IDENTIFICATION AND SEQUENCING OF THE SIMIAN HERPESVIRUS SA8 HOMOLOG OF THE HUMAN HERPES SIMPLEX VIRUS TYPE 1 gE GENE

By

MIN ZHANG

Bachelor of Medicine

Shanghai Second Medical University

Shanghai, P. R. China

1985

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July,1991

IDENTIFICATION AND SEQUENCING OF THE

SIMIAN HERPESVIRUS SA8 HOMOLOG

OF THE HUMAN HERPES SIMPLEX

VIRUS TYPE 1 gE GENE

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

1398545

This thesis is dedicated to my wife, Hua,
for her encouragement, sacrifice,
and love during this study.

PREFACE

Over twenty years ago an IgG Fc-binding glycoprotein (gE) was identified in HSV1. Since then homologs of the gE gene have been identified in several other α-herpesviruses. The purpose of this study was to search for a homolog of the HSV1 gE in SA8, a simian α-herpesvirus, and to determine its relationship to the HSV1 gE gene.

I would like to express my deepest appreciation and esteem to my major advisor, Dr. R. Eberle, for his tolerated guidence, financial support, great help in this study, and excellent editorial guidence in this paper. I would also like to express my thanks to Mrs. D. Black for her excellent technical assistance and sincere cooperation.

I would like to express my gratitude to my committee members, Dr. R.W. Fulton, Dr. J. d'Offay, and Dr. E. Short for their guidence and help.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The Herpesvirus

Herpesviruses have been isolated from humans and numerous other animal species (1,2,3,4,5,6,12,13,14). The herpesvirus family is divided into three subfamilies, the α-, ß-, and γ-herpesviruses, based on their biological and pathogenic properties (7) (Table 1). Some α-herpesviruses are human herpes simplex virus types 1 and 2 (HSV1, HSV2), varicella-zoster virus (VZV), simian agent 8 (SA8), *Herpesvirus simiae* (B virus), pseudorabies virus (PRV), bovine herpesviruses 1 and 2 (BHV1, BHV2), equine herpesviruses 1 and 4 (EHV1, EHV4), and channel catfish herpesvirus (CCV) (1,2,4,5,7,8,10). The α-herpesviruses are characterized by a rapid replication cycle which usually results in lysis of the infected cell and establishment of latent infections in sensory ganglionic neurons (7,9). The ß-herpesviruses are the cytomegaloviruses (CMV) which include human CMV, *H. aotus* 1 and 3, EHV2, and BHV4 (11,12,13,14). The ß-herpesviruses are characterized by their relatively long replication cycle and establishment of latent infections in salivary glands, kidneys, lymphoreticular tissue, and other tissues (11,15). The

TABLE I

GENERAL PROPERTIES OF MEMBERS OF THE HERPESVIRUS FAMILY AFFECTING HUMANS

Subfamily	Members	Site of Latency	Typical Diseases
α	HSV1, HSV2	Sensory neurons	Oro-facial, genital lesions, ocular infections, encephalitis
	VZV	Sensory neurons	Chickenpox in children (acute infection), shingles (zoster, recurrent infection)
ß	CMV	Secretory glands, kidneys, lympho-reticular cells	Hepatitis, lymphocytosis, respiratory infections, interstitial pneumonia in organ transplant patients, mental retardation and birth defects in infants
γ	EBV	B lymphocytes	Infectious mononucleosis, nasopharyngeal carcinoma, Burkett's lymphoma

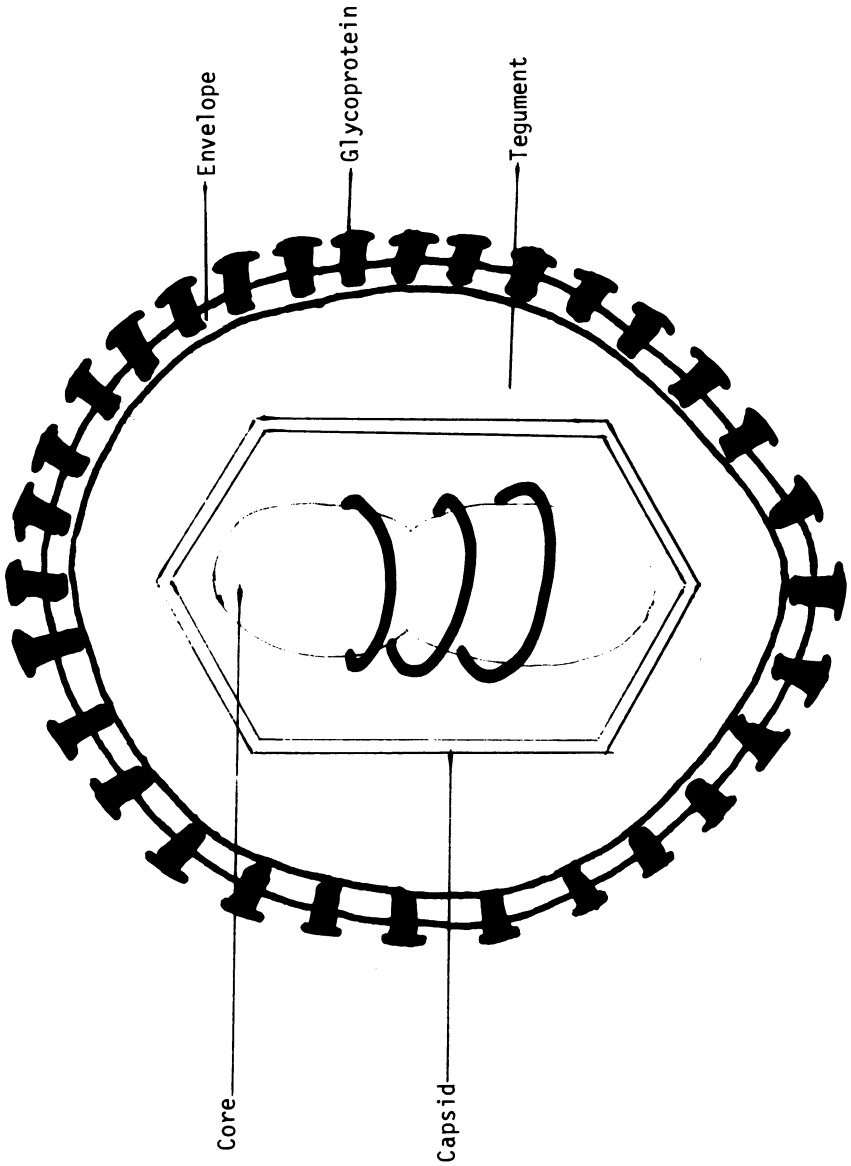
γ-herpesviruses include Epstein-Barr virus, *H. saimiri 2*, *H. ateles 2*, *H. aotus 2*, and gallid herpesvirus 2, (3,6,16,17). The γ-herpesviruses commonly cause oncogenic transformation of T or B lymphocytes and establish latent infections in lymphoid tissue (3,17).

HSV1 and HSV2 belong to the α-herpesvirus group (1,7). Both HSV1 and HSV2 are common infectious agents in human beings and usually cause asymptomatic infection or clinical oro-facial lesions (18). HSV1 is usually acquired from family members during childhood by close personal contact (1,18). HSV2 causes primarily genital infections in individuals and is acquired by venereal contact (1,19). HSV2 can also cause severe generalized disease in newborns when acquired at the time of birth by contact with the mother's infected birth canal (18).

As is typical of herpesvirus infections, HSV infection can be divided into three types: acute, latent and recurrent (18,20,21). Although many primary (acute) infections are asymptomatic, local lesions are also common (18). After primary infection, HSV can remain latent in a noninfectious state in the body for the life of its host (18,21). Since the latent infection is commonly in sensory nerve ganglia, recurrences are often caused by the reactivation of the virus due to stress, trauma or immunosuppression (20). Although HSV infections are usually asymptomatic or are clinically manifested as vesicles or

blisters on the face or other mucosal membrane of body, severe infections may also occur (18). These include ocular infections, encephalitis, and disseminated disease in newborns (18). In addition, HSV2 infection has been associated with cervical carcinoma (22).

Herpesvirus virions consist of a core structure, an icosahedral capsid, a pleomorphic tegument, and an envelope (Figure 1). The core of the virion is composed of the viral DNA and basic and/or phosphoproteins (23). The structure of the core has been suggested to be of toroidal form, but this has not been confirmed so far (57). The herpesvirus icosahedral capsid is characteristically composed of 162 capsomeres which enclose the core structure (24). The tegument is located between the envelope and nucleocapsid and appears in the electron microscopy as an amorphous structure (25). The tegument is probably involved in maintaining the shape or structure of the virion (25). Recent studies suggest that the tegument also includes some important proteins that regulate the transcription of viral DNA (28). The envelope of herpesviruses consists of a lipid bilayer derived from the host cell in which are anchored viral glycoproteins (27). The surface glycoproteins appear as spikes when virions are examined by electron microscopy (26). The number and quantity of glycoproteins vary among different herpesviruses (26,84,85). HSV encodes at least 7 glycoproteins


Envelope
Glycoprotein
Tegument
Core
Capsid

designated gB, gC, gD, gE, gG, gH and gI (15,26,58,60).

The HSV genome is a linear double-stranded DNA molecule (1). The genome is about $155x10^6$ base pairs with a 67% G+C composition (37). The genome is composed of two covalently linked parts, the large (L) region and the short (S) region (about 82% and 18% of the viral genome, respectively) (1,15,37). Both L and S consist largely of unique sequences enclosed by a pair of short inverted repeat sequences at either end (37). These repeat sequences allow the L and S regions to invert their orientations during viral DNA replication (1,15,37). Consequently, four isomers of the HSV genome are found in infected cells which differ in the orientation of unique (non-repeat) sequences within the L and/or S regions (1,12,15,37).

