are similar. In terms of gross antigenicity, analyses predict that the N-terminal and central regions of the glycoprotein are more immunogenic than the C-terminal region

with amino acids 20-200 and 300-500 possessing the highest scores in surface accessibility and antigenicity predictions (Fig. 6b,c).

The HSV-1 gH extracellular domain has a potential membrane associated region located close to the transmembrane domain sequence (Gompels and Minson, 1986) , amino acids 780-803 , which contains charged residues but which is enriched in hydrophobic residues. EBV gp85 also has a potential membrane-associated region spanning amino acids 538-554 (Heineman et al., 1988). With the exception of the transmembrane domain and signal peptide no other regions of significant enrichment of hydrophobic residues were identified in EHV-4 gH.

**iii) Hydrophobic transmembrane domain** - a characteristic feature of membrane glycoproteins is the existence of a stretch of around 20 or more hydrophobic amino acids towards the C terminus within which polar amino acids such as Ser and Thr may be tolerated but not charged residues such as Arg or Glu (Wiley, 1985). This has been hypothesised to constitute the membrane spanning portion of the glycoprotein and may consist of one or more alpha helical loops traversing the virion envelope or infected cell membrane. Amino acids 817-836 of the EHV-4 gH are enriched in hydrophobic residues (Fig. 4) and are predicted to constitute the gH transmembrane domain comprising a single traverse of the membrane.

**iv) Cytoplasmic domain** - the proposed cytoplasmic domain of EHV-4 gH, stretches from amino acids 837-855 . Residues are largely hydrophilic in character and the domain possesses an overall charge of -1. A proposed function of this domain is to anchor the transmembrane domain within the lipid bilayer and thus to fix the C-terminal end of the glycoprotein to the membrane.

Comparison of the predicted EHV-4 gH polypeptide to other gH-type glycoproteins revealed identities of 85%, 32%, 26%, 18%, and 17% with

the gH-type glycoproteins of EHV-1, VZV, HSV-1, HCMV, and HVS respectively . Dot matrix comparisons of EHV-4 gH with the gH-type glycoproteins of HSV-1 , VZV, HCMV, and EBV are presented in Figure 7. The C-terminal region of the protein is clearly conserved to a greater degree than the N-terminal region. Previous comparison of the amino acid sequence of the gH proteins of alpha, beta and gamma herpesviruses by Gompels et al. (1988b) and Cranage et al. (1988) had demonstrated the greater conservation of sequence in the C-terminal region of the protein and highlighted several features of the gH protein conserved throughout the herpesvirus family - i) an unusually short cytoplasmic domain of 14 or 15 amino acids in alphaherpesviruses and of 7 or 8 amino acids in beta and gammaherpesviruses, ii) four conserved cysteine residues at similar positions relative to the putative transmembrane domain and within conserved local sequence, and iii) a conserved glycosylation site sequence NGTV 13-18 amino acids N-terminal to the transmembrane domain.

EHV-4 gH exhibits all the above features as shown in an alignment of herpesvirus gH-type glycoproteins reported to date (Fig. 8): the proposed cytoplasmic domain is under 20 amino acids in length, the four conserved cysteines are present at positions 556, 591, 663 and 716, and the C-terminal glycosylation site is located within the sequence NGTV (amino acids 796-799) which is positioned 19 amino acids N-terminal to the putative EHV-4 transmembrane domain. A further site of cysteine conservation between the EHV-4 and HSV-1 gHs occurs within the N-terminal portion of the protein within the conserved sequence GCD (amino acids 245-247 in EHV-4 , 257-259 in HSV-1 ). The Cys residues at 737 and 740 in the EHV-4 gH occur at sites of cysteine conservation throughout most herpesvirus gHs, with the exception of HSV-1. It is clear from the multiple alignment that conservation of sequence is

limited to two parts of the polypeptide - a short domain of limited conservation within the N-terminal half of the protein and a larger domain commencing at the first site of cysteine conservation and ending at the conserved C-terminal NGTV sequence. Between these conserved regions there is no apparent identity between the beta/gammaherpesvirus gH group and the alphaherpesvirus gH group and intergroup alignment is arbritary. Of the 11 putative glycosylation sites of EHV-4 gH and the 7 of HSV-1, only three are conserved, one within the N-terminal portion and two within the C-terminal portion. Conservation of N-X-S/T sites throughout the gHs is apparent only at two sites within the conserved C-terminal domain with a further site of conservation within only beta and gammaherpesvirus gHs.

Figure 4.1

Localisation of the EHV-4 gH Gene Homologue

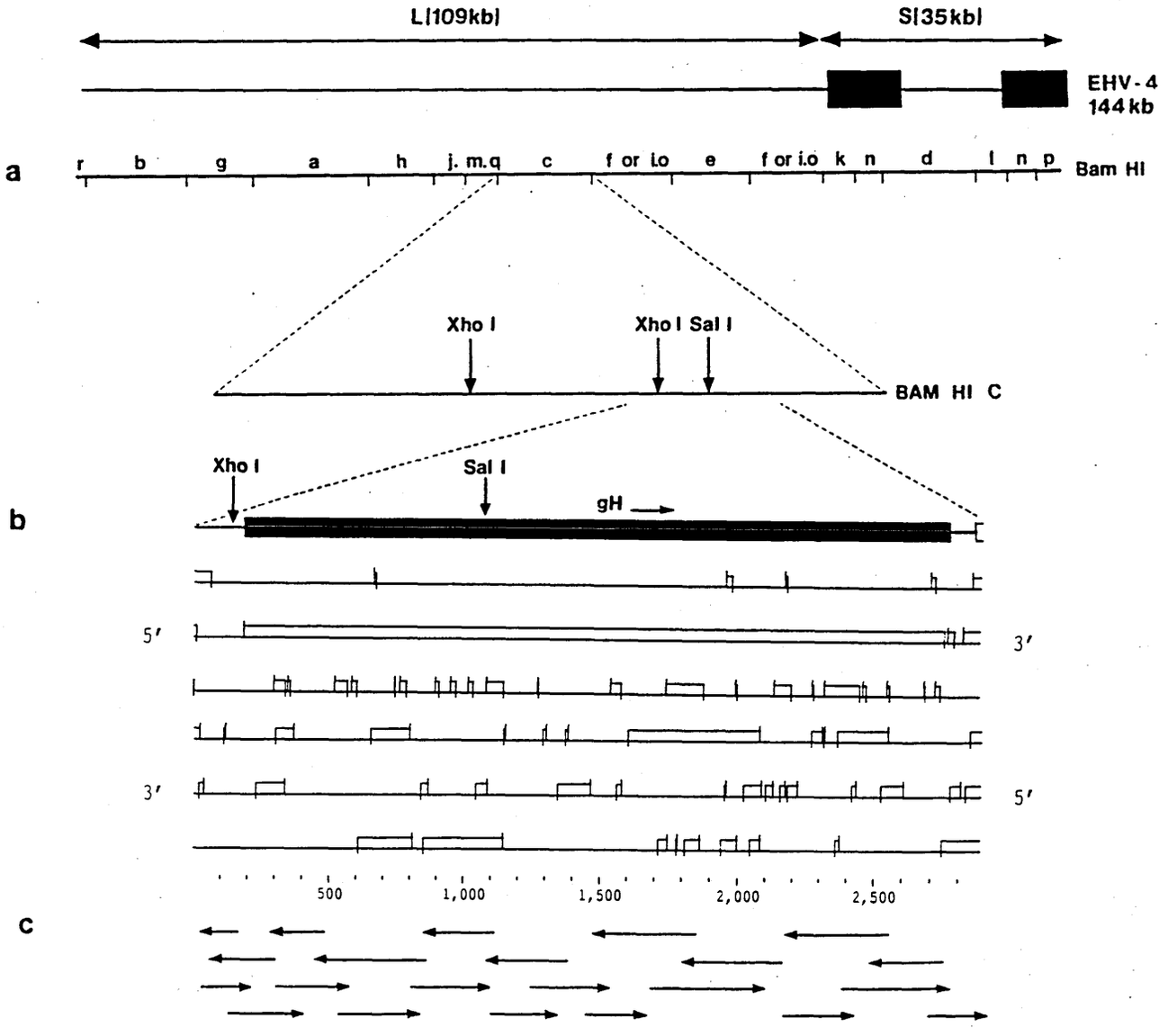



Figure 4.1 a) BamHI restriction map of EHV-4 genome (1942 strain: Cullinane et al., 1988) with expansion of BamHI C showing positions of XhoI and SalI restriction sites. b) Analysis of potential open reading frames within a 2.88 kbp region of BamHI C. The major open reading frame was identified as a gH gene homologue. c) Sequencing strategy.

Figure 4.2 Nucleotide sequence of 2.88 kbp EHV-4 DNA mapping immediately downstream of the TK gene. The predicted amino acid sequence of the product of a major open reading frame, EHV-4 gH, is indicated below the coding sequence. Putative polyadenylation signals (PS) and TATA (T) boxes are underlined as are potential N-linked glycosylation sites within the gH polypeptide.

# Figure 4.2

## Nucleotide Sequence of the EHV-4 gH Gene Homologue

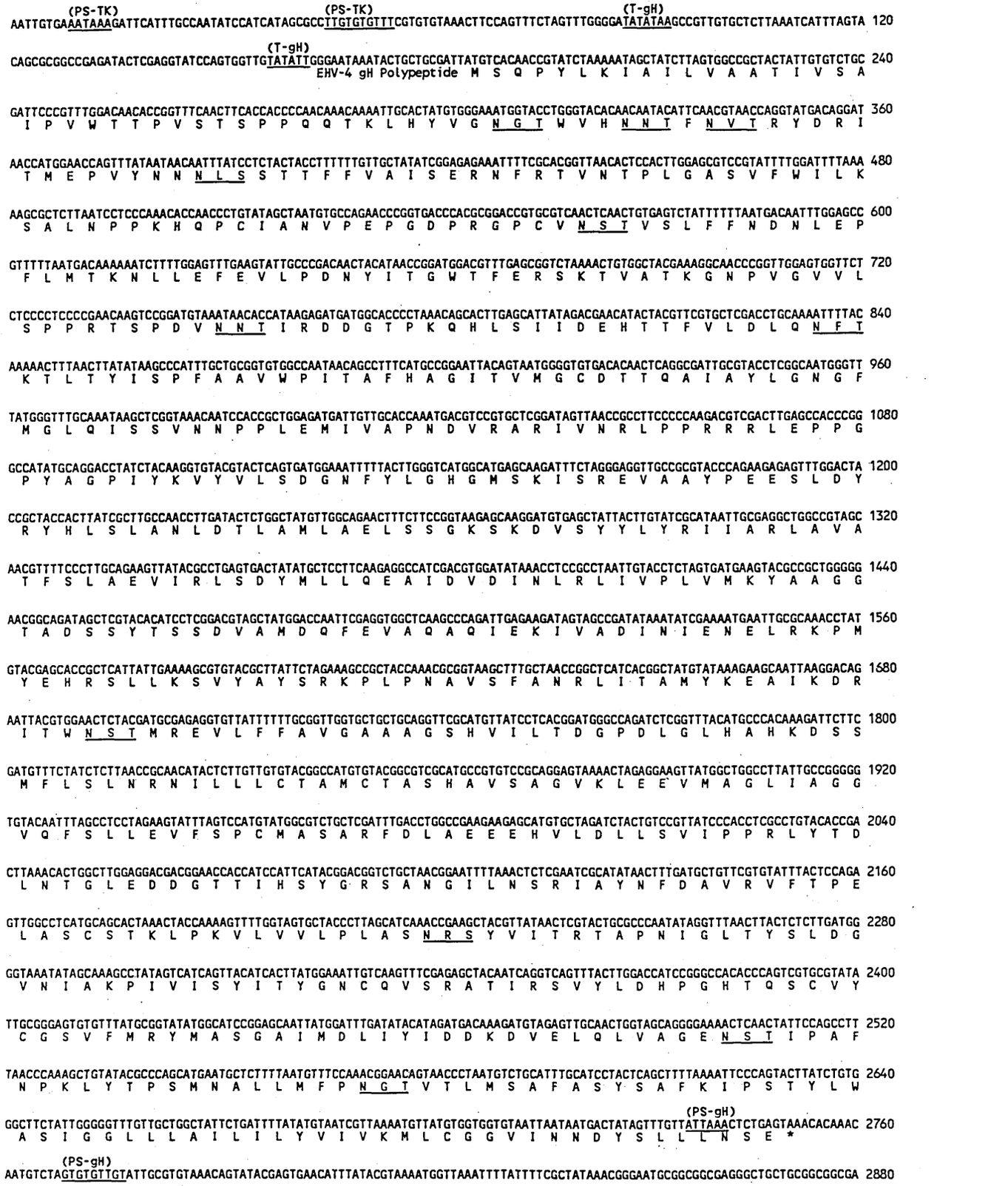


Figure 4.3

Putative Structural Domains within the EHV-4 gH Polypeptide

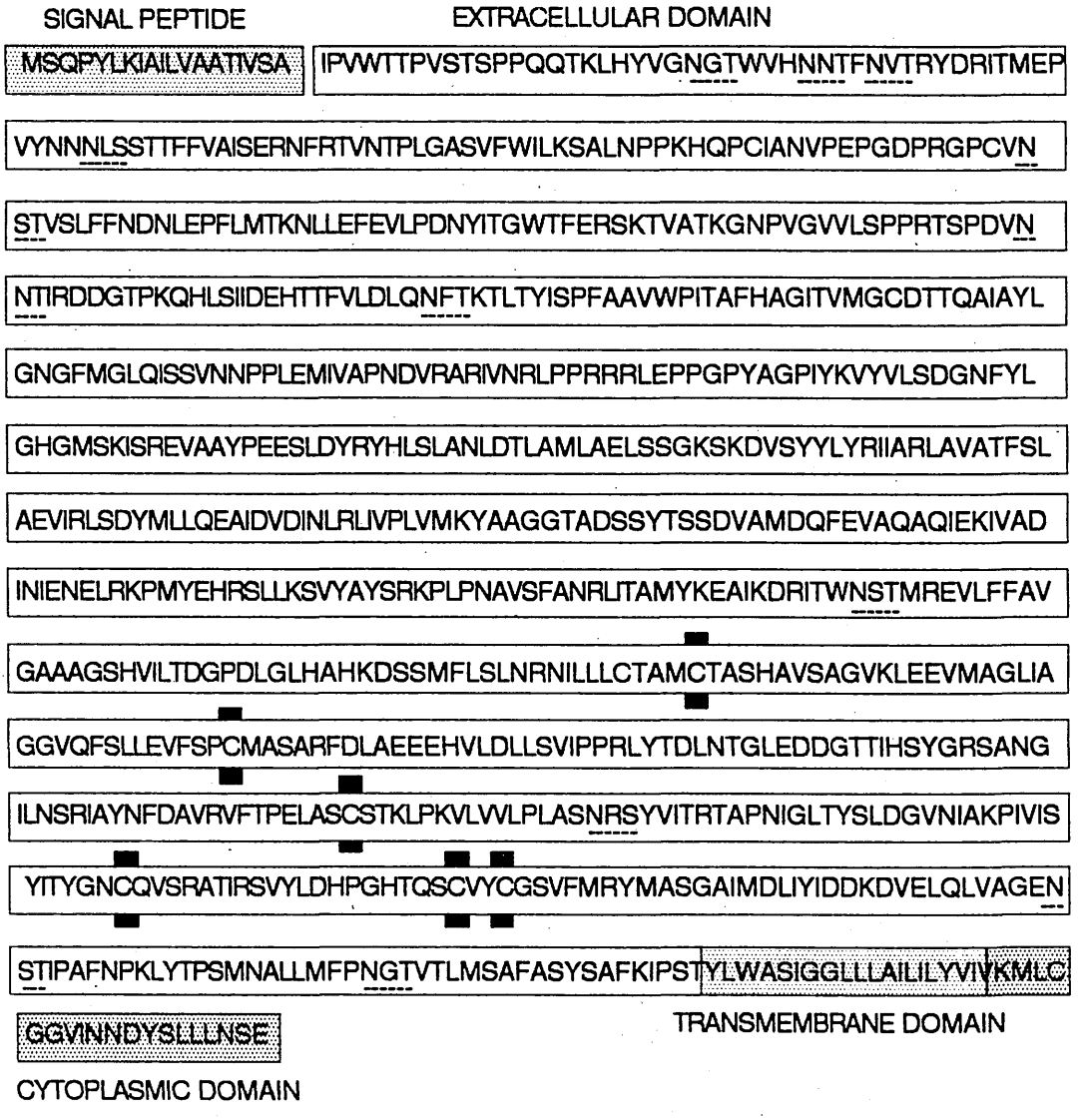



Figure 4.3 Predicted distribution of four structural domains within EHV-4 gH on the basis of analyses of local hydrophobicity scores along the polypeptide (Hopp and Woods, 1981).

Figure 4.4

Hydropathic Plot of the EHV-4 gH Polypeptide

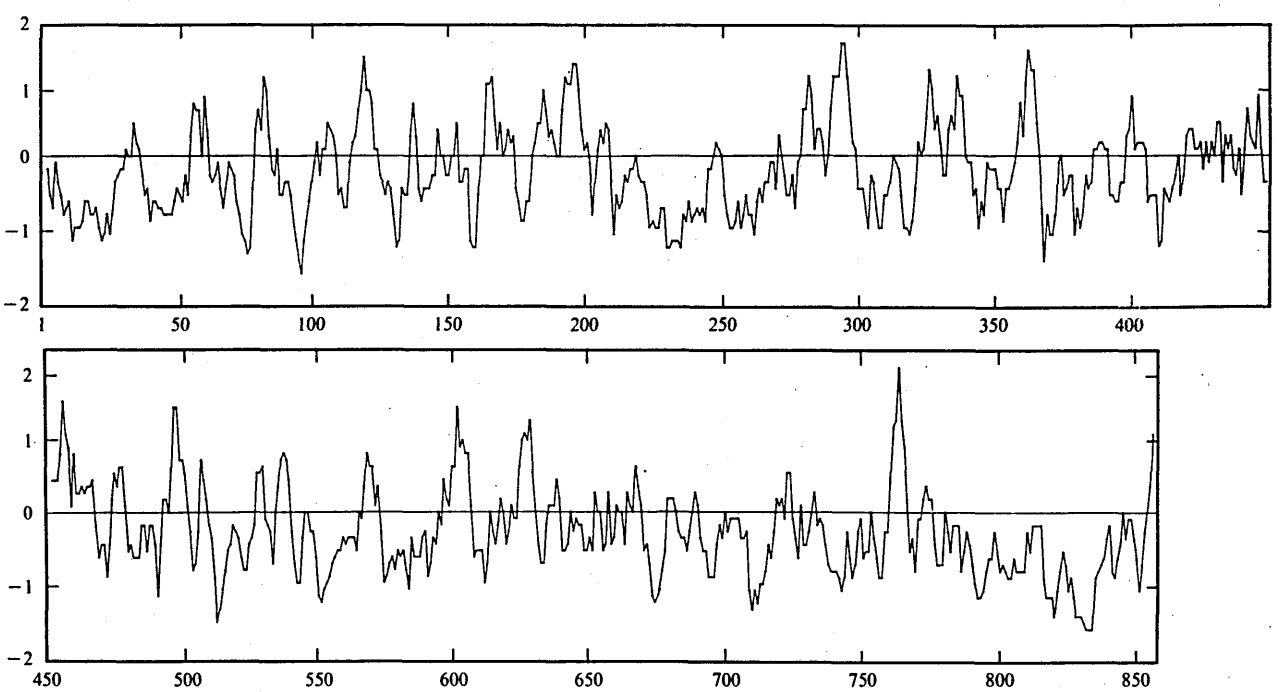



Figure 4.4 Hopp and Woods (1981) hydropathic plot comprising analysis of the local hydrophilicity of each residue of the EHV-4 gH polypeptide using a six amino acid window. Positive values indicate a local hydrophilic environment: negative values a local hydrophobic environment.

Figure 4.5

Secondary Structure Analysis of EHV-4 gH

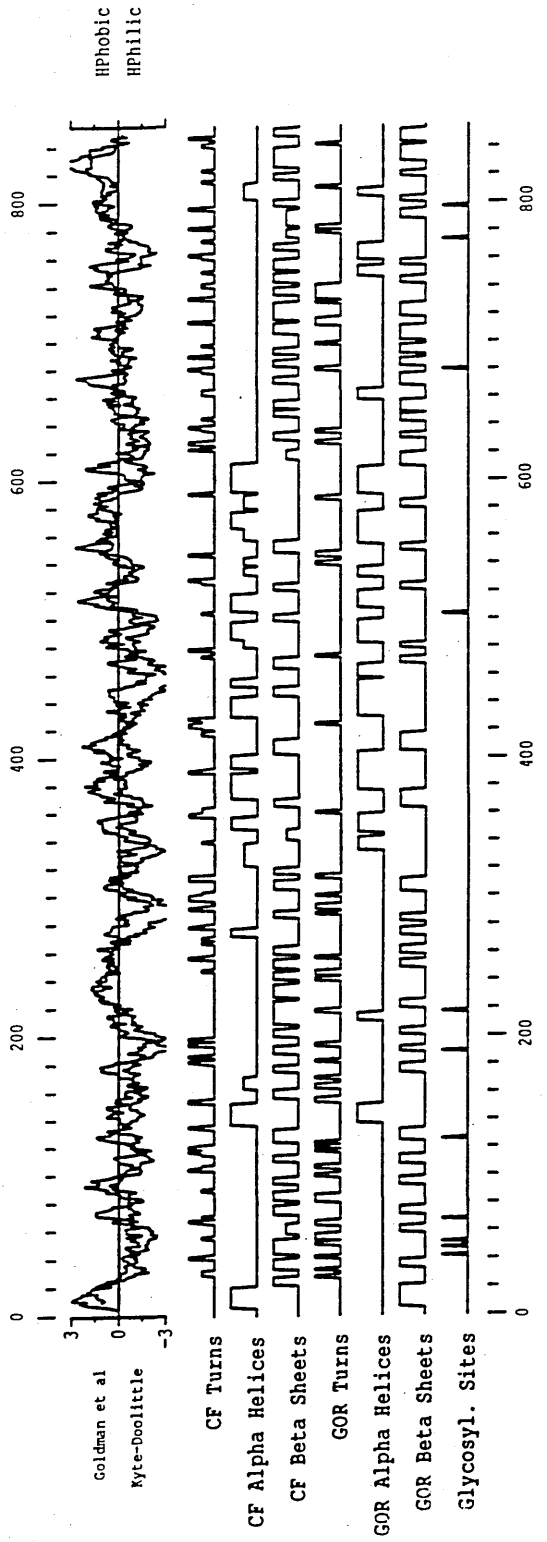



Figure 4.5 Hydropathic and secondary structure analyses of EHV-4 gH using algorithms available on VAX software (Devereux et al., 1984).

Figure 4.6 Analysis of EHV-4 gH using VAX software (Devereux et al., 1984). Overlaid on Chou-Fasman secondary structure predictions (a)) are motifs indicating regions of the polypeptide with high b) surface accessibility, and c) antigenicity scores.

Figure 4.6

Schematic Predictions of EHV-4 gH Secondary Structure, Surface Accessibility, and Antigenicity.

Figure 4.7

Dot Matrix Plots of Herpesvirus gH Polypeptides

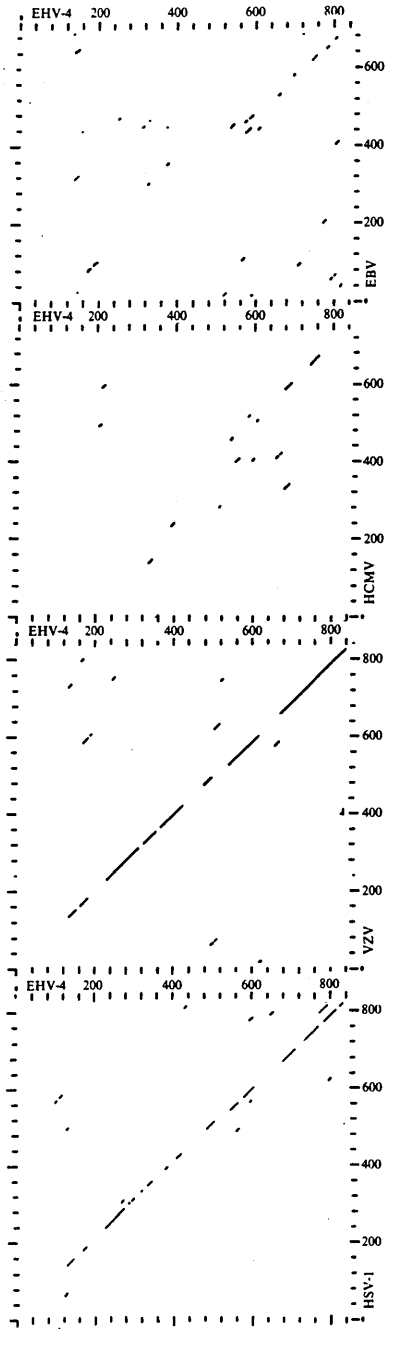



Figure 4.7 Comparison of EHV-4 gH with its counterparts in HSV-1 (McGeoch and Davison, 1986), VZV (Keller et al., 1987), and EBV (Heineman et al., 1988) by dot matrix analyses. Parameters used were a 30 amino acid window and a stringency of 53%.

# Figure 4.8

## Multiple Alignment of Herpesvirus gH-type Polypeptides

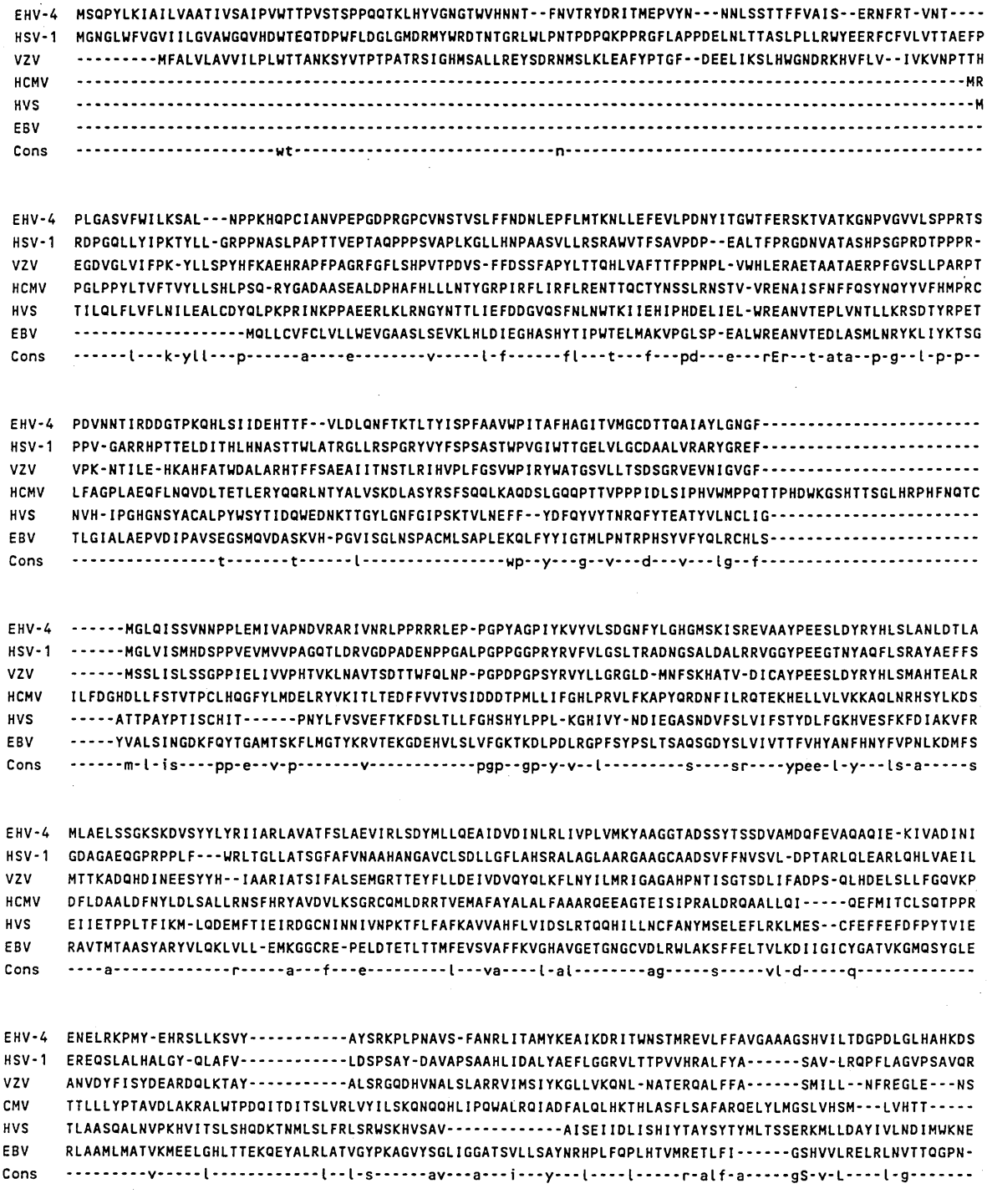


Figure 4.8 (cont.)

```
EHV-4  SMFLSLNRNILLLCTAMCTASHAVSAGVKLEEVMAGLIAGGVQFSLLEVFSPCMASARFDLAEEEHVLDLLSVIPPRLYTDLNTGLEDDGTTIHSYGRSA
HSV-1  ERA----RRSLLIASALCTSDVAAATNADLRTALARADHQKTLFWLPDHFSPCAASLRFDLDESVFILDALAQATRSETPVEVLAQQTHGLAST------
VZV    SRVLDG-RTTLLLMTSMCTAAHATQAALNIQEGLAYLNPSKHMFTIPNVYSPCMGSLRTDLTEEIHVMNLLSAIPTRPGLNEVLHTQLDESEIFDAAFKT
CMV    -----ERREIFIVETGLCSLAELS-------HFTQLLAHPHHEYLSDL-YTPCSSSGRRDHSLERLTRLFPDATVPATVPAALSILSTMQPSTLETFPDL
HVS    TVKKQDLLPY-VLSSSMCTSLEIG-------NLLL-HFGQKDVLDVYETFSPCYLSLRFDFTKEKLITEFPQSSLIAQKEINLGTNGFFQTLHMRHHTSL
EBV    ------LALYQLLSTALCSALEIG-------EVLRGLALGTESGL----FSPCYLSLRFDLTRDKLLSMAPQEATLDQAAVSNAVDGFLGRLSLERED--
Cons   -------R--lll-TamCta-ea--------e-la-l---k--f-----fSPC--SlRfDltee-l----l-----------------g----------

EHV-4  NGILNSRIAYNFDAVRVFTPELAS-CSTKLP---KVLVVLPLASNRSYVITRTAPNIGLTYSL-DGVNIAKPIVISYIT-YGNCQVSRATIRSVYLDHPG
HSV-1  ---LTRWAHYNALI-RA-VPEASHRCGGQSANVEPRILVPITH-NASYVVTHSPLPRGIGYKL-TGVDVRRPLFLTYLT--ATCEGSTRDIESKRLVRTQ
VZV    MMIFTTWTAKDLHILHTHVPEVFT-CQDAAARNGEYVLILPAVQGHSYVITRNKPQRGLVYSL-ADVDVYNPISVVYLS-RDTCVSEHGVIETVALPHPD
HCMV   ----------------------FCLPLGESFSALTVSEHV----SYVVTNQYLIKGISYPV-STTVVGQSLIITQTDSQTKCELTRNMHTTHSITAAL
HVS    EILPII-----------------KCIKSLS--TDIILSIPLK-NITYVISTKPVPNSKIYDV-SEVFLKTSMIISAV--NNDCKPYQGGSAAHQIPVIY
EBV    ---------------RDAWHLPAYKCVDRLD---KVLMIIPLI-NVTFIISSDREVRGSALYEASTTYLSSSLFLSPV-IMNKCSQGAVAGEPRQIPKIQ
Cons   ---------------r---pe----C---l-------l--pl--n-sYVit-----rG--Y-l-s-v-v--pl-isy------C------ie---lp---

EHV-4  HTQS----CVYCGSVFMRYMASGAIMDLIYIDDKDVELQLVAGENSTIPAFNPKLYTPSMNALLMFPNGTVTLMSAFASYSAFKIPSTYLWASIGGLLLA
HSV-1  NQRD----LGLVGAVFMRYTPAGEVMSVLLVDTDNTQQQIAAGPTEGAPSVFSSDVP--STALLLFPNGTVIHLLAFDTQPVAAIAPGFLAASALGVVMI
VZV    NLKE----CLYCGSVFLRYLTTGAIMDIIIIDSKDTERQLAAMGNSTIPPFNPDMHGDDSKAVLLFPNGTVVTLLGFERRQAIRMSGQYLGASLGGAFLA
HCMV   NISLEN--CAFCQSALLEYDDTQGVINIMYMHDSDDVLFALDPYNEVVVSSPRTHY------LMLLKNGTVLEVTDVVVDATDSRLLMMSVYALSAIIGI
HVS    NVTVPRRGCPYCSSVVLSYDESQGFQSMMYITDTYVQENLFT-EHS--PFFGDGNLHIHY--LILMNNGTVIEVRGAYRARLVN-FIIVIMVFILFLVGL
EBV    NFTRTQKSCIFCGFALLSYDEKEGLETTTYITSQEVQNSILSSN-----YFDFDNLHVHY--LLLTTNGTVMEIAGLYEERAHV-VLAIILYFIAFALGI
Cons   N-------C-yCgsvflrYd--gg-m---yidd-dvq-ql-a--ns--p-f----------aLlLfpNGTV-e--gf------------l-asi-g--g-

EHV-4  ILILYVIVKMLCGGVINNDYSLLLNSE* 855
HSV-1  TAALAGILKVLRTSVPFFWRRE*     838
VZV    VVGFGIIGWMLCGNSRLREYNKIPLT* 841
HCMV   YLLYRMLKTC*                  743
HVS    YLLYKLFVYLT*                 717
EBV    FLVHKIVMFFL*                 707
       -l----i---l
```

Figure 4.8 Multiple alignment of the predicted product of the HSV-1 gH gene (McGeoch and Davison, 1986) and its counterpart in EHV-4, VZV (Keller et al., 1987), HCMV (Cranage et al., 1988), HVS (Gompels et al., 1988b), and EBV (Baer et al., 1984). A consensus line is detailed in which residues conserved in 3 or 4 'gH's are represented by lower case, residues conserved in 5 or 6 'gH's by upper case.

## Table 4.1

## Codon Usage of EHV-4 gH Gene Homologue

| 1st \ 2nd | A (26%) | C (27%) | G (15%) | T (32%) | 3rd | Residue | f | (%) |
|---|---|---|---|---|---|---|---|---|
| A | Lys 19 | Thr 13 | Arg 6 | Ile 23 | A (25%) | Ala | 72 | ( 8.4) |
| | Asn 26 | Thr 12 | Ser 14 | Ile 13 | C (23%) | Arg | 37 | ( 4.3) |
| (30%) | Lys 10 | Thr 13 | Arg 5 | Met 25 | G (22%) | Asn | 50 | ( 5.8) |
| | Asn 24 | Thr 20 | Ser 10 | Ile 22 | T (30%) | Asp | 39 | ( 4.6) |
| | | | | | | Cys | 11 | ( 1.3) |
| C | Gln 12 | Pro 22 | Arg 6 | Leu 13 | A | Gln | 16 | ( 1.9) |
| | His 8 | Pro 13 | Arg 11 | Leu 11 | C | Glu | 34 | ( 4.0) |
| (19%) | Gln 4 | Pro 9 | Arg 5 | Leu 13 | G | Gly | 46 | ( 5.4) |
| | His 9 | Pro 8 | Arg 4 | Leu 14 | T | His | 17 | ( 2.0) |
| | | | | | | Ile | 58 | ( 6.8) |
| G | Glu 15 | Ala 17 | Gly 16 | Val 25 | A | Leu | 92 | (10.8) |
| | Asp 21 | Ala 17 | Gly 7 | Val 3 | C | Lys | 29 | ( 3.4) |
| (30%) | Glu 19 | Ala 12 | Gly 10 | Val 20 | G | Met | 25 | ( 2.9) |
| | Asp 18 | Ala 26 | Gly 13 | Val 21 | T | Phe | 32 | ( 3.7) |
| | | | | | | Pro | 52 | ( 6.1) |
| T | End 0 | Ser 11 | End 0 | Leu 18 | A | Ser | 71 | ( 8.3) |
| | Tyr 21 | Ser 10 | Cys 4 | Phe 3 | C | Thr | 58 | ( 6.8) |
| (21%) | End 0 | Ser 10 | Trp 7 | Leu 23 | G | Trp | 7 | ( 0.8) |
| | Tyr 19 | Ser 16 | Cys 7 | Phe 29 | T | Tyr | 40 | ( 4.7) |
| | | | | | | Val | 69 | ( 8.1) |

Table 4.1 Codon usage of the EHV-4 gH gene homologue and amino acid frequencies within its predicted product. The frequencies of each base within codon positions 1,2, and 3 are presented as percentages.

## DISCUSSION

The gH gene of EHV-4 maps immediately downstream of the TK gene and specifies a protein product of 855 amino acids in length. On the basis of sequence data , the EHV-1 gH mRNA transcript termination is signalled by a standard AATAAA polyadenylation signal sited approximately 80 bp downstream of the termination codon (Robertson et al., 1991). A consensus polyadenylation signal is not located at a corresponding position within EHV-4 although a match to the less common ATTAAA polyadenylation signal occurs within the extreme 3' end of the coding region of the EHV-4 gH gene. It is possible, therefore, that the fine details of transcription of the EHV-1 and EHV-4 gH genes differ.

With the exception of pseudorabies virus gH which is only 686 amino acids in length (Petrovskis et al., 1988), the gHs of alpha herpesviruses tend to be larger than their beta and gamma herpesvirus counterparts ranging from 838 (HSV-1) to 882 (BHV-1) as compared to 706-743 for the EBV, HVS and HCMV gH-type polypeptides. It has been hypothesised that this length difference is a consequence of expansion of the N-terminal region of alphaherpesvirus 'gH's and/or a contraction of beta/gamma herpesvirus 'gH's (Gompels et al., 1988b). At 855 amino acids in length the EHV-4 gH is of comparable size to most of the alphaherpesvirus gHs and is only 7 amino acids longer than the EHV-1 gH . The predicted amino acid sequence of the EHV-4 gH shares identity with the gH-type glycoproteins of EHV-1, HSV-1, VZV (gplll), HCMV (gp86), HVS, and EBV (gp85). The identity is greater with the 'gH's of other members of the alphaherpesvirus subgroup - 85% with EHV-1, 26% with HSV-1 and 32% with VZV - than with those of members of the beta and gamma subgroups - 17% with HCMV and 18% with HVS and EBV. That the gH of EHV-4 is more closely related to VZV gplll

than to HSV-1 gH is consistent with the phylogenetic tree of herpesvirus evolution constructed by Whalley et al. (1989) on the basis of the relatedness of the gB glycoproteins of several alpha, beta and gamma herpesviruses (Chapter 1; Fig. 3c). Conservation in N-terminal region of gH-type proteins tends to be limited to gH proteins from the same herpesvirus subgroup whereas conservation in C-terminal region extends across 'gH's of different subgroups.

The strong conservation of gH throughout the three herpesvirus subgroups implies it has an essential function and that functional constraints have restricted the divergence of its C-terminal primary amino acid sequence in the course of herpesvirus evolution. The strong conservation of cysteine residues throughout the alpha, beta and gamma herpesvirus gHs investigated implies some degree of conservation of the secondary and tertiary structure of these proteins presumably involving disulphide bonding (Gompels et al., 1988b) which is essential to this/these conserved function(s).

A tryptophan residue at 232 in EHV-4 gH coaligns with $Trp_{244}$ of the HSV-1 gH which is the site of a ts mutation which renders progeny virus noninfectious (Desai et al., 1988). This Trp residue is conserved in other alphaherpesvirus gHs including EHV-4 and EHV-1 . Due to the divergence of N-terminal beta/gammaherpesvirus gH sequences it is not possible to speculate whether a Trp residue might be critical for beta/gammaherpesvirus gH function : alignment of beta/gammaherpesvirus gHs with alphaherpesvirus gHs across this region of the polypeptide is arbitrary. If a functionally important Trp is conserved in beta/gammaherpesvirus gHs it is clearly not situated within a block of conserved residues.

From studies on other herpesvirus gHs glycosylation includes complex and high-mannose N-linked oligosaccharide chains ( Edson and

Thorley-Lawson, 1983; Strnad et al., 1983). Comparison of the apparent molecular weight of mature gH as isolated from virion extracts to the unglycosylated molecular weight predicted from the amino acid sequence indicates that glycosylation is not as extensive as that for gC-type glycoproteins. For example the apparent molecular weight of HSV-1 gH is 110 kDa (Showalter et al., 1981; Buckmaster et al., 1984) whereas that predicted for the amino acid sequence minus signal sequence is 88.5 kDa (McGeoch and Davison, 1986).

It remains to be determined whether the native, processed EHV-4 gH corresponds to any of the previously reported EHV-4 glycoproteins. The only feasible candidate is a 138K protein identified in EHV-4 (Meredith et al., 1989). It is possible that, as is the case with gH of HSV-1, EHV-4 gH is undetectable by SDS-PAGE due to low abundance within the virion envelope and low incorporation of radioactively labelled sugars such as N-acetylglucosamine .

Several approaches could be utilised to identify the protein product of the gH gene . Antibodies generated against a peptide sequence derived from the predicted gH protein product could be used to immunoprecipitate the protein from an EHV-4 protein preparation. Alternatively antibodies raised against a suspected EHV-4 gH protein could be tested for their ability to immunoprecipitate an in vitro translation product from mRNA selected by the gH ORF DNA. The former technique was successfully used to confirm identity of the VZV gpIII gene product (Keller et al., 1987) and of the EBV BXLF2 product (Oba and Hutt-Fletcher, 1988). The latter technique produced evidence linking EBV gene BXLF2 with protein gp85 ( Heineman et al., 1988; Seibl and Wolf, 1985). A further technique utilised was the transfection of the gH ORF into cells and testing of the expressed protein for reactivity with a monoclonal antibody against the suspected protein

product (Heineman et al., 1988). It is anticipated that application of such techniques to the investigation of the EHV-4 gH gene product should lead to the detection of a novel EHV-4 glycoprotein in infected cell extracts.

# CHAPTER 5

# DIRECT REPEAT ELEMENTS AND PUTATIVE ORIGIN OF DNA REPLICATION OF EHV-4

## INTRODUCTION

The genomes of different types of herpesvirus are distinct from each other in terms of size, base composition, repeat element distribution, coding characteristics, and cis-acting element distribution. Nucleotide sequence analysis of non-coding DNA immediately downstream of the EHV-4 gH gene revealed two features identified in the genomes of other sequenced herpesviruses : direct repeat elements and a putative origin of replication.

Alphaherpesvirus genomes possess regions of short direct repeats arranged in tandem at apparently random positions throughout the genome in addition to major repeat elements the arrangement of which determines herpesvirus genome classification (Chapter 1). The most thorough mapping of direct repeat elements exists for the completely sequenced HSV-1, VZV, and EBV genomes ( Baer et al., 1984 ; Davison and Scott, 1986; McGeoch et al., 1988a ) with the caveat that some elements may have deleted on cloning (Straus et al., 1983). It has been hypothesised that the elements are parasitic sequences which have accumulated, possibly by recombination, in coding and non-coding regions of the genome where their presence does not cause a selective disadvantage to the virus (Davison and Scott, 1986). The only reported functions for a set of such repeats in non-coding DNA of alphaherpesviruses are site-specific recombination (Chou and Roizman, 1985) and a role in maturation of unit sized genomic DNA (Hammerschmidt et al, 1988). Direct repeat regions within HCMV and EBV have been associated with enhancer activity ( Sugden and Warren, 1989; Weston, 1988). However, no enhancer activity has been reported within alphaherpesviruses . The only published EHV-4 minor direct repeat element is a region of 31 direct repeats of an 8bp sequence which has been mapped to the $TR_S$ component of the genome (Cullinane et al.,

1988).

Herpesvirus origins of replication are sites within the genome at which DNA replication is initiated. These origins are, like cleavage and packaging signals and regulatory elements such as promoters and enhancers, cis-acting elements the distribution of which varies between different herpesvirus genomes . The number and approximate sites of the origins of HSV-1 was originally determined by electron microscopy studies of replicating viral DNA (Friedmann et al, 1977, Hirsch et al., 1977) and by analysis of the genetic composition of replicative defective genomes (Vlazny and Frenkel, 1981). Fine mapping of the origins was subsequently determined by the analysis of the capacity of genomic fragments to support replication of plasmids in cells permissive for viral replication (Stow, 1982). This technique has become a standard approach to mapping novel functional herpesvirus origins of DNA replication.

The replication origins of five herpesviruses have been characterised and sequenced. These include the following - EHV-1 $ori_S$ (Baumann et al., 1989), HSV-1 $ori_S$ (Stow, 1982; Stow and McMonagle, 1983), HSV-1 $ori_L$ (Gray and Kaerner, 1984; Weller et al., 1985 ), HSV-2 $ori_S$ (Whitton and Clements, 1984), HSV-2 $ori_L$ (Lockshon and Galloway, 1986), VZV $ori_S$ (Stow and Davison, 1986), and EBV ori-P (Baer et al., 1984; Yates et al., 1984). The origins located within the S components of HSV-1, HSV-2, and VZV are positioned within the inverted repeats and are thus present in two copies per genome. All herpesvirus origins of replication reported to date exhibit dyad symmetry. However, the internal structure of the alphaherpesvirus and gammaherpesvirus origins differ.

Alphaherpesvirus origins are characterised by a palindromic sequence element of variable length the central region of which is rich

in A/T residues. To the left of the poly A/T tract is a conserved sequence element CGTTCGCAC, termed site I, which is located either completely or partially within the palindrome (EHV-1 $ori_S$, HSV-1 and HSV-2 $ori_S$ and $ori_L$) or outwith the palindrome (VZV $ori_S$) (Baumann et al., 1989; Stow and Davison, 1986; Weir and Stow, 1990). To the right, in all herpesvirus origins except VZV $ori_S$, is a region of DNA, termed site II, which is complementary in sequence to site I. Mutational analyses and binding studies have revealed that site I is critical for origin activity while site II and the poly A tract contribute to the efficiency of origin activity ( Deb and Deb, 1989; Stow and Davison, 1986; Weir and Stow, 1990). The conserved sequence, CGTTCGCAC of HSV-1 origins of replication is directly involved in the binding of the product of gene UL9 of HSV-1 ( Olivo et al., 1988), termed the origin binding protein (OBP) (Elias et al., 1986 ). The OBP binds to the sequence 5′-GT(G/T)CG-3′ which occurs within both strands of an 8bp binding domain in CGTTCGCAC (Koff and Tegtmeyer, 1988). Thus two OBP molecules bind asymmetrically to site I. This asymmetry of binding has been hypothesised to play a role in the uni-directionality of DNA replication possibly by restricting access of a multi-enzyme complex or melting events to one strand (Koff and Tegtmeyer, 1988). OBP molecules also associate with 5bp OBP-binding motifs within site II although binding affinity is significantly lower than that at site I.

In addition to a cis-acting origin of DNA replication, initiation and maintenance of DNA replication requires seven virus-specified products in HSV-1 including DNA polymerase and the major DNA binding protein (McGeoch et al., 1988b; Olivo et al., 1990; Wu et al., 1988). Although HSV-1 possess three origins of replication, genomes from which $ori_L$ and one copy of $ori_S$ have been deleted are replication competent and specify biologically active progeny virions ( Longnecker and

Roizman, 1986; Polvino-Bodnar et al., 1987).

Whereas the alphaherpesvirus origins of replication identified to date possess the discussed conserved features - poly A/T tract and CGTTCGCAC - the gammaherpesvirus EBV origin of replication , ori-P , has neither of these features suggesting that different mechanisms may be active in the initiation of alpha and gammaherpesvirus DNA replication (Yates et al., 1984). Two regions of the genome have been implicated in origin activity : a dyad symmetry element and, 1kbp upstream, a region of direct repeat elements which also act as a transcriptional enhancer (Reisman and Sugden, 1986; Reisman et al., 1985).

This chapter presents the nucleotide sequence of 1725bp of DNA mapping downstream of the EHV-4 gH gene termination codon and compares features of this sequence to similar features in other alphaherpesviruses.

## RESULTS

### 1. Sequence Analysis of EHV-4 Genome, Map Units 0.51-0.52

The nucleotide sequence of the extreme right end of EHV-4 BamHI C , determined according to the strategy presented in Figure 1, is detailed in Figure 2. Analysis of the nucleotide sequence of EHV-4 DNA mapping immediately downstream of the EHV-4 gH gene revealed three features of interest - i) a region of tandemly repeated direct repeat elements, ii) a putative origin of DNA replication, and iii) an open reading frame (Fig. 1). Sequence analysis of a limited region of the repeat region is presented in Figure 3.

**1) EHV-4 direct repeat elements**

The region of direct repeat elements maps 101 bp downstream of the gH termination codon and spans 423 base pairs (bp 2853 to 3275, Fig. 2). It comprises 20 complete repeat elements and a 3′ terminal partial repeat sequence. The complete elements range in size from 15 to 24bp and possess a conserved 15bp G/C rich core sequence -GCGGCGGCGAGGGCT - flanked by variable numbers of 5′ GCG elements and 3′ GCT elements such that the consensus repeat sequence is of the form - $(GCG)_{1-3}GCGGCGAGG(GCT)_{1-3}$ (Table 1). The %GC of the repeat region is 86.3% . In contrast, the %GC of the coding region of the gH gene is 45%.

A region of direct repeats is located at a similar position within the EHV-1 genome although the size and sequence of the elements is distinct . The repeat region is shorter than that of EHV-4 comprising two repeats of a 32bp element and a single 20 bp partial repeat (Robertson et al., 1991). The EHV-1 and EHV-4 repeat elements are, however, related in sequence as shown in Table 1. The EHV-4 repeat element bears little homology with the 8bp repeat located in EHV-4 $TR_S$ -GGAGGTGG- which is present in 31 complete copies and 1 partial copy with no variation repeat size (Cullinane et al., 1988).

The EHV-4 repeat elements are stringent multiples of three in size. An analysis of potential open reading frames within this part of the EHV-4 genome indicated two potential ORFs of 153 and 173 amino acids (Fig. 1b) spanning the repeats , 141 amino acids of which are encoded within the repeat region .

**2) Putative EHV-4 origin of DNA replication**

A putative EHV-4 origin of DNA replication was localised downstream of the repeat elements on the basis of similarity in sequence to that of reported HSV-1, HSV-2, VZV, and EHV-1 herpesvirus origins. The putative origin comprises an A/T rich region upstream of which is an 11bp element - 5′-CGTTCGCACTT-3′ - identical in sequence to an element conserved within all reported human herpesvirus origins and , across the 5′ 9bp , to one within EHV-1 $ori_S$ ( Baumann et al., 1989; Stow and Davison, 1986). A sequence element complementary to part of this conserved sequence, 5′-GTGCGAACG-3′, is positioned downstream of the EHV-4 poly A tract. A hypothetical stem and loop structure can thus be deduced for the putative origin given the extent of base pairing which might occur between the complementary flanking elements and within the A/T rich region . The hypothetical stem-and-loop structures of the putative EHV-4 $ori_L$, HSV-1 $ori_S$ , HSV-1 $ori_L$ , and EHV-1 $ori_S$ are presented in Figure 4. The EHV-4 palindrome differs from that of HSV-1 $ori_L$ in two major characteristics - i) it is considerably shorter, and ii) it possesses fewer potential intrastrand base pairing events. The EHV-4 $ori_L$ palindrome is thus apparently more similar in size to origins identified within the S components of herpesvirus genomes . It shares with EHV-1 $ori_S$ its lower degree of palindrome perfection in terms of potential base-pairing events.

An alignment of all functionally characterised origins of herpesvirus DNA replication and of the putative EHV-4 origin is

indicated in Figure 5. The sequence of the GT(T/G)CG motifs of characterised origins are compared to that of EHV-1 and EHV-4 in Table 2. With the exception of the VZV $ori_S$ , GT(T/G)CG motifs are present in both left and right arms of the palindrome. A difference in the second base exists within the motif in the right arm of the HSV-1 and HSV-2 $ori_S$ and has been speculated to account for the much lower affinity of OBP for the right end of the HSV-1 palindrome (Elias and Lehman, 1988; Koff and Tegtmeyer, 1988). The putative OBP-binding sites of the EHV-4 origin possess the sequences 5′-GTTCG-3′ (upper strand) and 5′-GTGCG-3′ (lower strand ) (left arm), and 5′-GTGCG-3′ (upper strand) and 5′-GTTCG-3′ (lower strand) (right arm). These motifs are identical to those within the EHV-1 $ori_S$ thus neither EHV origin has a non-consensus second base. The two flanking bases of the motif and the third position residue are both purines or pyrimidines (Koff and Tegtmeyer, 1988) . Since the 3rd position base and flanking residues are different in the upper and lower strand binding sites this may be an important determinant of the unidirectionally of DNA synthesis from the origin site. The putative EHV-4 $ori_L$ motifs are consistent with this observation : third base T is coupled to flanking CC and third base G to flanking AA or GA.

**3) Open reading frame**

Analysis of the coding capacity of all six reading frames within DNA downstream of the gH gene indicated that the C-terminal coding region of a gene maps within the extreme right hand end of BamHI C (Fig. 1b). This gene, transcribed from right to left, maps from bp 3715 through to the BamHI site at 4605 (Fig. 2 ) and presumably into the adjacent BamHI fragment, F. Analysis of the C-terminal 263 amino acids encoded by this open reading frame indicated sufficient similarity to the products of genes UL21 of HSV-1 and 38 of VZV to identify the EHV-4

gene as a homologue of these genes. A comparison of the predicted protein products of these three genes is outlined in Figure 6. The degrees of identity of the EHV-4 ORF3-specified C-terminal sequence with the corresponding regions of its HSV-1 and VZV counterparts are 22% and 30% respectively.

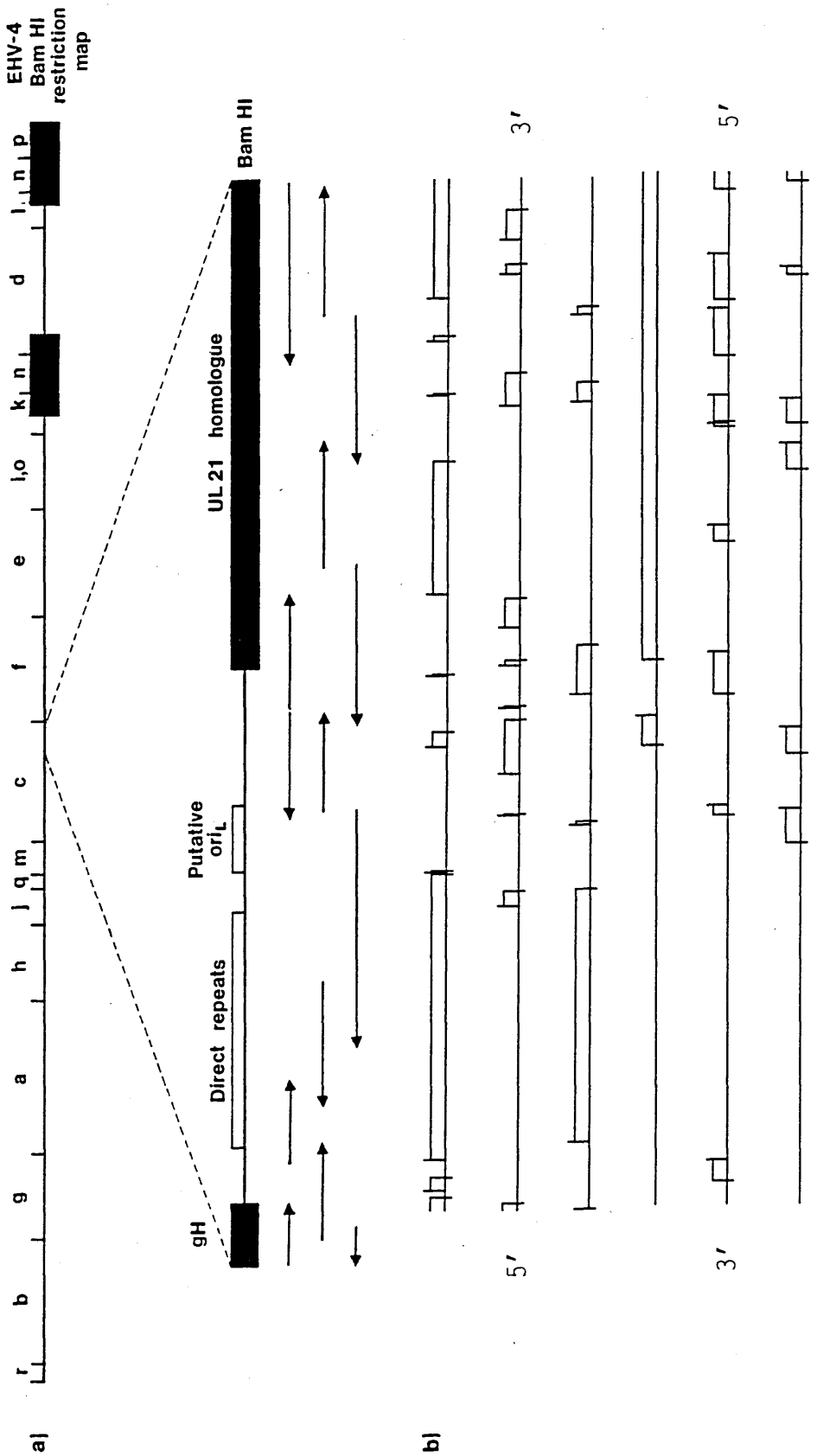
Figure 5.1 Positions of a direct repeat region, a putative origin of DNA replication, and a UL21 homologue within the right end of BamHI C.

a) BamHI restriction map of EHV-4 (strain 1942; Cullinane et al., 1988) with expansion of the sequenced region to show features of interest. The sequencing strategy is indicated below.

b) VAX analysis of potential ORF within all six reading frames.

Figure 5.1

EHV-4 Genome Map Units 0.51-0.52



## Figure 5.2

## Nucleotide Sequence of EHV-4 DNA , Map Units 0.51-0.52

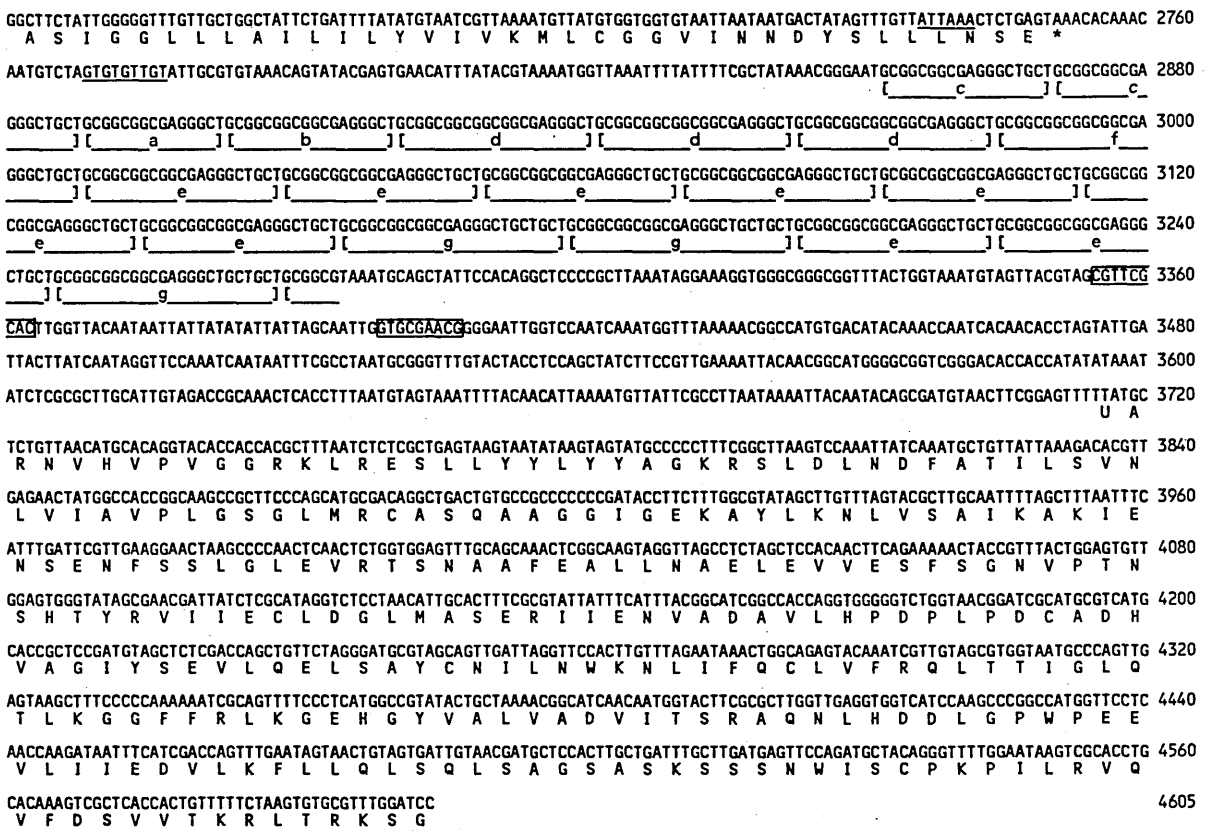


Figure 5.2 Nucleotide sequence of 1965bp of EHV-4 DNA . The numbering system used in Chapter 4 has been adopted such that the C-terminal coding region of the gH gene homologue is numbered as before, terminating at bp 2749. Individual repeat elements are bracketed and letters a to g assigned in order of increasing size (see Table 1) . The conserved element CGTTCGCAC and its complement within the putative EHV-4 $ori_L$ are boxed. The predicted C-terminal amino acid sequence of the product of the UL21 gene homologue is indicated below the nucleotide sequence.

Figure 5.3

Sequencing Reaction in Repeat Region

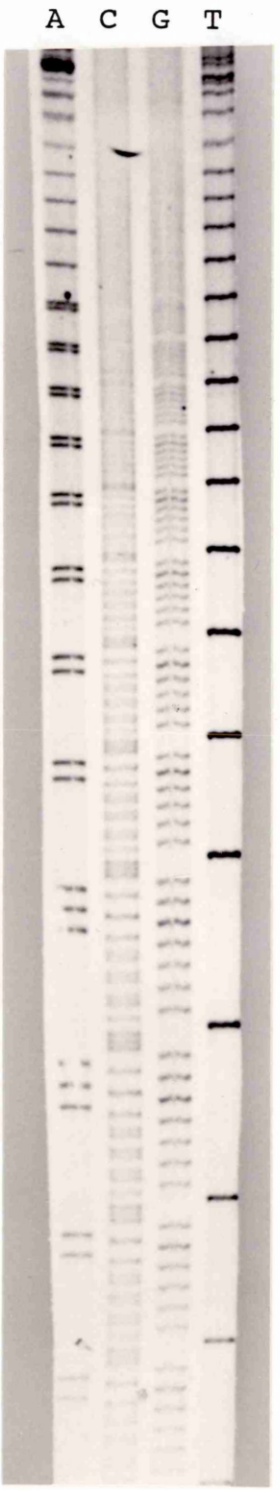



Figure 5.3 Autoradiograph of sequencing reaction using primer HL13 (bp 3467-3451) and denatured pBSSB4 as template.

## Figure 5.4

## Putative Stem-and-Loop Structures of HSV-1, EHV-1, and EHV-4 Origins of DNA Replication

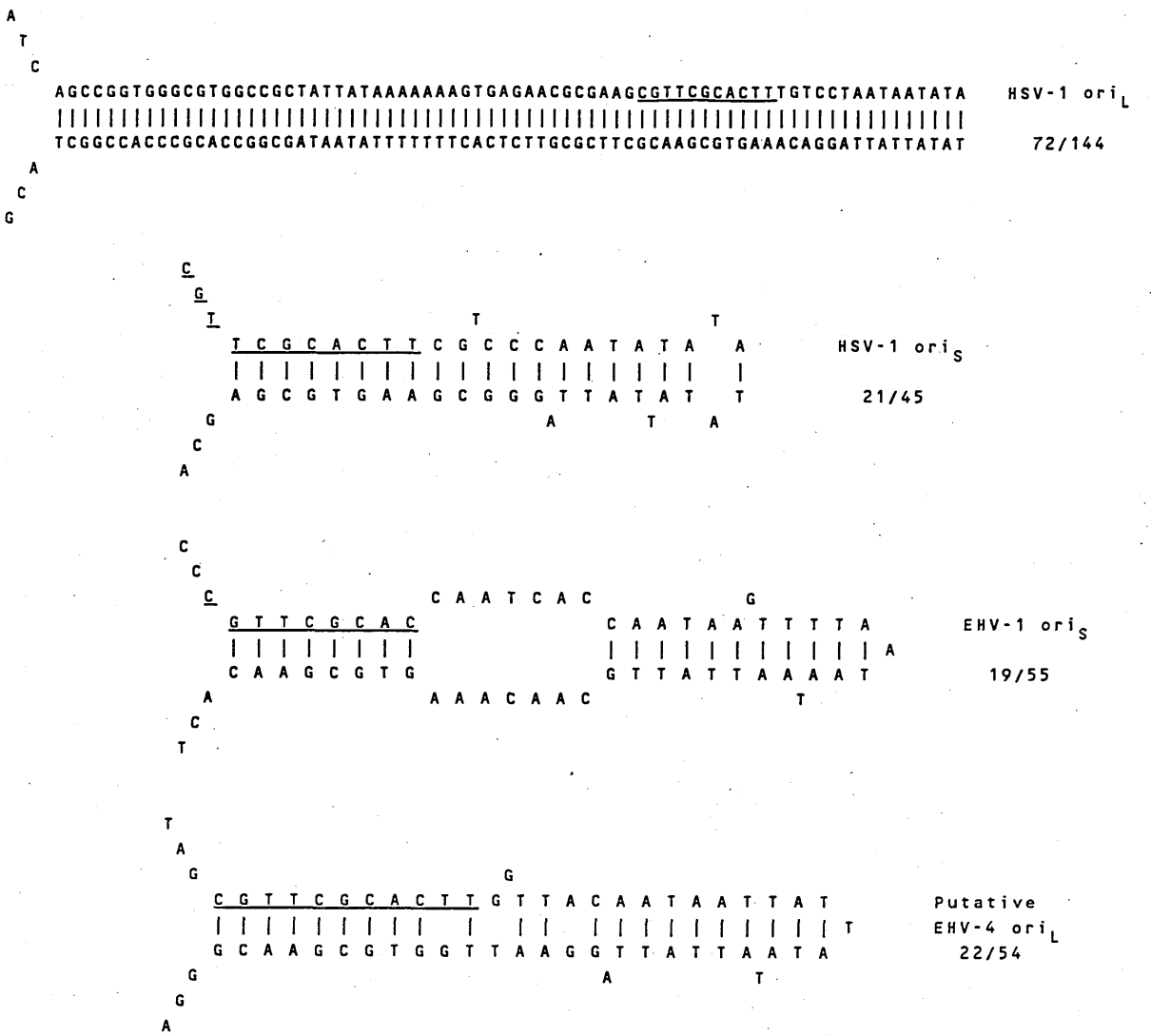



Figure 5.4 Potential base-pairing events within three functionally characterised origins of DNA replication - HSV-1 $ori_L$ (Weller et al., 1985), HSV-1 $ori_S$ (Stow and McMonagle, 1983), EHV-1 $ori_S$ (Baumann et al., 1989) - and within a putative EHV-4 origin . Figures assigned to each structure detail the number of potential base matches against the length of the palindrome. The position of the conserved sequence element CGTTCGCAC(TT) within each palindrome is indicated by underlining.

Figure 5.5

Alignment of A/T rich DNA and Flanking DNA

of Herpesvirus Origins of DNA Replication

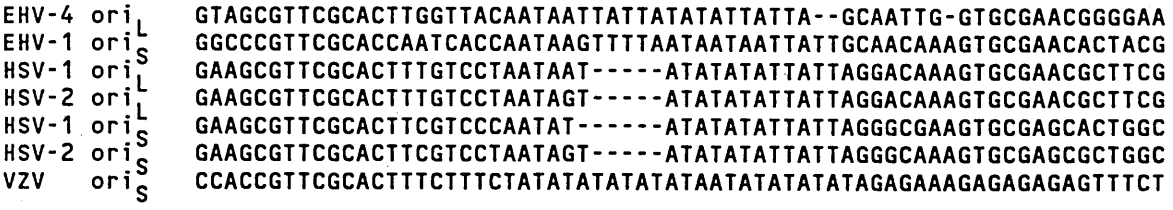



Figure 5.5 Nucleotide sequence similarity of DNA spanning the conserved CGTTCGCAC(TT) element and A/T rich DNA within the following herpesvirus origins of replication - EHV-4 $ori_L$ (putative), EHV-1 $ori_S$ (Baumann et al., 1989), HSV-1 $ori_L$ (Weller et al., 1985), HSV-2 $ori_L$ (Lockshon and Galloway, 1986), HSV-1 $ori_S$ (Stow and McMonagle, 1983), HSV-2 $ori_S$ (Whitton and Clements, 1984), and VZV $ori_S$ (Stow and Davison, 1986).

## Figure 5.6

## Alignment of the Predicted Products of HSV-1 UL21 and its VZV and EHV-4 Counterparts

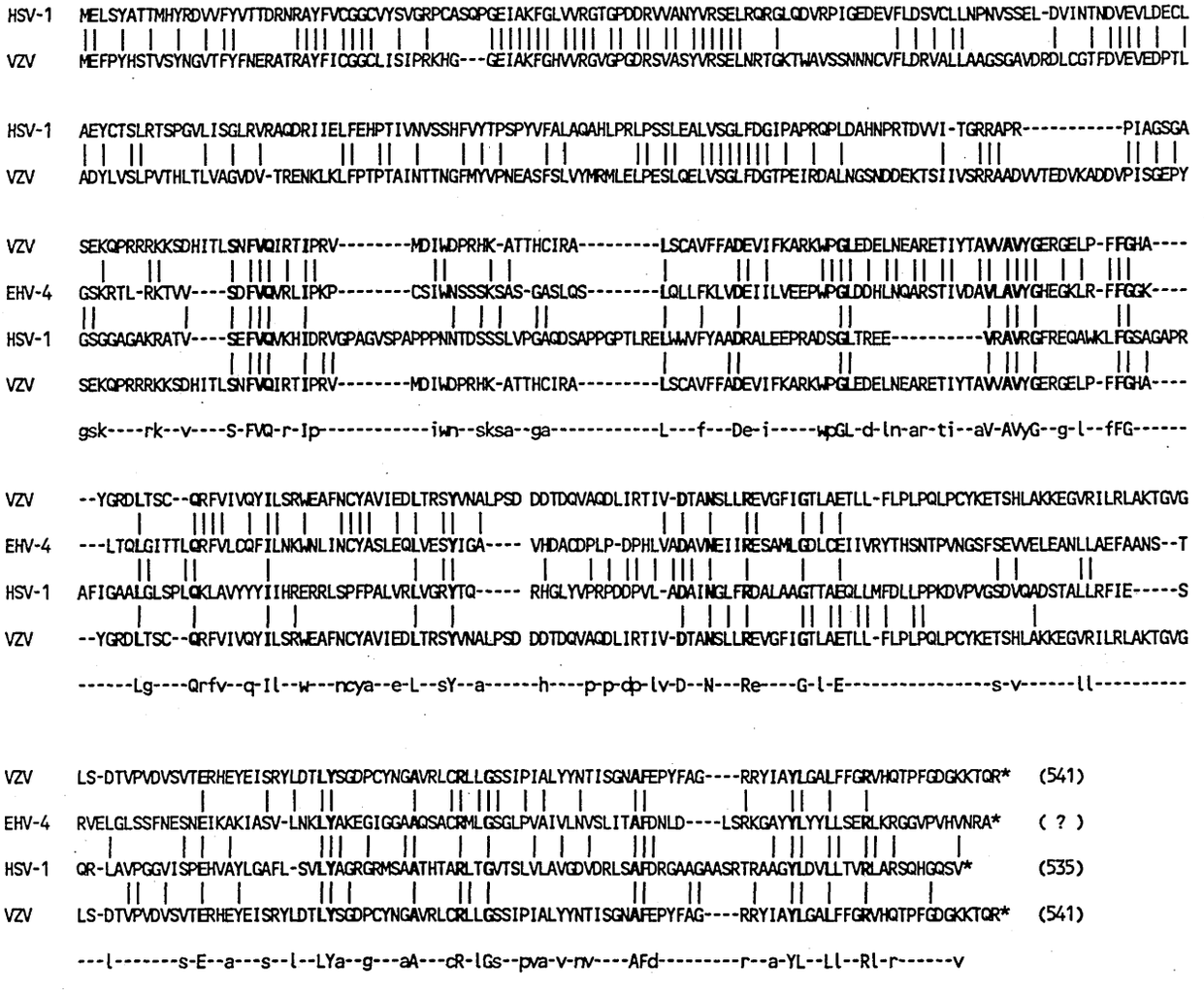



Figure 5.6 Alignment of the predicted products of gene UL21 of HSV-1, gene 38 of VZV, and the rightward flanking gene of the EHV-4 gH gene homologue (C-terminal region only). A consensus line is presented below the alignment in which lower case represents residues conserved in two polypeptides and upper case residues conserved in all three.