The earliest stages of the herpesvirus infectious cycle, attachment and penetration, are mediated by the viral glycoproteins (43,44,45,59). Attachment to the host cell occurs by the formation of complexes between viral glycoproteins and appropriate receptors on the host cell membrane (59). Penetration then occurs by fusion of the viral envelope with the host cell plasma membrane (59). In HSV1, mutation of either the gB or gD glycoproteins causes the virus to lose the ability to penetrate cells (43,44). The gB and gD glycoproteins therefore appear to be somehow involved in viral penetration into host cell.

It has also been suggested that the gC and gB glycoproteins are involved in viral attachment to the host cells (15,59).

After penetration, the nucleocapsid is transported through the cytoplasm and the viral DNA is released into the nucleus (1). Transcription of the viral DNA occurs in the nucleus, and viral proteins are synthesized in the cytoplasm (1,8). The immediate-early (IE) genes, the first genes to be expressed, are transcribed by the host enzyme RNA polymerase II (61). The IE proteins stimulate the transcription of other viral genes which encode the early (E) proteins (61,62). The IE genes are also self-regulated (63). Some IE proteins may bind to the promoters of the IE genes, thereby inhibiting their transcription (63). In addition, some E proteins are also involved in inhibiting IE gene transcription (63). Most of the viral E proteins are involved in replication of the viral DNA, such as the viral DNA polymerase and thymidine kinase (61). Other E proteins function to stimulate transcription of late (L) genes expressed after viral DNA replication (62,63). L proteins are primarily viral structural components such as capsid proteins and viral glycoproteins (62). Some L proteins are also involved in inhibiting transcription of E genes (63). Since some E and L proteins are synthesized at the relatively early times folowing infection (62,63,64,65), the designation of α, β, and γ

genes has been widely used in herpesvirus research instead of the designations of IE, E and L genes (64,65). Following nucleocapsid assembly in the nucleus, the envelope is obtained when the nucleocapsid buds through the inner nuclear membrane of the host cell (27). Mature virions accumulate in the membrane systems of the host cell, and may be released slowly by vacuolar membrane fusion and exocytosis or by lysis of the infected cells (21,27).

As a result of herpesvirus infection, both nonspecific and specific immune responses are induced against the virus and infected cells (71,66,67). The nonspecific immune response involves not only production of high levels of interferon (IFN-α, IFN-ß, IFN-γ), but activation of macrophage, neutrophile, and NK cells as well (71). Macrophage and neutrophile mediate cytotoxicity of virus-infected cells by binding the Fc region of anti-viral antibodies bound to infected cells (71,41,42). In the specific immune response, both humoral and cell-mediated immune mechanisms are involved (66,67). Both neutralizing and non-neutralizing antibodies are induced as a result of herpesvirus infection and neutralizing antibodies are directed against the viral glycoproteins (67,68,69). It has been shown that neutralizing antibodies are directed against and bind to epitopes on gB, gD, gC, and gG glycoproteins allowing antibody-dependent cellular cytotoxic (ADCC) and antibody-dependent complement-mediated cytotox-

icity (ADCMC) processes to occur (69,70,71).

The gE glycoprotein

Over twenty years ago, it was observed that HSV1 infected cells and virions could agglutinate sheep erythrocytes coated with rabbit anti-sheep erythrocyte antibodies (29). In addition, nonimmune rabbit IgG exhibited increased binding to HSV-infected over uninfected cells (29,30). Hemadsorption and IgG binding were also demonstrated to occur with cells infected by human CMV and VZV (31,32). Further study showed that the F(ab')2 fragment produced by pepsin digestion of IgG bound to infected cells much less than intact IgG, suggesting that IgG binding is associated with the Fc region (30).

The IgG Fc receptor induced in cells following HSV infection could be encoded either by the host cell or by the virus. At intervals of addition of actinomycin D to the HSV1 infected cells resulted in a relative decrease in the amount of IgG bound to cells in the infective late stage by comparing with that in the infected cells of no addition of actinomycin D (30). This further supports the idea that the Fc receptor is encoded by the virus genome (30). It has also been shown that the Fc receptor is glycosylated (33). A [^{14}C]-glucosamine-labeling polypeptide extracted from HSV1 infected cells was bound to IgG-Sepharose columns and this glycoprotein had the same

electrophoretic mobility as that of [^{35}S]-methionine-labeled IgG binding proteins (33). Since this was the fifth glycoprotein identified in HSV1-infected cells, it was designated gE (33).

Although HSV1 and HSV2 encode antigenically related Fc receptors, both vary in size based on SDS-PAGE analysis (34). Virus-specific monoclonal antibodies against the HSV1 and HSV2 gEs were used to identify which gE gene was present in various HSV1xHSV2 recombinant viruses (34,35). These studies mapped the position of the Fc receptor (gE) gene to the S region of the HSV genome (34,35).

The length of unique sequences of the S region (Us) of HSV1 is about 13 Kbp (37). The S region contains 12 genes based on mRNA mapping and DNA sequencing studies (36,37). To precisely map the gE gene, gE mRNA was isolated using gE monoclonal or monospecific antibodies to immunoprecipitate polyribosomes actively translating the gE polypeptide <u>in vitro</u> (36). By hybridization of gE mRNA to DNA subfragments of the Us region, it was determined that gE gene is located between map coordinates 0.924 to 0.951 (36). From DNA sequencing, it was further determined that the length of the gE open reading frame (ORF) is 1650 bp (37). Computer analysis of the predicted amino acid sequence of the gE gene revealed that gE has features typical of membrane proteins. These features include a signal sequence composed mostly of hydrophobic amino acids

near the N-terminus, a hydrophilic extra-cellular domain, an α-helical hydrophobic transmembrane region composed mostly of hydrophobic amino acids, and an internal cytoplasmic tail at the C-terminus. The exact three-dimensional structure has not been determined.

The function of the IgG Fc receptor (gE) in the infected cell membrane is not known completely (30). It has been suggested that it may serve to interfere with immune attack by the host, since aggregated HSV-immune rabbit IgG on HSV1 infected cells inhibits both ADCC and ADCMC activities in vitro (39). Recent studies have shown that HSV infected cells and virions may evade immune cytolysis when the Fc region of anti-viral IgG is bound by gE glycoprotein (38). In other words, the F(ab') region of immune IgG binds to its target glycoprotein (gB, gC or gD) while CH_3 portion of Fc region of the IgG is bound by the gE glycoprotein (38,40). In this way binding of C1q to CH_2 portion of Fc region of the immune IgG, the first step of activation of classical complement pathway, is inhibited due to steric restriction (38). Consequently, activation of the classical complement pathway by anti-viral IgG is greatly inhibited (38,41). The ADCC process may be similarly inhibited, since the gE glycoprotein may compete with cytotoxic cells for binding of the Fc region of anti-viral antibodies (41).

In addition to its ability to bind IgG Fc, the gE

glycoprotein may also be involved in induction of cell fusion (42). It has been demonstrated that the gB glycoprotein is responsible for penetration of the virion into host cells by inducing fusion of the virion envelope with the cell plasma membrane (43). Both gD and gH have also been suggested to have a role in the fusion process (44,45). Recently it was found that monoclonal antibodies against gE inhibit multinucleate cell formation following HSV1 infection, suggesting that gE is involved in syncytium formation (42). However, the significance of the apparent fusion activity of gE glycoprotein remains unknown at present (42).

Although gE has been identified as the Fc receptor on HSV-infected cells, another glycoprotein (gI) forms a complex with gE. Formation of this gE-gI complex is necessary for Fc receptor activity (46). Mutation of gI has been shown to greatly decrease the ability of gE to bind the Fc region of antiviral IgG <u>in vitro</u> (46,42). It is of interest that gI itself does not bind Fc region of IgG at all (46). However, the exact role of gI in the gI-gE complex has not been established so far.

Simian herpesvirus SA8

Simian herpesvirus SA8, an α-herpesvirus indigenous to African monkeys, causes asymptomatic or oro-genital infections in its natural host species (4,47). SA8 has

also been associated with stillbirths and congenital abnormalities in baboon breeding colonies (48,50). Unlike B virus, the pathogenic potential of SA8 for human beings is unknown (4,47,49). As an α-herpesvirus, SA8 has a rapid replication cycle and characteristically induces syncytium formation following infection (4). Previous serological studies have shown that SA8 is very closely related to HSV1 and HSV2 (55,56,83). DNA sequencing studies have shown that several genes of SA8 have significant homology to analogous genes of HSV1 (81).

Purpose of the study

Since SA8 is evolutionarily intermediate between HSV and other α-herpesvirus of lower mammals for which gE gene homologs have been identified and sequenced (54,53,52,51), identification and sequencing of the entire gE gene of SA8 could be of value in studies on the molecular evolution of the herpesviruses. The goals of this study are 1) to determine if SA8 has a homolog of the HSV1 gE gene; 2) to locate the position of this gene in the SA8 genome; and 3) to clone and sequence the SA8 gE gene homolog.