Table 5.1

Repeat Elements of EHV-1 and EHV-4

| | Sequence | Size | Number | Designation |
|---|---|---|---|---|
| EHV-4 Repeats | GCGGCGGCGAGGGCT GCT | 18 | 2 | c |
| | GCGGCGGCGAGGGCT | 15 | 1 | a |
| | GCG GCGGCGGCGAGGGCT | 18 | 1 | b |
| | GCGGCG GCGGCGGCGAGGGCT | 21 | 3 | d |
| | GCGGCG GCGGCGGCGAGGGCT GCT | 24 | 1 | f |
| | GCG GCGGCGGCGAGGGCT GCT | 21 | 7 | e |
| | GCG GCGGCGGCGAGGGCT GCTGCT | 24 | 2 | g |
| | GCG GCGGCGGCGAGGGCT GCT | 21 | 2 | e |
| | GCG GCGGCGGCGAGGGCT GCTGCT | 24 | 1 | g |
| | GCGGCG | 6 | 1 | partial |
| EHV-1 Repeats | GCGGGAGCGAGGGCTGCTGCGGCGGCGGCGCG | 32 | 2 | |
| | GCGGGAGCGAGGGCTGCTGC | 20 | 1 | partial |
| Comparison | GCGGGAGCGAGGGCTGCT GCGGCGGCGGCGCG EHV-1 intact repeat<br>\|\|\|\| \|\|\|\|\|\|\|\|\|\|\| \|\|\|\|\|\|\|\|\|\|\|<br>GCGGCGGCGAGGGCTGCT GCGGCGGCGGCGGC EHV-4 c,d (partial) repeats | | | |

Table 5.1 The nucleotide sequence, size, number, and order of repeat elements mapping downstream of the EHV-1 and EHV-4 gH gene homologues are indicated in the upper two sections of this table. The 'designation' column identifies repeat elements of EHV-4 as detailed in Figure 2. The third section details a comparison of the intact EHV-1 repeat element to two EHV-4 elements, 'c' and part of 'd'.

## Table 5.2

## Binding Motifs within Herpesvirus Origins of DNA Replication

| | Left arm | Right arm |
|---|---|---|
| EHV-4 $ori_L$ | C G T T C G C<br>A G T G C G A | G G T G C G A<br>C G T T C G C |
| EHV-1 $ori_S$ | C G T T C G C<br>G G T G C G A | A G T G C G A<br>T G T T C G C |
| HSV-1 $ori_S$ | C G T T C G C<br>A G T G C G A | A G T G C G A<br>T G <u>C</u> T C G C |
| HSV-1 $ori_L$ | C G T T C G C<br>A G T G C G A | A G T G C G A<br>C G T T C G C |
| HSV-2 $ori_S$ | C G T T C G C<br>A G T G C G A | A G T G C G A<br>C G <u>C</u> T C G C |
| HSV-2 $ori_L$ | C G T T C G C<br>A G T G C G A | A G T G C G A<br>C G T T C G C |
| VZV $ori_S$ | C G T T C G C<br>A G T G C G A | |

Table 5.2 Nucleotide sequences of the asymmetric 5bp GT(G/T)CG binding motifs within the right and left arms of the palindromes of herpesvirus origins of replication . 5′ and 3′ flanking residues are indicated to the left and right of the 5bp motifs. Nucleotides which do not conform to consensus rules (see text) are underlined. Data from Koff and Tegtmeyer (1988) and Baumann *et al.* (1989).

## DISCUSSION

Sequence analysis of 1855kbp of DNA immediately downstream of the EHV-4 gH gene indicated that EHV-4 encodes a gene homologue of genes UL21 of HSV-1 and 38 of VZV which are arranged in a similar genomic position , immediately proximal to the gH gene in tail-to-tail orientation. The function of the gene products of these genes is as yet unknown (Davison and Scott, 1986; McGeoch et al., 1988a). Comparison of the protein coding capacity of the portion of BamHI C sequenced in this study - UL24 homologue-TK-gH-UL21 homologue - indicates conservation of this gene block the HSV-1, VZV, and EHV-4 genomes.

Though gene arrangement might be conserved, the sequence analysis of EHV-4 revealed two elements - a region of direct repeats and a putative origin of replication - which are not located at a similar position within the genomes of HSV-1 or VZV. The direct repeat elements are of limited interest at present since there is little evidence as yet to suggest that any such elements have a functional role within alphaherpesviruses. However, it is interesting that direct repeat elements have been located in juxtaposition to other herpesvirus origins of replication, such as VZV $ori_S$. Analysis of the effect of repeat elements on HSV-1 ori-driven replication of plasmids in herpesvirus infected cells is in progress. It is possible the repeat elements might have an enhancer-type role in initiation of DNA synthesis at the origin but this is extremely speculative.

Direct repeat regions in herpesvirus genomes often occur at regions of heterogeneity between different viral isolates, with different numbers of repeats present in distinct isolates. Allen et al. (1983a) have reported that multiple passage of EHV-1 in non-equine cell lines leads to differences in genome size due to differences in the sizes of

copies of short repeat sequences occurring at specific regions of the genome and Whalley and coworkers (1981) have reported that EHV-1 EcoRI fragments K and L are heterogeneous in size.

Of the full complement of minor repeat elements within VZV and HSV-1 none are conserved in position or, to a significantly high degree, in sequence. There is evidence for the existence of a direct repeat region in the genome of EHV-1 at a similar position to that of EHV-4. However, this region differs from that of EHV-4 in that the direct repeat elements are larger and present in only two copies and a partial. It is possible, however, that deletion of part of this region in EHV-1 and/or EHV-4 may have occurred during passaging of the virus *in vitro* or on cloning and propagation in *E. coli*. A region of tandem repeats has been reported downstream of the PRV gH gene though sequence data has not been published (Petrovskis *et al.*, 1988). Further PRV repeat regions have been reported within the S component and adjacent to the gB gene homologue ( Simon *et al.*, 1989; Zhang and Leader, 1990). There is no evidence for repeats in the latter position in EHV-4, at least in EHV-4 strain 1942 (Riggio *et al.*, 1989). The only other reported EHV-4 tandem repeat element is an 8bp sequence - GGAGGTGG - 31 copies of which are located within $TR_S$ between gene homologues of HSV-1 genes US9 and US10 . A region of direct repeats is located at a similar position within HSV-1 (McGeoch *et al.*, 1985) and PRV (Zhang and Leader, 1990) but not within VZV (Davison, 1983). Given the number of repeats in HSV-1 and VZV it seems most likely that further sequence analysis of EHV-1 and EHV-4 will reveal further repeat regions especially within restriction fragments which are known to be heterogeneous.

The EHV-4 repeat elements were strict multiples of three in size which might imply constraints placed on their sequence by a protein coding function. Open reading frames spanning the repeat region were

almost exclusively represented by repeating codons and thus it seems most unlikely that any of these are functional protein coding genes. Furthermore, given the extent of gene conservation between EHV-4 and HSV-1 and VZV reported in this thesis and that of Riggio (1990) a gene exclusive to EHV-4 at this position would seem unlikely. It seems more likely , particularly since repeats in noncoding DNA of VZV and HSV are often multiples of three in size, that repeat sizes are determined by mechanisms involved in their generation .

In view of the sequence similarity of the putative EHV-4 origin to those of human and EHV-1 origins it would appear to be an excellent candidate for an EHV-4 $ori_L$. The only other origins of replication definitively mapped within the $U_L$ component of alphaherpesvirus genomes are the $ori_L$ of HSV-1 and of HSV-2. These are considerably larger than the putative EHV-4 $ori_L$ and are susceptible to deletion on cloning into plasmids (Gray and Kaerner, 1984; Weller et al., 1985). It seems likely that the sequence presented in this thesis is the intact element since deletion of the HSV-1 $ori_L$ results in a loss of the polyA/T tract.

A continuing debate within research into DNA replication origins is whether the palindromic sequence of origins adopts an intrastrand base paired conformation. There is as yet no experimental evidence that cruciform or stem-and-loop origin structures exist in vivo. Furthermore, mutational studies have indicated that although the left arm of the palindrome is crucial for origin activity, the right arm can be deleted with little effect (Deb and Doelberg, 1988). It is interesting that the EHV-1 $ori_S$ and our putative $ori_L$ possess fewer intrastrand base pairing events than do the human herpesvirus origins indicating a perfect palindrome is not required for origin activity.

The HSV-1 $ori_L$ maps in the intervening DNA between genes encoding UL29 and 30 which encode the major DNA binding protein and DNA

polymerase respectively (McGeoch et al., 1988a). Given the localisation of the origin and the function of the two flanking gene products in DNA replication it has been postulated that such an arrangement may be important to the regulation of initiation of DNA replication. The putative EHV-4 $ori_L$ mapping between gH and a gene of unknown function would seem to be removed from such control. A putative EHV-1 $ori_L$ sequence of very similar sequence to that of EHV-4 maps to a similar genomic position (Robertson et al., 1991). Whether an origin maps between the DNA pol and DNA binding protein genes of EHV-1 or EHV-4 has yet to be inferred either indirectly by sequence analysis or directly, by functional studies. The existence of two origins of replication within PRV, one at the extreme left end of the long component, the other downstream of gH, has been reported ( Kupershmidt et al., 1988; Petrovskis et al., 1988; Wu et al., 1986a,b) but sequence or functional characterisation has yet to be published. It might be expected from the comparative sequence analyses of PRV and EHV sequence data generated by ourselves and others that, with the exception of the 0.1-0.4 m.u. inversion of PRV (Davison and Wilkie, 1983), the EHV and PRV genomes would be more similar in terms of genome layout of coding regions and cis-acting elements than the EHV and HSV or VZV genomes. Consequently, it may be that the internal $U_L$ ori of PRV maps to a comparable position to that of EHV-1 and EHV-4 particularly in the light of the mapping of PRV tandem repeats at a similar position to those of EHV-1 and EHV-4 (Petrovskis et al., 1988).

# CHAPTER 6

# SEQUENCE ANALYSIS OF EHV-4 gp13 GENE AND FLANKING GENES

## INTRODUCTION

Genes encoding glycoproteins with sequence similarity to HSV-1 gC have been identified in several herpesviruses including HSV-2, VZV, PRV, BHV-1, MDV and EHV-1. Unlike gH and gB-type glycoproteins a gC-type glycoprotein seems to be specified only by alphaherpesviruses and not by beta and gammaherpesviruses, at least on current information.

The gC-type glycoproteins of HSV-1, MDV , and PRV are nonessential for viral replication *in vitro* yet gC- field isolates have not been isolated suggesting a critical role for gC *in vivo* (Churchill *et al.*, 1969; Holland *et al.*, 1984a; Mettenleiter *et al.*, 1988; Robbins *et al.*, 1986b). Functions assigned to the gC-type glycoproteins of herpesviruses include a role in pathogenicity (Kumel *et al.*, 1985; Mettenleiter *et al.*, 1988; Schreurs *et al.*, 1988), in binding of complement factor C3b (Bielefeldt Ohmann and Babiuk, 1988; Smiley and Friedman, 1985), and in the mediation of adsorption of the virion to target cells (Kuhn *et al.*, 1990; Mettenleiter *et al.*, 1988; Okazaki *et al.*, 1987; Zuckermann *et al.*, 1989). Comparative analysis of HSV-1 gC and PRV gIII has indicated that these glycoproteins might have diverse functional roles within their respective viral replicative cycles (Whealy *et al.*, 1989). In contrast, common functional roles have been reported for glycoproteins such as gH and gB which exhibit a greater degree of conservation in terms of amino acid sequence and representation in herpesviruses of different subgroups (Rauh and Mettenleiter, 1990) . An example of the diverse function of different 'gCs' is C3b binding activity which has only been demonstrated for HSV-1 gC and EHV-1 gp13 presented on infected cell membranes and for HSV-2 gC in a purified form ( Bielefeldt Ohmann and Babiuk, 1988; Seidel-Dugan *et al.*, 1988). The MDV gC-type glycoprotein, A antigen or gp57-65, differs from other 'gCs' in that it is

predominantly a secreted antigen and in this form may be involved in immunoevasion by protecting virus infected cells from attack through binding antibody (Isfort et al., 1986, 1987). The gC-type glycoproteins of BHV-1 and HSV-1 seem to assemble into homodimers which are responsible for specific spikes within the virion envelope (Kikuchi et al., 1990; Stannard et al., 1987; van Drunen Littel-van den Hurk and Babiuk, 1986).

An important function of all herpesvirus 'gCs', and the reason for our interest in the EHV-4 glycoprotein, is their interaction with the cellular and humoral immune system of the infected host. gC tends to be a major target of host humoral and cellular immune mechanisms (Ben-Porat et al., 1986; Glorioso et al., 1985 ).

Following mapping of the major glycoprotein genes of EHV-1, gp13 was predicted to be a gC-type glycoprotein since its gene mapped at a position colinear with that of the gC gene of HSV-1 (Allen and Yeargan, 1987). Analysis of the amino acid sequence specified by the gp13 gene confirmed that EHV-1 gp13 , alternatively known as gp88 (Meredith et al., 1989), is a gC-type glycoprotein : gp13 shares sequence identity with gC and exhibits positional conservation of 4 cysteine residues suggesting a degree of structural similarity. Polyclonal antisera and a proportion of monoclonal antibodies raised against EHV-1 gp13 are crossreactive with the EHV-4 glycoprotein identified as gp13 (Allen and Bryans, 1986; Allen et al., 1988) or gp92 (Meredith et al., 1989). Such evidence indicates shared antigenicity of these proteins which might be expected to be reflected in similarity of their respective amino acid sequences and gene nucleotide sequences. On the basis of these studies, there is compelling, albeit indirect, evidence to suggest that gp13 of EHV-4 is a gC-type glycoprotein and would thus be the expected product of an EHV-4 gC gene homologue. The term gp13 is used

throughout this thesis to refer to the product of the EHV-4 gC gene homologue. As yet all that is known of the EHV-1 and EHV-4 'gCs' is the C3b binding activity of the former and the capacity of both to elicit antibodies, some of which neutralise infectivity in the presence of complement ( Allen and Bryans, 1986; Allen and Coogle, 1988; Bielefeldt Ohmann and Babiuk, 1988). The glycoproteins have also been implicated in interaction with cellular immune mechanisms (Bridges et al., 1988). Whether either glycoprotein exhibits any of the other functions assigned to gC-type glycoproteins remains to be determined.

To date the amino acid sequences of the gC-type glycoprotein of seven herpesviruses have been published - HSV-1 gC (Dowbenko and Lasky, 1984; Draper et al., 1984; Frink et al., 1983; McGeoch et al., 1988a), HSV-2 gC (Dowbenko and Lasky, 1984; Swain et al., 1985), PRV gIII (Robbins et al., 1986a), VZV gpV (Kinchington et al., 1986), BHV-1 gIII (Fitzpatrick et al., 1989), MDV A antigen (Binns and Ross, 1989; Coussens and Velicer, 1988; Ihara et al., 1989 ), and EHV-1 gp13 (Allen and Coogle, 1988). Since the $U_L$ components of EHV-1 and EHV-4 are generally colinear (Cullinane et al., 1988) the EHV-4 gC gene homologue was expected to map at a comparable position to the EHV-1 gp13 gene (m.u. 0.114-0.136). This chapter describes the localisation of the EHV-4 gp13 gene and flanking genes by Southern analysis and sequence data, and reports on the similarity of the predicted EHV-4 gp13 protein with other herpesvirus gC-type proteins (Nicolson and Onions, 1990).

## MATERIALS AND METHODS

### 1) Generation of EHV-1 gp13 gene hybridisation probe

A BamHI library of EHV-1 strain HVS-25 in pBR322 was kindly provided by J.M. Whalley of the Macquarie Institute, Australia. The EHV-1 gp13 gene is localized within a BamHI/EcoRI fragment mapping between 0.114 and 0.148 (Allen and Yeargan, 1987) - a subfragment of EHV-1 BamHI H (Whalley et al., 1981). Recombinant plasmid pBR322:EHV-1 BamHI H was digested with BamHI or BamHI and EcoRI, electrophoresed through a 0.7% agarose gel and the intact EHV-1 H fragment or BamHI/EcoRI subfragment excised and purified using Costar spin tubes (NBL). These probes were $^{32}$P-labelled by nick translation in the presence of $\alpha$-$^{32}$P-dCTP (Amersham Radiochemicals Ltd.).

### 2) Construction of EHV-4 recombinant plasmids

An EHV-4 strain 1942 BamHI library in pUC9 was constructed as previously reported (Cullinane et al., 1988). Recombinant plasmid pUC9:EHV-4 BamHI G was digested with a series of restriction enzymes to generate subfragments of EHV-4 G which were then isolated from 0.7% agarose gels and cloned into a Bluescript M13+ plasmid vector (Stratagene) by standard techniques (Maniatis et al., 1982).

### 3) Southern Analysis

Recombinant plasmids containing EHV-4 BamHI fragments B and G - approximate map coordinates 0.01 - 0.17 - were digested with a series of restriction enzymes and the digest products electrophoresed through a 0.75% agarose gel to which $^{32}$P-labelled lambda Hind III fragments had been added as size markers. The gel was alkali denatured and neutralised and DNA fragments transferred to a Gene Screen membrane (New England Nuclear) by capillary action. The membrane was baked for 2 hours at 80$^{o}$C, prehybridised overnight at 42$^{o}$C in 45% formamide, 4.2 x SSC, 4.2 X Denhardts Solution, 8% dextran sulphate, 0.04M NaPPi,

80ug/ml salmon sperm DNA, 0.08% SDS and hybridised for a similar period following addition of 2 x $10^7$ cpm of denatured $^{32}$P-labelled EHV-1 H BamHI or BamHI/EcoRI probe DNA. The blot was rinsed with 2 x SSC, washed for three 30 minute periods at 65$^{o}$C in 1 x SSC, 0.1% SDS, rinsed in 0.1 x SSC , dried and exposed to Fuji RX film overnight at -70$^{o}$C in the presence of a Cronex intensifying screen.

## RESULTS

### 1. Localisation of the EHV-4 gC gene homologue

**1) Southern hybridisation**

The EHV-1 gC gene homologue (gp13) is localised within a BamHI/EcoRI subfragment ( map units 0.114 - 0.148) of EHV-1 BamHI H ( m.u. 0.114-0.164) (Allen and Yeargan, 1987; Whalley et al., 1981) (Fig. 1a). Given the reported colinearity between the EHV-1 and EHV-4 genomes (Cullinane et al., 1988), we expected the EHV-4 gC gene homologue to be located within BamHI G (Fig. 1b). Southern hybridisation experiments were performed in which fragments of EHV-1 DNA spanning the gp13 gene were hybridised to fragments BamHI B and G of the EHV-4 genome in order to confirm colinearity prior to sequence analysis. Such analysis confirmed EHV-4 BamHI G as the probable site of the EHV-4 gp13 gene since hybridisation was limited to G and to subfragments of G (Figs. 2,3). A restriction map of BamHI G was determined by conventional techniques to aid in the identification of the crosshybridising region of G (Fig. 3c). EHV-1 H DNA hybridised to all EHV-4 G restriction fragments tested with the exception of a 1.9kbp BamHI/SalI fragment mapping to the left end of G (Figs. 2,3c). Hybridisation to a 0.3kbp EcoRI/EcoRI fragment mapping between two hybridising EcoRI fragments was not detected presumably as a consequence of the amount of radiolabelled probe bound being below the detection limit . The BamHI/EcoRI subfragment of EHV-1 H (m.u. 0.114-0.148) failed to hybridise to the 0.95kbp EcoRI terminal fragment (Fig. 3). As hybridisation to the 0.3kbp internal EcoRI fragment was not detected with either probe , it was not clear whether the smaller BamHI/EcoRI probe shared homology with this fragment. Since the right terminal 1.9kbp BglI/BamHI fragment and the adjacent 1.9kbp BglI/BglI fragment co-migrated it was not possible to definitively conclude whether the

probe hybridised to this fragment. However, comparison of binding of label to the 1.9kbp region of the BamHI/EcoRI-probed blot to that of the BamHI/BamHI-probed blot suggested lesser binding to the doublet by the smaller probe. We concluded from these analyses that the EHV-4 gp13 gene maps between the innermost SalI and EcoRI sites, possibly extending to the EcoRI site at 0.3 (Fig. 3c).

**2) Sequence Analysis**

In order to precisely localise the EHV-4 gp13 gene, subfragments of the cross-hybridising region of EHV-4 G were cloned into a Bluescript plasmid vector and the nucleotide sequence of the termini of each fragment determined using commercially available Bluescript primers. Such analysis revealed a significant degree of homology between DNA spanning the rightward BglI site and the central region of the EHV-1 gp13 gene. Sequence analysis was extended from this region as detailed below.

## 2. Nucleotide Sequence of EHV-4 - Map Units 0.15-0.17

The nucleotide sequence of 2.88 kbp of EHV-4 BamHI G (approximate map units 0.15-0.17) was determined according to the strategy outlined in Figure 4c and is detailed in Figure 5. Two major open reading frames , ORF2 and ORF3, and a partial ORF, ORF1, were detected in transcriptional sense right to left (Fig. 4b).

ORF1 extends from the EcoRI site at bp 0 through to bp 233 and encodes the C-terminal region of a polypeptide. A potential polyadenylation signal, ATTAAA, is positioned across the termination codon. ORF2 , which extends from an ATG at 412 to a TAA termination codon at 1867, was identified as the EHV-4 gC gene homologue through the sequence similarity of its 485 amino acid product to the gC-type glycoproteins of other herpesviruses. Comparison of ORF2 to the gp13

gene of EHV-1 indicated 76% identity at the DNA level and 79% at the amino acid level. On the basis of the shared antigenicity of the gp13 glycoproteins of EHV-1 and EHV-4 the predicted product of ORF2 is gp13 and the gene and gene product of ORF2 are identified as such hereafter.

The DNA sequence upstream and downstream of ORF2 was searched for characteristic gene control elements. The 5′ non-coding region of many eukaryotic and viral genes contains two sequences required for precise or efficient transcription of the gene by RNA polymerase II - the TATA [TATA(A/T)A(T/A)] and CAAT [GC(C/T)CAATCT] boxes, positioned 26-34 and 70-80 bp, respectively, upstream of the RNA pol II initiation site (Corden et al., 1980). Several TATA homologues are located in the immediate upstream of the EHV-4 gC gene. Of these, only 2 (at 345 and 220) with a greater than 70% fit to the consensus sequence, have potential CAAT boxes at an appropriate distance upstream. Alignment of the EHV-1 and EHV-4 gp13 DNA sequences revealed perfect alignment of the TATA box at 345 and its corresponding CAAT box (bp 297) with the putative promoter elements of the EHV-1 gp13 gene (Fig. 6). However, given the extent of sequence homology within this region we cannot infer the conservation of these control element homologues is due to sequence constraint as a result of promoter functions. No consensus polyadenylation signal - AATAAA (Proudfoot and Brownlee, 1976) - was located within the immediate downstream region. However, two minor polyadenylation signal sequences -AATACA and ATTAAA (Birnstiel et al., 1985) - were located 40 and 100bp downstream of the TAA stop codon. It is possible that in the EHV-4 gp13 transcription termination and polyadenylation process a relatively uncommon core signal sequence is utilised. However, no equivalent polyadenylation signals are located within the corresponding region of the EHV-1 gp13 gene (Fig. 6) suggesting that, in EHV-1 at least and perhaps in EHV-4, the gp13 gene

may be transcribed as part of a multigenic RNA with the RNA polymerase II transcribing downstream of gC, possibly through ORF3.

Two ATG codons are located within the first 100 bp of ORF2. The ATG at 412-414 seems the more likely initiation codon given that i) its local nucleotide sequence - CAGCAATGG - gives a more favourable fit to the optimal mRNA translation initiation codon sequence of Kozak (1984, 1986) - CC(A/G)CCATG(G)- than the ATG at 430-432 and ii) alignment of the EHV-4 and EHV-1 gp13 sequences brings the EHV-1 ATG initiation codon, into alignment with EHV-4 ATG 412-414 (Fig. 6).

The codon usage of the EHV-4 gp13 gene is indicated in Table 1 with the predicted amino acid composition of its product. The %GC of the gene is 48.4% which is similar to that of the EHV-4 gB gene (Riggio et al., 1989). All 61 amino acid specifying codons are utilised.

ORF3 , or dsgC, maps downstream of the gp13 gene and extends for 678 nucleotides . It shares no homology with VZV or HSV-1 genes in a comparable genomic position. TATA and CAAT homologues are located upstream of this ORF at 2065 and 2020 respectively. Codon usage is presented in Table 2.

## 3. Predicted Protein Products of ORF1, ORF2, and ORF3

### 1) EHV-4 UL43 gene homologue product

The 76 amino acid sequence specified by partial ORF1 shares identity with the C-terminal region of the protein products of genes UL43 of HSV-1 (McGeoch et al., 1988a) and 15 of VZV (Davison and Scott, 1986) to a level of 23% and 21% respectively (Fig. 7).

### 2) EHV-4 gp13

At 485 amino acids in length with molecular weight 52513, the EHV-4 gp13 primary translation product is of comparable length to the equivalent protein in EHV-1 (468) (Allen and Coogle, 1988), PRV (479)

(Robbins et al., 1986a), HSV-1, HSV-2 (511,479) (Dowbenko and Lasky, 1984; Swain et al., 1985), MDV (501)(Binns and Ross, 1989; Coussens and Velicer, 1988 ) and BHV-1 (521)(Fitzpatrick et al., 1989). The 591 amino acid VZV gpV is longer due to the presence of direct repeat elements within its coding sequence (Kinchington et al., 1986).

A putative signal peptide was identified by a stretch of hydrophobic amino acids at the N-terminus evident within hydropathic analysis of gp13 using the algorithm of Hopp and Woods (1981) or Kyte and Doolittle (1982 ) (Figs. 8,9). Using the eukaryotic weight matrix of von Heijne (1986) the EHV-4 gp13 signal sequence is predicted to be 32 amino acids in length with cleavage occurring between the Ala and Ser residues at positions 32 and 33 respectively. This predicted cleavage position is consistent with the `-3,-1` rule in that the side chains of the amino acids at positions -3 (Ala) and -1 (Ala) are small and uncharged. Furthermore, the existence of a glycine residue at -6 is consistent with the observation that strong helix breaking amino acids are often positioned between the hydrophobic core of the signal sequence and the -3 to -1 region.

The extracellular domain of gp13 is predicted to span residues 33 to 444 (Fig. 10). This 412 amino acid domain possesses 11 putative N-linked glycosylation sites ( N-X-S/T (Hubbard and Ivatt, 1981) ) and has a predicted secondary structure of predominantly beta sheets with a few alpha helices and turn regions (Fig. 9). Immediately contiguous glycosylation sites occur at two positions reducing the potentially functional N-linked glycosylation sites to 9. Critical antigenic determinants with respect to elicitation of virus neutralising antibody are likely to be located in this domain given its exposure on the surface of the virion or plasma membrane. Charged, hydrophilic amino acids are common features of antigenic determinants and the peak of

hydrophilicity in a Hopp and Woods hydropathic plot often correlates with the position of an immunogenic epitope (Hopp and Woods, 1981). EHV-4 gp13 may thus contain an antigenic determinant spanning residue 409 (Asn) since this lies within the peak hydrophilicity point of the EHV-4 gp13 plot (Fig. 8).

Amino acids 446 to 472 are enriched in hydrophobic residues and may constitute the glycoprotein transmembrane domain the function of which is to anchor the protein to the virion envelope or infected cell plasma membrane (Fig. 10). The predicted C-terminal structural organisation of the gp13 C-terminal domain contrasts with that of the EHV-4 gB polypeptide in which a predicted 69 amino acid transmembrane domain comprising three loops traversing the membrane, is followed by a cytoplasmic domain of 109 amino acids (Riggio et al., 1989).

The putative C-terminal cytoplasmic domain spans residues 473 to 485 , is hydrophilic in character and possesses a net positive charge of 2.

EHV-4 gp13 is most closely related to EHV-1 gp13 (79%) followed by PRV gIII (29%), BHV-1 (27%), HSV-1 gC, HSV-2 gC, and VZV gpV (25-26%), and MDV A antigen (22%). It is evident from multiple alignments and other comparative analyses of herpesvirus gCs (Allen and Coogle, 1988; Binns and Ross, 1989; Fitzpatrick et al., 1989) that the N-terminal half of the protein has been conserved to a lesser extent than the C-terminal half in the course of evolution. This is reflected in dot matrix comparisons of EHV-4 gp13 with EHV-1 gp13, BHV-1 gII, HSV-1 gC, and VZV gpV as presented in Figure 11. Features common to gC-type glycoproteins include a short, positively charged cytoplasmic domain and six sites of cysteine conservation (Allen and Coogle, 1988; Binns and Ross, 1989; Fitzpatrick et al., 1989). A few of the putative N-linked glycosylation sites exist in similar positions but are not

strictly conserved. For the purposes of comparing the EHV-4 gp13 to other gCs in terms of these specifically conserved features an alignment of eight herpesvirus gC-type polypeptides is presented in Figure 12. Although single and double amino acid insertions and deletions occur throughout each sequence relative to the other, the majority of the misaligned residues are located within the N-terminal portion of the protein to EHV-4 residue 70. It is particularly evident between the two EHV gp13s : identity within the region EHV-1 68-468, EHV-4 84-485 is 86% compared with 42% within the rest of the protein. EHV-4 gp13 possesses cysteine residues at each of the six conserved positions - amino acids 256, 318, 357, 361, 390, and 416. Nine EHV-4 gp13 glycosylation sites are conserved in EHV-1 gp13 and three in PRV.

A structural model has been proposed for BHV-1 gIII on the basis of similarity of certain regions of its extracellular domain to domains within members of the immunoglobulin super-family (Fitzpatrick et al., 1989). Greatest conservation was found across amino acids 261-370 of gIII and class II MHC constant domains. This part of gIII corresponds to amino acids 236-330 of EHV-4 gp13 which show a degree of similarity to the MHC domain (Fig. 13). The model also takes into account the probable role of disulphide bond formation in the determination of secondary and tertiary structure given the positional conservation of six cysteine residues within the extracellular domain.

**3) EHV-4 ORF3**

ORF3 encodes a 226 amino acid polypeptide, 'dsgc', which has no apparent identity with proteins encoded by HSV-1 and VZV genes mapping at a similar genomic position. Analysis of the predicted amino acid sequence indicates two interesting features of the polypeptide : a lysine and arginine enriched region ( amino acids 54-65 ) and 6 potential N-linked glycosylation sites, N-X-S/T. Comparison of the

predicted EHV-4 protein to proteins within the NBRF data bank using VAX software (Devereux et al., 1984) indicated limited regions of similarity to domains within a variety of proteins. The highest score was attained in a comparison of the EHV-4 protein and a probable membrane antigen, p140, of HVS: 22% identity between amino acids 85-165 and 345-422 of the EHV-4 and HVS proteins respectively. No single region of the EHV-4 protein was overrepresented in the comparisons . It is not possible therefore to speculate on any function the EHV-4 protein might have. Similar analyses of gC and thymidine kinase species have highlighted conserved functional domains such as the MHC class II-type domain and a nucleotide binding domain within these proteins and unrelated proteins, the similarity in the latter case being due to convergent rather than divergent evolution.

Figure 6.1 a) Schematic representation of the map positions of 6 EHV-1 glycoprotein genes as determined by Allen and Yeargan, 1987. BamHI and EcoRI restriction maps of EHV-1 strain HVS25 (Whalley et al., 1981; Robertson and Whalley, 1988) are detailed. The gp13 gene maps within BamHI H hence the use of two hybridisation probes , BamHI H and BamHI/EcoRI H, in Southern analyses of EHV-4 DNA.
b) BamHI restriction map of EHV-4 (strain 1942; Cullinane et al., 1988) presented in such a way that the EHV-1 and EHV-4 genomes align in terms of map units. On the basis of colinearity of the EHV-1 and EHV-4 $U_L$ components the EHV-4 gp13 gene would be expected to map within BamHI G .

Figure 6.1

Map Position of the EHV-1 gp13 Gene and of DNA Fragments used as Southern Hybridisation Probes

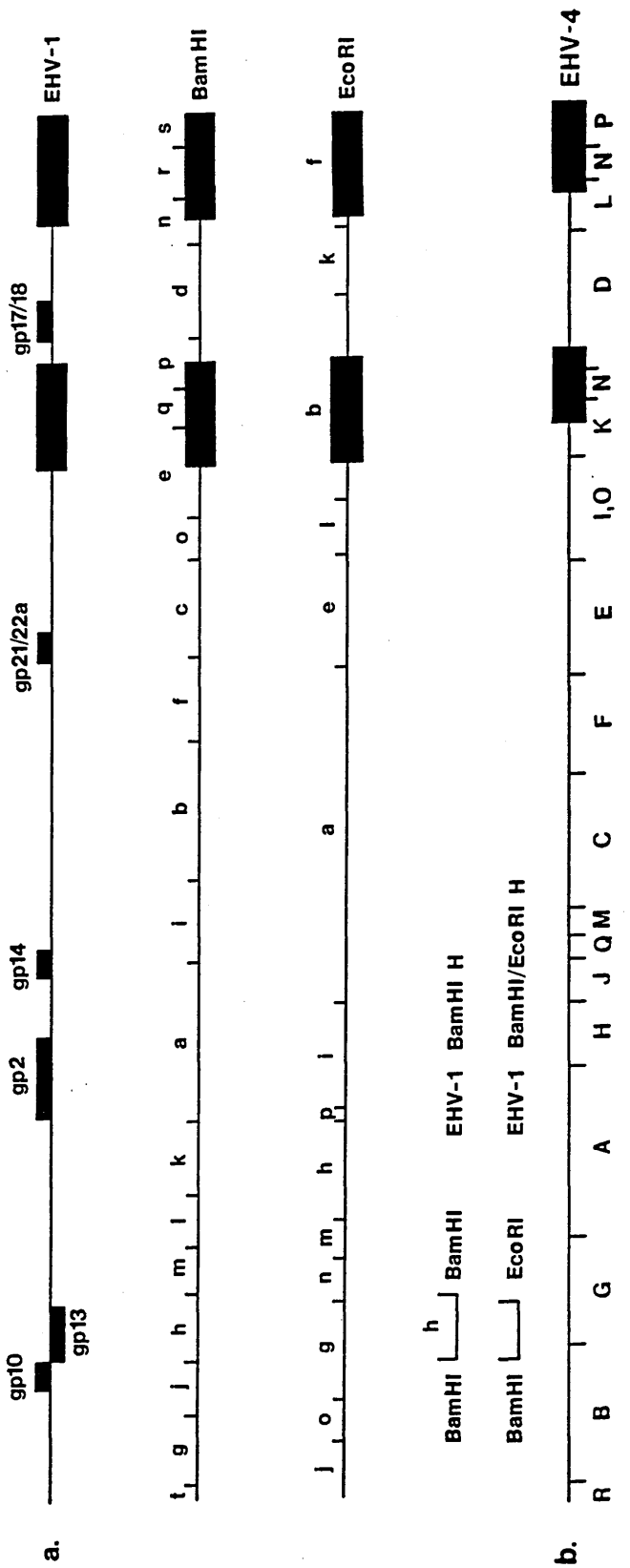


Figure 6.2 Hybridisation of EHV-1 BamHI H DNA to BamHI G and BamHI B of EHV-4.

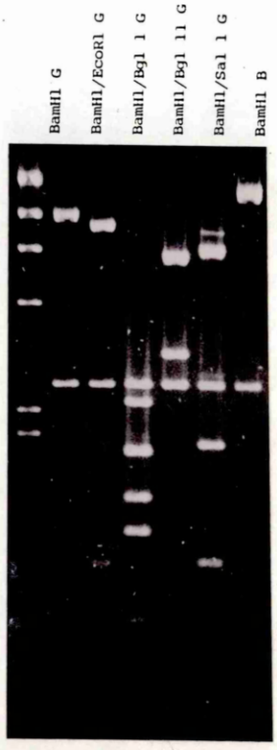
a) Agarose gel loaded with lambda HindIII markers, BamHI , BamHI/EcoRI, BamHI/BglI, BamHI/BglII, and BamHI/SalI digests of pUC9:G, and BamHI digest of pUC9:B.

b) Autoradiograph of membrane prepared from above gel probed with EHV-1 BamHI H.

Figure 6.2

Hybridisation of EHV-1 BamHI H to EHV-4 DNAs

a)




b)

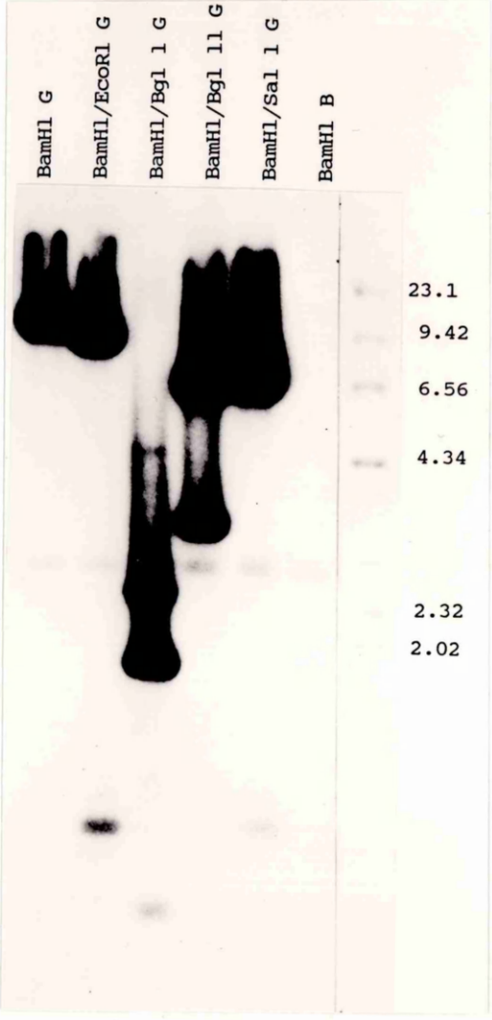


Figure 6.3 Hybridisation of EHV-1 BamHI/EcoRI H to BamHI G and BamHI B of EHV-4.

a) Autoradiograph of membrane prepared from gel (Figure 6.2 a)) probed with EHV-1 BamHI/EcoRI H.

b) Region of cross-hybridisation of the probe with EHV-4 G.

Figure 6.3

Hybridisation of EHV-1 BamHI/EcoRI H to EHV-4 DNA

a)


b)

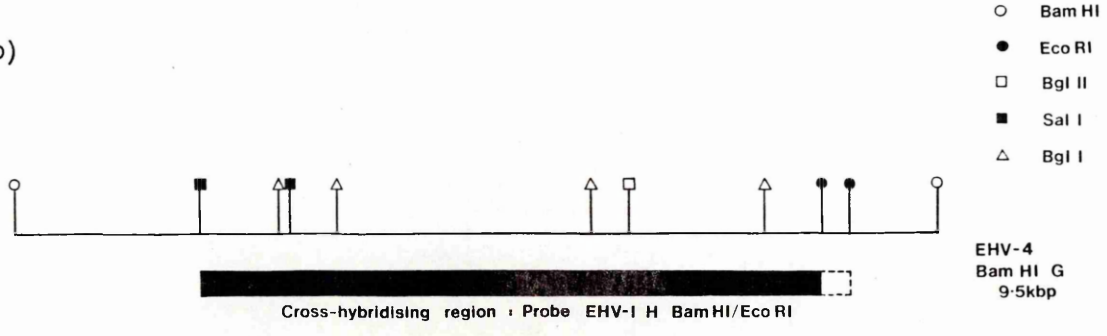


Figure 6.4 Localisation of three ORF within a 2.88kbp region of EHV-4 BamHI G.

a) BamHI restriction map of EHV-4 (strain 1942; Cullinane et al., 1988) with expansion of BamHI G to show the positions of SalI, EcoRI, BglI, and BglII restriction sites.

b) Expansion of the sequenced region of BamHI G with the three ORF identified by VAX analysis of start and stop codon positions (shown below) indicated by black boxes.

c) Sequencing strategy utilised. Arrows indicate regions of sequence elucidated using different primers and/or templates. Their direction indicates whether upper or lower strand sequence was determined.

Figure 6.4

Localisation of the EHV-4 gp13 Gene and Flanking ORFs

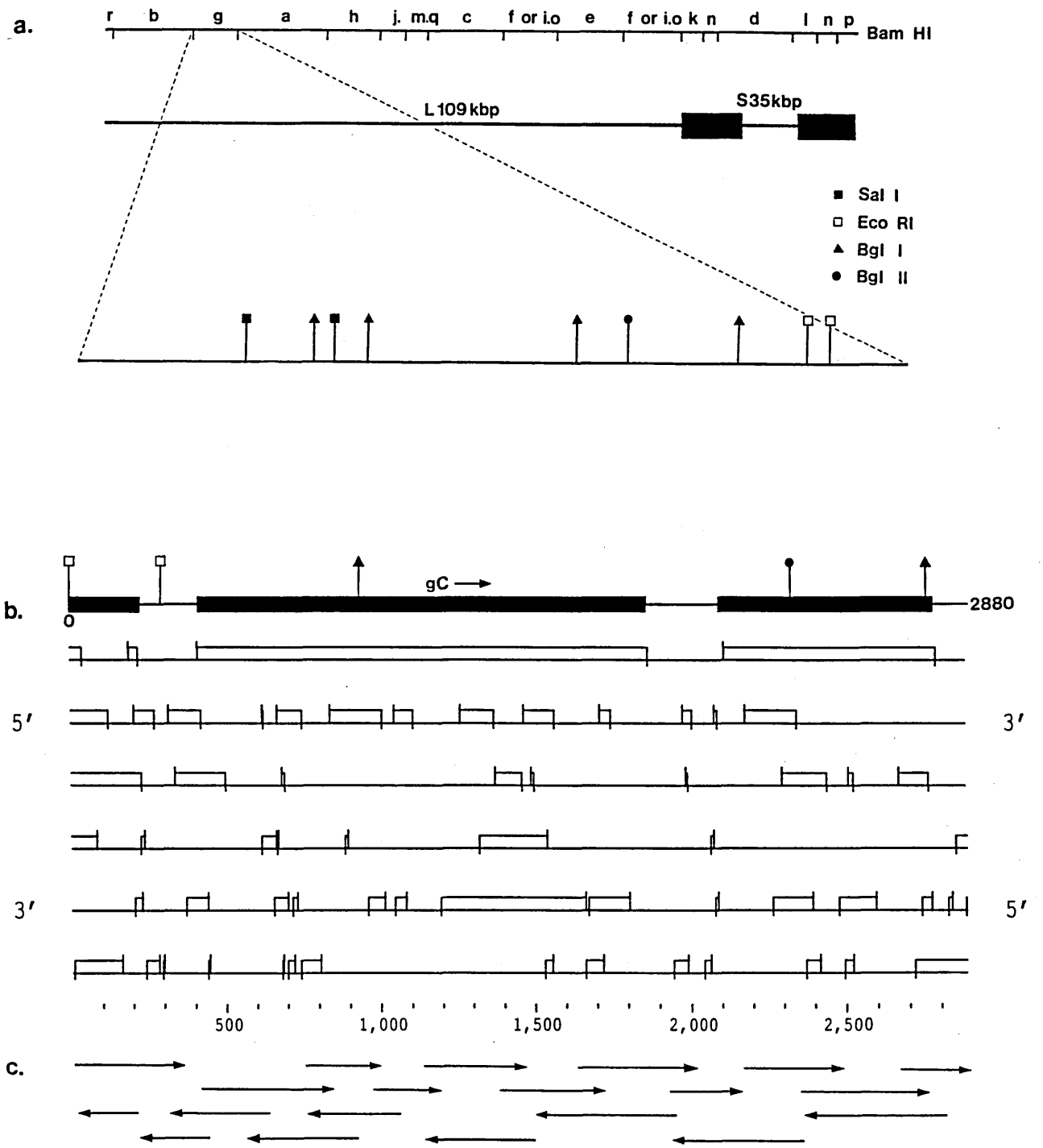
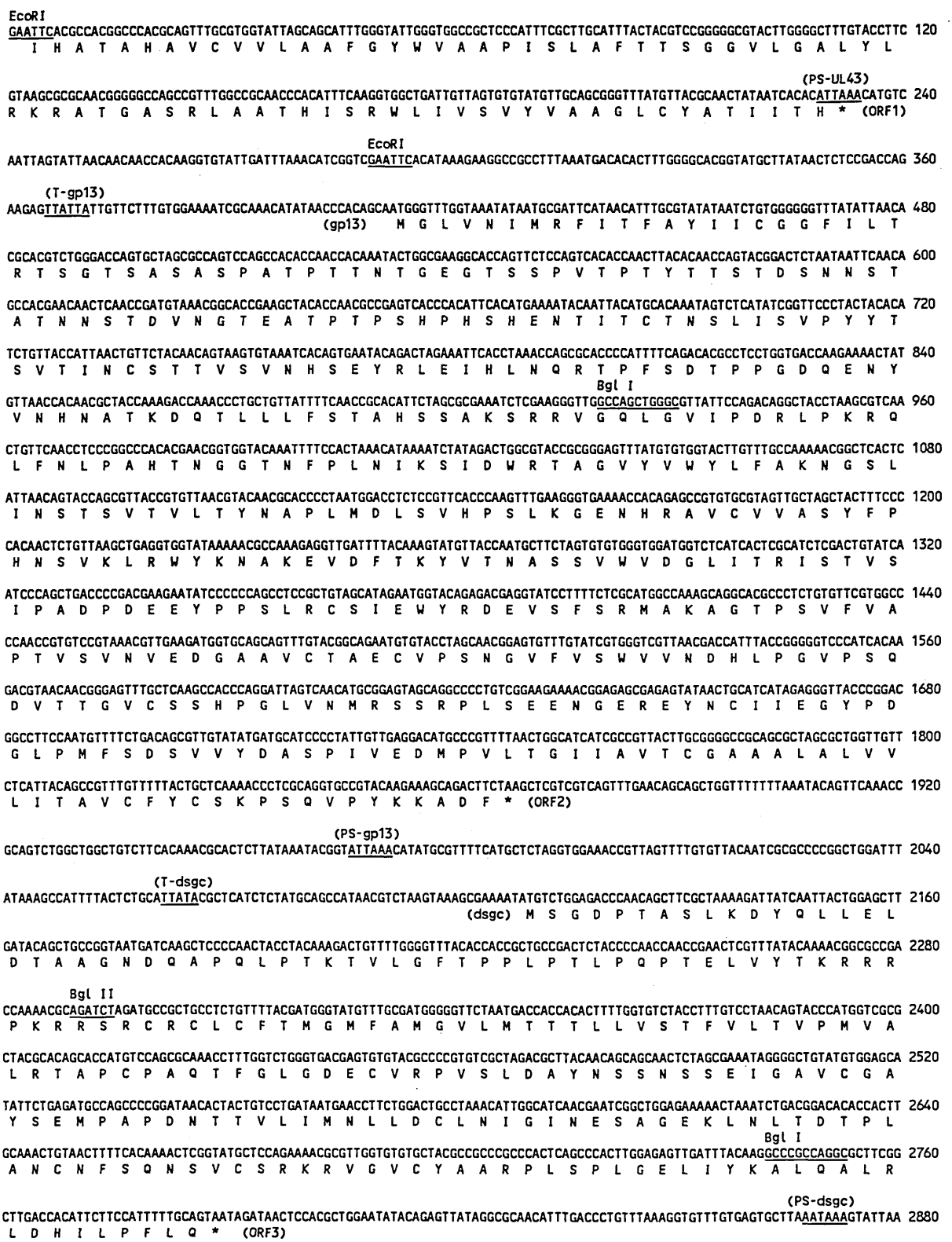


Figure 6.5 Nucleotide sequence of 2.88 kbp of EHV-4 BamHI G. The predicted amino acid sequences of the products of three open reading frames are indicated below the coding sequence. Putative polyadenylation signals (PS) and TATA (T) boxes are underlined.

# Figure 6.5

## Nucleotide Sequence of 2.88kbp of EHV-4 BamHI G



## Figure 6.6

## Alignment of the EHV-1 and EHV-4 gp13 Genes

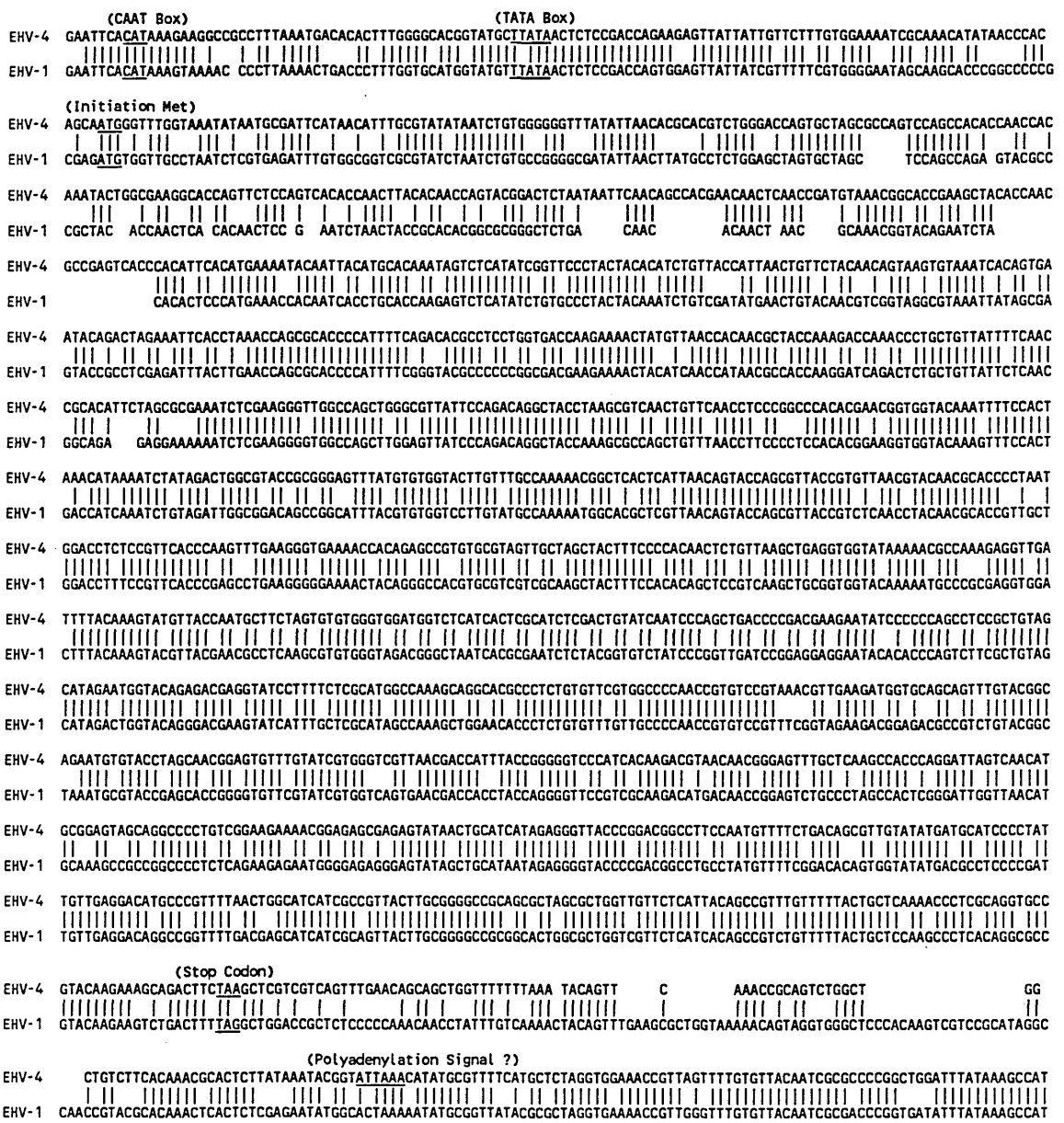



Figure 6.6 Alignment of the nucleotide sequences of the EHV-4 and EHV-1 gp13 genes and of 5′ and 3′ noncoding DNA. Putative promoter and polyadenylation signal sequences are underlined as are the initiation Met and termination codons.

## Figure 6.7

## Alignment of the Predicted Products of HSV-1 UL43 and its VZV and EHV-4 Counterparts

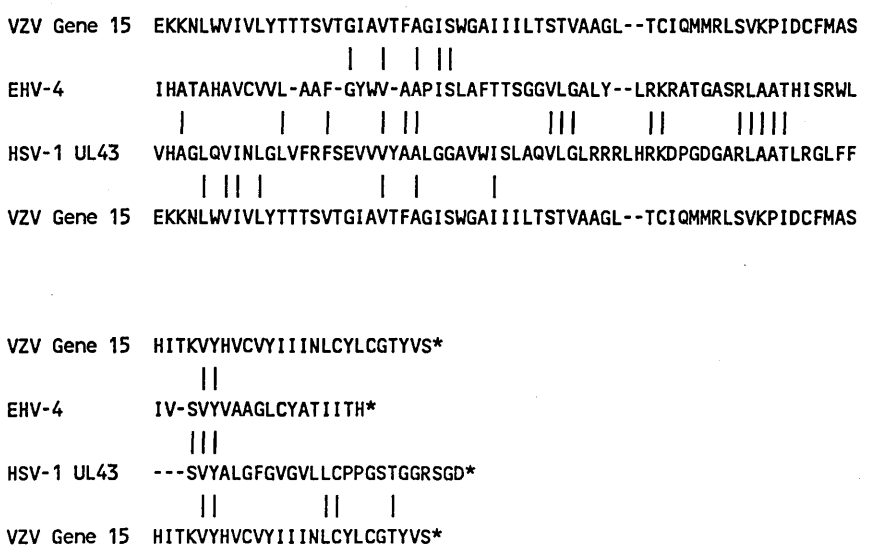



Figure 6.7 Alignment of the predicted products of gene UL43 of HSV-1, gene 15 of VZV, and the rightward flanking gene (in genome terms) of the EHV-4 gC gene homologue (C-terminal region only). Conserved residues are indicated by vertical lines.

Figure 6.8

Hydropathic Plot of EHV-4 gp13

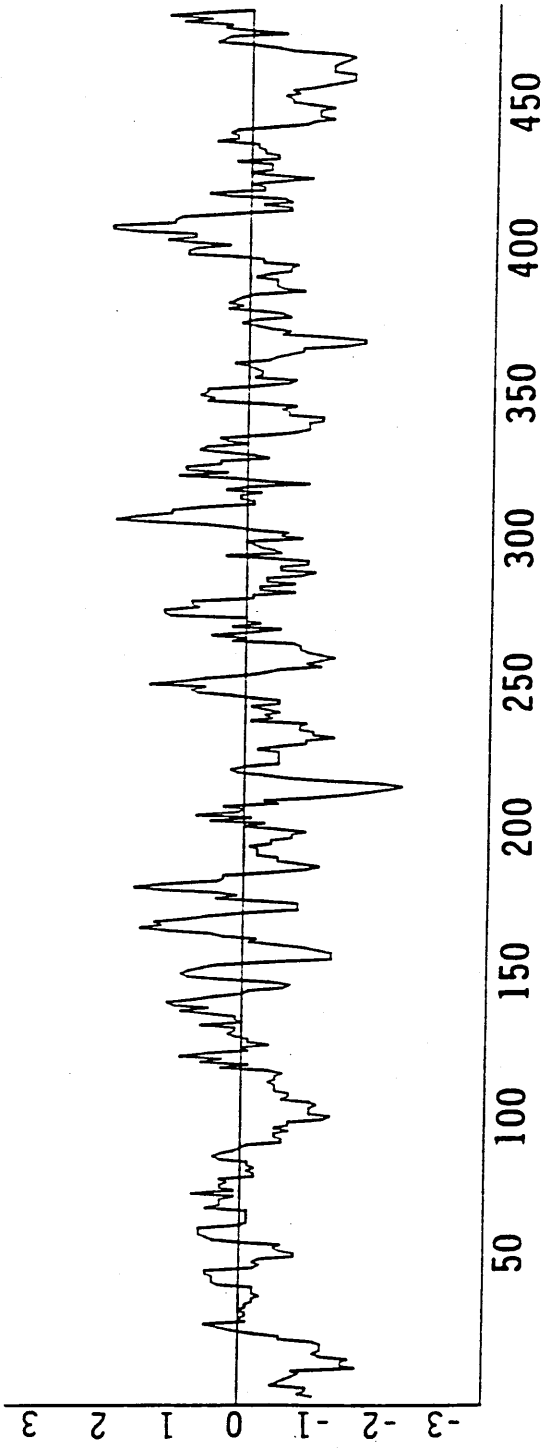



Figure 6.8 Hopp and Woods (1981) hydropathic plot comprising analysis of the local hydrophilicity of each residue of the predicted EHV-4 gp13 polypeptide using a six amino acid window. Positive values indicate a local hydrophilic environment: negative values a local hydrophobic environment.

Figure 6.9 Hydropathic and secondary structure analyses of EHV-4 gp13 using algorithms available on VAX software (Devereux et al., 1984). Hydropathic plots and predicted secondary structure features such as turns, alpha helices , and beta sheets are presented. The position of potential N-linked glycosylation sites along the primary sequence is indicated.

Figure 6.9

Secondary Structure Analysis of EHV-4 gp13

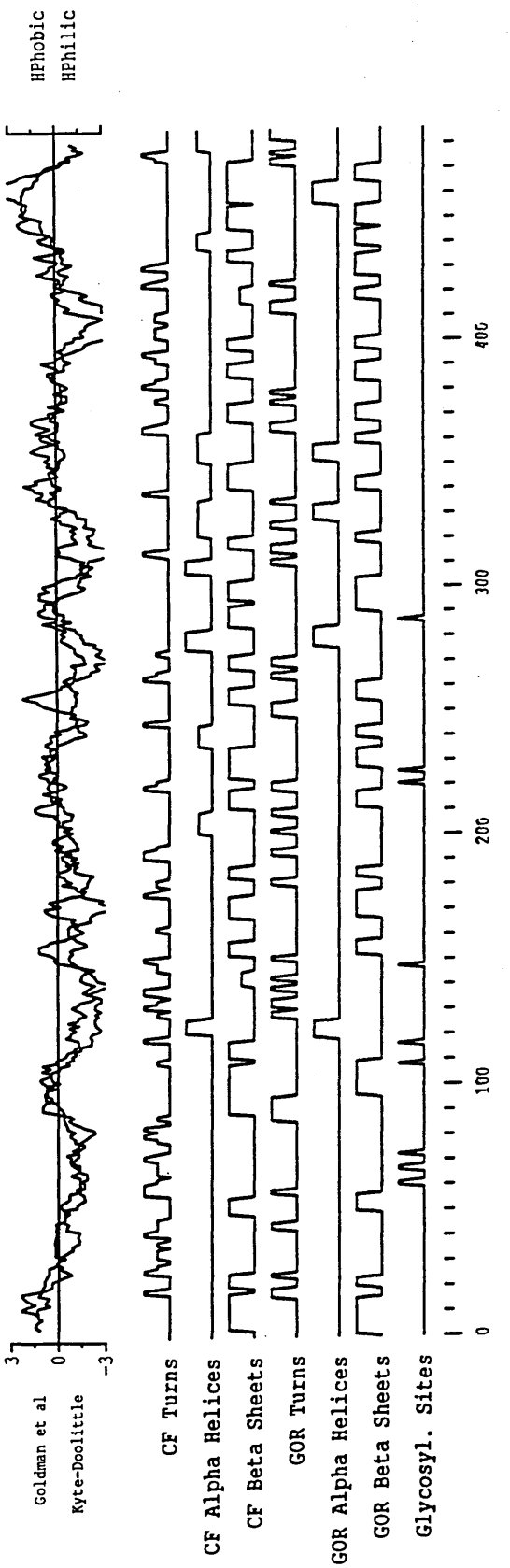


Figure 6.10 Predicted distribution of four structural domains within EHV-4 gp13 on the basis of analyses of local hydrophobicity scores along the polypeptide (Hopp and Woods, 1981). The 6 sites of cysteine conservation are indicated by black boxes (see Fig. 6.12). Potential N-linked glycosylation sites are underlined.

Figure 6.10

Putative Structural Domains within the EHV-4 gp13 Polypeptide

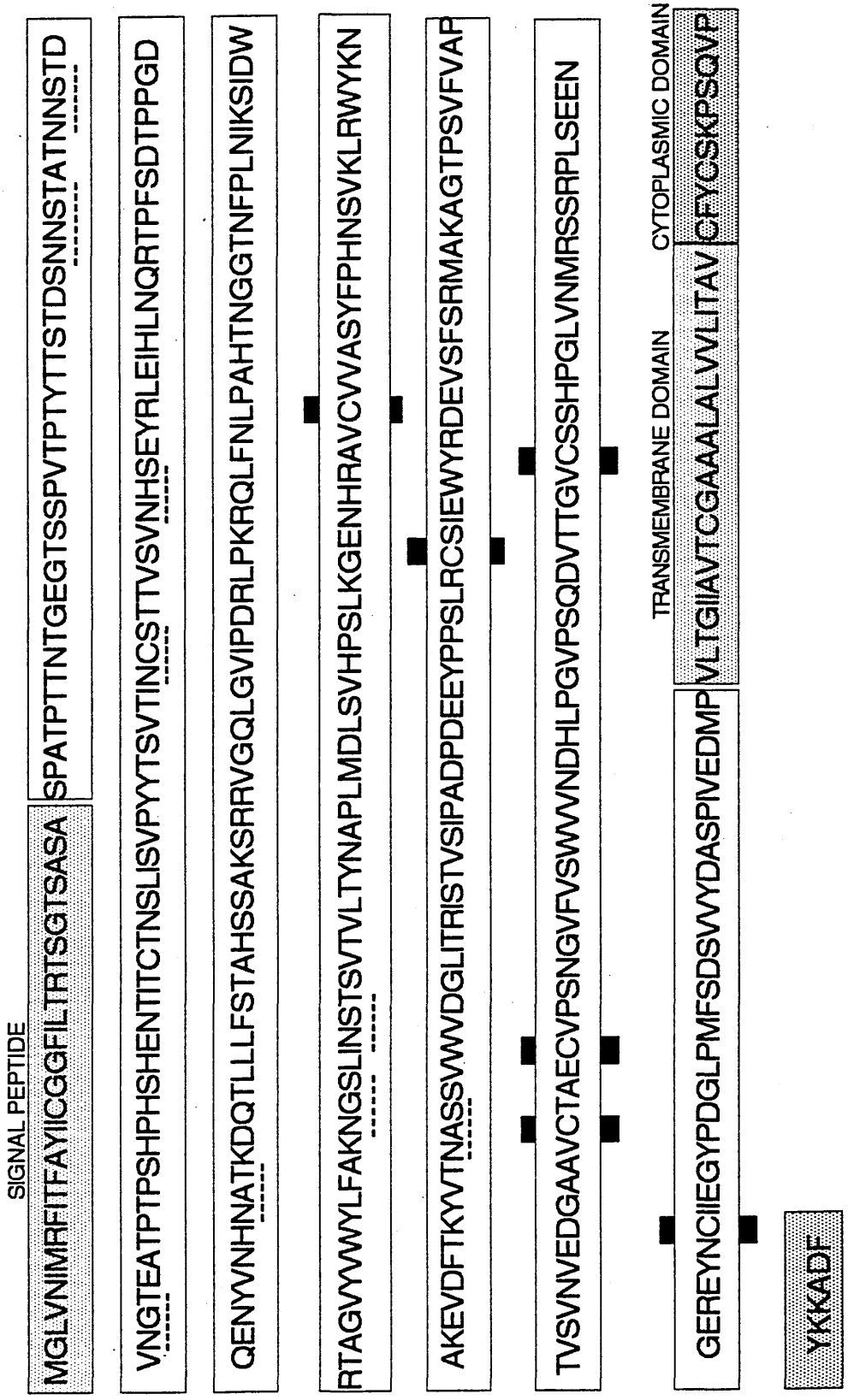


Figure 6.11

Dot Matrix Plots of Herpesvirus gC-Type Polypeptides

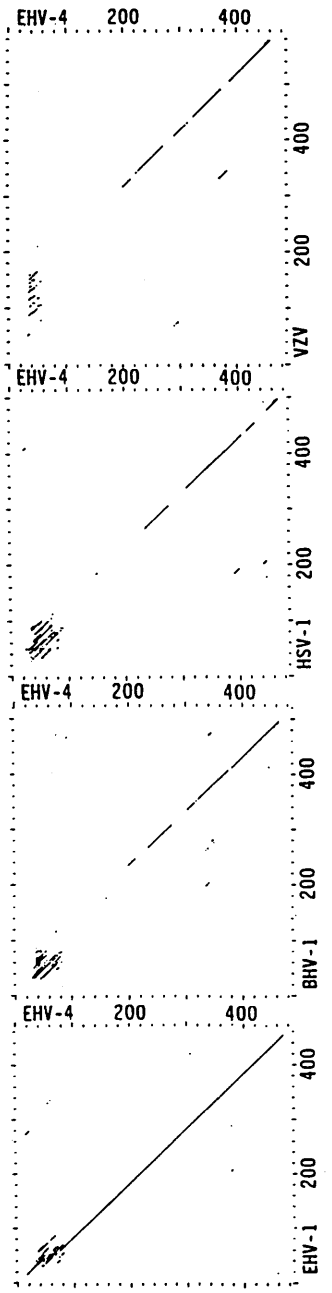



Figure 6.11 Comparison of EHV-4 gp13 with its counterparts in EHV-1 (Allen and Coogle, 1988), BHV-1 (Fitzpatrick et al., 1989), VZV (Kinchington et al., 1986), and HSV-1 (Dowbenko and Lasky, 1984; Frink et al., 1983) by dot matrix analyses. Parameters used were a 30 amino acid window and a stringency of 53%.

## Figure 6.12

## Multiple Alignment of Herpesvirus gC-Type Glycoproteins

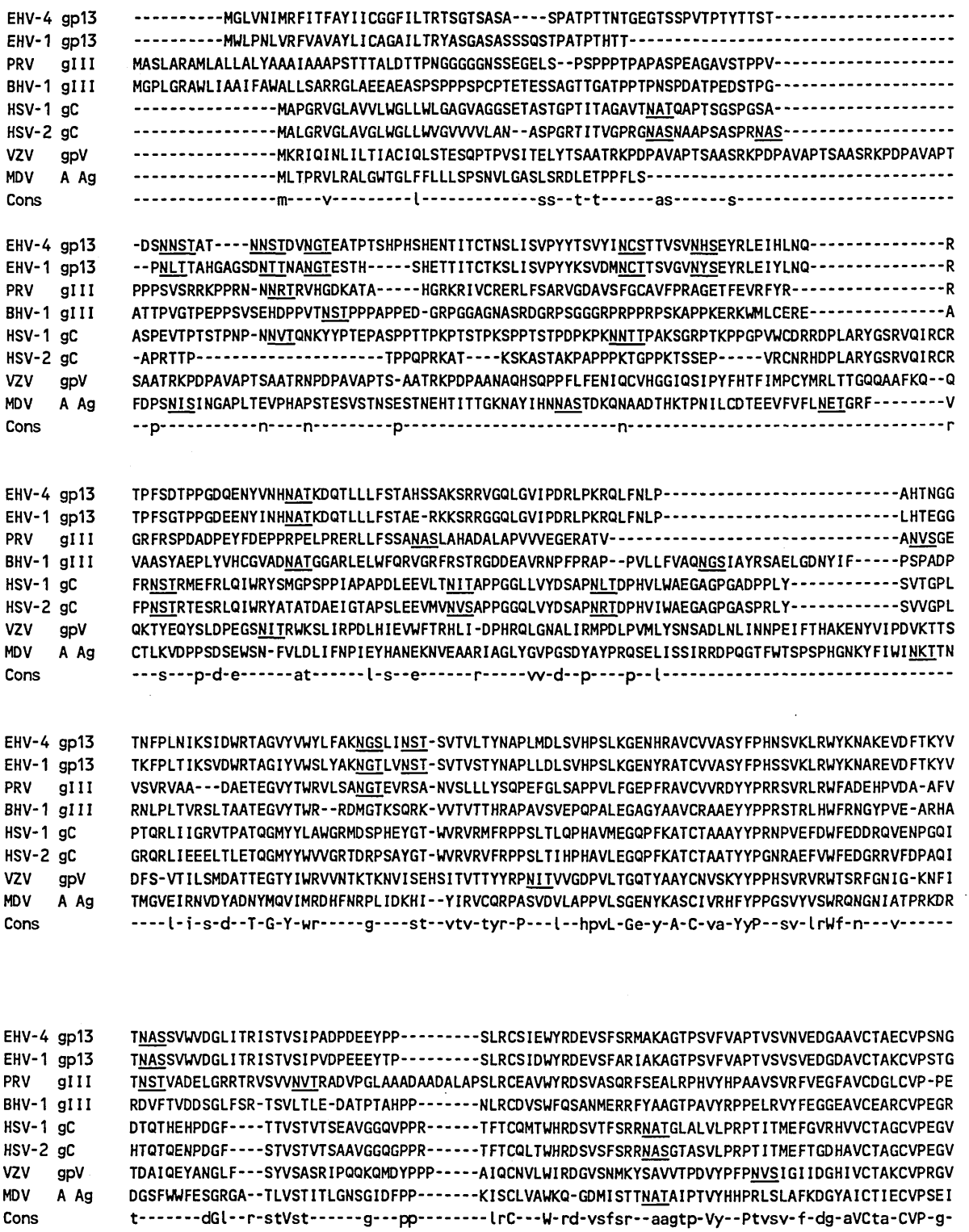


Figure 6.12 (cont.)

```
EHV-4 gp13   VFVSWVVNDHLPGVPSQDVTTGVCSSHPGLVNMRSSRPLSEENGEREYNCIIEGYPDGLPMFSDSVVYDASPIVEDMPVLTGIIAVTCGAA
EHV-1 gp13   VFVSWSVNDHLPGVPSQDMTTGVCPSHSGLVNMQSRRPLSEENGEREYSCIIEGYPDGLPMFSDTVVYDASPIVEDRPVLTSIIAVTCGAA
PRV   gIII   ARLAWS--DHAADTVYHL---GACAEHPGL-NVRSARPLSDLDGPVDYTCRLEGLPSQLPVFEDTQRYDASPASVSWPVVSSMIVVIAGIG
BHV-1 gIII   VSLRWTVRD--GIAPSRTEQTGVCAERPGLVNLRGVRLLSTTDGPVDYTCTATGYPAPLPEFSATATYDASPGLIGSPVLVSVVAVACGLG
HSV-1 gC     TFAWFLGDDPSPAAKSAVTAQESC-DHPGLATVRSTLPIS-YDYS-EYICR-TGYPAGIPVLEHHGSHQPPPRDPTERQVIEAIEWV-GIG
HSV-2 gC     TFAWFLGDDSSPAEKVAVASQTSC-GRPGTATIRSTLPVS-YEQT-EYICRLAGYPDGIPVLEHHGSHQPPPRDPTERQVIRAVEG-AGIG
VZV   gpV    VHFVWWVND-SPINHENSEITGVCDQNKRFVNMQSSCPTSELDGPITYSCHLDGYPKKFPPFSAVYTYDASTYATTFSVVAVIIGAISILG
MDV   A Ag   TVRWLVHDEAQPNTTYNTVVTGLCRTIDRHRNLLSRIPVWDNWTKTKYTCRLIGYPFDEDKFQDSEYYDATPSARGTPMVITVTAVL-GLA
Cons         vf--w-v-D--p---s----tgvC--hpglvn-rS-rPlS--dg--eY-Crl-GYP-glP-fsd---ydasP-----pvv---iav--G-g

EHV-4 gp13   ALALVVLITAVCFYCSKPSQVPYKKADF*
EHV-1 gp13   ALALVVLITAVCFYCSKPSQAPYKKSDF*
PRV   gIII   ILAIVLVIMATCVYYRQAGP*
BHV-1 gIII   AVGLLLV-AASCLRRKARVIQPGLTRARALGSAP*
HSV-1 gC     IGVLAAGVLVVTAIVYVVRTSQSRQRHRR*
HSV-2 gC     VAVLVAVVLAGTAVVYLTHASSVRYRRLR*
VZV   gpV    TLGLIAVIATLCIRCCS*
MDV   A Ag   VILGMGIIMTALCLYNSTRKNIRI*
Cons         -l-lv-vi-a-c-----------
```

Figure 6.12 Multiple alignment of the predicted products of the HSV-1 gC gene (Dowbenko and Lasky, 1984; Frink et al., 1983) and its counterpart in EHV-4, EHV-1 (Allen and Coogle, 1988), PRV (Robbins et al., 1986), BHV-1 (Fitzpatrick et al., 1989), HSV-2 (Dowbenko and Lasky, 1984; Swain et al., 1985), VZV (Kinchington et al., 1986) and MDV (Binns and Ross, 1989). A consensus line is detailed in which residues conserved in 4 to 6 gCs are represented by lower case, residues conserved in 7 or 8 gCs by upper case.