CHAPTER II

MATERIALS AND METHODS

Experimental Approach

The BamHI D fragment (10 Kbp) which contains most of the Us region of the SA8 genome was already cloned into plasmid pUC19 in this laboratory (designated pSBD). To map and to subclone pSBD, the plasmid was first transfected into DH5α cells and grown on agar plates (72,75). Colonies were picked, grown up in liquid culture overnight, and the plasmid DNA isolated by an alkaline lysis" miniprep" method (74). Plasmid DNAs were characterized by restriction enzyme digestion analysis on agarose gels (72). Subcloning of restriction fragments was performed by shotgun cloning or electroelution of fragments from agarose gels, ligation into pUC19, and transfection into DH5α cells (73,72). The double stranded plasmid DNA used for sequencing template was purified by a modified miniprep method or by use of QIAGEN anion-exchange columns. DNAs were sequenced using a modified dideoxynucleotide chain-termination procedure (76,77). DNA sequence data were assembled and analyzed with the aid of a computer.

Specific Methods

Plasmids and Cells

E.coli DH5α cells (BRL, Gaithersburg, MD) were used for propagation of plasmids. Plasmid pUC19 (75) and pLH1 (78) were used to clone all DNA fragments. pLH1 is the same plasmid as pUC19, differing only in the polylinker sequence. Both of these plasmids carry an ampicillin resistance (Amp^r) gene with the polylinker cloning sites located between the promoter and coding sequences of the ß-galactosidase (ß-gal) gene. Cloning of a foreign DNA fragment into these cloning sites results in inactivation of the ß-gal gene. Recombinant plasmids can therefore be selected for by Amp^r (only DH5α cells which have acquired the plasmid can grow in the presence of amplicillin) and X-gal (which when cleaved by the ß-gal enzyme produces a blue color; bacterial colonies containing plasmids with a foreign DNA fragment appear white rather than blue).

Tansformation of Plasmids into *E. coli*

Transformation-competent DH5α cells were prepared using a modified $CaCl_2$ procedure (72). DH5α cells were grown in LB media overnight and 100 ul of this culture were inoculated into 10 ml LB broth to grow up until bacteria produced an apparent absorbance of 0.2 at 600 nm. After centrifugation the pellet was resuspended in 20 ml of transformation buffer (50 mM $CaCL_2$, 15% glycerol, 10 mM

MOPS, pH 6.6). After setting on ice 20 min, cells were pelleted and resuspended in 2 ml transformation buffer. Cells were then distributed in 100 ul aliquots (about $6x10^6$ cells) into 1.5 ml microfuge tubes, immersed briefly in liquid nitrogen (about 5 sec), and stored at -70°C.

One microfuge tube of competent DH5α cells was used for each transformation. Cells were thawed at room temperature and placed on ice for 30 min. The ligated plasmid-viral DNA (about 1 ug) was added to the bacteria and held on ice for 30 min. Cells were then heat shocked at 37° C in a water bath for 2 min. 900 ul of prewarmed SOC media (2.5% LB base, 10 mM $MgSo_4$, 10 mM $MgCl_2$, 20 mM glucose, pH 7.5) were added to the tube and incubated at 37°C for about 45 mins. Aliquots of cells (150 ul) were plated on agar plates (1.5% Bacto-Agar in LB broth, 25 ug/ml ampicillin and 40 ug/ml X-gal). The plates were inverted and incubated at 37° C about 16 hrs.

Screening and Plasmid DNA Recovery

Individual white colonies were picked from plates and grown up overnight in LB media containing 25 ug/ml ampicillin at 37°C with constant shaking. Plasmid DNA was isolated by an alkaline lysis miniprep method as follows (74): 1.5 ml of overnight bacterial culture was added to a microfuge tube and centrifuged in a microfuge (10,000 rpm) for 15 sec. The supernatant was removed and the tube

refilled with another 1.5 ml of the bacterial culture. This was then centrifuged and the supernatant again removed. The pellet was resuspended in 100 ul of Solution I (50 mM glucose, 25 mM Tris, 10 mM EDTA, pH 7.5). After 5 min at room temperature, 200 ul Solution II (0.2 N NaOH, 1% SDS) was added, mixed by inversion, and put on ice for 5 min. 150 ul cold Solution III (3 M KOAc, 11.5% glacial acetic acid) was then added, mixed by inverting gently and put on ice for 5 min. The bacterial chromosomal DNA was then precipitated and removed by centrifugation at 10,000 rpm for 5 min. The plasmid DNA in the supernatant was then precipitated by adding an equal volume of isopropanol and centrifugation at 10,000 rpm for 15 min. The pellet was rinsed with 70% EtOH, dried, and resuspended in 10T/0.1E buffer (10 mM Tris, 0.1 mM EDTA, pH 8.0). Since the activity of restriction enzymes could be affected by contaminating proteins in the DNA sample, the plasmid DNA was further purified by adding 1/2 volume NH_4Ac (7.5 M), incubating on ice for 10 min, and centrifugation for 10 min at 10,000 rpm. DNA in the supernatant was then precipitated by adding 2 volumes of cold 100% EtOH. After centrifugation, the DNA pellet was dried and resuspended in 10T/0.1E or distilled water.

Restriction Enzyme Digestion

For restriction enzyme analysis of DNA, 0.5-1.0 ug

samples of plasmid DNA were digested with 0.5-2.0 units of restriction enzyme in a 10-15 ul volume for about 1-2 hr at 37°C (72). If digestion with more than one enzyme was necessary, digestion with the enzyme requiring the lowest salt buffer was performed first, followed by digestion for an additional 1 hr with enzyme requiring a higher salt buffer (72).

Gel Electrophoresis

Agarose gels, usually between 0.6-1.2%, were made by dissolving agarose in Tris-acetate buffer [40 mM Tris-acetate, 2 mM EDTA, pH 8.0; (TAE)] containing 10 ug/ml EtBr by heating in a microwave oven (72). Prior to running on the gels, contaminating RNA in DNA samples was removed by addition of 0.5 ug RNase A together with tracking dye/glycerol to 2% final concentration. Samples were then run by electrophoresis in TAE buffer containing 10 ug/ml EtBr at 60-110 volts, depending on the size of the DNA fragments. Phage DNA cut by EcoRI and by HindIII+EcoRI was used as a size marker. The size of restriction fragments was determined by visualization of DNA under ultraviolet light and determination of their mobility relative to the phage marker DNA.

Subcloning of Restriction Fragments

Restriction enzyme-digested DNA samples from which a

fragment was to be electroeluted were run on inhibitor-free agarose gels (molecular biology grade, BioRad Laboratories; Richarmond, CA) (72). After electrophoresis to separate DNA fragments, a gel plug containing the desired fragment was cut out of the gel and inserted into a dialysis bag (72) with TAE buffer (without EtBr) and electrophoresed in TAE buffer (also without EtBr) for about 45 min at 100 volts. The polarity was then reversed for about 10 sec, the gel removed from the dialysis tubing, and the buffer containing the DNA fragment collected. Plasmid DNA which had been previously digested with the appropriate restriction enzyme was added to the electroeluted DNA fragment, both DNAs were co-precipitated by NH_4Ac/EtOH precipitation as described above, and the pellet resuspended in 17.5 ul distilled water. Where restriction fragments were shotgun cloned, the parent plasmid DNA and vector DNA were mixed and digested with 0.5-2.0 units restriction enzyme for 2 hr at 37°C followed by NH_4Ac/EtOH precipitation, and the DNA resuspended in 17.5 ul distilled water.

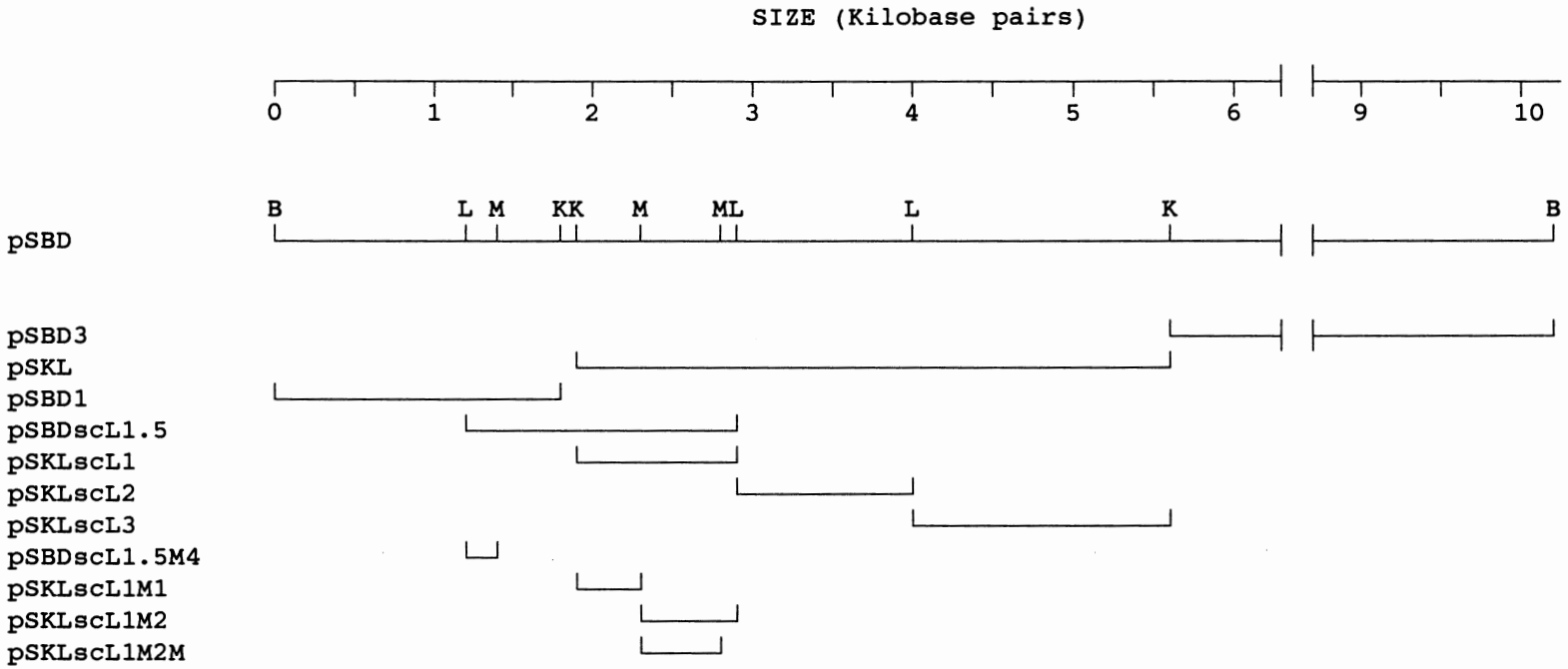
For ligation of inserts into plasmid DNA, 2 ul 10x ligase buffer and 0.5 ul (1.5 U) T4 ligase (Promega Biotech, Madison, WI) were added, mixed, and put at 4°C overnight. The following day the DNA ligation mixture was transfected into DH5α cells and plated on Amp-Xgal agar as described above. Colonies were picked and grown up in 3.5

ml LB media with 25 ug/ml ampicillin. Plasmid DNA was isolated by the miniprep method described above and checked with the appropriate restriction enzymes to confirm that the plasmid contained a single copy of the desired restriction fragment.

DH5α cells carrying plasmids were prepared for long term storage as follows: 700 ul of an overnight liquid culture were transferred to a sterile microfuge tube and 300 ul of sterile 50% glycerol added. These were mixed and stored at -70°C. All cell lines were frozen in triplicate. All the clones and subclones produced and used in this study are shown in Figure 2.

DNA Sequence Analysis

Double stranded template DNA was isolated from 20 ml overnight cultures using the miniprep method described above. DNA was then further purified by treatment with RNase A for 10 min at 37°C and precipitated with NH_4Ac /EtOH. Double stranded plasmid DNA templates were also purified from 100 ml overnight cultures by passing the product of an alkaline lysis procedure over a QIAGEN anion-exchange column (QIAGEN Inc., Chatsworth, CA). The DNA was then recovered by EtOH precipitation. About 4 ug template DNA was denatured in 0.4 M NaOH/4 mM EDTA for 5 min, neutralized by adding NH_4Ac to a final concentration of 2.5 M, and precipitated with 2 volumes of EtOH. The


SIZE (Kilobase pairs)
0
1
2
3
4
5
6
9
10
B
L
M
KK
M
ML
L
K
B
pSBD
pSBD3
pSKL
pSBD1
pSBDscL1.5
pSKLscL1
pSKLscL2
pSKLscL3
pSBDscL1.5M4
pSKLscL1M1
pSKLscL1M2
pSKLscL1M2M

DNA was rinsed with 70% EtOH and dried.

Sequencing gels (18x40 cm) of 6.0%-7.5% acrylamide containing 9 M urea were cross-linked with bis using ammonium persulfate (25 mg/ml) and TEMED. The gel solution was added to the gel unit with a syringe while held at a 45^{o} angle. The gel was allowed to polymerize with the unit laying flat for at least 1 hr. Gels were pre-electrophoresed for 1 hr at 1600 volts. All gels were run in Tris-borate buffer (working solution: 89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA, pH 8.3).

The high G+C content of SA8 DNA presented problems in sequencing the gE gene. A high G+C content often causes problems due to formation of secondary structures in the template DNA. This can result in poor annealling of primers with the template DNA, band compression on gels, and lack of fidelity of sequence product in these regions. Since a high reaction temperature will help prevent secondary structure formation, a modified dideoxynucleotide chain termination reaction was used (79,77). First, the Taq polymerase was used for DNA sequencing. The Taq DNA polymerase isolated from *Thermus aquaticus* is thermostable up to 95^{o}C, allowing sequencing to be carried out at 56-60^{o}C rather than the usually 37^{o}C. Second, 7-deaza dGTP was used in place of dGTP in all sequencing reaction to help alleviate band compression caused by self annealing of product DNAs, particularly when the ends have a high

G+C content.

Universal pUC/M13 forward and reverse primers (24 mers) were purchased from Promega Biotech (Madison, WI). Synthetic primers (24 mers) complementary to SA8 sequences were purchased from National Biosciences, Inc. (Hamel, MN). Primers were annealed with template DNA at 56°C for 10 min in 5x sequencing buffer (500 mM Tris-HCl, 100 mM $MgCl_2$, 50 mM DTT, 1.0 mM spermidine, pH 7.5) and 7.5 uM each of dGTP,dTTP, and dCTP. 2 ul α-[^{32}P]-dATP (20 uCi) and 1.5 ul Taq polymerase (0.75 u) were added and DNA synthesis extension carried out at 56°C for about 5 min. Aliquots of 1 ul of this extension mixture were then transferred to four microfuge tubes each containing one of the four different deoxy/dideoxy nucleotide mixtures. The termination reaction was then allowed to proceed for 15 min at 65°C. In some instances, reaction products were tailed using terminal deoxynucleotide transferase for 30 min at 37°C. All reactions were stopped by addition of stop solution (10 mM NaOH, 95% formamide, 0.05% bromophenol blue, 0.05% xylene cyanole). Samples were heated at 100°C for 5 min immediately before loading on sequencing gels. Samples were usually loaded at multiple times to obtain maximal sequence information from each sequencing reaction. Sequencing gels were run at a constant voltage of 1600 volts. After electrophoresis, gels were transferred to Whatman 3 mm filter paper and dried at 80°C

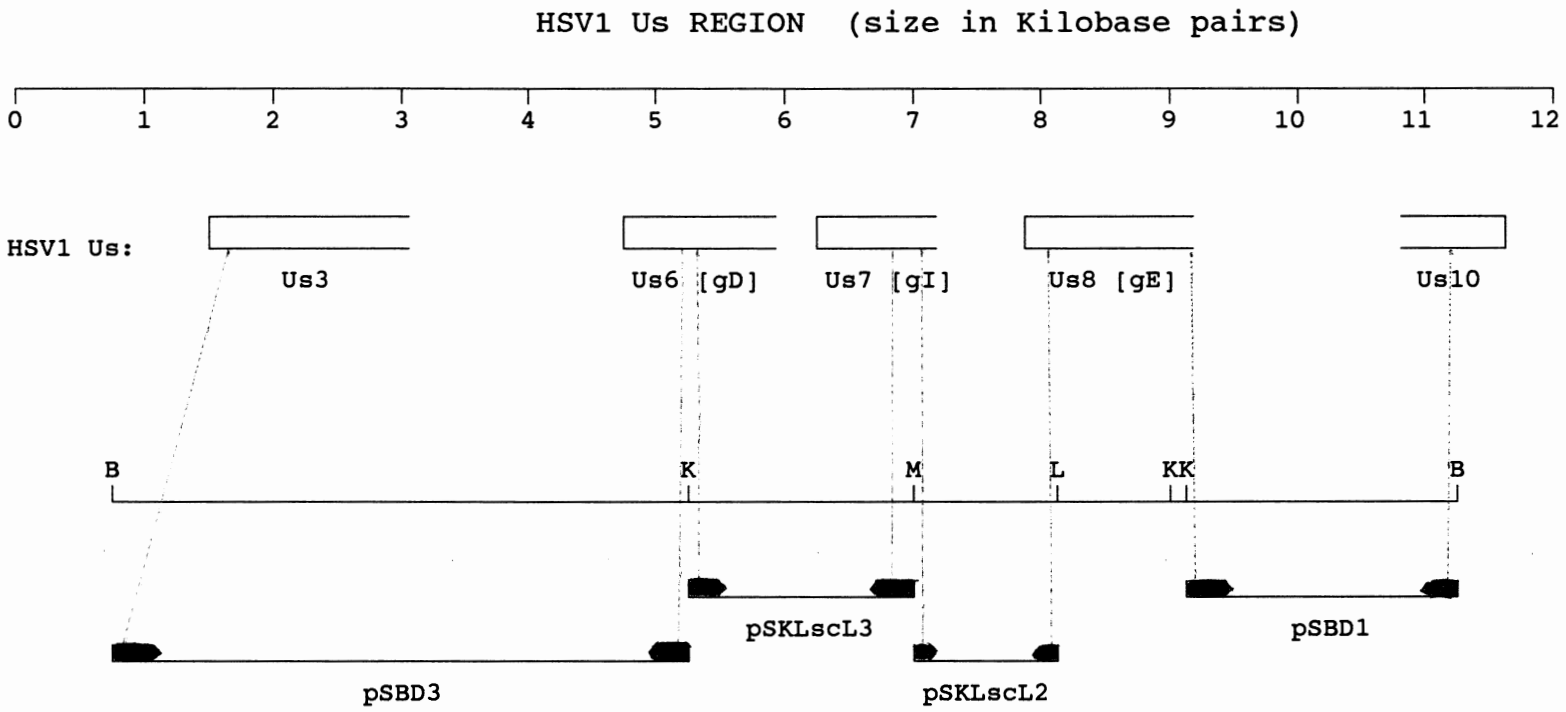
under vacuum for 30 min. DNA sequencing products were detected by autoradiography on Kodak XAR film for 1-18 hrs.

DNA sequence data were recorded and assembled on a PC using the IBI Pustell DNA analysis programs. These programs were also used for DNA homology searches (80). The UWGCG programs running on a VAX 6320 were used to predict and analyze amino acid sequences. The MULTALIN program of Corpet (81) was used for multiple sequence alignment and hierarchical cluster analysis.