Figure 6.13

Comparison of Class II MHC Domain Sequences with BHV-1 and EHV-4 gp13 Domains

```
MHCIIb2hu  e p k v t v s p s k t q p l q h h n l l v c s v t d f y p g s i -
MHCIIb2ro  e p t v t v y p t r t q p l e h h n l l v c s v t d f y p g n i -
MHCIIa2hu  p p e v t v f s k s p v e l g q p n t l i c l v d n f f p p v v -
MHCIIa2ro  a p q a t v f p k s p v l l g q p n t l i c f v d n i f p p v i -
BHV-1 gIII a p a v s v e p q - p a l e g a g y a a v c r a a e y y p p r s t
EHV-4 gp13 A P L M D L S V H - P S L K G E N H R A V C V V A S Y F P H N S V
CONS       - P - v t v - p - - p - - l g - - n - l v c - v - - - - P - - - -

MHCIIb2hu  e v r w f r n g q e e t a g v v s t g l i r n g d w t - f q i l v
MHCIIb2ro  e v r w f r n g q e e k t g v v s t g l i r n g d w t - f q t l v
MHCIIa2hu  n i t w l r n g k p v t e g v s e t s f l s r s d h s - f r k f h
MHCIIa2ro  n i t w l r n s k s v t d g v y e t s f l v n r d h s - f h k l s
BHV-1 gIII r l h w f r n g y p v - e a r h a r d v f t v d d s g l f s r t s
EHV-4 gp13 K L R W Y K N A K E V D F T K Y V T N A S S V W V D G L I T R I S
CONS       - - - W - r N g - - v - - g v - - t - - - - - - d - - - f - - l -

MHCIIb2hu  m l e m t p q s g d v y t - - - - - c q v e h p s l t s p l t v
MHCIIb2ro  m l e m t p q s g e v y t - - - - - c q v e h p s l t s p v t v
MHCIIa2hu  y l t f l p s a e d v y d - - - - - c r v e h w g l d e p l l k
MHCIIa2ro  y l t f i p s d d d i y d - - - - - c k v e h w g l e e p v l k
BHV-1 gIII v l t l e d a t p t - a h p p n l r c d v s - w f q s a n m e r
EHV-4 gp13 T V S - I P A D P D E E Y P P S L R C S I E - W Y R D E V S F S
CONS       - l - - - p - - - d - y - - - - - - C - v e h w - l - - p - - -
```

Figure 6.13 Comparison of a region of the predicted EHV-4 gp13 polypeptide to the corresponding region of BHV-1 gIII which bears sequence similarity to MHC Class II domains (Fitzpatrick et al., 1989). MHC domain sequences detailed are consensus human and mice class IIa and b sequences (Figueroa and Klein, 1986; Kaufman et al., 1984).

Table 6.1

Codon Usage of the EHV-4 gp13 Gene

| 1st \ 2nd | A (26%) | C (31%) | G (17%) | T (26%) | 3rd | Residue | f | (%) |
|---|---|---|---|---|---|---|---|---|
| A | Lys 9 | Thr 18 | Arg 3 | Ile 9 | A (26%) | Ala | 32 | ( 6.6) |
| | Asn 23 | Thr 16 | Ser 10 | Ile 7 | C (31%) | Arg | 18 | ( 3.7) |
| (32%) | Lys 5 | Thr 10 | Arg 4 | Met 7 | G (17%) | Asn | 32 | ( 6.6) |
| | Asn 9 | Thr 6 | Ser 12 | Ile 7 | T (26%) | Asp | 21 | ( 4.3) |
| | | | | | | Cys | 12 | ( 2.5) |
| C | Gln 4 | Pro 15 | Arg 3 | Leu 6 | A | Gln | 7 | ( 1.4) |
| | His 9 | Pro 10 | Arg 5 | Leu 7 | C | Glu | 20 | ( 4.1) |
| (18%) | Gln 3 | Pro 5 | Arg 1 | Leu 7 | G | Gly | 27 | ( 5.6) |
| | His 4 | Pro 5 | Arg 2 | Leu 1 | T | His | 13 | ( 2.7) |
| | | | | | | Ile | 23 | ( 4.7) |
| G | Glu 14 | Ala 9 | Gly 5 | Val 12 | A | Leu | 30 | ( 6.2) |
| | Asp 16 | Ala 12 | Gly 9 | Val 4 | C | Lys | 14 | ( 2.9) |
| (31%) | Glu 6 | Ala 5 | Gly 5 | Val 10 | G | Met | 7 | ( 1.4) |
| | Asp 5 | Ala 6 | Gly 8 | Val 23 | T | Phe | 16 | ( 3.3) |
| | | | | | | Pro | 35 | ( 7.2) |
| T | End 1 | Ser 10 | End 0 | Leu 6 | A | Ser | 54 | (11.1) |
| | Tyr 11 | Ser 4 | Cys 6 | Phe 4 | C | Thr | 50 | (10.3) |
| (19%) | End 0 | Ser 5 | Trp 6 | Leu 3 | G | Trp | 6 | ( 1.2) |
| | Tyr 8 | Ser 13 | Cys 6 | Phe 12 | T | Tyr | 19 | ( 3.9) |
| | | | | | | Val | 49 | (10.1) |

Table 6.1 Codon usage of the EHV-4 gp13 gene and amino acid frequencies within its predicted product. The frequencies of each base within codon positions 1,2, and 3 are presented as percentages.

## Table 6.2
## Codon Usage of ORF3

| 1st \ 2nd | A (23%) | C (30%) | G (18%) | T (29%) | 3rd | Residue | f | (%) |
|---|---|---|---|---|---|---|---|---|
| A | Lys 5 | Thr 9 | Arg 3 | Ile 2 | A (25%) | Ala | 72 | ( 8.4) |
| | Asn 9 | Thr 5 | Ser 4 | Ile 1 | C (23%) | Arg | 37 | ( 4.3) |
| (26%) | Lys 2 | Thr 2 | Arg 0 | Met 8 | G (22%) | Asn | 50 | ( 5.8) |
| | Asn 2 | Thr 4 | Ser 0 | Ile 3 | T (30%) | Asp | 39 | ( 4.6) |
| | | | | | | Cys | 11 | ( 1.3) |
| C | Gln 6 | Pro 11 | Arg 1 | Leu 9 | A | Gln | 16 | ( 1.9) |
| | His 1 | Pro 4 | Arg 8 | Leu 3 | C | Glu | 34 | ( 4.0) |
| (29%) | Gln 2 | Pro 3 | Arg 2 | Leu 6 | G | Gly | 46 | ( 5.4) |
| | His 0 | Pro 1 | Arg 0 | Leu 8 | T | His | 17 | ( 2.0) |
| | | | | | | Ile | 58 | ( 6.8) |
| G | Glu 4 | Ala 3 | Gly 4 | Val 4 | A | Leu | 92 | (10.8) |
| | Asp 6 | Ala 5 | Gly 1 | Val 3 | C | Lys | 29 | ( 3.4) |
| (28%) | Glu 4 | Ala 4 | Gly 3 | Val 3 | G | Met | 25 | ( 2.9) |
| | Asp 4 | Ala 6 | Gly 5 | Val 4 | T | Phe | 32 | ( 3.7) |
| | | | | | | Pro | 52 | ( 6.1) |
| T | End 1 | Ser 1 | End 0 | Leu 1 | A | Ser | 71 | ( 8.3) |
| | Tyr 3 | Ser 1 | Cys 5 | Phe 0 | C | Thr | 58 | ( 6.8) |
| (17%) | End 0 | Ser 4 | Trp 0 | Leu 4 | G | Trp | 7 | ( 0.8) |
| | Tyr 3 | Ser 5 | Cys 5 | Phe 7 | T | Tyr | 40 | ( 4.7) |
| | | | | | | Val | 69 | ( 8.1) |

Table 6.2 Codon usage of ORF3 and amino acid frequencies within its predicted product. The frequencies of each base within codon positions 1,2, and 3 are presented as percentages.

## DISCUSSION

Sequence analysis of 2.88kbp of EHV-4 BamHI G revealed three open reading frames, two complete and one partial. The central ORF (ORF2) was identified as the EHV-4 gp13 gene, a homologue of the HSV-1 gC gene. The partial ORF mapping to the right of the gp13 gene (in genome terms)(ORF1) was identified as a homologue of HSV-1 gene UL43. Mapping to the left of the gp13 gene was an ORF (ORF3) which could not be identified as a homologue of any HSV-1 or VZV genes. Comparison of the entire VZV and HSV-1 genomes has indicated this region of the genome corresponds to one of the few regions of the $U_L$ component at which the two viral genomes encode nonhomologous proteins: HSV-1 gene UL45 encodes a 172 amino acid protein of unknown function transcribed in the same sense as gC ; VZV gene 13 encodes thymidylate synthetase which is transcribed in the opposite sense to gpV (Davison and Scott, 1986; McGeoch et al., 1988a). It is thus possible the function specified by ORF3 may be unique to EHV-4 and its closest evolutionary relatives at least at this particular genome position. Given the high frequency of N-X-S/T sites within the predicted ORF3 product it is tempting to speculate that ORF3 encodes one of the EHV-1- and EHV-4-specific glycoprotein species for which a genome coding position has yet to be established. However, the predicted polypeptide lacks a characteristic N-terminal signal peptide and C-terminal transmembrane domain . A short stretch of hydrophobic residues interspersed with Ser and Thr residues positioned N-terminal to three internal Met codons, $Met_{71,73,76,}$ could potentially serve as a signal peptide with one of the internal Met codons acting as initiation codon. Alternatively, the first Met in the ORF could be used as the initiation codon since uncharacteristically long signal peptides have been identified in other glycoproteins, including EHV-4 gB (Riggio et al., 1989). However, in

gB-type glycoproteins the long signal peptide is hypothesised to result from sequence constraints imposed by the overlap of the gB gene with the ICP 18.5 gene homologue and no overlap exists in this situation. In the absence of data linking the product of the ORF3 gene to a specific EHV polypeptide the function of the ORF3 product remains unknown.

The EHV-4 gp13 gene maps between coordinates 0.15 and 0.17 within genomic fragment BamHI G. This coding position is consistent with that of VZV gpV (0.11; Davison and Scott, 1986; Kinchington et al., 1986), EHV-1 gp13 (0.136-0.148; Allen and Coogle, 1988 ), and BHV-1 gIII (0.122-0.135; Fitzpatrick et al., 1989) and inverted relative to that of HSV-1 and HSV-2 gC (0.62-0.64; Frink et al., 1983; Dowbenko and Lasky, 1984 ) and MDV (Binns and Ross, 1989). In contrast, the PRV gIII maps within the central region of the $U_L$ component at 0.4, due to a 0.1-0.4 m.u. inversion ( Ben-Porat et al., 1983; Robbins et al., 1984 ). Accordingly, in EHV-1, EHV-4 and VZV the gene is transcribed from right to left, whereas in HSV ($P_L$ isomer), MDV, and PRV it is transcribed in the opposite direction.

The coding region of the gp13 gene of EHV-4 has a %GC composition of 48% which compares to 53%, 73%, 75% and 47% with the gC gene homologues of EHV-1, BHV-1, PRV, and VZV (Allen and Coogle, 1988; Fitzpatrick et al., 1989 ; Kinchington et al., 1986 ; Robbins et al., 1986a).

The gp13 coding region is defined by an open reading frame of 1455 base pairs or 485 amino acids. Although putative TATA and CAAT sequences are located upstream of the open reading frame it remains to be determined which of these are functionally operative. The HSV-1 gC gene is regulated as a $gamma_2$ gene through signals specified within a 15bp sequence element which includes the gC TATA box (Homa et al.

1988). A CAAT box does not therefore seem to be required for HSV true late gene expression . The sequence ATTAAA, which serves as a polyadenylation signal in a minority of HSV-1 transcripts, is positioned 100bp downstream of the stop codon . The EHV-4 gp13 may thus be transcribed, like PRV gIII (Robbins et al., 1986a), as an independent unit. No putative polyadenylation signal is specified within the corresponding region of the EHV-1 genome (Allen and Coogle, 1988) leading to the suggestion that the EHV-1 gp13 gene may be transcribed as part of a multigenic RNA in an analogous manner to HSV-1 gC transcription in which the gC mRNA encodes both gC and a 172 amino acid open reading frame, UL45, located just downstream of the gC gene (Frink et al., 1983; McGeoch et al., 1988a). The HSV-1 UL45 gene is transcribed as an independent 730 bp mRNA under the control of its own promoter in addition to its transcription as part of the gC transcript (Frink et al., 1983). Whatever the form of the EHV-4 gp13 transcript, whether uni- or multigenic, the existence of putative TATA and CAAT boxes upstream of EHV-4 ORF3 suggests it may similarly be transcribed under control of its own promoter. The protein encoded by ORF3 shares no significant homology with VZV or HSV-1($I_L$) genes encoded immediately left of the gC gene.

The predicted EHV-4 gp13 protein is 485 amino acids in length with a molecular weight of 49117 on cleavage of the signal peptide. In contrast, gp13 isolated from virions has an apparent molecular weight, as determined by SDS-PAGE, of approximately 92000 (Meredith et al., 1989) indicating that gp13 is subject to co/post-translational processing. Such processing is likely to include cleavage of the putative 32 amino acid signal sequence, N-linked glycosylation at up to nine positions per molecule , and possibly, as in HSV-1 gC, O-linked glycosylation and sulphation.

Computer-assisted analysis of the primary sequence by hydropathic plot and secondary structure prediction algorithms (Queen and Korn, 1984) permits the distinction of four structural domains i.e. an N-terminal signal sequence (amino acids 1-32), a large extracellular domain (amino acids 33-444) across which 11 putative glycosylation sites are distributed, a transmembrane domain (amino acids 445-472) and a polar C-terminal domain (amino acids 473-485). Although there is no biological evidence for the functional roles of specific domains within the EHV-4 gp13, the essential role of the extreme N-terminal region for entry into the export pathway and the involvement of the putative transmembrane domain and C-terminal cytoplasmic domain in the efficiency of anchoring of the glycoprotein to the plasma membrane have been reported for the homologous PRV gIII and HSV-1 gC proteins (Enquist et al., 1988; Holland et al., 1984a, 1988; Robbins et al., 1989; Ryan et al., 1987).

The degree of identity of EHV-4 gp13 protein with other alphaherpesvirus gCs ranged from 25-29% with HSV-1 gC, HSV-2 gC, PRV gIII, VZV gpV, MDV A antigen, and BHV-1 gIII to 79% with EHV-1 gp13.

Fitzpatrick and coworkers (1989) have proposed, on the basis of their finding significant homology within a domain of BHV-1 gIII to that of class II MHC domains , that the herpesvirus gC-type glycoprotein may have been derived from a host cell gene encoding a member of the Ig superfamily. They have further proposed that the similar proposed structure of gIII and Ig superfamily domains may be involved in a role for gC in immune modulation.

Multiple alignment or matrix plot comparisons of the predicted protein sequence of herpesvirus gCs highlights a greater constraint on the sequence of the 3′ end of the gene, encoding C-terminal amino

acids. Allen and Coogle (1988) have proposed that the N-terminal domain of EHV-1 gp13 may specify epitopes which elicit a protective immune response in which case the greater divergence of this domain may be a consequence of the selective advantage of N-terminal gp13 variants in the infection of horses previously exposed to EHV. The greater conservation of the C-terminal domain may be due to sequence constraint imposed by some function conserved throughout the herpesvirus gCs. In HSV-1 gC the N-terminal domain is not restricted to an antigenic role since C3b binding activity of the protein is apparently dependent on structural elements within the N and C terminal regions of the protein (Huemer et al., 1989). EHV-1 is known to induce C3b binding activity on the surface of infected cells (Bielefeldt Ohmann and Babiuk, 1988) although this has yet to be assigned to a particular protein. EHV-4 has yet to be tested in this respect but it might be expected that it too would specify C3b binding activity at least in the purified gp13 protein if not on the surface of infected cells. By binding C3b gC may act to inhibit the effect of the complement cascade such that virus or virus-infected cells are protected form complement-mediated damage (Fries et al., 1986; McNearney et al., 1987).

Allen and Coogle (1988) suggest the function of gC is adaptation to infection of the host species which is supported by evidence that the PRV, BHV and HSV 'gCs' at least are involved in adsorption to the host cell and might therefore specify tissue tropism. However, although gC is non-essential for the replication of herpesviruses such as HSV-1 in vitro, field gC$^-$ viruses have not been isolated suggesting an additional important role in vivo perhaps in pathogenicity as in PRV . Whatever the function of the gC-type glycoprotein of herpesviruses it is clear that regions of the protein can tolerate greater divergence than others so the hypothesis of Allen and Coogle is likely to be true

to an extent: parts of the polypeptide specifying immunodominant antigenic determinants , tissue tropism and adaptation to host niche exhibiting divergent sequence and other parts being more conserved to retain important functional conformations.

An immunodominant antigenic site of EHV-1 gp13 comprising six continuous epitopes has been mapped to amino acids 136-191, corresponding to the major hydrophilic region within a Hopp and Woods hydropathic plot. Although the corresponding region of EHV-4 (amino acids 152-207) is hydrophilic in nature and may therefore be exposed on the surface of the protein constituting an antigenic domain, the major hydrophilic peak in the EHV-4 hydropathic plot is located around amino acid 409. Given the 79% identity between the EHV-1 and EHV-4 gp13 amino acid sequences it might be expected that the proteins possess type-common epitopes as is demonstrated by the cross reactivity of sera raised against EHV-1 or 4 gp13 with the heterologous gp13 (Allen and Bryans, 1986). However, Allen and coworkers have reported that 85% of immunodominant EHV-1 gp13 antigens are type-specific (Allen et al., 1988). Given the 79% homology between the predicted primary translation products of the EHV-1 and EHV-4 gCs it seems unlikely that the protein would contain predominantly type-specific epitopes although the immunodominant epitopes may be specific to one or other type.

Studies on the gCs of herpesviruses are encouraging with regard to the potential of this protein as a vaccine component (see next chapter). However, the role of individual EHV-4 glycoproteins in the elicitation of a humoral and cell-mediated immune response has not yet been extensively studied.

# CHAPTER 7

# ANALYSIS OF THE IMMUNOGENICITY OF EHV-4 gp13 AND gH PEPTIDES IN HAMSTERS

## INTRODUCTION

This chapter outlines preliminary work designed to apply sequence data on the EHV-4 gp13 and gH-type glycoprotein genes to the investigation of immunogenicity of their products.

Herpesvirus gC-type glycoproteins have been implicated in the elicitation of both humoral and cellular immune responses within the host and are often a major target of these mechanisms. PRV gIII is the immunodominant protein with respect to a PRV neutralising reponse (Ben-Porat et al., 1986), and HSV-1 gC, VZV gpV, PRV gIII, and BHV-1 gIII are targets for cytotoxic T-cells in animal models or the natural host (Fitzpatrick et al., 1989; Giller et al., 1990; Glorioso et al., 1985; Rosenthal et al., 1987; Zuckermann et al., 1990). Recombinant vaccinia virus expressing HSV-1 gC and anti-gC monoclonal antibodies protect mice against HSV-1 challenge (Kumel et al., 1985; Weir et al., 1989) suggesting that gC might be a useful constituent of a subunit vaccine. Studies on the gp13 glycoproteins of EHV-1 and EHV-4 have similarly indicated an important role in interaction with host immune mechanisms. gp13 is a major target of the humoral immune response and at least four neutralising epitopes (Allen and Bryans, 1986; Allen et al., 1988). Involvement in cellular immunity has been implied through i) demonstration of lymphocyte transformation against an EHV-1 and EHV-4 protein fraction which contains gp13 (Bridges et al., 1988), and ii) complement mediated lysis of EHV-1 infected cells by anti-EHV gp13 monoclonal antibodies (MAbs) (Stokes et al., 1991). Anti-EHV-1 gp13 monoclonal antibodies and a recombinant gp13 vaccinia virus protect hamsters against challenge (Guo et al., 1989, 1990; Shimizu et al., 1989; Stokes et al., 1989, 1991).

Mapping of epitopes or epitope domains on herpesvirus glycoproteins has been approached in several ways including i) sequence analysis of

monoclonal antibody resistant (mar) glycoprotein mutants , ii) analysis of the reactivity of short peptide sequences or protein fragments with anti-glycoprotein antibody , and iii) analysis of competitive binding of different monoclonal antibodies (Allen et al., 1988; Brocade Wu et al., 1990; Fitzpatrick et al., 1990; Jacobs et al., 1990). The effectiveness of the analysis of the antigenicity of a given glycoprotein will depend on the source of the target material whether native glycoprotein purified from virion extract, intact or partial polypeptide expressed in a prokaryotic or eukaryotic expression system, or linear fragments produced by peptide synthesis.

A factor which might hinder the antigenic analysis of gC-type glycoproteins expressed in non-eukaryotic systems is the extensive glycosylation of the protein. Studies on HSV-1 gC have indicated that a proportion of epitopes are dependent on carbohydrate for their expression presumably because either the carbohydrate alters the local polypeptide conformation and/or is itself antigenic (Marlin et al., 1985; Olofsson et al., 1990; Sjoblom et al., 1987). Current evidence suggests the former option . Two antigenic sites exist - a carbohydrate-dependent site (II) in the N-terminal half of the protein, and a carbohydrate-independent site (I) mapping from amino acids 297-360 (Marlin et al., 1985) (Figure 1). The carbohydrate dependency of site II epitopes is apparently due to N-linked rather than O-linked oligosaccharides (Olofsson et al., 1990). No data is currently available on the glycosylation state of EHV-1 and EHV-4 gp13 and the dependency of their antigenic structures on carbohydrate residues other than that both glycoproteins possess potential N-linked glycosylation sites and both mature glycoproteins have an apparent molecular weight twice that predicted from the amino acid sequence. Although it seems likely that the HSV-1 gC and the EHV gp13

glycoproteins share some structural features given the conservation of several cysteine residues it is impossible to predict from existing data whether the EHV proteins will exhibit a similar distribution of carbohydrate-dependent and -independent epitope domains.

Current data on epitope domains within EHV glycoproteins is limited to studies on EHV-1 gp14 and gp13 (Allen and Coogle, 1988; Allen et al., 1988; Sinclair et al., 1990; G. Allen, personal communication). An immunodominant domain has been identified within EHV-1 gp13 spanning amino acids 136 to 191 (Allen and Coogle, 1988). The hydrophilic peak of the polypeptide sequence as predicted by the Hopp and Woods algorithm lies within this domain at amino acids 145-150.

Allen and coworkers (1988) analysed 42 EHV-1 gp13-specific monoclonal antibodies for their type specificity and reactivity with a series of EHV-1 isolates to investigate the epitopic characteristics of the protein. The Mabs collectively interacted with at least 16 epitopes. 85% of the Mabs were type-specific, 20% elicited the production of virus-neutralising antibody in the presence of complement, and 90% varied between isolates. Four neutralising monoclonals were isolated, one of which was cross reactive with EHV-4 gp13. Sinclair and coworkers (1989) isolated three anti-gp13 (EHV-1) monoclonal antibodies which recognise conformational epitopes and which neutralise EHV-1 infectivity in vitro in the presence of complement. Of these MAbs, two crossreact with EHV-4 but do not neutralise the virus . Sinclair and coworkers (1990) mapped epitope domains to EHV-1 and EHV-4 gp13 by competitive binding of monoclonal antibodies. Their results indicate that EHV-1 gp13 possesses an epitope domain composed of both type-specific and type-common epitopes which elicit the production of antibody which neutralises EHV-1. Four type-specific domains were mapped, three to EHV-4 gp13 and one to EHV-1 gp13 which

did not elicit production of neutralising antibody.

A second EHV-4 glycoprotein of interest in immunogenic terms is the gH-type glycoprotein. On the basis of work done on the gH-type glycoproteins of other herpesviruses including HSV-1, VZV, and EBV, this glycoprotein species is apparently either low in abundance and/or immunogenicity compared to several other glycoprotein species . Nevertheless gH-type glycoproteins specify at least one epitope that elicits the production of antibody which neutralises viral infectivity in the absence of complement.

HSV-1 gH can been produced in a mammalian transient expression system and in a recombinant vaccinia virus system to levels greater than that obtained in viral infection (Forrester et al., 1990; Gompels and Minson, 1989 ). However, the glycoprotein produced differs biologically from native viral gH in terms of its size, antigenic nature, and cellular localisation with no gH expressed on the infected cell plasma membrane. These differences can be partially resolved by superinfection of the transfected cells with wild type virus. A block in the transport of gH to the cell surface has also been reported for HCMV gp58 and EBV gp85 transfected into mammalian cells ( Cranage et al., 1988; Heineman et al., 1988). This altered targetting of gH could be due to the absence of herpesvirus-specified proteins critical to the intracellular transport of gH. Alternatively, perhaps gH transport is dependent on gross effects of virus infection. Whichever is the case the analysis of gH immunogenicity is obviously hampered by its incorrect expression in eukaryotic transfection systems and by its low abundance within virions. The failure of PRV gH expressed in baculovirus to elicit neutralising antibody may be similarly due to an aberrant expression process (Petrovskis et al., 1988).

As mentioned in Chapter 4, the gH-type glycoproteins of EHV-1 and

EHV-4 have yet to be identified and as a result the involvement of these glycoproteins in eliciting an immune response is unknown. The fact that no monoclonal antibodies were raised to a product of the EHV-1 gH gene in the glycoprotein gene mapping study of Allen and Yeargan (1987) suggests that, like other gH-type glycoproteins, it may be low in abundance and/or immunogenicity.

Since sequence analysis of the EHV-4 gp13 and gH genes provided us with predicted amino acid sequences for gp13 and gH we set about to investigate the immunogenicity of peptides derived from the external domain of these glycoproteins. Three gp13 and two gH peptide sequences were synthesised and their immunogenicity was determined in a hamster model by our collaborators, Dr. Anne Stokes, Dr. Keith Murray, and Dr. Tim Doel at the Institute for Animal Health , Pirbright.

## MATERIALS AND METHODS

### 1. Selection of Peptide Sequences for Investigation

Peptides were selected from regions of the polypeptide predicted to be oriented on the surface of the molecule and therefore likely to be accessible to the immune system of the infected host. In the first instance the Hopp and Woods algorithm (1981) was used. As mentioned in Chapter 6 in connection with hydropathic analyses, different algorithms and predictive programs provide different outputs. In order to increase the chances that the tested peptides would be immunogenic, the polypeptide sequence was subjected to multiple predictive programs and peptides selected from those regions of the extracellular domain which were most frequently represented as putative surface regions. Peptide sequences of 17-20 amino acids in length were synthesised and a cysteine residue attached to the N or C terminus of each in case conjugation to carrier protein might be required at a later date (Fig. 2).

### 2. Immunisations and Immunological Assays

Techniques utilised in this chapter are described in Stokes et al., 1989, 1991. Briefly, the experimental animals used were 6-8 week old Syrian hamsters (Mesocricetus auratus) of the DSN inbred strain. 500ug doses of peptide were administered subcutaneously in an equal volume of complete Freunds adjuvant, to hamsters on day 0. Further immunisations were performed, in incomplete Freunds adjuvant as follows - days 14 and 34 (peptides 15/89, 15/89, and 16/89), days 13 and 52 (peptide 25/89), and day 14 (peptide 28/89).

Virus neutralising assays were conducted by incubating serial dilutions of hamster sera with guinea pig complement and 100 $TCID_{50}$ units of virus at $37^{o}C$ for 1 hour. The preparation was then added to

equine foetal kidney (EFK) cells in monolayer culture and the cytopathic effect determined.

Complement-mediated antibody lysis (CDL) assays entailed the incubation of EHV-1-infected EFK cells radiolabelled with $^{51}Cr$ with heat inactivated hamster serum in the presence of guinea pig complement. Lysis of target cells was monitored by determining the amount of $^{51}Cr$ released into the medium from lysed cells.

Peptides were coated on ELISA plates at a concentration of 3ug/ml as solid phase for peptide-specific ELISAs.

Hamsters immunised with peptides 14/89, 15/89, and 16/89 were challenged with $2x10^{5.3}$ $TCID_{50}$ units of EHV-1 strain Kentucky D (KyD) on day 64 ( administered intraperitoneally (i.p.)) . Hamsters inoculated with peptides 25/89 and 28/89 (gH peptides) were challenged i.p. with $1x10^{6}$ $TCID_{50}$ units of EHV-1 (KyD) on day 132 (peptide 25/89) or day 79 (peptide 28/89).

## RESULTS

### 1. Investigation of the Immunogenicity of EHV-4 gp13 Peptides

A comparison of the HSV-1 gC, EHV-1 gp13 (Allen and Coogle, 1988) and EHV-4 gp13 glycoproteins is detailed in Figure 1 with the selected EHV-4 peptides underlined. On the basis of Hopp and Woods' algorithm (1981) the peak hydrophilic value of EHV-4 gp13 occurred at amino acids 406-411 (Fig.6.8). One peptide was selected spanning this region, 16/89 (amino acids 400-415), and two further peptides, 14/89 (37-54), 15/89 (157-176). These peptide sequences are compared to the corresponding regions of EHV-1 gp13 in Figure 2. Peptide 16 is 'type-common' possessing 15 of 17 residues identical to those of the corresponding peptide in EHV-1 gp13. Peptides 14 and 15 are more 'type-specific' in nature with 3/18 and 15/20 residues in common with the corresponding EHV-1 gp13 peptides.

The immunogenicity of these peptides was investigated by the immunisation of hamsters with 500ug of peptide at 0, 14, and 34 days. Hamster sera was investigated at intervals for its reactivity against the peptide used in the immunisation (Table 1) and against whole virus (EHV-4) (Table 2). Each peptide elicited the production of anti-peptide antibody as demonstrated by the peptide-specific ELISA results (Table 1). Anti-virus responses were evident within some, but not all, of the immunised hamsters (Table 2). Peptide 15 induced the highest titres of antibody against peptide and against whole virus. It is interesting that this peptide maps within part of the EHV-4 gp13 polypeptide which shares sequence identity with the immunodominant domain of EHV-1 gp13 (Fig. 1). None of the hamster sera were capable of virus neutralisation *in vitro* (data unshown) and immunisation with the peptides did not protect hamsters against challenge with EHV-1 (Table 3).

Peptide 15 was reactive with an anti-EHV-1 gp13 MAb p17 produced by R. Killington and coworkers in Leeds mapping a linear epitope to amino acids 157-176 of EHV-4 and 141-159 of EHV-1 gp13. Assuming the reactivity with the EHV-4 peptide is not spurious this suggests the epitope maps within the C-terminal end of the peptide where there is a run of six identical amino acids (Fig. 2).

2. Investigation of the Immunogenicity of EHV-4 gH Peptides

Two peptides - amino acids 287-303 and 620-639 were selected from regions of the EHV-4 gH polypeptide predicted to be surface oriented. These peptides are compared to similar regions of the EHV-1 gH polypeptide in Figure 2 (Robertson et al., 1991).

Immunisation of hamsters with 500ug of peptides 25/89 and 28/89 induced production of anti-peptide antibody and, in some hamsters (121,123,152,154,155) anti-virus (EHV-1 and EHV-4) antibody (Table 4 ). Neither peptide induced antibody which neutralised EHV-4 in vitro (data unshown). Sera with anti-EHV-4 antibody activity exhibited anti-EHV-1 activity indicating the induction of antibody to both EHV-1 and EHV-4 proteins by immunisation with EHV-4 gH peptides. This is to be expected given the extent of identity between the peptides and their counterpart sequences in EHV-1 gH (Fig. 2).

Immunisation of hamsters with peptides 25/89 or 28/89 did not appear to induce protective immunity against EHV-1: virus titre in the liver and lungs of challenged hamsters was comparable to that of control (non-immunised) hamsters (Table 3).

## Figure 7.1

## Alignment of the gC-Type Glycoproteins of EHV-1, EHV-4, and HSV-1

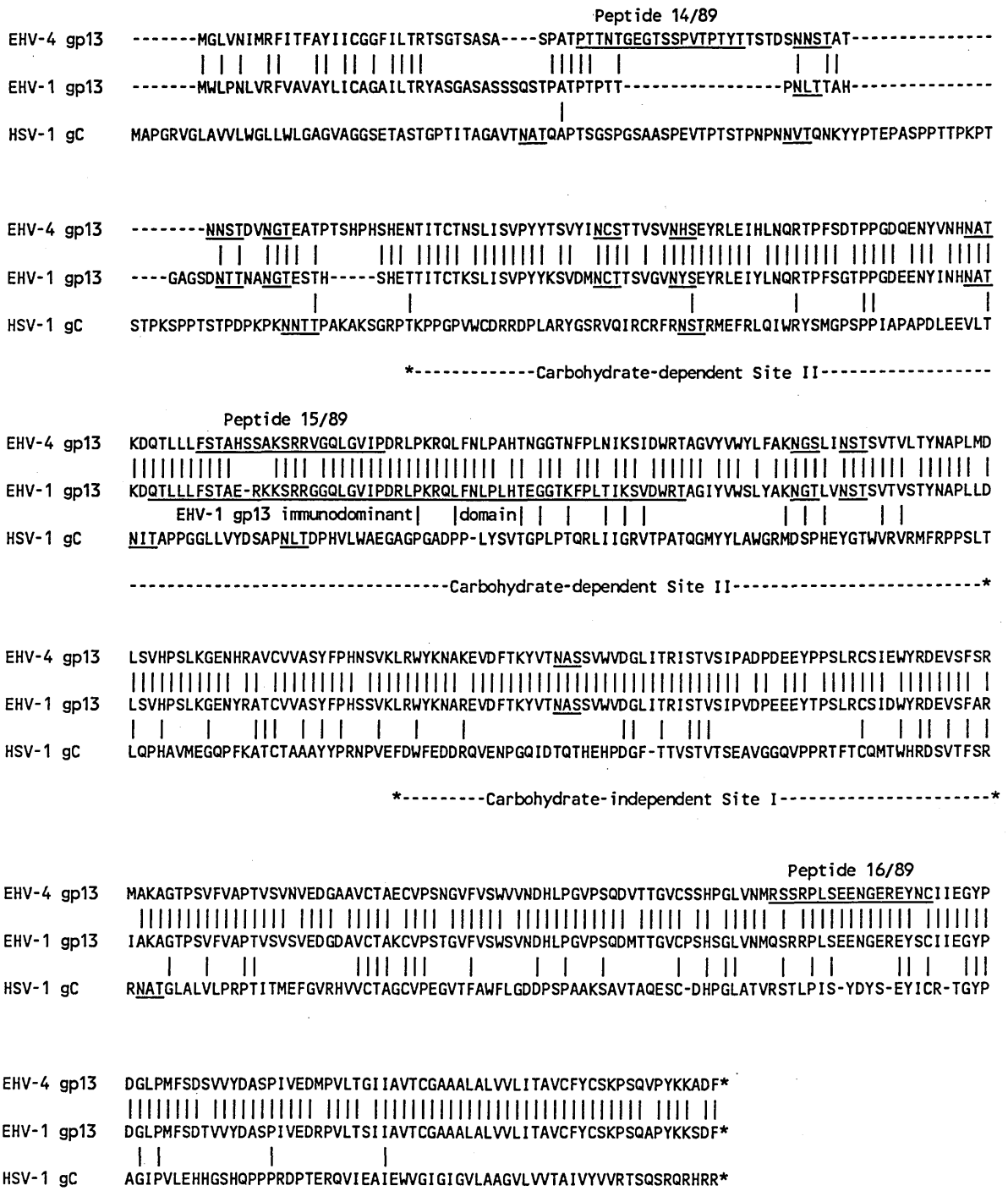



Figure 7.1 Alignment of the gC-type glycoproteins of EHV-4, EHV-1, and HSV-1 with mapped antigenic sites of the latter two polypeptides indicated. Three EHV-4 peptides selected for immunogenic analysis are underlined.

## Figure 7.2
## EHV-4 gp13 and gH Peptides

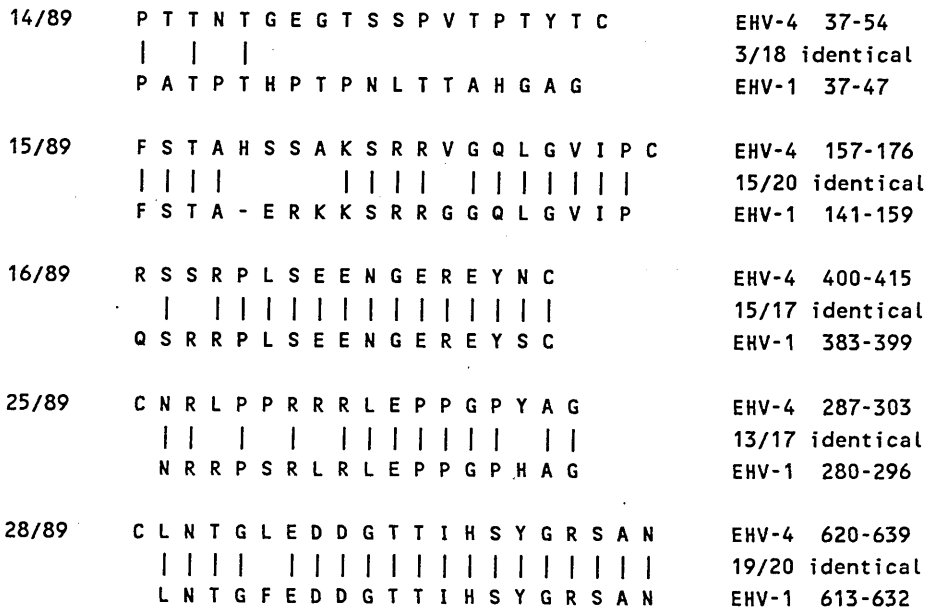



Figure 7.2 Identity of selected EHV-4 gp13 and gH peptides with the corresponding regions of the EHV-1 gp13 and gH glycoproteins (Allen and Coogle, 1988; J.M. Whalley, personal communication).