CHAPTER III

RESULTS

The BamHI D fragment of SA8 viral DNA had previously been mapped to the short region of the genome and was presumed to comprise most of the Us region of SA8. A number of subclones of this 10 Kbp fragment were made using KpnI, SalI, SmaI/XmaI, and EcoRI (Figure 2). End sequencing of a number of these cloned fragments yielded several hundred base pairs of DNA sequence each. A computer search was made for homology with sequences in the Us region of HSV1 using the IBI Pustell DNA analysis programs. Several regions of the SA8 BamHI D fragment were found to have good homology with genes in the HSV1 Us region. These regions of homology were located in the US3, US6(gD), US7(gI), US8 (gE) and US10 genes of HSV1. Comparison of the distances between these homologous sequences in HSV1 (based on sequences data; 37) and SA8 (based on size estimates of cloned restriction fragments as determined by agarose gel electrophoresis) indicated that distances between the regions of homology were similar in the two viruses (Figure 3). These results imply that 1) the order and orientation of these five genes are the same in SA8 as in HSV1, and 2) no major deletions or insertions are


HSV1 Us REGION (size in Kilobase pairs)
0
1
2
3
4
5
6
7
8
9
10
11
12
HSV1 Us:
Us3
Us6 [gD]
Us7 [gI]
Us8 [gE]
Us10
B
K
M
L
KK
B
pSKLscL3
pSBD1
pSBD3
pSKLscL2

present in the SA8 Us region relative to the HSV1 Us region.

Since the SA8 KpnI L fragment (pSKL) and the adjacent BamHI-KpnI fragment (pSBD1) both had good homology with the gE gene of HSV1, a number of subclones were made of these two fragments (Figure 2). Subclones spanning the two adjacent KpnI sites within the gE gene were completely sequenced on both strands utilizing both universal pUC/M13 primers and synthetic primers. The entire gE sequence of SA8 was assembled based on overlapping sequences and, to a lesser extent, on alignment with the HSV1 gE sequence. The complete DNA sequence of SA8 gE ORF and flanking non-coding sequences is shown in Figure 4. The SA8 gE ORF was 1602 nucleotides in length and had a base composition of 76.2% G+C. This compares to the 1650 nucleotide gE ORF of HSV1 which has a 66.6% G+C (37). The predicted amino acid sequence of the SA8 gE gene product is 534 residues in length as compared to the 550 amino acid HSV1 gE polypeptide.

Alignment of the SA8 gE ORF sequence with the HSV1 gE ORF indicated that the DNA sequence had 58% homology. Figure 5 shows the predicted SA8 gE amino acid sequence aligned with the HSV1 gE sequence. It was necessary to introduce gaps into SA8 and HSV1 sequences in order to obtain maximal alignment. The introduced gaps occur mostly in the N- and C-terminal regions of the polypep-

```
        10         20         30         40         50         60         70
    *    *    *    *    *    *    *    *    *    *    *    *    *    *    *
CGGAACGCGCGCGTAGGCAGGCTGCGCGGGCCTATTAAGGCCGCGGCCGGCCGGCGTCTTCGCGTTCCGAACGCCCGGT

          90        100        110        120        130        140        150
*    *    *    *    *    *    *    *    *    *    *    *    *    *    *      *
GGATCGTTCCCCGCTCCGGTCGCTAGCGACGCGCCCGTGGCCAGGAAGAGGGGAGACGACGAGAAGCGGGG ATG GCT
                                                                          Met Ala

  160          170          180          190          200          210
    *     *      *      *     *      *     *     *      *      *     *      *
CTC GCC CGG GCC CCG CGG GGG CTC CTG GCG GCG TGG ATT CTC GCG GCG TGG GTC GGC GTC
Leu Ala Arg Ala Pro Arg Gly Leu Leu Ala Ala Trp Ile Leu Ala Ala Trp Val Gly Val

  220          230          240          250          260          270
    *     *      *      *     *      *     *     *      *      *     *      *
GCG GCC GTG GAG ACG ACG TGG AAG CAC GCG AGC GCC GGC GAC GAC GTG GTG TTC TTC GTC
Ala Ala Val Glu Thr Thr Trp Lys His Ala Ser Ala Gly Asp Asp Val Val Phe Phe Val

  280          290          300          310          320          330
    *     *      *      *     *      *     *     *      *      *     *      *
CTT CCC GCG GGG CGC CCG GGC GGA CCC CCG CGC GAG CTG GCG TGG GAG TTC GCT TCT ATG
Leu Pro Ala Gly Arg Pro Gly Gly Pro Pro Arg Glu Leu Ala Trp Glu Phe Ala Ser Met

  340          350          360          370          380          390
    *     *      *      *     *      *     *      *     *      *     *      *
CGG AAC TGC GGG CCG CTG CGG CCG TCG TGG GTC TCC CTC CAC CCC CCC GGG CAG GTG CTG
Arg Asn Cys Gly Pro Leu Arg Pro Ser Trp Val Ser Leu His Pro Pro Gly Gln Val Leu
```

```
  400          410          420          430          440          450
    *    *       *    *       *    *       *    *       *    *       *
GAG ACC GTC GTC GAC GCC CAG TGC GTC GGC GCC CCC CCG GTT CAT GGC GGC GTG GTA CGG
Glu Thr Val Val Asp Ala Gln Cys Val Gly Ala Pro Pro Val His Gly Gly Val Val Arg

  460          470          480          490          500          510
    *    *       *    *       *    *       *    *       *    *       *
CCG ACC CGA CGG GGC CCC CGC CCC GGG CCC CGA GGC GGC GTG GCC CCC GCG CGT GAC GTG
Pro Thr Arg Arg Gly Pro Arg Pro Gly Pro Arg Gly Gly Val Ala Pro Ala Arg Asp Val

  520          530          540          550          560          570
    *    *       *    *       *    *       *    *       *    *       *
AGC AAC GGG ACG CTG ACG CTC CGC GAG GCC CGA CCG AGC GAC AGC GGG ATG TAC GTC CTG
Ser Asn Gly Thr Leu Thr Leu Arg Glu Ala Arg Pro Ser Asp Ser Gly Met Tyr Val Leu

  580          590          600          610          620          630
    *    *       *    *       *    *       *    *       *    *       *
ACG GTG TCG CGC GCC CCC AAC TCC ACG GCG GCC CGT CGG GTG GTG TTC CTG ACC GTC GGG
Thr Val Ser Arg Ala Pro Asn Ser Thr Ala Ala Arg Arg Val Val Phe Leu Thr Val Gly

  640          650          660          670          680          690
    *    *       *    *       *    *       *    *       *    *       *
CCG CGG GTC GCG GCG GCG GTC CCC GGC GGC CCG CCC CCT CTC GCG GAG GGG GCG GGA GCC
Pro Arg Val Ala Ala Ala Val Pro Gly Gly Pro Pro Pro Leu Ala Glu Gly Ala Gly Ala

  700          710          720          730          740          750
    *    *       *    *       *    *       *    *       *    *       *
GAG GCC GGG GCC GCC GCG ACC CGG GCC CCC GCC GCC CAC CCC TAC CCC CAC CCC CAC CCC
Glu Ala Gly Ala Ala Ala Thr Arg Ala Pro Ala Ala His Pro Tyr Pro His Pro His Pro
```

```
  760          770          780          790          800          810
    *     *      *      *     *      *      *     *      *      *     *      *
ATC GCC GAG GTG GCG CAC GTG CAC GGC GTC ACG GTC TCC CTG CGC ACC CAG ACG GCG ATC
Ile Ala Glu Val Ala His Val His Gly Val Thr Val Ser Leu Arg Thr Gln Thr Ala Ile

  820          830          840          850          860          870
    *     *      *      *     *      *      *     *      *      *     *      *
CTG TTT AGC CCG GGG GAC ACC GTC CAC ACC GCG GTC TCC ATC GTG CCG TTC GCC CAC GAC
Leu Phe Ser Pro Gly Asp Thr Val His Thr Ala Val Ser Ile Val Pro Phe Ala His Asp

  880          890          900          910          920          930
    *     *      *      *     *      *      *     *      *      *     *      *
GAC GAC CCC TAC GTC ATG GAG GTG GTC TGG GTG CGG TTC GAC GTC CCC GAG GAG TGC GGG
Asp Asp Pro Tyr Val Met Glu Val Val Trp Val Arg Phe Asp Val Pro Glu Glu Cys Gly

  940          950          960          970          980          990
    *     *      *      *     *      *      *     *      *      *     *      *
GAG ATG CGG ATC TAC GAG CCC TGC CTG TAC CAC CCG CGG CTG CCC GAG TGC CGC TCG CCC
Glu Met Arg Ile Tyr Glu Pro Cys Leu Tyr His Pro Arg Leu Pro Glu Cys Arg Ser Pro

 1000         1010         1020         1030         1040         1050
    *     *      *      *     *      *      *     *      *      *     *      *
GCC GAC GCC CCG TGC GCG GCA AGC GTC TGG ACG GAG CGC CTG GCA GTG CGG CGG TAC GGG
Ala Asp Ala Pro Cys Ala Ala Ser Val Trp Thr Glu Arg Leu Ala Val Arg Arg Tyr Gly

 1060         1070         1080         1090         1100         1110
    *     *      *      *     *      *      *     *      *      *     *      *
CCC TGC TCG CGC CAC GTG CCG CCC CCC CGG TGC CCC ACG GAC GCC GCC ATG GAG GCC AGG
Pro Cys Ser Arg His Val Pro Pro Pro Arg Cys Pro Thr Asp Ala Ala Met Glu Ala Arg
```

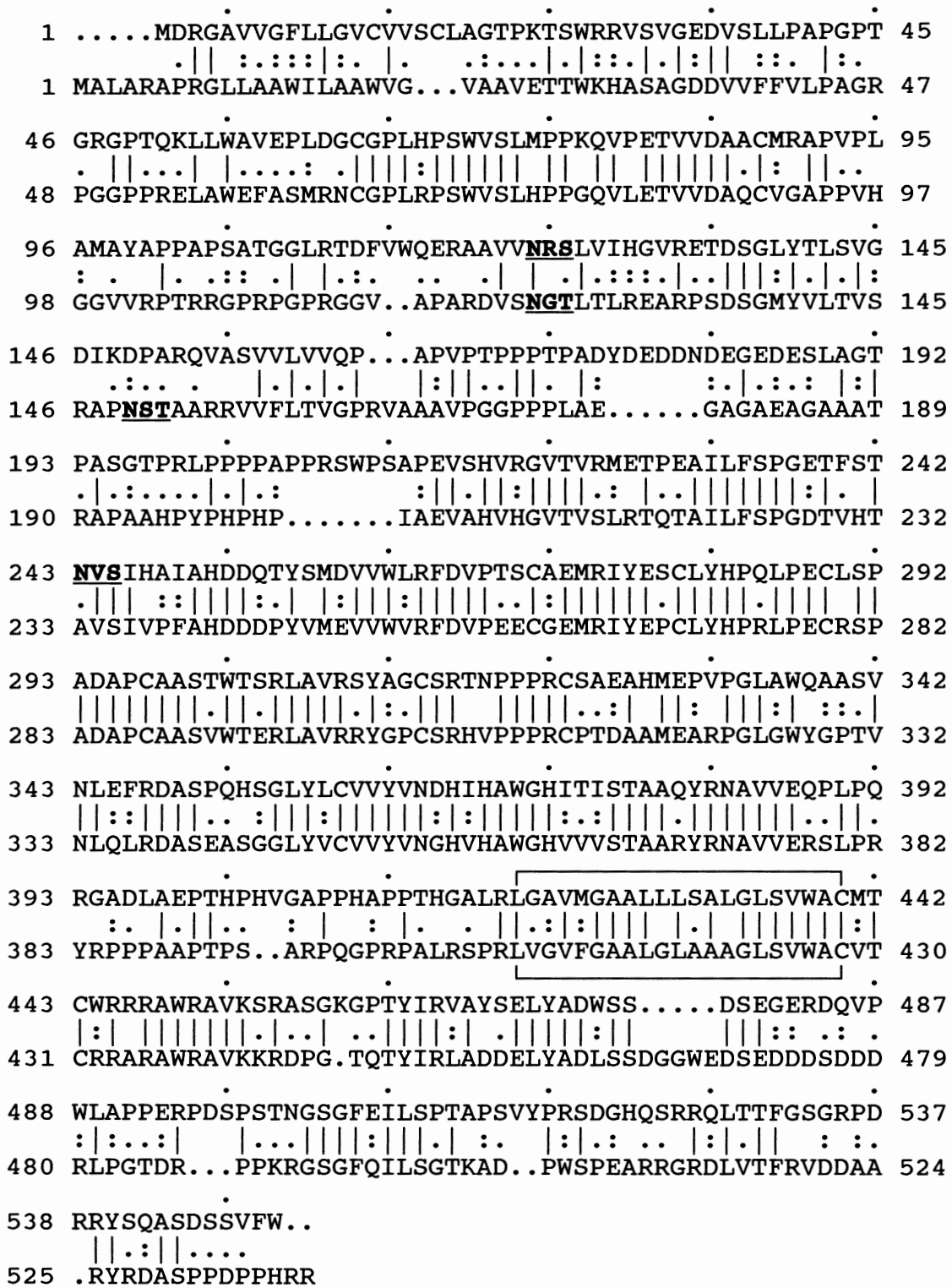
```
 1120         1130         1140         1150         1160         1170
    *    *       *    *    *    *       *    *       *    *    *    *
CCC GGC CTG GGG TGG TAC GGG CCC ACC GTT AAC CTG CAG CTC CGC GAC GCC TCC GAG GCG
Pro Gly Leu Gly Trp Tyr Gly Pro Thr Val Asn Leu Gln Leu Arg Asp Ala Ser Glu Ala

 1180         1190         1200         1210         1220         1230
    *    *       *    *    *    *       *    *       *    *    *    *
AGC GGC GGG CTC TAC GTG TGC GTC GTG TAC GTC AAC GGG CAC GTG CAC GCC TGG GGC CAC
Ser Gly Gly Leu Tyr Val Cys Val Val Tyr Val Asn Gly His Val His Ala Trp Gly His

 1240         1250         1260         1270         1280         1290
    *    *       *    *    *    *       *    *       *    *    *    *
GTC GTC GTC AGC ACC GCG GCG AGG TAC CGG AAC GCG GTC GTG GAG CGG TCC CTG CCG CGG
Val Val Val Ser Thr Ala Ala Arg Tyr Arg Asn Ala Val Val Glu Arg Ser Leu Pro Arg

 1300         1310         1320         1330         1340         1350
    *    *       *    *    *    *       *    *       *    *    *    *
TAC CGC CCC CCT CCG GCC GCA CCC ACC CCT TCG GCG CGG CCC CAG GGG CCG CCC GCG CTC
Tyr Arg Pro Pro Pro Ala Ala Pro Thr Pro Ser Ala Arg Pro Gln Gly Pro Pro Ala Leu

 1360         1370         1380         1390         1400         1410
    *    *       *    *    *    *       *    *       *    *    *    *
CGC TCG CCG CGC CTC GTG GGC GTC TTC GGC GCG GCC CTG GGG CTG GCC GCC GCC GGC CTC
Arg Ser Pro Arg Leu Val Gly Val Phe Gly Ala Ala Leu Gly Leu Ala Ala Ala Gly Leu

 1420         1430         1440         1450         1460         1470
    *    *       *    *    *    *       *    *       *    *    *    *
TCC GTG TGG GCC TGC GTC ACC TGC CGG CGC GCG CGG GCG TGG CGA GCC GTT AAA AAG CGG
Ser Val Trp Ala Cys Val Thr Cys Arg Arg Ala Arg Ala Trp Arg Ala Val Lys Lys Arg
```

```
 1480          1490          1500          1510          1520          1530
    *     *       *     *       *     *     *       *     *       *     *
GAC CCG GGA ACG CAG ACT TAC ATT CGC CTG GCG GAC GAC GAG CTC TAC GCC GAC CTG AGC
Asp Pro Gly Thr Gln Thr Tyr Ile Arg Leu Ala Asp Asp Glu Leu Tyr Ala Asp Leu Ser

 1540          1550          1560          1570          1580          1590
    *     *       *     *     *       *     *     *       *     *     *       *
TCC GAC GGC GGC TGG GAG GAC TCC GAG GAC GAC GAC TCC GAC GAC GAC CGC CTG CCG GGA
Ser Asp Gly Gly Trp Glu Asp Ser Glu Asp Asp Asp Ser Asp Asp Asp Arg Leu Pro Gly

 1600          1610          1620          1630          1640          1650
    *     *       *     *     *       *     *     *       *     *     *       *
ACG GAC CGG CCT CCG AAG CGG GGC TCC GGG TTC CAG ATC CTC TCC GGG ACG AAG GCG GAC
Thr Asp Arg Pro Pro Lys Arg Gly Ser Gly Phe Gln Ile Leu Ser Gly Thr Lys Ala Asp

 1660          1670          1680          1690          1700          1710
    *     *       *     *     *       *     *     *       *     *     *       *
CCG TGG TCC CCC GAG GCG CGC CGG GGC CGC GAC CTC GTC ACC TTC CGC GTG GAT GAC GCG
Pro Trp Ser Pro Glu Ala Arg Arg Gly Arg Asp Leu Val Thr Phe Arg Val Asp Asp Ala

 1720          1730          1740          1750          1760          1770
    *     *       *     *     *       *     *     *       *     *     *     *
GCC AGA TAT CGC GAC GCC TCG CCC CCC GAT CCA CCG CAC CGG CGC TGA CGCCCGTGCGCGGGC
Ala Arg Tyr Arg Asp Ala Ser Pro Pro Asp Pro Pro His Arg Arg  —

       1790      1800      1810      1820      1830      1840      1850
*    *    *    *    *    *    *    *    *    *    *    *    *    *    *
CCGAGGCGCCCGACGTTCCCATCGGCGTCGCCGAAGACCTGCGGGCCCTCGCAAGTACGTGCCGCGCCGCGCGCGCT

        1870      1880      1890      1900
 *    *    *    *    *    *    *    *    *
CCGCCGTCGGCCCCGCCGCCCCTCCCCCTGGAGGCGGTCGAC
```

tides. Based on this alignment, the SA8 and HSV1 gE amino acid sequences are 49.1% identical (based on the percent of positions where the same amino acid occurs in the same position in both sequences). When conservative amino acid substitutions are taken in consideration, the SA8 and HSV1 gE sequence are 64.7% identical and/or similar.

Several parameters of the predicted secondary structure of SA8 and HSV1 gE polypeptides generated using UWGCG programs PEPTIDESTRUCTURE\PLOTSTRUCTURE are shown in Figure 6. Plots are shown for the hydrophilicity and the occurrence of α-helices, ß-sheets, and ß-turns in both polypeptides. Prediction of secondary structural features for the SA8 gE indicate that the polypeptide has a number of properties typical of membrane proteins. These include a short strongly hydrophobic region at the N-terminus, a long hydrophilic region which has 2 potential sites for addition of N-linked carbohydrate residues, a hydrophobic region of about 20-25 amino acids near the C-terminus which could serve as a transmembrane anchor sequence, and a charged hydrophilic C-terminus.