Table 7.1

Anti-gp13 Peptide ELISA

| Hamster | Peptide | Time (days) 0 | 14 | 20 | 34 | 44 |
|---|---|---|---|---|---|---|
| 96 | 14/89 | <50 | <50 | <50 | 400 | 1600 |
| 97 | 14/89 | <50 | <50 | 200 | >6400 | >6400 |
| 98 | 14/89 | <50 | <50 | 100 | 400 | 800 |
| 99 | 14/89 | <50 | <50 | 50 | 1600 | 1600 |
| 100 | 14/89 | <50 | <50 | <50 | 200 | 800 |
| 101 | 15/89 | <50 | <50 | <50 | 1600 | >6400 |
| 102 | 15/89 | <50 | <50 | 100 | | 3200 |
| 103 | 15/89 | <50 | >6400 | >6400 | >6400 | >6400 |
| 104 | 15/89 | <50 | 50 | 400 | >6400 | >6400 |
| 105 | 15/89 | <50 | <50 | <50 | 3200 | >6400 |
| 106 | 16/89 | <50 | <50 | 400 | 1600 | 1600 |
| 107 | 16/89 | <50 | 50 | 50 | 50 | 400 |
| 108 | 16/89 | <50 | <50 | <50 | 800 | 3200 |
| 109 | 16/89 | <50 | <50 | 50 | 3200 | >6400 |
| 110 | 16/89 | <50 | <50 | 200 | 800 | 1600 |

Table 7.1 Anti-peptide antibody titres in hamster serum from animals inoculated with EHV-4 gp13 peptides, 14/89, 15/89, and 16/89. Values are given as the reciprocal ELISA titre. For < and > read $\leq$ and $\geq$. Data from Stokes et al., 1991.

Table 7.2

Anti-EHV-4 ELISA

| Hamster | Peptide | Time (days) 0 | 14 | 20 | 34 | 44 |
|---|---|---|---|---|---|---|
| 96 | 14/89 | 0.00 | 0.00 | 0.05 | 0.06 | 0.05 |
| 97 | 14/89 | 0.00 | 0.08 | 0.03 | 0.32 | 0.14 |
| 98 | 14/89 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 99 | 14/89 | 0.00 | 0.09 | 0.10 | 0.14 | 0.17 |
| 100 | 14/89 | 0.00 | 0.00 | 0.00 | 0.09 | 0.25 |
| 101 | 15/89 | 0.00 | 0.00 | 0.07 | 0.10 | 0.54 |
| 102 | 15/89 | 0.00 | 0.01 | 0.25 | 0.34 | 0.36 |
| 103 | 15/89 | 0.00 | 0.00 | 0.00 | 0.00 | 0.13 |
| 104 | 15/89 | 0.00 | 0.00 | 0.01 | 0.06 | 0.30 |
| 105 | 15/89 | 0.00 | 0.00 | 0.00 | 0.04 | 0.02 |
| 106 | 16/89 | 0.00 | 0.06 | 0.03 | 0.07 | 0.13 |
| 107 | 16/89 | 0.00 | 0.00 | 0.07 | 0.07 | 0.05 |
| 108 | 16/89 | 0.00 | 0.02 | 0.00 | 0.13 | 0.01 |
| 109 | 16/89 | 0.00 | 0.05 | 0.07 | 0.24 | 0.08 |
| 110 | 16/89 | 0.00 | 0.03 | 0.07 | 0.15 | 0.24 |

Table 7.2 Anti-EHV-4 activity of sera from hamsters inoculated with gp13 peptides. Values are given as optical density with day 0 value taken as background. Serum samples were diluted 1/100. Data from Stokes et al., 1991.

Table 7.3

Virus Titres in Tissues of Challenged Hamsters

| Peptide | Tissue | Amount of virus in tissue ($log_{10}TCID_{50}/g$) | | | | |
|---|---|---|---|---|---|---|
| 14/89 | Hamster | 96 | 97 | 98 | 99 | 100 |
| | Liver | 4.5 | 2.5 | 3.0 | 5.0 | 3.5 |
| | Lung | 2.0 | <1.5 | <1.5 | <1.5 | <1.5 |
| 15/89 | Hamster | 101 | 102 | 103 | 104 | 105 |
| | Liver | 2.0 | 3.5 | 3.0 | 2.0 | 3.0 |
| | Lung | 2.0 | <1.5 | <1.5 | <1.5 | <1.5 |
| 16/89 | Hamster | 106 | 107 | 108 | 109 | 110 |
| | Liver | 4.0 | 3.5 | 4.5 | <1.5 | 2.5 |
| | Lung | <1.5 | <1.5 | 3.0 | 2.0 | <1.5 |
| 25/89 | Hamster | 121 | 122 | 123 | 124 | 125 |
| | Liver | died | 3.5 | 2.5 | ND | 4.0 |
| | Lung | died | <1.5 | <1.5 | ND | <1.5 |
| 28/89 | Hamster | 151 | 152 | 153 | 154 | 155 |
| | Liver | 2.5 | 3.5 | 3.5 | <1.5 | 2.5 |
| | Lung | <1.5 | 2.0 | <1.5 | <1.5 | <1.5 |
| Control | Hamster | 1 | 2 | 3 | 4 | 5 |
| | Liver | 3.0 | 3.5 | <1.5 | 3.0 | 4.0 |
| | Lung | <1.5 | <1.5 | <1.5 | <1.5 | <1.5 |

Table 7.3 Post challenge (EHV-1 KyD) virus titres in the liver and lungs of hamsters immunised with EHV-4 gp13 and gH peptides. Control and gH peptide-immunised hamsters were challenged with $10^6$ $TCID_{50}$ units of virus and gp13 peptide-immunised hamsters with $2x10^{5.3}$ units of virus. For < and > read $\leqslant$ and $\geqslant$.

## Table 7.4

## Anti-gH Peptide and Anti-Virus ELISA Data

| | Hamster | Time (days) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 0 | 13 | 27 | 41 | 52 | 66 | 132 |
| Peptide 25/89 Anti-Peptide | 121 | <50 | 3200 | 3200 | >6400 | >6400 | 10240 | died |
| | 122 | <50 | <50 | 200 | 800 | 3200 | 200 | >6400 |
| | 123 | <50 | 200 | 1600 | 3200 | >6400 | 3200 | >6400 |
| | 124 | <50 | <50 | 200 | 1600 | >6400 | 3200 | |
| | 125 | <50 | 100 | 1600 | >6400 | >6400 | >6400 | >6400 |
| Peptide 25/89 Anti-EHV-1 | 121 | <50 | <50 | 800 | <50 | <50 | | |
| | 122 | <50 | <50 | <50 | <50 | <50 | | |
| | 123 | <50 | <50 | 800 | 200 | <50 | | |
| | 124 | <50 | <50 | <50 | <50 | <50 | | |
| | 125 | <50 | <50 | <50 | <50 | <50 | | |
| Peptide 25/89 Anti-EHV-4 | 121 | <50 | <50 | 400 | 400 | <50 | | |
| | 122 | <50 | <50 | <50 | <50 | <50 | | |
| | 123 | <50 | <50 | 800 | 200 | <50 | | |
| | 124 | <50 | <50 | <50 | <50 | <50 | | |
| | 125 | <50 | <50 | <50 | <50 | <50 | | |

| | Hamster | Time (days) | | | | | |
|---|---|---|---|---|---|---|---|
| | | 0 | 14 | 22 | 29 | 38 | 79 |
| Peptide 28/89 Anti-Peptide | 151 | <50 | <50 | 50 | 200 | 800 | 1600 |
| | 152 | <50 | <50 | 100 | 800 | 3200 | 3200 |
| | 153 | <50 | <50 | <50 | 800 | 3200 | 1600 |
| | 154 | <50 | <50 | <50 | <50 | 50 | >6400 |
| | 155 | <50 | <50 | <50 | 50 | 800 | >6400 |
| Peptide 28/89 Anti-EHV-1 | 151 | <50 | <50 | | <50 | | 50 |
| | 152 | <50 | <50 | | <50 | | 200 |
| | 153 | <50 | <50 | | <50 | | <50 |
| | 154 | <50 | <50 | | <50 | | 200 |
| | 155 | <50 | <50 | | <50 | | 400 |
| Peptide 28/89 Anti-EHV-4 | 151 | <50 | <50 | | <50 | | <50 |
| | 152 | <50 | <50 | | <50 | | 100 |
| | 153 | <50 | <50 | | <50 | | <50 |
| | 154 | <50 | <50 | | <50 | | 50 |
| | 155 | <50 | <50 | | <50 | | 50 |

Table 7.4 Reciprocal titres of anti-peptide, anti-EHV-1, and anti-EHV-4 antibody in sera of hamsters immunised with peptides 25/89 and 28/89 (gH) as determined by ELISA. For < and > read $\leqslant$ and $\geqslant$. Data from A. Stokes, personal communication.

## DISCUSSION

Immunisation of hamsters with peptides derived from regions of the predicted EHV-4 gp13 and gH polypeptides resulted in the production of anti-peptide antibody in all hamsters. This indicates that the peptides were of reasonable immunogenicity, in the presence of adjuvant, without attachment to a carrier protein. It is possible that coupling of the peptides to hamster proteins by virtue of the terminal cysteine residue might have occurred thus increasing their effective immunogenicity (T. Doel, personal communication). It is interesting that all peptides elicited the production of anti-virus antibody in at least some hamsters. Thus antibody raised against linear peptide epitopes apparently crossreacts with gp13 or gH as presented in the solid phase of the ELISA assay. With the caveat that the proteins might be partially denatured in that context, this suggests the methods used to locate parts of the polypeptide that might be surface oriented and therefore targets of the immune system are of reasonable predictive value. None of the peptides induced production of neutralising antibody or protection against challenge. Of course it is possible that the majority of neutralising epitopes of EHV-4 gp13 and gH are conformational in nature. Production of neutralising antibody by linear peptides would then be dependent on the existence of sufficient linear sequence within the conformational epitope such that anti-peptide antibody would crossreact with the native neutralising epitope at a detectable level (van Regenmortel, 1989) .

Further investigations are required before any linear epitopes possessed by the EHV-4 peptides are ruled out of any involvement in protection for a number of reasons -

i) Several reports have detailed elicitation of a protective immune response on administration of two or more monoclonal antibodies which

individually were nonprotective (Lussenhop et al., 1988; Roberts et al., 1985; van Drunen Littel-van den Hurk, 1985). Thus antibodies binding to different epitopes may act synergistically to induce protection. It is of interest that immunisation of guineapigs with gp13 and gp14 of EHV-1 elicited a greater immune response than would be predicted from the responses to gp13 and gp14 separately (Guo et al., 1990). It is therefore possible that combinations of peptides derived from EHV-4 glycoproteins might induce a protective immune response.

ii) Given that we are ultimately interested in epitopes which stimulate a protective immune response it is important not to focus too strongly on a neutralising antibody response since in vitro neutralising activity and in vivo protection are not necessarily correlated. For example, non-neutralising monoclonal antibodies have been shown to induce protection on passive transfer (Stokes et al., 1989). Furthermore, a hamster immunised with a portion of EHV-4 gp14 (gB) expressed as a beta-galactosidase fusion product was protected against EHV-1 challenge yet analysis of sera at intervals indicated no neutralising response to EHV-1 (A. Stokes, M.P. Riggio, personal communication). In HSV-1 the protective immune response is comprised of contributions from not only antibody neutralisation but also from cellular mechanisms which recognise specific glycoproteins (Hall and Katrak, 1986). It is probable that the same applies to the EHV situation with cellular immune mechanisms perhaps playing a dominant role (Bryans, 1969).

iii) A final caveat with regard to the peptide studies is that the immune response of a small animal model such as the hamster does not necessarily correlate with that of the natural equine host : the entire surface of a non-self protein or peptide is likely to be antigenic in any animal yet the immunogenicity of a specific site may vary among

different animals (Geysen et al., 1987).

Studies on the immunogenicity of peptides derived from the gC and gH-type glycoproteins of other herpesviruses have identified several epitopes. Five non-neutralising epitopes have been mapped to BHV-1 gIII (gC) (Fitzpatrick et al., 1990). Three peptides derived from the predicted VZV gpIII sequence have also been tested - amino acids 152-173, 793-803, and 828-841 - all of which elicited antibody reactive with native gpIII (Keller et al., 1987). Analysis of the immunogenicity of an EBV peptide derived from amino acids 518-533 of the gH-type glycoprotein mapped a neutralising epitope to this part of the polypeptide although the affinity of polyclonal anti-peptide antibody for the epitope was low (Oba and Hutt-Fletcher, 1988).

The fact that a single immunodominant linear epitope domain has been mapped to the EHV-1 gp13 glycoprotein is most encouraging with respect to the potential development of a peptide vaccine against EHV-1 or EHV-4. A more comprehensive analysis of linear B-cell epitopes, and ideally also of T-cell epitopes, would seem a worthwhile exercise. Such studies coupled to monoclonal antibody studies and immunogenicity studies of intact gp13 should shed light on the apparent type specificity of EHV-1 gp13 (Allen et al., 1988) despite the 79% conservation of amino acid sequence between EHV-1 and EHV-4 gp13. gC-type glycoproteins have been implicated in immune evasion through three mechanisms : binding of complement component C3b (Smiley and Friedman, 1985); binding of antibody by free extracellular glycoprotein (Isfort et al., 1987); molecular mimicry of self epitopes (through homology to immunoglobulin superfamily members) (Fitzpatrick et al., 1989). The ideal gC subunit or peptide vaccine would possess only B- and T-cell epitopes important in stimulating, as opposed to suppressing, anti-viral immune responses and would render immune

evasion mechanisms elicited by gC in the context of lytic infection inactive. Of course the ultimate test of such a vaccine would be its potential to elicit a protective immune response as compared to that of other vaccine designs such as live attenuated virus.

Whether or not EHV-4 gp13 or gH peptides are eventually identified as eliciting a neutralising and/or protective immune response, analysis of the immunogenicity of peptides derived from these glycoproteins should provide some interesting data on their antigenic structure.

# CHAPTER 8

# GENERAL DISCUSSION

# GENERAL DISCUSSION

## Presented Data

The primary genes of interest in this study were the EHV-4 gene homologues of HSV-1 genes encoding glycoproteins gC and gH and the TK gene. Sequence data is presented for these three genes and for a UL24 gene homologue and a gene apparently unique to EHV-4. Partial sequence data is presented for UL21 and UL43 gene homologues. Analyses of these data provide information which can be used to deduce the amino acid sequences of the polypeptide products of these genes and which can be related, in a limited sense, to possible functional roles of the proteins. Additionally these data provide more general information of use in the evaluation of i) the relatedness of EHV-4 genes to their counterparts in other herpesviruses in terms of genome position and composition, and ii) the %GC content of the EHV-4 genome.

A summary of the alignment and relatedness of EHV-4 genes mapped in this project to genes in HSV-1 and VZV is summarised in Figure 1. One EHV-4 gene was identified which has no apparent homologue in HSV-1 or VZV, a gene , dsgC, mapping to the left of the gp13 gene. Sequence data for the corresponding region of EHV-1 genome has yet to be published but one might expect , given the results of this thesis, that EHV-1 would possess a counterpart of EHV-4 dsgC. The only other major difference between the sequenced parts of the EHV-4 genome and equivalent regions in the genomes of HSV-1 and VZV is the existence of a region of direct repeat elements and a putative origin of DNA replication in the EHV-4 genome between the gH gene and the UL21 gene homologue. Functional confirmation of the putative origin of replication of EHV-4 could be achieved using the plasmid amplification assay developed by Stow (1982).

Figure 8.1

Comparison of EHV-4, HSV-1, and VZV genome organisation across two parts of $U_L$

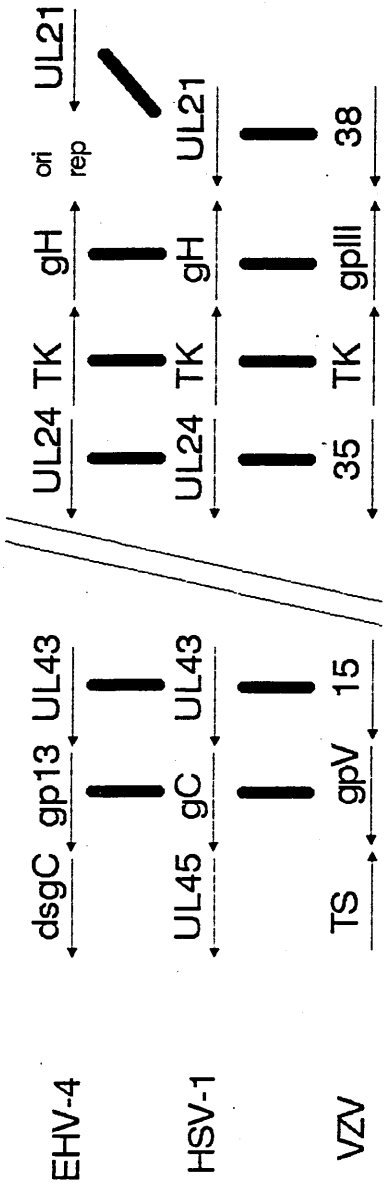



Figure 8.1 Comparison of the genome organisation of the EHV-4, HSV-1, and VZV genomes around the gC and gH gene loci. Transcriptional direction of genes is indicated by arrows and homologies ( at the amino acid level) by vertical bars.

The %GC content of the coding region of the gC, UL24, TK, gH, and dsgC genes of EHV-4 are compared to those of their gene counterparts in EHV-1, HSV-1, and VZV in Table 8.1. The overall %GC for the complete sequence data presented in this thesis is 48%. Comparison of the EHV-1 and EHV-4 data suggests that the %GC of the complete EHV-4 genome might be slightly lower than the experimentally determined %GC of EHV-1, perhaps in the order of 50%.

Comparison of the DNA identity between EHV-1 and EHV-4 DNA indicated identities greater than the 17% homology derived by liquid hybridisation data (Allen and Turtinen, 1982). For example, comparison of 2070bp spanning the EHV-4 gp13 gene with the corresponding region of EHV-1 DNA indicated a nucleotide sequence identity of 66%. It is most likely that similar comparisons within regions of the $U_S$ component would indicate a lower degree of nucleotide identity since this part of the genome is more divergent than $U_L$. Nevertheless, one would predict an overall sequence homology in the order of that between HSV-1 and HSV-2 (50%) (Kieff et al., 1971).

The amino acid identities of the EHV-4 gp13, UL24, TK, and gH gene products with counterparts in EHV-1, HSV-1, VZV, HCMV, and EBV are presented in Table 2. The gC-type glycoprotein is the least well conserved of the 5 gene products in terms of representation within herpesvirus genomes and in amino acid sequence. The TK and UL24 genes are the most conserved using the criteria employed.

Given the relatively high levels of identity, at the amino acid level, of the genes studied in this thesis to their EHV-1 counterparts it could be argued that the redesignation of these two viruses as separate types is unjustified. Certainly the sequence data currently available has suggested a lesser degree of divergence than might have been predicted from the completely distinct restriction endonuclease

Table 8.1

Percentage GC content of the coding region of selected EHV-4, EHV-1, HSV-1, and VZV genes

| | dsgC | gC | UL24 | TK | gH | Genome |
|---|---|---|---|---|---|---|
| EHV-4 | 50.4 | 48.4 | 49.7 | 51.8 | 44.5 | ? |
| EHV-1 | | 53.6 | 55.3 | 58.1 | 51.1 | 57 |
| HSV-1 | — | 68.3 | 64.6 | 65.3 | 66.7 | 67 |
| VZV | — | 47.2 | 46.0 | 45.3 | 44.4 | 46 |

Table 8.1 Comparison of the %GC composition of the coding portion of selected herpesvirus genes. Experimentally-derived values of overall %GC content are detailed for EHV-1, HSV-1, and VZV in the final column (Kieff et al., 1971; Ludwig et al., 1972; Soehner et al., 1965)

Table 8.2

Comparative identity of 4 EHV-4 genes with counterparts in five herpesviruses

| VIRUS | Type | % Identity (at amino acid level) to EHV-4 gene | | | |
|---|---|---|---|---|---|
| | | gC | UL24 | TK | gH |
| EHV-1 | alpha | 79 | 83 | 89 | 85 |
| HSV-1 | alpha | 25 | 36 | 36 | 26 |
| VZV | alpha | 26 | 37 | 35 | 32 |
| HCMV | beta | - | 23 | - | 18 |
| EBV | gamma | - | 23 | 25 | 17 |

Table 8.2 Percentage identity , at the amino acid level, of the gC, UL24, TK, and gH genes of EHV-4, EHV-1, HSV-1, VZV, HCMV, and EBV. Entire polypeptides were analysed with the exception of the EHV-4/EBV TK comparison in which amino acids 27-326 of EHV-4 TK and 286-586 of EBV TK were analysed.

patterns manifested by EHV-1 and EHV-4 genomes. However, EHV-1 and EHV-4 seem to be related to a similar degree as that of HSV-1 and HSV-2 which are regarded as separate types. Furthermore, EHV-1 and EHV-4 have distinct though overlapping disease sequelae and may differ in their latent characteristics. Thus, redesignation seems warranted.

**Applications of the Data**

With the long term aims of the group being the development of novel recombinant EHV vaccines and improved diagnostic techniques, the nucleotide sequence data presented in this thesis was obtained with the intention of utilising the data in three ways - i) to predict the amino acid sequences of the EHV-4 gp13 and gH-type glycoproteins and to investigate the immunogenicity of peptides derived from these sequences, ii) to enable derivation of suitable primer sequences for use in polymerase chain reaction (PCR) amplification of EHV DNA, and iii) to precisely localise the boundaries of the gp13, gH, and TK genes to aid in the production of TK-deleted EHV-4 and in the development of EHV-4 glycoprotein expression systems or cloning cassettes.

The immunogenicity of three gp13 and two gH peptides in hamsters was investigated. Anti-viral antibodies were raised against all 5 peptides. However, no virus neutralising or protective immune response was elicited. This work leads on to a more comprehensive analysis of peptide immunogenicity. Of particular interest would be the analysis of the ability of antibodies in convalescent equine sera to bind epitopes within a range of gp13 or gH peptides.

PCR-based diagnostic systems offer an alternative to current methods of identification of EHV-1 and EHV-4 which include virus isolation, serological analyses, DNA fingerprinting and/or Southern hybridisation, and monoclonal antibody typing (Allen et al. 1983b; Chowdhury et al.,

1986; Morris and Field, 1988; Yeargan et al., 1985). PCR is an extremely sensitive technique whereby specific DNA sequences are amplified using minute amounts of target DNA as template. It has the potential to offer increased sensitivity and decreased sample analysis time over existing diagnostic techniques. Comparison of EHV-4 gene sequences presented in Chapters 4 and 6 to their EHV-1 counterparts permits the localisation of conserved sequence regions from which primers can be derived. These primers could be used to amplify both EHV-1 and EHV-4 in clinical samples such as nasal secretions. Typing of the viral DNA in the sample could be achieved by hybridisation of EHV-1- or EHV-4-specific probes to amplified DNA. The first report of PCR applied to EHV DNA amplification has recently been published (Ballagi-Pordany et al., 1990).

The current impetus for research into various aspects of EHV molecular biology and pathogenesis is a result of increased funding by the equine industry which is ultimately directed towards development of novel vaccines against EHV-1 and EHV-4. As detailed in Chapter 1, an approach to the development of a novel EHV vaccine is the production of attenuated recombinant EHV-1 or EHV-4 viruses. The determination of the precise boundaries of the EHV-4 TK gene permits the construction of TK-deleted plasmid constructs, as presented in Chapter 3. These could be used in the production of TK- EHV-4 which could then be tested for its virulence properties relative to that of TK+ virus. Recombinant virus could be produced by cotransfection of TK-deleted plasmid DNA with intact TK+ EHV-4 or by cotransfection of a cosmid library of EHV-4 containing a TK-deleted cosmid (Chapter 1, Fig. 7).

Information on the position and coding capacity of the gp13 and gH genes could be applied to the development of subunit vaccines. Peptide sequences or polypeptides expressed in a suitable vector could be

tested for their immunogenicity in small animals and ultimately in horses. Alternatively, these genes could be inserted into an attenuated EHV-1 genome in order to produce a bivalent vaccine strain which expressed immunodominant glycoproteins of both EHV-1 and EHV-4.

Whatever the outcome of EHV research in terms of novel vaccine development the studies underway should provide further insight into the pathology and molecular biology of these viruses.

# REFERENCES

Allen, G.P. and Bryans, J.T. (1986). Molecular epizootiology, pathogenesis, and prophylaxis of equine herpesvirus 1 infections. In 'Progress in Veterinary Microbiology and Immunology', vol. 2, pp.78-144. Edited by R. Pandey. S. Karger, Basel.

Allen, G.P. and Coogle, L.D. (1988). Characterization of an equine herpesvirus type 1 gene encoding a glycoprotein (gp13) with homology to herpes simplex virus glycoprotein C. J. Virol. **62**, 2850-2858.

Allen, G.P. and Turtinen, L.W. (1982). Assessment of the base sequence homology between the two subtypes of equine herpesvirus 1. J. Virol. **44**, 249-255.

Allen, G.P. and Yeargan, M.R. (1987). Use of $\lambda$ gt11 and monoclonal antibodies to map the genes for the six major glycoproteins of equine herpesvirus 1. J. Virol. **61**, 2454-2461.

Allen, G.P., McGowan, J.J., Randall, C.C., Mancini, W., Cheng, Y.-C. and Gentry, G.A. (1979). Purification and characterization of deoxythymidine kinase (dTK) induced in $dTK^-$ 3T3 mouse cells by equine herpesvirus type 1 ( EHV-1 ). Virology **92**, 367-374.

Allen, G., Yeargan, M.R. and Bryans, J.T. (1983a). Alterations in the genome of equine herpesvirus-1 after in vitro and in vivo virus passage. Infect. Immun. **40**, 436-439.

Allen, G.P., Yeargan, M.R., and Coogle, L.D. (1988). Equid herpesvirus-1 glycoprotein 13 (gp13) : epitope analysis, gene structure, and expression in E. coli. In 'Equine Infectious Diseases V - Proceedings of the Fifth International Conference' , pp. 103-110. Edited by D.G. Powell. University Press of Kentucky, Kentucky.

Allen, G.P., Yeargan, M.R., Turtinen, L.W., Bryans, J.T. and McCollum, W.H. (1983b). Molecular epizootiologic studies of equine herpesvirus-1 infections by restriction endonuclease fingerprinting of viral DNA. Am. J. Vet. Res. **44**, 263-271.

Anon. (1987). Compendium of data sheets for veterinary products 1987-1988. Compiled by G. Walker. Datapharm Publications Ltd., UK.

Anon. (1989). EHV-1 : a recurrent problem. Vet. Rec. **124**, 443-444.

Audonnet, J.-C., Winslow, J., Allen, G. and Paoletti, E. (1990). Equine herpesvirus type 1 unique short fragment encodes glycoproteins with homology to herpes simplex virus type 1 gD, gI and gE. J. Gen. Virol. **71**, 2969-2978.

Awan, A.R., Chong, Y.-C. and Field, H.J. (1990). The pathogenesis of equine herpesvirus type 1 in the mouse : a new model for studying host responses to the infection. J. Gen. Virol. **71**, 1131-1140.

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, London **310**, 207-211.

Baker, T.S., Newcomb, W.W., Booy, F.P., Brown, J.C. and Steven, A.C. (1990). Three-dimensional structures of maturable and abortive capsids of equine herpesvirus 1 from cryoelectron microscopy. J. Virol. **64**, 563-573.

Ballagi-Pordany, A., Klingeborn, B., Flensburg, J. and Belak, S. (1990). Equine herpesvirus type 1 : detection of viral DNA sequences in aborted fetuses with the polymerase chain reaction. Vet. Microbiol. **22**, 373-381.

Baumann, R.P., Staczek, J. and O'Callaghan, D.J. (1987). Equine herpesvirus type 1 defective-interfering (DI) particle DNA structure : the central region of the inverted repeat is deleted from DI DNA. Virology **159**, 137-146.

Baumann, R.P., Sullivan, D.C, Staczek, J. and O'Callaghan, D.J. (1986). Genetic relatedness and colinearity of genomes of equine herpesvirus types 1 and 3. J. Virol. **57**, 816-825.

Baumann, R.P., Yalamanchili, V.R.R. and O'Callaghan, D.J. (1989). Functional mapping and DNA sequence of an equine herpesvirus 1 origin of replication. J. Virol. **63**, 1275-1283.

Bell, C.W., Boyle, D.B. and Whalley, J.M. (1990). Transcript analysis of the equine herpesvirus 1 glycoprotein B gene homologue and its expression by a recombinant vaccinia virus. J. Gen. Virol. **71**, 1119-1129.

Bello, L.J., Whitbeck, J.C. and Lawrence, W.C. (1987). Map location of the thymidine kinase gene of bovine herpesvirus 1. J. Virol. **61**, 4023-4025.

Ben-Porat, T., De Marchi, J.M., Lomniczi, B. and Kaplan, A.S. (1986). Role of glycoproteins of pseudorabies virus in eliciting neutralizing antibodies. Virology **154**, 325-334.

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-Porat, T., Veach, R.A. and Ihara, S. (1983). Localization of the regions of homology between the genomes of herpes simplex virus, type 1, and pseudorabies virus. Virology **127**, 194-204.

Bielefeldt Ohmann, H. and Babiuk, L.A. (1988). Induction of receptors for complement and immunoglobulins by herpesviruses of various species. Virus Res. **9**, 335-342.

Binns, M.M. and Ross, N.L.J. (1989). Nucleotide sequence of the Marek's disease virus (MDV) RB-1B A antigen gene and the identification of the MDV A antigen as the herpes simplex virus-1 glycoprotein C homologue. Virus Res. **12**, 371-382.

Birnboim, H.C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**, 1513-1523.

Birnstiel, M.L., Busslinger, M. and Strub, K. (1985). Transcription termination and 3′ processing : the end is in site! Cell **41**, 349-359.

Bishop, G.A., Kumel, G., Schwartz, S.A. and Glorioso, J.C. (1986). Specificity of human natural killer cells in limiting dilution culture for determinants of herpes simplex virus type 1 glycoproteins. J. Virol. **57**, 294-300.

Blakeslee Jr., J.R. , Olsen, R.G., McAllister, E.S., Fassbender, J. and Dennis, R. (1975). Evidence of respiratory tract infection induced by equine herpesvirus, type 2, in the horse. Can. J. Microbiol. **21**, 1940-1946.

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.

Bradshaw Jr., H.D. and Deininger, P.L. (1984). Human thymidine kinase gene : molecular cloning and nucleotide seqeunce of a cDNA expressible in mammalian cells. Mol. Cell. Biol. **4**, 2316-2320.

Bridges , C.G. and Edington, N. (1987a). The proteins of equid herpesvirus 1 (EHV 1) recognised by equine antisera and their ability to promote antibody-dependent cell-mediated cytotoxicity. Tierarztl. Prax. Suppl. **2**, 47-49.

Bridges, C.G. and Edington, N. (1987b). Genetic restriction of cytolysis during equid herpesvirus 1 subtype 2 infection. Clin. Exp. Immunol. **70**, 276-282.

Bridges, C.G., Ledger, N. and Edington, N. (1988). The characterization of equine herpes virus-1-infected cell polypeptides recognized by equine lymphocytes. Immunol. **63**, 193-198.

Brocade Wu, C.-T., Levine, M., Homa, F., Highlander, S.L. and Glorioso, J.C. (1990). Characterization of the antigenic structure of herpes simplex virus type 1 glycoprotein C through DNA sequence analysis of monoclonal antibody-resistant mutants. J. Virol. **64**, 856-863.

Browning, G.F. and Studdert, M.J. (1987a). Genomic heterogeneity of equine betaherpesviruses. J. Gen. Virol. **68**, 1441-1447.

Browning, G.F. and Studdert, M.J. (1987b). Epidemiology of equine herpesvirus 2 ( equine cytomegalovirus ). J. Clin. Microbiol. **25**, 13-16.

Browning, G.F. and Studdert, M.J. (1989). Physical mapping of a genome of equine herpesvirus 2 ( equine cytomegalovirus ). Arch. Virol. **104**, 77-86.

Browning, G.F., Bulach, D.M., Ficorilli, N., Roy, E.A., Thorp, B.H. and Studdert, M.J. (1988). Latency of equine herpesvirus 4 ( equine rhinopneumonitis virus). Vet. Rec. **123**, 518-519.

Bryans, J.T. (1969). On immunity to disease caused by equine herpesvirus 1. J. Am. Vet. Med. Assoc. **155**, 294-300.

Bryans, J.T. and Allen, G.P. (1973). In vitro and in vivo studies of equine 'coital' exanthema . In 'Proceedings of the 3rd International Conference on Equine Infectious Diseases', pp. 322-336. Edited by J.T. Bryans and H. Gerber. S. Karger, Basel.

Bryans, J.T. and Prickett, M.E. (1972). A consideration of the pathogenesis of abortigenic disease caused by equine herpesvirus-1. In 'Proceedings of the 2nd International Conference on Equine Infectious Diseases', pp. 34-40. Edited by J.T. Bryans and H. Gerber. S. Karger, Basel.

Bublot, M., Van Bressem, M.-F., Thiry, E., Dubuisson, J. and Pastoret, P.-P. (1990). Bovine herpesvirus 4 genome : cloning, mapping and strain variation analysis. J. Gen. Virol. **71**, 133-142.

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.

Buckmaster, A.E., Scott, S.D., Sanderson, M.J., Boursnell, M.E.G., Ross, N.L.J. and Binns, M.M. (1988). Gene sequence and mapping data from Marek's disease virus and herpesvirus of turkeys : implications for herpesvirus classification. J. Gen. Virol. **69**, 2033-2042.

Burrows, R. and Goodridge, D. (1973). In vivo and in vitro studies of equine rhinopneumonitis virus strains. In' Proceedings of the Third International Conference on Equine Infectious Diseases', pp. 306-321. Edited by J.T. Bryans and H. Gerber. Karger, Basel.

Burrows, R. and Goodridge, D. (1984). Studies of persistent and latent equid herpesvirus-1 and herpesvirus-3 infections in the Pirbright pony herd. In 'Wittman Latent Herpesvirus Infections in Veterinary Medicine', pp. 307-319. Martinus Nijhoff, Boston.

Burrows, R., Goodridge, D. and Denyer, M.S. (1984). Trials of an inactivated equid herpesvirus 1 vaccine : challenge with a subtype 1 virus. Vet. Rec. **114**, 369-374.

Buscaglia, C. and Calneck, B.W. (1988). Maintenance of Marek's disease herpesvirus latency in vitro by a factor found in conditioned medium. J. Gen. Virol. **69**, 2809-2818.

Cai, W., Gu, B. and Person, S. (1988). Role of glycoprotein B of herpes simplex virus type 1 in viral entry and cell fusion. J. Virol. **62**, 2596-2604.

Cameron, J.M., McDougall, I., Marsden, H.S., Preston, V.G., Ryan, D.M. and Subak-Sharpe, J.H. (1988). Ribonucleotide reductase encoded by herpes simplex virus is a determinant of the pathogenicity of the virus in mice and a valid antiviral target. J. Gen. Virol. **69**, 2607-2612.

Caughman, G.B., Robertson, A.T., Gray, W.L., Sullivan, D.C. and O'Callaghan, D.J. (1988). Characterization of equine herpesvirus type 1 immediate early proteins. Virology **163**, 563-571.

Caughman, G.B., Staczek, J. and O'Callaghan, D.J. (1985). Equine herpesvirus type 1 infected cell polypeptides : evidence for immediate early/early/late regulation of viral gene expression. Virology **145**, 49-61.

Cebrian, J., Kaschka-Dierich, C., Berthelot, N. and Sheldrick, P. (1982). Inverted repeat nucleotide sequences in the genomes of Marek's disease virus and the herpesvirus of the turkey. Proc. Natl. Acad. Sci. USA **79**, 555-558.

Challberg, M.D. and Kelly, T.J. (1989). Animal virus DNA replication. Ann. Rev. Biochem. **58**, 671-717.

Chee, M.S., Bankier, A.T., Beck, S., Bohni, R., Brown, C.M., Cerny, R., Horsnell, T., Hutchison III, C.A., Kouzarides, T., Martignetti, J.A., Preddie, E., Satchwell, S.C., Tomlinson, P., Weston, K.M. and Barrell, B.G. (1990). Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr. Top. Microbiol. Immunol. **154**, 125-169.

Chen, M.S. and Prusoff, W.H. (1978). Association of thymidylate kinase activity with pyrimidine deoxyribonucleoside kinase induced by herpes simplex virus. J. Biol. Chem. **253**, 1325-1327.

Chen, M.S., Summers, W.P., Walker, J., Summers, W.C. and Prusoff, W.H. (1979). Characterization of pyrimidine deoxyribonucleoside kinase (thymidine kinase) and thymidylate kinase as a multifunctional enzyme in cells transformed by herpes simplex virus type 1 and in cells infected with mutant strains of herpes simplex virus. J. Virol. **30**, 942-945.

Cheung, A.K. (1989). Detection of pseudorabies virus transcripts in trigeminal ganglia of latently infected swine. J. Virol. **63**, 2908-2913.

Chou, P.Y. and Fasman, G.D. (1978). Prediction of the secondary structure of proteins from their amino acid sequence. In 'Advances in Enzymology', vol. 47, pp.45-148. Edited by A. Meister. Wiley and Sons, New York.

Chou, J. and Roizman, B. (1985). Isomerization of 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.

Chowdhury, S.I., Buhk, H.-J., Ludwig, H. and Hammerschmidt, W. (1990). Genomic termini of equine herpesvirus 1. J. Virol. **64**, 873-880.

Chowdhury, S.I., Hammerschmidt, W., Ludwig, H., Thein, P. and Buhk, H.-J. (1986). Rapid method for the identification and screening of herpesviruses by DNA fingerprinting combined with blot hybridization. J. Vir. Meth. **14**, 285-291.

Chowdhury, S.I., Ludwig, H. and Buhk, H.J. (1988). Molecular biological characterization of equine herpesvirus type 1 (EHV-1) isolates from ruminant hosts. Virus Res. **11**, 127-139.

Churchill, A.E., Chubb, R.C. and Baxendale, W. (1969). The attenuation, with loss of oncogenicity, of the herpesvirus-type virus of Marek's disease (strain HPRS-16) on passage in cell culture. J. Gen. Virol. **4**, 557-563.

Coen, D.M. and Schaffer, P.A. (1980). Two distinct loci confer resistance to acycloguanosine in herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **77**, 2265-2269.

Coen, D.M., Irmiere, A.F. Jacobson, J.G. and Kerns, K.M. (1989). Low levels of herpes simplex virus thymidine-thymidylate kinase are not limiting for sensitivity to certain antiviral drugs or for latency in a mouse model. Virology **168**, 221-231.

Colacino, J.M., Flowers, C.C., Menna, J., O'Callaghan, D.J. and Staczek, J. (1989). Physical structure and molecular cloning of equine cytomegalovirus DNA. Virology **173**, 566-580.

Colacino, J., Wolcott, M., Robertson, A., O'Callaghan, D. and Staczek, J. (1988). Induction of deoxypyrimidine nucleoside kinase in cells permissively infected by equine cytomegalovirus. Abstract 132, 13th International Herpesvirus Workshop. Irvine, Ca, USA.

Collins, P. (1983). The spectrum of antiviral activities of acyclovir *in vitro* and *in vivo*. J. Antimicrob. Chemother. **12**, Suppl. B, 19-27.

Cook, R.F., O'Neill, T., Strachan, E., Sundquist, B. and Mumford, J.A. (1990). Protection against lethal equine herpes virus type 1 (subtype 1) infection in hamsters by immune stimulating complexes (ISCOMs) containing the major viral glycoproteins. Vaccine **8**, 491-496.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. and Chambon, P. (1980). Promoter sequences of eukaryotic protein-coding genes. Science **209**, 1406-1414.

Cornick, J., Martens, J., Martens, R., Crandell, R., McConnell, S. and Kit, S. (1990). Safety and efficacy of a thymidine kinase negative equine herpesvirus-1 vaccine in young horses. Can. J. Vet. Res. **54**, 260-266.

Coussens, P.M. and Velicer, L.F. (1988). Structure and complete nucleotide sequence of the Marek's disease herpesvirus gp57-65 gene. J. Virol. **62**, 2373-2379.

Cranage, M.P., Kouzarides, T., Bankier, A.T., Satchwell, S., Weston, K., Tomlinson, P., Barrell, B., Hart, H., Bell, S.E., Minson, A.C. and Smith, G.L. (1986). Identification of the human cytomegalovirus glycoprotein B gene and induction of neutralizing antibodies via its expression in recombinant vaccinia virus. EMBO J. **5**, 3057-3063.

Cranage, M.P., Smith, G.L., Bell, S.E., Hart, H., Brown, C., Bankier, A.T., Tomlinson, P., Barrell, B.G. and Minson, A.C. (1988). Identification and expression of a human cytomegalovirus glycoprotein with homology to the Epstein-Barr virus BXLF2 product, varicella-zoster virus gpIII, and herpes simplex virus type 1 glycoprotein H. J. Virol. **62**, 1416-1422.

Crandell, R.A., Ichimura, H. and Kit, S. (1988). Isolation and comparative restriction endonuclease DNA fingerprinting of equine herpesvirus-1 from cattle. Am. J. Vet. Res. **49**, 1807-1813.

Cullinane, A.A., Rixon, F.J. and Davison, A.J. (1988). Characterization of the genome of equine herpesvirus 1 subtype 2. J. Gen. Virol. **69**, 1575-1590.

Dales, S. and Silverberg, H. (1969). Viropexis of herpes simplex virus by HeLa cells. Virology **37**, 475-480.

Darby, G., Larder, B.A. and Inglis, M.M. (1986). Evidence that the 'active centre' of the herpes simplex virus thymidine kinase involves an interaction between three distinct regions of the polypeptide. J. Gen. Virol. **67**, 753-758.

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. and McGeoch, D.J. (1986). Evolutionary comparisons of the S segments in the genomes of herpes simplex virus type 1 and varicella-zoster virus. J. Gen. Virol. **67**, 597-611.

Davison, A.J., and Scott, J.E. (1986). The complete DNA sequence of varicella-zoster virus. J. Gen. Virol. **67**, 1759-1816.

Davison, A.J. and Taylor, P. (1987). Genetic relations between varicella-zoster virus and Epstein-Barr virus. J. Gen. Virol. **68**, 1067-1079.

Davison, A.J. and Wilkie, N.M. (1983). Location and orientation of homologous sequences in the genomes of five herpesviruses. J. Gen. Virol. **64**, 1927-1942.

Davison, A.J., Edson, C.M., Ellis, R.W., Forghani, B., Gilden, D., Grose, C., Keller, P.M., Vafai, A., Wroblewska, Z. and Yamanishi, K. (1986). New common nomenclature for glycoprotein genes of varicella-zoster virus and their glycosylated products. J. Virol. **57**, 1195-1197.

Davison, A.J., Waters, D.J. and Edson, C.M. (1985). Identification of the products of a varicella-zoster virus glycoprotein gene. J. Gen. Virol. **66**, 2237-2242.

Deb, S. and Deb, S.P. (1989). Analysis of Ori-S sequence of HSV-1 : identification of one functional DNA binding domain. Nucleic Acids Res. **17**, 2733-2752.

Deb, S. and Doelberg, M. (1988). A 67-base-pair segment from the ori-S region of herpes simplex virus type 1 encodes origin function. J. Virol. **62**, 2516-2519.

De Clercq, E. (1982). Selective antiherpes agents. Trends Pharmacol. Sci. **3**, 492-495.

Desai, P.J., Schaffer, P.A. and Minson, A.C. (1988). Excretion of non-infectious virus particles lacking glycoprotein H by a temperature-sensitive mutant of herpes simplex virus type 1 : evidence that gH is essential for virion infectivity. J. Gen. Virol. **69**, 1147-1156.

Dever, T.E., Glynias, M.J. and Merrick, W.C. (1987). GTP-binding domain : three consensus sequence elements with distinct spacing. Proc. Natl. Acad. Sci. USA **84**, 1814-1818.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**, 387-395.

Dimock, W.W. and Edwards, P.R. (1933). Is there a filterable virus of abortion in mares? Suppl. Ky. Agr. Exp. Sta. Bull. **333**, 297-301.

Doll, E.R. (1952). Seasonal incidence and fetal age in equine virus abortion. Cornell Vet. **42**, 505-509.

Doll, E.R. and Bryans, J.T. (1963). Epizootiology of equine viral rhinopneumonitis. J. Am. Vet. Med. Assoc. **142**, 31-37.

Doll, E.R. and Wallace, M.E. (1954). Cultivation of the equine influenza virus in suckling Syrian hamsters. Its similarity to the equine abortion virus. Cornell Vet. **44**, 133-138.

Dowbenko, D.J. and Lasky, L.A. (1984). Extensive homology between the herpes simplex virus type 2 glycoprotein F gene and the herpes simplex virus type 1 glycoprotein C gene. J. Virol. **52**, 154-163.

Draper, K.G., Costa, R.H., Lee, G.T.-Y., Spear, P.G. and Wagner, E.K. (1984). Molecular basis of the glycoprotein-C-negative phenotype of herpes simplex virus type 1 macroplaque strain. J. Virol. **51**, 578-585.

Ecker, J.R. and Hyman, R.W. (1982). Varicella zoster virus DNA exists as two isomers. Proc. Natl. Acad. Sci. USA **79**, 156-160.

Edington, N. and Bridges, C.G. (1990). One way protection between equid herpesvirus 1 and 4 in vivo. Res. Vet. Sci. **48**, 235-239.

Edington, N., Bridges, C.G. and Broad, S.C. (1987). Rapid diagnosis and characterization of equid herpesvirus 1 using monoclonal antibodies. Tierarztl. Prax. Suppl. **2**, 37-40.

Edington, N. , Bridges, C.G. and Huckle, A. (1985). Experimental reactivation of equid herpesvirus 1 ( EHV-1 ) following the administration of corticosteroids. Equine Vet. J. **17**, 369-372.

Edson, C.M. and Thorley-Lawson, D.A. (1983). Synthesis and processing of the three major envelope glycoproteins of Epstein-Barr virus. J. Virol. **46**, 547-556.

Efstathiou, S., Kemp, S., Darby, G. and Minson, A.C. (1989). The role of herpes simplex type 1 thymidine kinase in pathogenesis. J. Gen. Virol. **70**, 869-879.

Elias, P. and Lehman, I.R. (1988). Interaction of origin binding protein with an origin of replication of herpes simplex virus 1. Proc. Natl. Acad. Sci. USA **85**, 2959-2963.

Elias, P., O'Donnell, M.E., Mocarski, E.S. and Lehman, I.R. (1986). A DNA binding protein specific for an origin of replication of herpes simplex virus type 1. Proc. Natl. Acad. Sci. USA **83**, 6322-6326.

Elion, G.B. (1983). The biochemistry and mechanism of action of acyclovir. J. Antimicrobiol Chemother. **12**, Suppl. B, 9-17.

Enquist, L.W., Keeler Jr., C.L., Robbins, A.K., Ryan, J.P. and Whealy, M.E. (1988). An amino-terminal deletion mutation of pseudorabies virus glycoprotein gIII affects protein localization and RNA accumulation. J. Virol. **62**, 3565-3573.

Estes, J.E. and Huang, E.-S. (1977). Stimulation of cellular thymidine kinases by human cytomegalovirus. J. Virol. **24**, 13-21.

Everett, R.D. (1987). The regulation of transcription of viral and cellular genes by herpesvirus immediate-early gene products. Anticancer Res. **7**, 589-604.

Fenwick, M.L. and McMenamin, M. (1984). Synthesis of $\alpha$ (immediate-early) proteins in vero cells infected with pseudorabies virus. J. Gen. Virol. **65**, 1449-1456.

Field, H.J. and Awan, A.R. (1990). Effective chemotherapy of equine herpesvirus 1 by phosphonylmethoxyalkyl derivatives of adenine demonstrated in a novel murine model for the disease. Antimicrob. Ag. Chemother. **34**, 709-717.

Field, H.J. and Wildy, P. (1978). The pathogenicity of thymidine kinase-deficient mutants of herpes simplex virus in mice. J. Hyg. **81**, 267-277.

Figueroa, F. and Klein, J. (1986). The evolution of MHC class II genes. Immunol. Today **7**, 78-81.

Fitzpatrick, D.R. and Studdert, M.J. (1984). Immunologic relationships between equine herpesvirus type 1 (equine abortion virus) and type 4 (equine rhinopneumonitis virus). Am. J. Vet. Res. **45**, 1947-1952.

Fitzpatrick, D.R., Babiuk, L.A., and Zamb, T.J. (1989). Nucleotide sequence of bovine herpesvirus type I glycoprotein gIII, a structural model for gIII as a new member of the immunoglobulin superfamily, and implications for the homologous glycoproteins of other herpesviruses. Virology **173**, 46-57.

Fitzpatrick, D.R., Redmond, M.J., Attah-Poku, S.K., van Drunen Littel-van den Hurk, S., Babiuk, L.A. and Zamb, T. J. (1990). Mapping of 10 epitopes on bovine herpesvirus type 1 glycoproteins gI and gIII. Virology **176**, 145-157.

Folkers, G. and Trumpp, S. (1987). Sequence and structural homologies of the substrate binding site of cellular and viral thymidine kinases to rabbit muscle adenylate kinase. Med. Sci. Res. **15**, 1495-1496.

Forghani, B., Dupuis, K.W. and Schmidt, N.J. (1984). Varicella-zoster viral glycoproteins analyzed with monoclonal antibodies. J. Virol. **52**, 55-62.

Forrester, A.J. and Minson, A.C. (1990). Expression of herpes simplex virus type 1 glycoprotein H in recombinant vaccinia virus and eukaryotic cells. Abstract 245 , 15th International Herpesvirus Workshop. Washington DC, USA.

Frenkel, N., Locker, H. and Vlazny, D.A. (1980). Studies of defective herpes simplex viruses. Ann. NY Acad. Sci. **354**, 347-370.

Friedman, H.M., Cohen, G.H., Eisenberg, R.J., Seidel, C.A. and Cines, D.B. (1984). Glycoprotein C of herpes simplex virus 1 acts as a receptor for the C3b complement component on infected cells. Nature, London **309**, 633-635.

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.

Fries, L.F., Friedman, H.M., Cohen, G.H., Eisenberg, R.J., Hammer, C.H. and Frank, M.M. (1986). Glycoprotein C of herpes simplex virus 1 is an inhibitor of the complement cascade. J. Immunol. **137**, 1636-1641.

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.

Fry, D.C., Kuby, S.A. and Mildvan, A.C. (1986). ATP-binding site of adenylate kinase : mechanistic implications of its homology with ras-encoded p21, $F_1$-ATPase, and other nucleotide-binding proteins. Proc. Natl. Acad. Sci. USA **83**, 907-911.

Fuller, A.O. and Spear, P.G. (1985). Specificities of monoclonal and polyclonal antibodies that inhibit adsorption of herpes simplex virus to cells and lack of inhibition by potent neutralizing antibodies. J. Virol. **55**, 475-482.

Fuller, A.O. and Spear, P.G. (1987). Anti-glycoprotein D antibodies that permit adsorption but block infection by herpes simplex virus 1 prevent virion-cell fusion at the cell surface. Proc. Natl. Acad. Sci. USA **84**, 5454-5458.

Fuller, A.O., Santos, R.E. and Spear, P.G. (1989). Neutralizing antibodies specific for glycoprotein H of herpes simplex virus permit viral attachment to cells but prevent penetration. J. Virol. **63**, 3435-3443.

Furman, P.A., St. Clair, M.H. and Spector, T. (1984). Acyclovir triphosphate is a suicide inactivator of the herpes simplex virus DNA polymerase. J. Biol. Chem. **259**, 9575-9579.

Fyfe, J.A. (1982). Differential phosphorylation of (E)-5-(2-bromovinyl)-2′-deoxyuridine monophosphate by thymidylate kinases from herpes simplex viruses types 1 and 2 and varicella zoster virus. Mol. Pharmacol. **21**, 432-437.

Fyfe, J.A., McKee, S.A. and Keller, P.M. (1983). Altered thymidine-thymidylate kinases from strains of herpes simplex virus with modified drug sensitivities to acyclovir and (E)-5-(2-bromovinyl)-2′-deoxyuridine. Mol. Pharmacol. **24**, 316-323.

Gao, Q. and Spear, P.G. (1990). The product of the US5 open reading frame of herpes simplex virus type 1. Abstract 237, 15th International Herpesvirus Workshop. Washington DC, USA.

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. USA **82**, 5265-5269.

Gentry, G.A. (1985). Locating a nucleotide-binding site in the thymidine kinase of vaccinia virus and of herpes simplex virus by scoring triply aligned protein sequences. Proc. Natl. Acad. Sci. USA **82**, 6815-6819.

Gentry, G.A., Lowe, M., Alford, G. and Nevins, R. (1988). Sequence analyses of herpesviral enzymes suggest an ancient origin for human sexual behaviour. Proc. Natl. Acad. Sci. USA **85**, 2658-2661.

Geysen, H.M., Tainer, J.A., Rodda, S.J., Mason, T.J., Alexander, H., Getzoff, E.D. and Lerner, R.A. (1987). Chemistry of antibody binding to a protein. Science **235**, 1184-1190.

Giller, R.H., Winistorfer, S. and Grose, C. (1990). Cellular and humoral immunity to varicella zoster virus (VZV) glycoprotein V (gpV) : studies using a recombinant viral antigen. Abstract 412, 15th International Herpesvirus Workshop. Washington DC, USA.

Ginsberg, H.S. (1980). Herpesviruses. In 'Virology', pp. 1061-1076. Edited by R. Dulbecco and H.S. Ginsberg. Harper and Row, Hagerstown, USA.

Gleeson, L.J. and and Coggins, L. (1980). Response of pregnant mares to equine herpesvirus 1 (EHV-1). Cornell Vet. **70**, 391-400.

Glorioso, J., Kees, U., Kumel, G., Kirchner, H. and Krammer, P.H. (1985). Identification of herpes simplex virus type 1 (HSV-1) glycoprotein gC as the immunodominant antigen for HSV-1-specific memory cytotoxic T lymphocytes. J. Immunol. **135**, 575-582.

Gompels, U. and Minson, A. (1986). The properties and sequence of glycoprotein H of herpes simplex virus type 1. Virology **153**, 230-247.

Gompels, U.A. and Minson, A.C. (1989). Antigenic properties and cellular localization of herpes simplex virus glycoprotein H synthesized in a mammalian cell expression system. J. Virol. **63**, 4744-4755.

Gompels, U.A., Craxton, M.A. and Honess, R.W. (1988a). Conservation of gene organization in the lymphotropic herpesvirus herpesvirus saimiri and Epstein-Barr virus. J. Virol. **62**, 757-767.

Gompels, U.A., Craxton, M.A. and Honess, R.W. (1988b). Conservation of glycoprotein H (gH) in herpesviruses: nucleotide sequence of the gH gene from herpesvirus saimiri. J. Gen. Virol. **69**, 2819-2829.

Graham, F.L. and van der Eb, A.J. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**, 456-467.

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.

Gray, W.L., Baumann, R.P., Robertson, A.T., Caughman, G.B., O'Callaghan, D.J. and Staczek, J. (1987a). Regulation of equine herpesvirus type 1 gene expression : characterization of immediate early, early, and late transcription. Virology **158**, 79-87.

Gray, W.L., Baumann, R.P., Robertson, A.T., O'Callaghan, D.J. and Staczek, J. (1987b). Characterization and mapping of equine herpesvirus type 1 immediate early, early, and late transcripts. Virus Res. **8**, 233-244.

Gretch, D.R., Kari, B., Rasmussen, L., Gehrz, R.C. and Stinski, M.F. (1988). Identification and characterization of three distinct families of glycoprotein complexes in the envelopes of human cytomegalovirus. J. Virol. **62**, 875-881.

Grose, C., Edwards, D.P., Friedrichs, W.E., Weigle, K.A. and McGuire, W.L. (1983). Monoclonal antibodies against three major glycoproteins of varicella-zoster virus . Infect. Immun. **40**, 381-388.

Grundy, F.J., Baumann, R.P. and O'Callaghan, D.J. (1989). DNA sequence and comparative analyses of the equine herpesvirus type 1 immediate early gene. Virology **172**, 223-236.

Guo, P., Goebel, S., Davis, S., Perkus, M.E., Languet, B., Desmettre, P., Allen, G. and Paoletti, E. (1989). Expression in recombinant vaccinia virus of the equine herpesvirus 1 gene encoding glycoprotein gp13 and protection of immunized animals. J. Virol. **63**, 4189-4198.

Guo, P., Goebel, S., Perkus, M.E., Taylor, J., Norton, E., Allen, G., Languet, B., Desmettre, P. and Paoletti, E. (1990). Coexpression by vaccinia virus recombinants of equine herpesvirus 1 glycoproteins gp13 and gp14 results in potentiated immunity. J. Virol. **64**, 2399-2406.

Haarr, L. and Flatmark, T. (1987). Evidence that deletion of coding sequences in the 5′ end of the thymidine kinase gene of herpes simplex virus type 1 affects the stability of the gene products. J. Gen. Virol. **68**, 2817-2829.

Hall, M.J. and Katrak, K. (1986). The quest for a herpes simplex virus vaccine : background and recent developments. Vaccine **4**, 138-150.

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 viruses using intertypic recombinants. J. Gen. Virol. **49**, 235-253.

Halpern, M.E. and Smiley, J.R. (1984). Effects of deletions on expression of the herpes simplex virus thymidine kinase gene from the intact viral genome: the amino terminus of the enzyme is dispensable for catalytic activity. J. Virol. **50**, 733-738.

Hammerschmidt, W., Ludwig, H. and Budk, H.J. (1988). Specificity of cleavage in replicative-form DNA of bovine herpesvirus 1. J. Virol. **62**, 1355-1363.

Hampl, H., Ben-Porat, T., Ehrlicher, L., Habermehl, K.-O. and Kaplan, A.S. (1984). Characterization of the envelope proteins of pseudorabies virus. J. Virol. **52**, 583-590.

Hartley, W.J. and Dixon, R.J. (1979) An outbreak of foal perinatal mortality due to equid herpesvirus type-1 : pathological observations. Equine Vet. J. **11**, 215-218.

Heil, A., Muller, G., Noda, L., Pinder, T., Schirmer, H., Schirmer, I. and von Zabern, I. (1974). The amino acid sequence of porcine adenylate kinase from skeletal muscle. Eur. J. Biochem. **43**, 131-144.

Heineman, T., Gong, M., Sample, J. and Kieff, E. (1988). Identification of the Epstein-Barr virus gp85 gene. J. Virol. **62**, 1101-1107.

Henderson, L.M., Levings, R.L., Davis, A.J., Katz, J.B., Sturz, D.R. and Mayfield, J.E. (1990). Vaccine strains of suid herpesvirus 1 recombine in swine to produce virulent progeny virus. Abstract 433, 15th International Herpesvirus Workshop. Washington DC, USA.

Henikoff, S. (1984). Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28**, 351-359.

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.

Highlander, S.L., Cai, W., Person, S., Levine, M. and Glorioso, J.C. (1988). Monoclonal antibodies define a domain on herpes simplex virus glycoprotein B involved in virus penetration. J. Virol. **62**, 1881-1888.

Hill, T.J. (1982). Herpes simplex virus latency. In 'The Herpesviruses', vol. 3, pp. 175-240. Edited by B. Roizman. Plenum Press, NY.

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.

Holland, T.C., Homa, F.L., Marlin, S.D., Levine, M. and Glorioso, J. (1984a). Herpes simplex virus type 1 glycoprotein C-negative mutants exhibit multiple phenotypes, including secretion of truncated glycoproteins. J. Virol. **52**, 566-574.

Holland, T.C., Lerch, R.J., and Earhart, K. (1988). The cytoplasmic domain of herpes simplex virus type 1 glycoprotein C is required for membrane anchoring. J. Virol. **62**, 1753-1761.

Holland, L.E., Sandri-Goldin, R.M., Goldin, A.L., Glorioso, J.C. and Levine, M. (1984b). Transcriptional and genetic analyses of the herpes simplex virus type 1 genome: coordinates 0.29 to 0.45. J. Virol. **49**, 947-959.

Homa, F.L., Glorioso, J.C. and Levine, M. (1988). A specific 15-bp TATA box promoter element is required for expression of a herpes simplex virus type 1 late gene. Genes and Dev. **2**, 40-53.

Honess, R.W. (1984). Herpes simplex and 'the herpes complex' : diverse observations and a unifying hypothesis. J. Gen. Virol. **65**, 2077-2107.

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 Roizman, B. (1975). Regulation of herpesvirus macromolecular synthesis : sequential transition of polypeptide synthesis requires functional viral polypeptides. Proc. Natl. Acad. Sci. USA **72**, 1276-1280.

Honess, R.W., Craxton, M.A., Williams, L. and Gompels, U.A. (1989a). A comparative analysis of the sequence of the thymidine kinase gene of a gammaherpesvirus, herpesvirus saimiri. J. Gen. Virol. **70**, 3003-3013.

Honess, R.W., Gompels, U.A., Barrell, B.G., Craxton, M., Cameron, K.R., Staden, R., Chang, Y.-N. and Hayward, G.S. (1989b). Deviations from expected frequencies of CpG dinucleotides in herpesvirus DNAs may be diagnostic of differences in the states of their latent genomes. J. Gen. Virol. **70**, 837-855.

Hopp, T.P. and Woods, K.R. (1981). Prediction of protein antigenic determinants from amino acid sequences. Proc. Natl. Acad. Sci. USA **78**, 3824-3828.

Hruby, D., Maki, R.A., Miller, D.B. and Ball, L.A. (1983). Fine structure analysis and nucleotide sequence of the vaccinia virus thymidine kinase gene. Proc. Natl. Acad. Sci. USA **80**, 3411-3415.

Hubbard, S.C. and Ivatt, R.J. (1981). Synthesis and processing of asparagine-linked oligosaccharides. Ann. Rev. Biochem. **50**, 555-583.

Hudson, G.C., Whalley, J.M., Robertson, G.R., Scott, N.A., Sabine, M.M. and Love, D.N. (1988). In 'Equine Infectious Diseases V - Proceedings of the Fifth International Conference' , pp. 111-117. Edited by D.G. Powell. University Press of Kentucky, Kentucky.

Huemer, H.P., Broker, M., Larcher, C., Lambris, J.D. and Dierich, M.P. (1989). The central segment of herpes simplex virus type 1 glycoprotein C (gC) is not involved in C3b binding: demonstration by using monoclonal antibodies and recombinant gC expressed in Escherichia coli. J. Gen. Virol. **70**, 1571-1578.

Ihara, T., Kato, A., Ueda, S., Ishihama, A. and Hirai, K. (1989). Comparison of the sequence of the secretory glycoprotein A (gA) gene in Md5 and BC-1 strains of Marek's disease virus type 1. Virus Genes **3**, 127-140.

Inglis, M.M. and Darby, G. (1987). Analysis of the role of the cysteine 171 residue in the activity of herpes simplex virus type 1 thymidine kinase by oligonucleotide-directed mutagenesis. J. Gen. Virol. **68**, 39-46.

Isfort, R.J., Kung, H.-J. and Velicer, L.F. (1987). Identification of the gene encoding Marek's disease herpesvirus A antigen. J. Virol. **61**, 2614-2620.

Isfort, R.J., Stringer, R.A., Kung, H.-J. and Velicer, L.F. (1986). Synthesis, processing, and secretion of the Marek's disease herpesvirus A antigen glycoprotein. J. Virol. **57**, 464-474.

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.

Jacobs, L., Meloen, R.H., Gielkens, A.L.J. and van Oirschot, J.T. (1990). Epitope analysis of glycoprotein I of pseudorabies virus. J. Gen. Virol. **71**, 881-887.

Jacobson, J.G., Leib, D.A., Goldstein, D.J., Bogard, C.L., Schaffer, P.A., Weller, S.K. and Coen, D.M. (1989a). A herpes simplex virus ribonucleotide reductase deletion mutant is defective for productive acute and reactivatable latent infections of mice and for replication in mouse cells. Virology **173**, 276-283.

Jacobson, J.G., Martin, S.L. and Coen D.M. (1989b). A conserved open reading frame that overlaps the herpes simplex virus thymidine kinase gene is important for viral growth in cell culture. J. Virol. **63**, 1839-1843.

Jamieson, A.T., Gentry, G.A. and Subak-Sharpe, J.H. (1974). Induction of both thymidine and deoxycytidine kinase activity by herpes viruses. J. Gen. Virol. **24**, 465-480.

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

Johnson, D.C., Frame, M.C., Ligas, M.W., Cross, A.M. and Stow, N.D. (1988). Herpes simplex virus immunoglobulin G Fc receptor activity depends on a complex of two viral glycoproteins, gE and gI. J. Virol. **62**, 1347-1354.

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.

Jongeneel, C.V. and Bachenheimer, S.L. (1981). Structure of replicating herpes simplex virus DNA. J. Virol. **39**, 656-660.

Kamps, M.P. and Sefton, B.M. (1986). Neither arginine nor histidine can carry out the function of lysine-295 in the ATP-binding site of $_{p}$60src. Mol. Cell. Biol. **6**, 751-757.

Kaufman, J.F., Auffray, C., Korman, A.J., Shackleford, D.A. and Strominger, J. (1984). The class II molecules of the human and murine major histocompatibility complex. Cell **36**, 1-13.

Keller, P.M., Davison, A.J., Lowe, R.S., Riemen, M.W. and Ellis, R.W. (1987). Identification and sequence of the gene encoding gpIII, a major glycoprotein of varicella-zoster virus. Virology **157**, 526-533.

Keller, P.M., Lonergan, K., Neff, B.J., Morton, D.A. and Ellis, R.W. (1986). Purification of individual varicella-zoster virus (VZV) glycoproteins gpI, gpII and gpIII and their use in ELISA for detection of VZV glycoprotein-specific antibody. J. Virol. Methods **14**, 177-188.

Keller, P.M., Neff, B.J. and Ellis, R.W. (1984). Three major glycoprotein genes of varicella-zoster virus whose products have neutralization epitopes. J. Virol. **52**, 293-297.

Kemeny, L. and Pearson, J.E. (1970). Isolation of herpesvirus from equine leukocytes : comparison with equine rhinopneumonitis virus. Can. J. Comp. Med. **34**, 59-65.

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.

Kikuchi, G.E., Glorioso, J.C. and Nairn, R. (1990). Cross-linking studies show that herpes simplex virus type 1 glycoprotein C molecules are clustered in the membrane of infected cells. J. Gen. Virol. **71**, 455-458.

Kinchington, P.R., Remenick, J., Ostrove, J.M., Straus, S.E., Ruyechan, W.T. and Hay, J. (1986). Putative glycoprotein gene of varicella-zoster virus with variable copy numbers of a 42-base-pair repeat sequence has homology to herpes simplex virus glycoprotein C. J. Virol. **59**, 660-668.

Kit, S. (1985a). Thymidine kinase. Microbiol. Sci. **2**, 369-375.

Kit, S. (1985b). U.S. Patent Application, no. 4514497.

Kit, S. and Dubbs, D.R. (1977). Regulation of herpesvirus thymidine kinase activity in LM(TK-) cells transformed by ultraviolet light-irradiated herpes simplex virus. Virology **76**, 331-340.

Kit, S., Ichimura, H. and De Clerq, E. (1987a). Phosphorylation of nucleoside analogs by equine herpesvirus type 1 pyrimidine deoxyribonucleoside kinase. Antiviral Res. **7**, 53-67.

Kit, S., Ichimura, H. and De Clerq, E. (1987b). Differential metabolism of (E)-5-(2-iodovinyl)-2′-deoxyuridine (IVDU) by equine herpesvirus type 1- and herpes simplex virus-infected cells. Antiviral Res. **8**, 41-51.

Kit, S., Kit, M., Ichimura, H., Crandell, R. and McConnell, S. (1986). Induction of thymidine kinase activity by viruses with group B DNA genomes : bovine cytomegalovirus (bovine herpesvirus 4). Virus Res. **4**, 197-212.

Kit, S., Kit, M. and Pirtle, E.C. (1985a). Attenuated properties of thymidine kinase-negative deletion mutants of pseudorabies virus. Am. J. Vet. Res. **46**, 1359-1367.

Kit, S., Kit, M. Qavi, H., Trkula, D. and Otsuka, H. (1983a). Nucleotide sequence of the herpes simplex virus type 2 (HSV-2) thymidine kinase gene and predicted amino acid sequence of thymidine kinase polypeptide and its comparison with the HSV-1 thymidine kinase gene. Biochim. Biophys. Acta **741**, 158-170.

Kit, S., Qavi, H., Dubbs, D.R. and Otsuka, H. (1983b). Attenuated marmoset herpesvirus isolated from recombinants of virulent marmoset herpesvirus and hybrid plasmids. J. Med. Virol. **12**, 25-36.

Kit, S., Qavi, H., Gaines, J.D., Billingsley, P. and McConnell, S. (1985b). Thymidine kinase-negative bovine herpesvirus type 1 mutant is stable and highly attenuated in calves. Arch. Virol. **86**, 63-83.

Kit, S., Sheppard, M., Ichimura, H. and Kit, M. (1987c). Second-generation pseudorabies virus vaccine with deletions in thymidine kinase and glycoprotein genes. Am. J. Vet. Res. **48**, 780-793.

Kit, S., Sheppard, M., Ichimura, H., Nusinoff-Lehrman, S., Ellis, M.N., Fyfe, J.A. and Otsuka, H. (1987d). Nucleotide sequence changes in thymidine kinase gene of herpes simplex virus type 2 clones from an isolate of a patient treated with acyclovir. Antimicrob. Agents Chemother. **31**, 1483-1490.

Koff, A. and Tegtmeyer, P. (1988). Characterization of major recognition sequences for a herpes simplex virus type 1 origin-binding protein. J. Virol. **62**, 4096-4103.

Kouzarides, A., Bankier, A.T., Satchwell, S.C., Weston, K., Tomlinson, P. and Barrell, B.G. (1987). Large-scale rearrangement of homologous regions in the genomes of HCMV and EBV. Virology **157**, 397-413.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**, 857-872.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44**, 283-292.

Kuhn, J.E., Kramer, M.D., Willenbacher, W., Wieland, U., Lorentzen, E.U. and Braun, R.W. (1990). Identification of herpes simplex virus type 1 glycoproteins interacting with the cell surface. J. Virol. **64**, 2491-2497.

Kumel, G., Kaerner, H.C., Levine, M., Schroder, C.H., and Glorioso, J.C. (1985). Passive immune protection by herpes simplex virus-specific monoclonal antibodies and monoclonal antibody-resistant mutants altered in pathogenicity. J. Virol. **56**, 930-937.

Kupershmidt, S., Lu, Z. , De Marchi, J.M. and Ben-Porat, T. (1988). Characterization of an origin of replication of pseudorabies virus present within the terminal region of the L component of the genome. Abstract 92, 13th International Herpesvirus Workshop. Irvine, Ca, USA.

Kyte, J. and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**, 105-132.

Larder, B.A., Cheng, Y.-C. and Darby, G. (1983). Characterization of abnormal thymidine kinases induced by drug-resistant strains of herpes simplex virus type 1. J. Gen. Virol. **64**, 523-532.

Lawrence , G.L., Chee, M., Craxton, M.A., Gompels, U.A., Honess, R.W. and Barrell, B.G. (1990). Human herpesvirus 6 is closely related to human cytomegalovirus. J. Virol. **64**, 287-299.

Lindahl, T., Adams, A., Bjursell, G., Bornkamm, G.W., Kaschka-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.

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, Q. and Summers, W.C. (1988). Site-directed mutagenesis of a nucleotide-binding domain in HSV-1 thymidine kinase: effects on catalytic activity. Virology **163**, 638-642.

Lockshon, D. and Galloway, D.A. (1986). Cloning and characterization of $\underline{ori}_{L2}$, a large palindromic DNA replication origin of herpes simplex virus type 2. J. Virol. **58**, 513-521.

Longnecker, R. and Roizman, B. (1986). Generation of an inverting herpes simplex virus 1 mutant lacking the L-S junction a sequences, an origin of DNA synthesis, and several genes including those specifying glycoprotein E and the $\alpha$ 47 gene. J. Virol. **58**, 583-591.

Longnecker, R. and Roizman, B. (1987). Clustering of genes dispensable for growth in culture in the S component of the HSV-1 genome. Science **236**, 573-576.

Ludwig, H., Biswal, N., Bryans, J.T. and McCombs, R.M. (1971). Some properties of the DNA from a new equine herpesvirus. Virology **45**, 534-537.

Ludwig, H., Haines, H.G., Biswal, N. and Benyesh-Melnick, M. (1972). The characterization of varicella-zoster virus DNA. J. Gen. Virol. **14**, 111-114.

Lussenhop, N.O., Goertz, R., Wabuke-Bunoti, M., Gehrz, R. and Kari, B. (1988). Epitope analysis of human cytomegalovirus glycoprotein complexes using murine monoclonal antibodies. Virology **164**, 362-372.

McDermott, M.R., Graham, F.L., Hanke, T. and Johnson, D.C. (1989). Protection of mice against lethal challenge with herpes simplex virus by vaccination with an adenovirus vector expressing HSV glycoprotein B. Virology **169**, 244-247.

McDougall, J.K., Masse, T.H. and Galloway, D.A. (1980). Location and cloning of the herpes simplex virus type 2 thymidine kinase gene. J. Virol. **33**, 1221-1224.

McGeoch, D.J. and Davison, A.J. (1986). DNA sequence of the herpes simplex virus type 1 gene encoding glycoprotein gH, and identification of homologues in the genome of varicella-zoster virus and Epstein-Barr virus. Nucleic Acids Res. **14**, 4281-4292.