The N-terminal hydrophobic region has a stretch of 21 hydrophobic amino acids which probably serves as a membrane insertion signal. This hydrophobic core region is preceded by a positively charged residue (Arg_8) and is terminated by another positively charged residue (Lys_{30}). Two potential sites for cleavage of the signal peptide



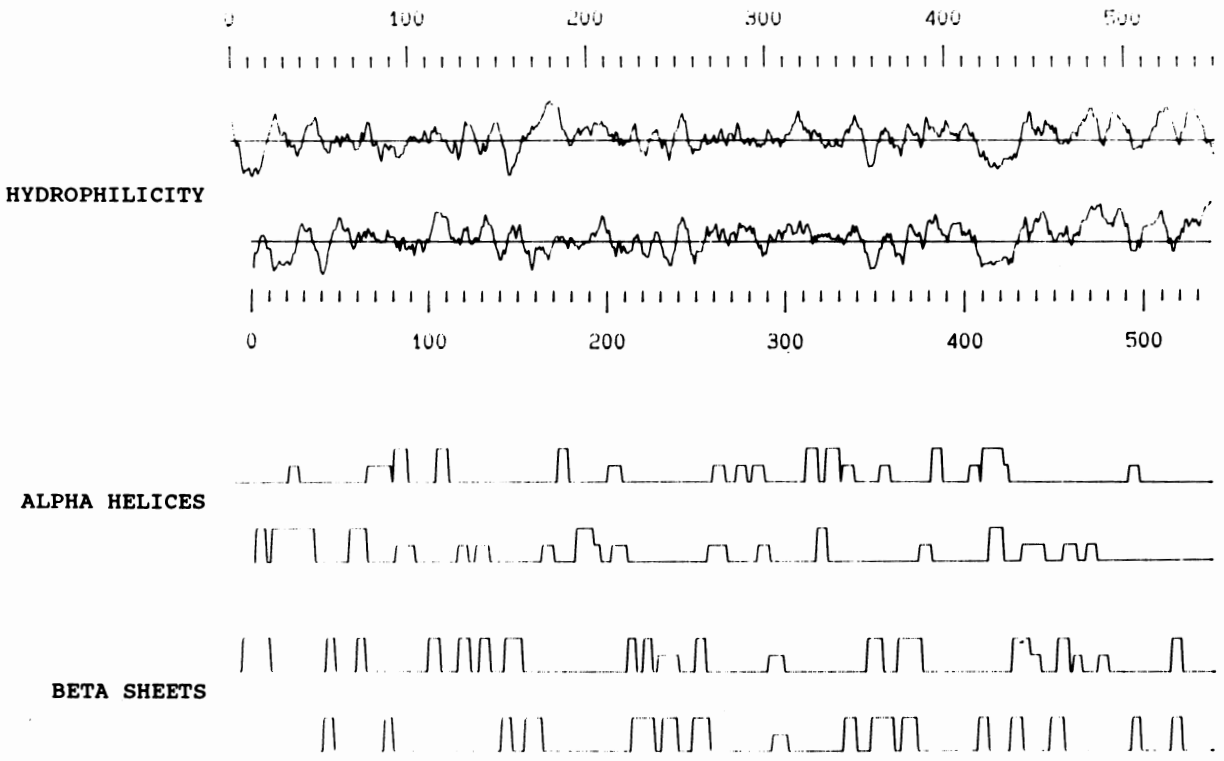

HYDROPHILICITY
0
100
200
300
400
500
ALPHA HELICES
BETA SHEETS

Figure 6

occur following the hydrophobic core sequence. One site is located at Ala_{32}-Ser_{33}-Ala_{34} and the other at Leu_{43}-Pro_{44}-Ala_{45}. The first potential cleavage site at Ala_{34} is four residues from the end of the hydrophobic core sequence. The second potential site at Ala_{45} is situated within a predicted ß-turn structure resulting from the presence of Pro residues at positions 44 and 48. The peptide Ala-X-Ala is the most frequent sequence for potential cleavage and an Ala at position +4 is one of the most frequent residues for peptidase cleavage (82). The presence of a Pro in the second potential site mostly reduces the probability of peptidase cleavage (82). Therefore the sequence Ala_{32}-Ser_{33}-Ala_{34} is favored as the probably site for cleavage of the SA8 gE signal peptide. This is similar to that of HSV1 gE in which the leader sequence is predicted to be cleaved at C-terminal of the Val_{32} residue (37).

The external hydrophilic region extends from SA8 gE amino acids 30 to 401. The C-terminal part of this region has strong homology with the HSV1 gE sequence (from HSV1 residues 214 to 414; see Figure 5). In this region the two polypeptides are 63% identical to each other. Introduction of only 1 gap was necessary to achieve maximal alignment of the sequences in this region. Of the seven Cys residues in this region of HSV1 gE, all are conserved in SA8. In addition, predicted secondary structural

features such as α-helices and ß-sheets are also very similar in this region of the two polypeptides (Figure 6). Two potential N-linked glycosylation sites (Asn-Xxx-Ser/Thr) are present in the gE polypeptides of both SA8 and HSV1. One of these is conserved while the other is located at a different but close site in the two polypeptides. These results imply that a substantial part of the extracellular domain of the SA8 gE is structurally very similar to that of the HSV1 gE polypeptide.

The region from amino acids 407-428 of the SA8 gE polypeptide is composed primarily of hydrophobic nonpolar amino acids and has good homology with the corresponding region of the HSV1 gE polypeptide (from amino acids 420-440). The predicted strong α-helical structure in this region of both viral gE polypeptides (Figure 6) supports the idea that this region of the SA8 gE polypeptide represent the transmembrane domain which may also serves to anchor the cell membrane.

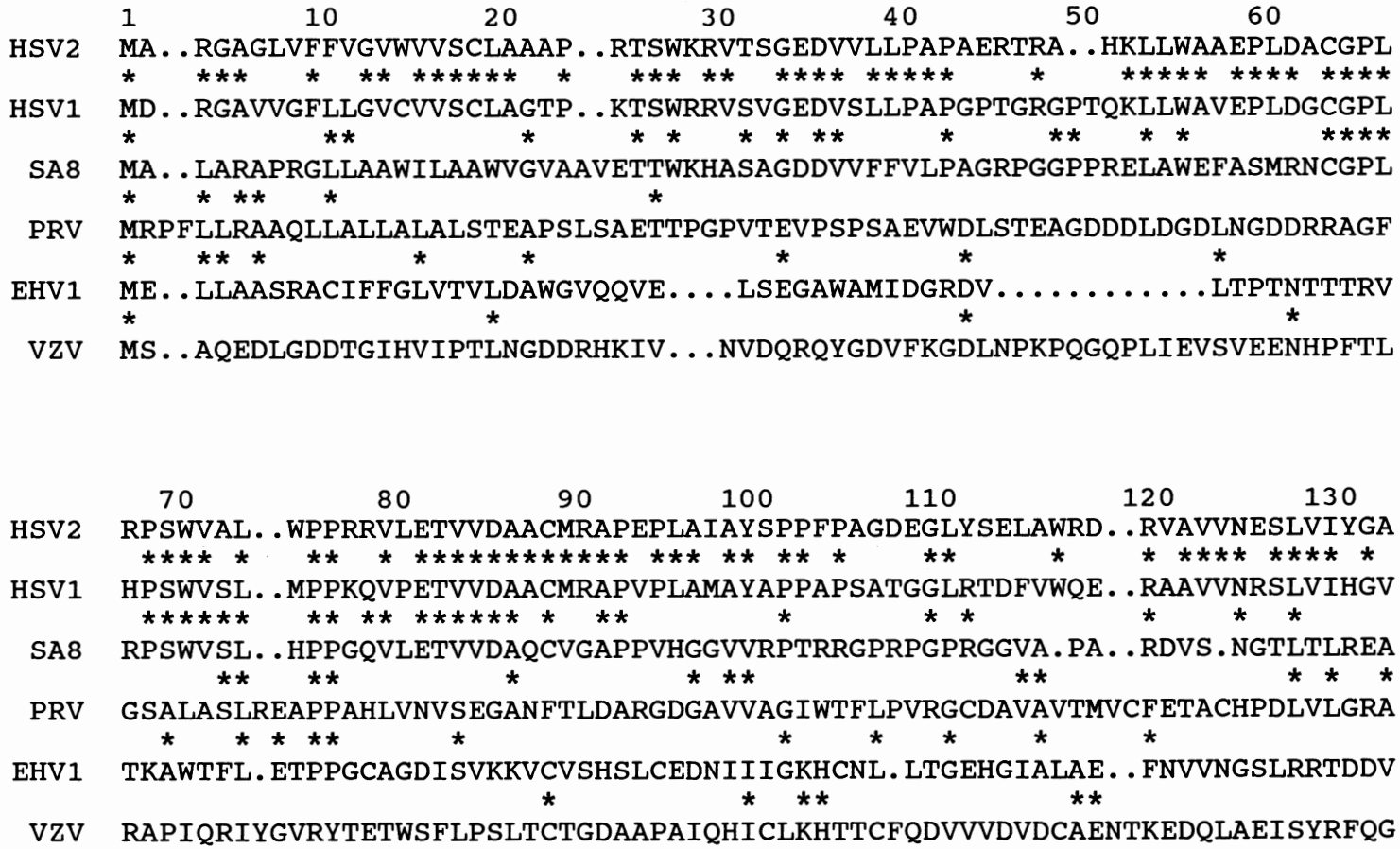
The HSV1 and SA8 gE C-terminii are predicted to extend from the cell membrane into the cell cytoplasm. Both the HSV1 and SA8 gE cytoplasmic tails are highly charged and hydrophilic. Both are also the same size (110 amino acids). However, there is limited sequence homology between the two peptides in this region, particulary at the extreme C-terminus. These results imply that this region may be 1) unimportant (non-functional), 2) involved

in altogether different functions in the two viruses, or 3) serves a common function in the two viruses but does so in some virus-specific manner (such as interaction with non-conserved viral proteins).