McGeoch, D.J., Dalrymple, M.A., Davison, A.J., Dolan, A., Frame, M.C., McNab, D., Perry, L.J., Scott, J.E. and Taylor, P. (1988a). The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69**, 1531-1574.

McGeoch, D.J., Dalrymple, M.A., Dolan, A., McNab, D., Perry, L.J., Taylor, P. and Challberg, M.D. (1988b). Structures of herpes simplex virus type 1 genes required for replication of DNA. J. Virol. **62**, 444-453.

McGeoch, D.J., Dolan, A., Donald, S. and Brauer, D.H.K. (1986). Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res. **14**, 1727-1745.

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 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. Nucleic 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.

McLauchlan, J., Gaffney, D., Whitton, J.L. and Clements, J.B. (1985). The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3′ termini. Nucleic Acids Res. **13**, 1347-1368.

McNearney, T.A., Odell, C., Holers, V.M., Spear, P.G. and Atkinson, J.P. (1987). Herpes simplex virus glycoproteins gC-1 and gC-2 bind to the third component of complement and provide protection against complement-mediated neutralization of viral infectivity. J. Exp. Med. **166**, 1525-1535.

Mancini, W.R., De Clercq, E. and Prusoff, W.A. (1983). The relationship between incorporation of E-5-(2-bromovinyl)-2'-deoxyuridine into herpes simplex virus type 1 DNA with virus infectivity and DNA integrity. J. Biol. Chem. **258**, 792-795.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). 'Molecular Cloning : A Laboratory Manual'. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Marlin, S.D., Holland, T.C., Levine, M. and Glorioso, J.C. (1985). Epitopes of herpes simplex virus type 1 glycoprotein gC are clustered in two distinct antigenic sites. J. Virol. **53**, 128-136.

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.

Martin, S.L., Aparisio, D.I. and Bandyopadhyay, P.K. (1989). Genetic and biochemical characterization of the thymidine kinase gene from herpesvirus of turkeys. J. Virol. **63**, 2847-2852.

Martin, S., Courtney, R.J., Fowler, G. and Rouse, B.T. (1988). Herpes simplex virus type 1-specific cytotoxic T lymphocytes recognise virus nonstructural proteins. J. Virol. **62**, 2265-2273.

Matsuo, T., Heller, M., Petti, L., O'Shiro, E. and Kieff, E. (1984). Persistence of the entire Epstein-Barr virus genome integrated into human lymphocyte DNA. Science **226**, 1322-1325.

Meigner, B., Longnecker, R., Mavromara-Nazos, P., Sears, A.E. and Roizman, B. (1988). Virulence of and establishment of latency by genetically engineered deletion mutants of herpes simplex virus 1. Virology **162**, 251-254.

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

Meredith, D.M., Stocks, J-M., Whittaker, G.R., Halliburton, I.W., Snowden, B.W. and Killington, R.A. (1989). Identification of the gB homologues of equine herpesvirus types 1 and 4 as disulphide-linked heterodimers and their characterization using monoclonal antibodies. J. Gen. Virol. **70**, 1161-1172.

Mettenleiter, T.C., Lukacs, N., Thiel, H.-J., Schreurs, C. and Rziha, H.-J. (1986). Location of the structural gene of pseudorabies virus glycoprotein complex gII. Virology **152**, 66-75.

Mettenleiter, T.C., Schreurs, C., Zuckermann, F., Ben-Porat, T. and Kaplan, A.S. (1988). Role of glycoprotein gIII of pseudorabies virus in virulence. J. Virol. **62**, 2712-2717.

Mettenleiter, T.C., Zsak, L., Zuckermann, F., Sugg, N., Kern, H. and Ben-Porat, T. (1990). Interaction of glycoprotein gIII with a cellular heparinlike substance mediates adsorption of pseudorabies virus. J. Virol. **64**, 278-286.

Meyer, H., Thein, P. and Hubert, P. (1987). Characterization of two equine herpesvirus (EHV) isolates associated with neurological disorders in horses. J. Vet. Med. B **34**, 545-548.

Miller, N. and Hutt-Fletcher, L.M. (1988). A monoclonal antibody to glycoprotein gp85 inhibits fusion but not attachment of Epstein-Barr virus. J. Virol. **62**, 2366-2372.

Mittal, S.K. and Field, H.J. (1989). Analysis of the bovine herpesvirus type 1 thymidine kinase (TK) gene from wild-type virus and TK-deficient mutants. J. Gen. Virol. **70**, 901-918.

Montalvo, E.A. and Grose, C. (1986). Neutralization epitope of varicella zoster virus on native viral glycoprotein gp118 (VZV glycoprotein gpIII). Virology **149**, 230-241.

Morgan, C., Rose, H.M. and Mednis, B. (1968). Electron microscopy of herpes simplex virus . I. Entry. J. Virol. **2**, 507-516.

Morris, C.M. and Field, H.J. (1988). Application of cloned fragments of equine herpesvirus type-1 DNA for detection of virus-specific DNA in equine tissues. Equine Vet. J. **20**, 335-340.

Mumford, J.A. (1985). Equid herpesvirus 1 (EHV-1) latency : more questions than answers. Equine Vet. J. **17**, 340-342.

Mumford, J.A. and Bates, J. (1984). Trials of an inactivated equid herpesvirus 1 vaccine : challenge with a subtype 2 virus. Vet. Rec. **114**, 375-381.

Newcomb, W.W., Brown, J.C., Booy, F.P. and Steven, A.C. (1989). Nucleocapsid mass and capsomer protein stoichiometry in equine herpesvirus 1 : scanning transmission electron microscopic study. J. Virol. **63**, 3777-3783.

Nicolson, L. and Onions, D.E. (1990) . The nucleotide sequence of the equine herpesvirus 4 gC gene homologue . Virology **179**, 378-387.

Nicolson, L., Cullinane, A.A. and Onions, D.E. (1990a). The nucleotide sequence of an equine herpes virus 4 (EHV-4) gene homologue of the herpes simplex virus 1 (HSV-1) glycoprotein H gene. J. Gen. Virol. **71**, 1793-1800.

Nicolson, L., Cullinane, A.A. and Onions, D.E. (1990b). The nucleotide sequence of the equine herpesvirus 4 thymidine kinase gene. J. Gen. Virol. **71**, 1801-1805.

Norrild, B. (1985). Humoral response to herpes simplex virus infections. In 'The Herpesviruses', vol. 4, pp. 69-86. Edited by B. Roizman and C. Lopez. Plenum Publishing Corp., New York.

Nunberg, J.H., Wright, D.K., Cole, G.E., Petrovskis, E.A., Post, L.E., Compton, T. and Gilbert, J.H. (1989). Identification of the thymidine kinase gene of feline herpesvirus : use of degenerate oligonucleotides in the polymerase chain reaction to isolate herpesvirus gene homologs. J. Virol. **63**, 3240-3249.

Oba, D.E. and Hutt-Fletcher, L.M. (1988). Induction of antibodies to the Epstein-Barr virus glycoprotein gp85 with a synthetic peptide corresponding to a sequence in the BXLF2 open reading frame. J. Virol. **62**, 1108-1114.

O'Callaghan, D.J. and Randall, C.C. (1976). Molecular anatomy of herpesviruses : recent studies. Prog. Med. Virol. **22**, 152-210.

O'Callaghan, D.J., Gentry, G.A. and Randall, C.C. (1983). The equine herpesviruses. In 'The Herpesviruses', vol.2, pp. 215-318. Edited by B. Roizman. Plenum Press, New York.

Okazaki, K., Honda, E., Minetoma, T. and Kumagai T. (1987). Bovine herpesvirus type 1 gp87 mediates both attachment of virions to susceptible cells and hemagglutination. Arch. Virol. **97**, 297-307.

Olivo, P.D., Nelson, N.J. and Challberg, M.D. (1988). Herpes simplex virus DNA replication : the UL9 gene encodes an origin-binding protein. Proc. Natl. Acad. Sci. USA **85**, 5414-5418.

Olivo, P.D., Nelson, N.J. and Challberg, M.D. (1989). Herpes simplex virus type 1 gene products required for DNA replication : identification and overexpression. J. Virol. **63**, 196-204.

Olofsson, S., Lundstrom, M. and Datema, R. (1985). The antiherpes drug (E)-5-(2-bromovinyl)-2'-deoxyuridine (BVdU) interferes with formation of N-linked and of O-linked oligosaccharides of the herpes simplex virus type 1 glycoprotein C. Virology **147**, 201-205.

Olofsson, S., Sjoblom, I. and Jeansson, S. (1990). Activity of herpes simplex virus type 1-specified glycoprotein C antigenic site II epitopes reversibly modulated by peripheral fucose or galactose units of glycoprotein oligosaccharides. J. Gen. Virol. **71**, 889-895.

Otsuka, H. and Kit, S. (1984). Nucleotide sequence of the marmoset herpesvirus thymidine kinase gene and predicted amino acid sequence of thymidine kinase polypeptide. Virology **135**, 316-330.

Otsuka, H., Hazen, M., Kit, M., Qavi, H. and Kit, S. (1981). Cloning of the marmoset herpesvirus thymidine kinase gene and analyses of the boundaries of the coding region. Virology **113**, 196-213.

Pai, E.F., Sachsenheimer, W., Schirmer, R.H. and Schulz, G.E. (1977). Substrate positions and induced-fit in crystalline adenylate kinase. J. Mol. Biol. **114**, 37-45.

Papp-Vid, G. and Derbyshire, J.B. (1978). The protective antigens of equine herpesvirus type 1. Can. J. Comp. Med. **42**, 219-226.

Papp-Vid, G. and Derbyshire, J.B. (1979). The virus neutralizing activity of antibodies specific to the envelope and nucleocapsid of equine herpesvirus type 1. Can. J. Comp. Med. **43**, 231-233.

Para, M.F., Goldstein, L. and Spear, P.G. (1982). Similarities and differences in the Fc-binding glycoprotein (gE) of herpes simplex virus types 1 and 2 and tentative mapping of the viral gene for this glycoprotein. J. Virol. **41**, 137-144.

Para, M.F., Parish, M.L., Noble, A.G. and Spear, P.G. (1985). Potent neutralizing activity associated with anti-glycoprotein D specificity among monoclonal antibodies selected for binding to herpes simplex virions. J. Virol. **55**, 483-488.

Patel, J.R. and Edington, N. (1983). The pathogenicity in mice of respiratory, abortion, and paresis isolates of equine herpesvirus-1. Vet. Microbiol. **8**, 301-305.

Patel, J.R., Edington, N. and Mumford, J.A. (1982). Variation in cellular tropism between isolates of equine herpesvirus-1 in foals. Arch. Virol. **74**, 41-51.

Perlman, D. and Halvorson, H.O. (1983). A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J. Mol. Biol. **167**, 391-409.

Perry, L.J. and McGeoch, D.J. (1988). The DNA sequence of the long repeat region and adjoining parts of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69**, 2831-2846.

Petrovskis, E.A. and Post, L.E. (1987). A small open reading frame in pseudorabies virus and implications for evolutionary relationships between herpesviruses. Virology **159**, 193-195.

Petrovskis, E.A, Duffus, P.W.H., Thomsen, D.R., Meyer, A.L. and Post, L.E. (1988). Sequence of pseudorabies virus and infectious bovine rhinotracheitis virus glycoprotein H genes. Abstract 217, 13th International Herpesvirus Workshop. Irvine, Ca, USA.

Petrovskis, E.A., Timmins, J.G., Armentrout, M.A., Marchioli, C.C., Yancey Jr., R.J., and Post, L.E. (1986a). DNA sequence of the gene for pseudorabies virus gp50, a glycoprotein without N-linked glycosylation. J. Virol. **59**, 216-223.

Petrovskis, E.A., Timmins, J.G. and Post, L.E. (1986b). Use of $\lambda$ gt11 to isolate genes for two pseudorabies virus glycoproteins with homology to herpes simplex virus and varicella-zoster virus glycoproteins. J. Virol. **60**, 185-193.

Plummer, G. and Waterson, A.P. (1963). Equine herpes virus. Virology **19**, 412-416.

Plummer, G. Goodheart, C.R. and Studdert, M.J. (1973). Equine herpesviruses : antigenic relationships and deoxyribonucleic acid densities. Infect. Immun. **8**, 621-627.

Polvino-Bodnar, M., Orberg, P.K. and Schaffer, P.A. (1987). Herpes simplex virus type 1 $\underline{ori}_L$ is not required for virus replication or for the establishment and reactivation of latent infection in mice. J. Virol. **61**, 3528-3535.

Proudfoot, N.J. and Brownlee, G.G. (1976). 3′ Non-coding region sequences in eukaryotic messenger RNA. Nature, London **263**, 211-214.

Queen, C. and Korn, L.J. (1984). A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res. **12**, 581-599.

Raab-Traub , N., Dambaugh, T. and Kieff, E. (1980). DNA of Epstein-Barr virus. VIII: B95-8, the previous prototype, is an unusual deletion derivative. Cell **22**, 257-267.

Rabkin, S.D. and Hanlon, B. (1990). Herpes simplex virus DNA synthesis at a preformed replication fork in vitro. J. Virol. **64**, 4957-4967.

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. Am. J. Pathol. **29**, 139-153.

Rasmussen, L.E., Nelson, R.M., Kelsall, D.C. and Merigan, T.C. (1984). Murine monoclonal antibody to a single protein neutralizes the infectivity of human cytomegalovirus. Proc. Natl. Acad. Sci. USA **81**, 876-880.

Rauh, I. and Mettenleiter, T.C. (1990). A glycoprotein gII-negative pseudorabies virus is functionally complemented by BHV-1 glycoprotein I. Abstract 179, 15th International Herpesvirus Workshop. Washington DC, USA.

Rea, T.J., Timmins, J.G., Long, G.W. and Post, L.E. (1985). Mapping and sequence of the gene for the pseudorabies virus glycoprotein which accumulates in the medium of infected cells. J. Virol. **54**, 21-29.

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

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

Riggio, M.P. (1990). Sequence analysis of selected regions of the equine herpesvirus 4 genome. PhD Thesis, University of Glasgow.

Riggio, M.P., Cullinane, A.A. and Onions, D.E. (1989). Identification and nucleotide sequence of the glycoprotein gB gene of equine herpesvirus 4. J. Virol. **63**, 1123-1133.

Robbins, A.K., Dorney, D.J., Wathen, M.W., Whealy, M.E., Gold, C., Watson, R.J., Holland, L.E., Weed, S.D., Levine, M., Glorioso, J.C. and Enquist, L.W. (1987). The pseudorabies virus gII gene is closely related to the gB glycoprotein gene of herpes simplex virus . J. Virol. **61**, 2691-2701.

Robbins, A.K., Ryan, J.P., Whealy, M.E. and Enquist, L.W. (1989). The gene encoding the gIII envelope protein of pseudorabies virus vaccine strain Bartha contains a mutation affecting protein localization. J. Virol. **63**, 250-258.

Robbins, A.K., Watson, R.J., Whealy, M.E., Hays, W.W. and Enquist, L.W. (1986a). Characterization of a pseudorabies virus glycoprotein gene with homology to herpes simplex virus type 1 and type 2 glycoprotein C. J. Virol. **58**, 339-347.

Robbins, A.K., Weis, J.H., Enquist, L.W. and Watson, R.J. (1984). Construction of E. coli expression plasmid libraries : localization of a pseudorabies virus glycoprotein gene . J. Mol. Appl. Genet. 2, 485-496.

Robbins, A.K., Whealy, M.E., Watson, R.J. and Enquist, L.W. (1986b). Pseudorabies virus gene encoding glycoprotein gIII is not essential for growth in tissue culture. J. Virol. **59**, 635-645.

Roberts, P.L., Duncan, B.E., Raybould, T.J.G and Watson, D.H. (1985). Purification of herpes simplex virus glycoproteins B and C using monoclonal antibodies and their ability to protect mice against lethal challenge. J. Gen. Virol. **66**, 1073-1085.

Robertson, G.R. and Whalley, J.M. (1988). Evolution of the herpes thymidine kinase gene : identification and comparison of the equine herpesvirus 1 thymidine kinase gene reveals similarity to a cell-encoded thymidylate kinase. Nucleic Acids Res. **16**, 11303-11317.

Robertson, A.T., Baumann, R.P., Staczek, J. and O'Callaghan, D.J. (1988). Molecular characterization of the gene products of the short region of the equid herpesvirus-1 genome. In 'Equine Infectious Diseases V - Proceedings of the Fifth International Conference' , pp. 132-139. Edited by D.G. Powell. University Press of Kentucky, Kentucky.

Robertson, G.R., Scott, N.A., Miller, J.M., Sabine, M., Zheng, M., Bell, C.W. and Whalley, J.M. (1991). Sequence characteristics of a gene in equine herpesvirus 1 homologous to glycoprotein H of herpes simplex virus. DNA Sequence, In Press.

Rock, D.L., Nesburn, A.B., Ghiasi, H., Ong, J., Lewis, T.L., Lokensgard, J.R. and Wechsler, S.L. (1987). Detection of latency-related viral RNAs in trigeminal ganglia of rabbits latently infected with herpes simplex virus type 1. J. Virol. **61**, 3820-3826.

Roizman, B. (1978). The herpesviruses. In 'Molecular Biology of Animal Viruses', vol. 2 , pp. 769-848. Edited by D.P. Nayak. Marcel Dekker inc., NY.

Roizman, B. (1982). The family herpesviridae: general description, taxonomy and classification. In ' The herpesviruses', vol. 1, pp. 1-23. Edited by B. Roizman. Plenum Press, New York.

Roizman, B. and Furlong, D. (1974). The replication of herpesviruses. In 'Comprehensive Virology', vol.3, pp. 229-403. Edited by H. Fraenkel-Conrat and R.R. Wagner. Plenum Press, NY.

Rosenthal, K.L., Smiley, J.R., South, S. and Johnson, D.C. (1987). Cells expressing herpes simplex virus glycoprotein gC but not gB, gD, or gE are recognised by murine virus-specific cytotoxic T lymphocytes. J. Virol. **61**, 2438-2447.

Ruyechan, W.T., Dauenhauer, S.A. and O'Callaghan, D.J. (1982). Electron microscopic study of equine herpesvirus type 1 DNA. J. Virol. **42**, 297-300.

Ryan, J.P., Whealy, M.E., Robbins, A.K. and Enquist, L. (1987). Analysis of pseudorabies virus glycoprotein gIII localization and modification by using novel infectious viral mutants carrying unique EcoRl sites. J. Virol. **61**, 2962-2972.

Rziha, H.-J., Mettenleiter, T.C., Ohlinger, V. and Wittmann, G. (1986). Herpesvirus (pseudorabies virus) latency in swine : occurrence and physical state of viral DNA in neural tissues. Virology **155**, 600-613.

Sabine, M. and Whalley, J.M. (1989). Towards a vaccine against equine herpesvirus 1. Aust. Vet. J. **66**, 403-404.

Sabine, M., Robertson, G.R. and Whalley, J.M. (1981). Differentiation of sub-types of equine herpesvirus 1 by restriction endonuclease analysis. Aust. Vet J. **57**, 148-149.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**, 5463-5467.

Sawyer, M.H., Inchauspe, G., Biron, K.K., Waters, D.J., Straus, S.E. and Ostrove, J.M. (1988). Molecular analysis of the pyrimidine deoxyribonucleoside kinase gene of wild-type and acyclovir-resistant strains of varicella-zoster virus. J. Gen. Virol. **69**, 2585-2593.

Sawyer, M.H., Ostrove, J.M., Felser, J.M. and Straus, S.E. (1986). Mapping of the varicella voster virus deoxypyrimidine kinase gene and preliminary identification of its transcript. Virology **149**, 1-9.

Schreurs, C., Mettenleiter, T.C., Zuckermann, F., Sugg, N. and Ben-Porat, T. (1988). Glycoprotein gIII of pseudorabies virus is multifunctional. J. Virol. **62**, 2251-2257.

Scott, S.D. (1989). Identification and sequence analysis of the thymidine kinase and flanking genes in two Marek's disease herpesviruses. PhD Thesis, University of Cambridge.

Scott, J.C., Dutta, S.K. and Myrup, A.C. (1983). In vivo harboring of equine herpesvirus-1 in leukocyte populations and subpopulations and their quantitation from experimentally infected ponies. Am. J. Vet. Res. **44**, 1344-1348.

Scott, S.D., Ross, N.L.J. and Binns, M.M. (1989). Nucleotide and predicted amino acid sequences of the Marek's disease virus and turkey herpesvirus thymidine kinase genes; comparison with thymidine kinase genes of other herpesviruses J. Gen. Virol. **70**, 3055-3065.

Seibl, R. and Wolf, H. (1985). Mapping of Epstein-Barr virus proteins on the genome by translation of hybrid-selected RNA from induced P3HR1 cells and induced Raji cells. Virology **141**, 1-13.

Seidel-Dugan, C., Ponce de Leon, M., Friedman, H.M., Fries, L.F., Frank, M.M., Cohen, G.H. and Eisenberg, R.J. (1988). C3b receptor activity on transfected cells expressing glycoprotein C of herpes simplex virus types 1 and 2. J. Virol. **62**, 4027-4036.

Sheldrick, P. and Berthelot, N. (1975). Inverted repetitions in the chromosome of herpes simplex virus. Cold Spring Harbor Symp. Quant. Biol. **39**, 667-678.

Sheppard, M. and May, J.T. (1989). Location and characterization of the bovine herpesvirus type 2 thymidine kinase gene. J. Gen. Virol. **70**, 3067-3071.

Shih, M.-F., Arsenakis, M., Tiollais, P. and Roizman, B. (1984). Expression of hepatitis B virus S gene by herpes simplex virus type 1 vectors carrying $\alpha$- and $\beta$ -regulated gene chimeras. Proc. Natl. Acad. Sci. USA **81**, 5867-5870.

Shimizu, T., Ishizaki, R., Ishii, S., Kawakami, Y., Kaji, R., Sugimura, K. and Matumoto, M. (1959). Isolation of equine abortion virus from natural cases of equine abortion in horse kidney cell culture. Japanese J. Exp. Med. **29**, 643-649.

Shimizu, M., Satou, K. and Nishioka, N. (1989). Monoclonal antibodies with neutralizing activity to equine herpesvirus 1. Arch. Virol. **104**, 169-174. **toc**

Showalter, S.D., Zweig, M. and Hampar, B. (1981). Monoclonal antibodies to herpes simplex virus type 1 proteins, including the immediate-early protein ICP4. Infect. Immun. **34**, 684-692.

Simon, A., Mettenleiter, T.C. and Rziha, H.-J. (1989). Pseudorabies virus displays variable numbers of a repeat unit adjacent to the 3′ end of the glycoprotein gII gene. J. Gen. Virol. **70**, 1239-1246.

Sinclair, R., Binns, M., Moult, B. and Mumford, J.A. (1990). Antigenic analysis of glycoprotein gp13 of euqid herpesvirus type-1 and type-4 using monoclonal antibodies. Abstract 422, 15th International Herpesvirus Workshop. Washington, DC, USA.

Sinclair, R., Cook, R.F. and Mumford, J.A. (1989). The characterization of neutralizing and non-neutralizing monoclonal antibodies against equid herpesvirus type 1. J. Gen. Virol. **70**, 455-459.

Sjoblom, I., Lundstrom, M., Sjogren-Jansson, E., Glorioso , J.C., Jeansson, S. and Olofsson, S. (1987). Demonstration and mapping of highly carbohydrate-dependent epitopes in the herpes simplex virus type 1-specified glycoprotein C. J. Gen. Virol. **68**, 545-554.

Smiley, M.L. and Friedman, H.M. (1985). Binding of complement component C3b to glycoprotein C is modulated by sialic acid on herpes simplex virus type 1-infected cells. J. Virol. **55**, 857-861.

Soehner, R.L., Gentry, G.A. and Randall, C.C. (1965). Some physiochemical characteristics of equine abortion virus nucleic acid. Virology **26**, 394-405.

Southern, E. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**, 503-517.

Spear, P.G. (1984). Glycoproteins specified by herpes simplex virus . In 'The Herpesviruses', vol. 3, pp.315-356. Edited by B. Roizman. Plenum Press, New York.

Staczek, J. , Atherton, S.S. and O'Callaghan, D.J. (1983). Genetic relatedness of the genomes of equine herpesvirus types 1, 2, and 3. J. Virol. **45**, 855-858.

Stannard, L.M., Fuller, A.O. and Spear, P.G. (1987). Herpes simplex virus glycoproteins associated with different morphological entities projecting from the virion envelope. J. Gen. Virol. **68**, 715-725.

Staschke, K., Colacino, J.M., Wolcott, R.M. and Staczek, J. (1990). Equine cytomegalovirus encoded thymidine kinase. Abstract 165, 15th International Herpesvirus Workshop. Washington, DC, USA.

Stein , S., Todd, P. and Mahoney, J. ( 1970). Infectivity of herpes simplex virus particles with and without envelopes. Can. J. Microbiol. **16**, 953-957.

Steiner, I., Spivack, J.G., Lirette, R.P., Brown, M.S., MacLean, A.R., Subak-Sharpe, J.H. and Fraser, N.W. (1989). Herpes simplex virus type 1 latency-associated transcripts are evidently not essential for latent infection. EMBO J. **8**, 505-511.

Stevens, J.G., Wagner, E.K., Devi-Rao, G.B., Cook, M.L. and Feldman, L.T. (1987). RNA complementary to a herpesvirus $\alpha$ gene mRNA is prominent in latently infected neurons. Science **235**, 1056-1059.

Stokes , A. and Wardley, R.C. (1988). ADCC and complement-dependent lysis as immune mechanisms against EHV-1 infection in the horse. Res. Vet. Sci. **44**, 295-302.

Stokes, A., Allen, G.P., Pullen, L.A. and Murray, P.K. (1989). A hamster model of equine herpesvirus-1 infection, passive protection by monoclonal antibodies to EHV-1 glycoproteins 13, 14 and 17/18. J. Gen. Virol. **70**, 1173-1183.

Stokes, A., Corteyn, A.H., Pullen, L.A., Doel, T.R., Meredith, D.M., Killington, R.A., Halliburton, I.W., Whittaker, G.R., Wheldon, L.A., Nicolson, L., Onions, D.E., Allen, G.P., and Murray, P.K. (1991). The antigenicity and protective immunogenicity of equid herpesvirus 1 and 4 glycoprotein 13 in a hamster model. J. Gen. Virol., In Press.

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 Davison, A.J. (1986). Identification of a varicella-zoster virus origin of DNA replication and its activation by herpes simplex virus type 1 gene products. J. Gen. Virol. **67**, 1613-1623.

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., Subak-Sharpe, J.H. and Wilkie, N.M. (1978). Physical mapping of herpes simplex virus type 1 mutations by marker rescue. J. Virol. **28**, 182-192.

Straus, S.E., Hay, J., Smith, H. and Owens, J. (1983). Genome differences among varicella-zoster virus isolates. J. Gen. Virol. **64**, 1031-1041.

Strnad, B.C., Adams, M.R. and Rabin, H. (1983). Glycosylation pathways of two major Epstein-Barr virus membrane antigens. Virology **127**, 168-176.

Strnad, B.C., Shuster, T., Klein, R., Hopkins III, R.F., Witmer, T., Neubauer, R.H. and Rabin, H. (1982). Production and characterization of monoclonal antibodies against the Epstein-Barr virus membrane antigen. J. Virol. **44**, 258-264.

Studdert, M.J. (1983). Restriction endonuclease DNA fingerprinting of respiratory, foetal and perinatal foal isolates of equine herpesvirus type 1. Arch. Virol. **77**, 249-258.

Studdert, M.J. and Blackney, M.H. (1979). Equine herpesviruses: on the differentiation of respiratory from foetal strains of type 1. Aust. Vet. J. **55**, 488-492.

Studdert, M.J., Simpson, T. and Roizman, B. (1981). Differentiation of respiratory and abortigenic isolates of equine herpesvirus 1 by restriction endonucleases. Science **214**, 562-564.

Sugden, B. and Warren, N. (1989). A promoter of Epstein-Barr virus that can function during latent infection can be transactivated by EBNA-1, a viral protein required for viral DNA replication during latent infection. J. Virol. **63**, 2644-2649.

Sugiura, T., Fukuzawa, Y., Kamada, M., Ando, Y. and Hirasawa, K. (1983). Isolation of equine herpesvirus type 2 from foals with pneumonitis. Bull. Equine Res. Inst. **20**, 148-153.

Sullivan, D.C., Allen, G.P. and O'Callaghan, D.J. (1989). Synthesis and processing of equine herpesvirus type 1 glycoprotein 14. Virology **173**, 638-646.

Sullivan, D.C., Atherton, S.S., Staczek, J. and O'Callaghan, D.J. (1984). Structure of the genome of equine herpesvirus type 3. Virology **132**, 352-367.

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). Characterization of the gene encoding herpes simplex virus type 2 glycoprotein C and comparison with the type 1 counterpart. J. Virol. **53**, 561-569.

Telford, E., Lankinen, H. and Marsden, H. (1990). Inhibition of equine herpesvirus type 1 subtype 1-induced ribonucleotide reductase by the nonapeptide YAGAVVNDL. J. Gen. Virol. **71**, 1373-1378.

Tenser, R.B. and Dunstan, M.E. (1979). Herpes simplex virus thymidine kinase expression in infection of the trigeminal ganglion. Virology **99**, 417-422.

Tenser, R.B., Hay, K.A. and Edris, W.A. (1989). Latency-associated transcript but not reactivatable virus is present in sensory ganglion neurons after inoculation of thymidine kinase-negative mutants of herpes simplex virus type 1. J. Virol. **63**, 2861-2865.

Turtinen, L.W. and Allen, G.P. (1982). Identification of the envelope surface glycoproteins of equine herpesvirus type 1. J. Gen. Virol. **63**, 481-485.

Turtinen, L.W., Allen, G.P., Darlington, R.W. and Bryans, J.T. (1981). Serological and molecular comparisons of several equine herpesvirus type 1 strains. Am. J. Vet. Res. **42**, 2099-2104.

van Drunen Littel-van den Hurk, S. and Babiuk, L.A. (1986). Synthesis and processing of bovine herpesvirus 1 glycoproteins. J. Virol. **59**, 401-410.

van Drunen Littel-van den Hurk, S., van den Hurk, J.V. and Babiuk, L.A. (1985). Topographical analysis of bovine herpesvirus type-1 glycoproteins : use of monoclonal antibodies to identify and characterize functional epitopes. Virology **144**, 216-227.

van Regenmortel, M.H.V. (1989). Structural and functional approaches to the study of protein antigenicity. Immunol. Today **10**, 266-272.

Vaughan, P.J., Banks, L.M., Purifoy, D.J.M. and Powell, K.L. (1984). Interactions between herpes simplex virus DNA-binding proteins. J. Gen. Virol. **65**, 2033-2041.

Vlazny, D.A. and Frenkel, N. (1981). Replication of herpes simplex virus DNA : localization of replication recognition signals within defective virus genomes. Proc. Natl. Acad. Sci. USA **78**, 742-746.

Vogelstein, B. (1987). Rapid purification of DNA from agarose gels by centrifugation through a disposable plastic column. Anal. Biochem. **150**, 115.

von Heijne, G. (1983). Patterns of amino acids near signal-sequence cleavage sites. Eur. J. Biochem. **133**, 17-21.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. **14**, 4683-4690.

Wagner, E.K. (1982). Individual HSV transcripts : characterization of specific genes. In ' The Herpesviruses', vol. 3, pp. 45-104. Edited by B. Roizman. Plenum Press, New York.

Wagner, E.K., Devi-Rao, G., Feldman, L.T., Dobson, A.T., Zhang, Y.-F., Flanagan, W.M. and Stevens, J.G. (1988). Physical characterization of the herpes simplex virus latency-associated transcript in neurons. J. Virol. **62**, 1194-1202.

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. USA **78**, 1441-1445.

Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982). Distantly related sequences in the $\alpha$ - and $\beta$ - subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1**, 945-951.

Walter, P. and Lingappa, V.R. (1986). Mechanism of protein translocation across the endoplasmic reticulum membrane. Ann. Rev. Cell Biol. **2**, 499-516.

Watari, E., Dietzschold, B., Szokan, G. and Heber-Katz, E. (1987). A synthetic peptide induces long-term protection from lethal infection with herpes simplex virus type 2. J. Exp. Med. **165**, 459-470.

Wathen, M.W. and Wathen, L.M.K. (1984). Isolation, characterisation, and physical mapping of a pseudorabies virus mutant containing antigenically altered gp50. J. Virol. **51**, 57-62.

Wathen, M.W. and Wathen, L.M.K. (1986). Characterization and mapping of a nonessential pseudorabies virus glycoprotein. J. Virol. **58**, 173-178.

Weber, P.C., Levine, M. and Glorioso, J.C. (1987). Rapid identification of nonessential genes of herpes simplex virus type 1 by Tn5 mutagenesis. Science **236**, 576-579.

Weir, H.M. and Stow, N.D. (1990). Two binding sites for the herpes simplex virus type 1 UL9 protein are required for efficient activity of the $ori_S$ replication origin. J. Gen. Virol. **71**, 1379-1385.

Weir, J.P., Bennet, M., Allen, E.M., Elkins, K.L., Martin, S. and Rouse, B.T. (1989). Recombinant vaccinia virus expressing the herpes simplex virus type-1 glycoprotein C protects mice against herpes simplex virus challenge. J. Gen. Virol. **70**, 2587-2594.

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 $\underline{ori}_L$, a herpes simplex virus type 1 origin of DNA synthesis. Mol. Cell. Biol. **5**, 930-942.

Weston, K. (1988). An enhancer element in the short unique region of human cytomegalovirus regulates the production of a group of abundant immediate early transcripts. Virology **162**, 406-416.

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 EcoRI, BglII and BamHI. J. Gen. Virol. **57**, 307-323.

Whalley, J.M., Robertson, G.R., Scott, N.A., Hudson, G.C., Bell, C.W. and Woodworth, L.M. (1989). Identification and nucleotide sequence of a gene in equine herpesvirus 1 analogous to the herpes simplex virus gene encoding the major envelope glycoprotein gB. J. Gen. Virol. **70**, 383-394.

Wharton, J.H., Henry, B.E. and O'Callaghan, D.J. (1981). Equine cytomegalovirus : cultural characteristics and properties of viral DNA. Virology **109**, 106-119.

Whealy, M.E., Robbins, A.K. and Enquist, L.W. (1989). Replacement of the pseudorabies virus glycoprotein gIII gene with its postulated homolog, the glycoprotein gC gene of herpes simplex virus type 1. J. Virol. **63**, 4055-4059.

Whitton, J.L. and Clements, J.B. (1984). 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. Nucleic Acids Res. **12**, 2061-2079.

Wigler, M., Silverstein, S., Lee, L.-S., Pellicer, A., Cheng, Y.-C. and Axel, R. (1977). Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. Cell **11**, 223-232.

Wildy, P. (1985). Herpes viruses : a background. British Med. Bull. **41**, 339-344.

Wildy, P., Russell, W.C. and Horne, R.W. (1960). The morphology of herpes virus. Virology **12**, 204-222.

Wiley, D.C. (1985). Viral Membranes. In 'Virology', pp. 45-67. Edited by B.N. Fields. Raven Press, NY.

Wu, C.A., Harper, L. and Ben-Porat, T. (1986a). Molecular basis for interference of defective interfering particles of pseudorabies virus with replication of standard virus. J. Virol. **59**, 308-317.

Wu, C.A., Harper, L. and Ben-Porat, T. (1986b). cis functions involved in replication and cleavage-encapsidation of pseudorabies virus. J. Virol. **59**, 318-327.

Wu, C.A., Nelson, N.J., McGeoch, D.J. and Challberg, M.D. (1988). Identification of herpes simplex virus type 1 genes required for origin-dependent DNA synthesis. J. Virol. **62**, 435-443.

WuDunn, D. and Spear, P.G. (1989). Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. J. Virol. **63**, 52-58.

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

Yeargan, M.R., Allen, G.P. and Bryans, J.T. (1985). Rapid subtyping of equine herpesvirus 1 with monoclonal antibodies. J. Clin. Microbiol. **21**, 694-697.

Zhang, G. and Leader, D.P. (1990). The structure of the pseudorabies virus genome at the end of the inverted repeat sequences proximal to the junction with the short unique region. J. Gen. Virol. **71**, 2433-2441.

Zuckermann, F.A., Mettenleiter, T.C., Schreurs, C., Sugg, N. and Ben-Porat, T. (1988). Complex between glycoproteins gI and gp63 of pseudorabies virus : its effect on virus replication. J. Virol. **62**, 4622-4626.

Zuckermann, F.A., Zsak, L., Mettenleiter, T.C. and Ben-Porat, T. (1990). Pseudorabies virus glycoprotein gIII is a major target antigen for murine and swine virus-specific cytotoxic T lymphocytes. J. Virol. **64**, 802-812.

Zuckermann, F., Zsak, L., Reilly, L., Sugg, N. and Ben-Porat, T. (1989). Early interactions of pseudorabies virus with host cells : functions of glycoprotein gIII. J. Virol. **63**, 3323-3329.