Since the SA8 gE amino acid sequence has significant homology with its HSV1 homolog, a multiple alignment of the SA8 gE amino acid sequence with that of other α-herpesvirus (PRV, VZV, EHV1, and part of EHV4, and HSV2) gE amino acid sequences (51,52,53,54) was made (Figure 7). The published VZV gE ORF was about 120 amino acid residues longer than all the other gE sequences which presented problems in performing multiple alignments. Since an internal Met residue at position of 120 of the VZV gE could possibly serve as an alternative initiation site, the N-terminal 119 amino acids were removed from the VZV gE sequence for alignment.

The multiple alignment of the seven gE sequences revealed that the region of the extracellular domain proximal to the transmembrane region was fairly conserved among all of these viruses. That the seven Cys residues in this region are all conserved in the SA8 and HSV1 gEs has already been described. Six of these seven Cys residues are also conserved in PRV, VZV, EHV1. These results further demonstrate the conserved structure of this region of the gE polypeptide.

The interrelationship among the gE polypeptides of all

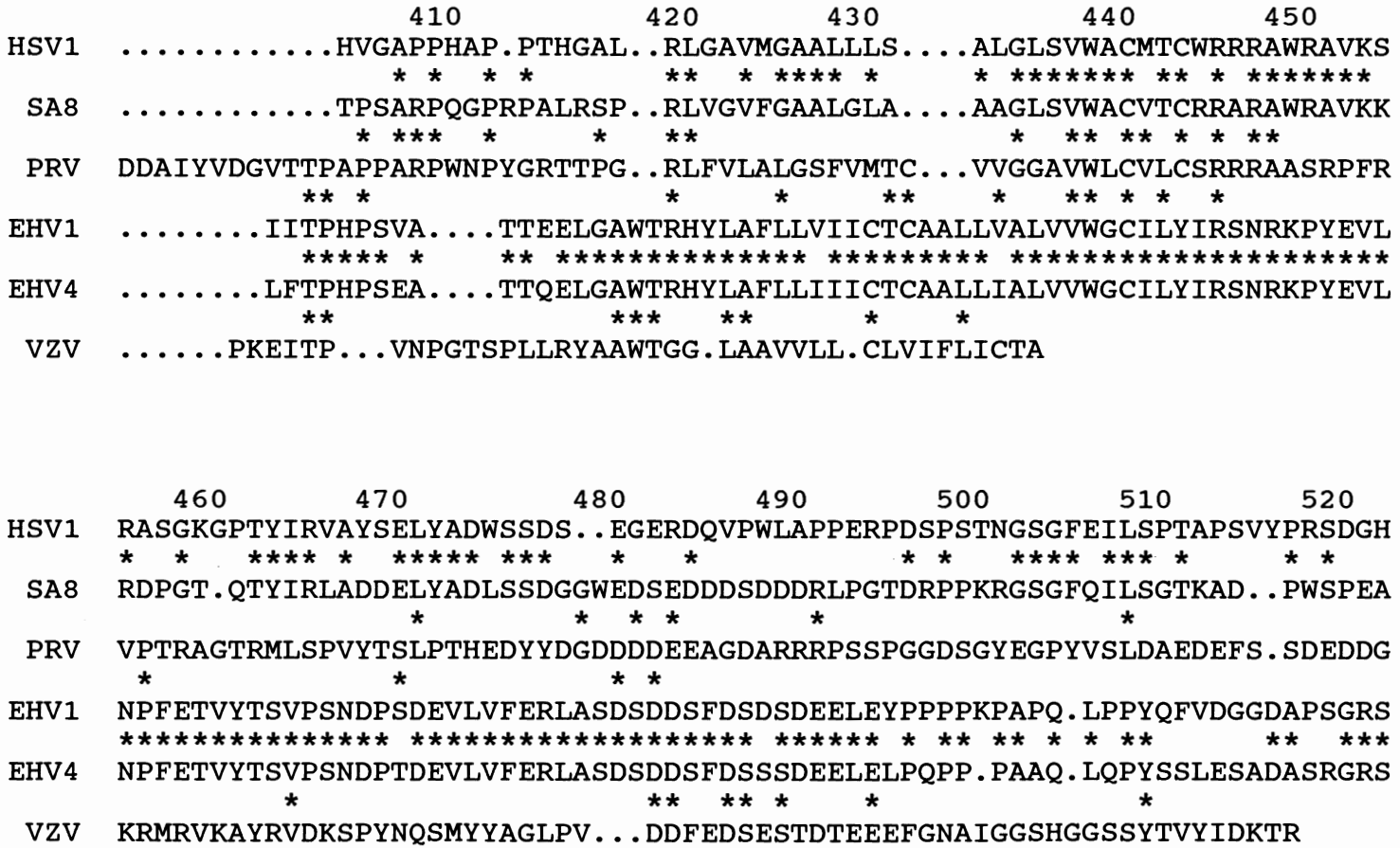

1 10 20 30 40 50 60
HSV2 MA..RGAGLVFFVGVWVVSCLAAAP..RTSWKRVTSGEDVVLLPAPAERTRA..HKLLWAAEPLDACGPL
HSV1 MD..RGAVVGFLLGVCVVSCLAGTP..KTSWRRVSVGEDVSLLPAPGPTGRGPTQKLLWAVEPLDGCGPL
SA8 MA..LARAPRGLLAAWILAAWVGVAAVETTWKHASAGDDVVFFVLPAGRPGGPPRELAWEFASMRNCGPL
PRV MRPFLLRAAQLLALLALALSTEAPSLSAETTPGPVTEVPSPSAEVWDLSTEAGDDDLDGDLNGDDRRAGF
EHV1 ME..LLAASRACIFFGLVTVLDAWGVQQVE....LSEGAWAMIDGRDV.....LTPTNTTTRV
VZV MS..AQEDLGDDTGIHVIPTLNGDDRHKIV...NVDQRQYGDVFKGDLNPKPQGQPLIEVSVEENHPFTL
70 80 90 100 110 120 130
HSV2 RPSWVAL..WPPRRVLETVVDAACMRAPEPLAIAYSPPFPAGDEGLYSELAWRD..RVAVVNESLVIYGA
HSV1 HPSWVSL..MPPKQVPETVVDAACMRAPVPLAMAYAPPAPSATGGLRTDFVWQE..RAAVVNRSLVIHGV
SA8 RPSWVSL..HPPGQVLETVVDAQCVGAPPVHGGVVRPTRRGPRPGPRGGVA.PA..RDVS.NGTLTLREA
PRV GSALASLREAPPAHLVNVSEGANFTLDARGDGAVVAGIWTFLPVRGCDAVAVTMVCFETACHPDLVLGRA
EHV1 TKAWTFL.ETPPGCAGDISVKKVCVSHSLCEDNIIIGKHCNL.LTGEHGIALAE..FNVVNGSLRRTDDV
VZV RAPIQRIYGVRYTETWSFLPSLTCTGDAAPAIQHICLKHTTCFQDVVVDVDCAENTKEDQLAEISYRFQG

```
          140        150        160        170        180        190        200
HSV2  LETDSGLYTLSVVGLSDEARQVASVVLVVEPAPVPTPTPD..DYDEED..DAGVTNARRSAFPPQPPPRR
       ***********    * ********** ******* *   **** *  * *             *
HSV1  RETDSGLYTLSVGDIKDPARQVASVVLVVQPAPVPTPPPTPADYDEDD.NDEGEDESLAGTPASGTPRLP
      *  *** * *       *    *  *      *       *            *     * * *      *
 SA8  RPSDSGMYVLT..VSRAPNSTAARRVVFLTVGP.RVAAAVPGGPPPLA.EGAGAEAGAAATRAPAAHPYP
                             *     *         *    *
 PRV  CVPEAPERGIG.DYLPPEVPRLQREPPIVTPERWSPHLTVRRATPNDT.GLYTLHDASGPRAVFFVAVGD
             *               *  *        *              * *
EHV1  YFVNGTVFPIL.AETRSVL.QIHRATPSIA.GVYTLHVSIDGMMKHSV.VLLTVKKPPKQPQPQPRLRVK
                    *           * *  **
 VZV  KKEADQPWIVV..NTSTLFDELELDPPEIEPGVLKVLRTEKQYLGVYIWNMRGSDGTSTYATFLVTWKGD

          210        220        230        240        250        260
HSV2  PPVAPPTHPRVIPEVSHVRGVTVHMETLEAILFAPGETFGTNVSIH.AIAHDDGPYAMDVVWMRFD.VPS
      ** ***      ********** *** ***** ***** ****** ******  * ***** *** **
HSV1  PPPAPPRSWPSAPEVSHVRGVTVRMETPEAILFSPGETFSTNVSIH.AIAHDDQTYSMDVVWLRFD.VPT
       *              ** ****   *  ******* *  * ***    ****  * * *** *** **
 SA8  HPHPIAEV.......AHVHGVTVSLRTQTAILFSPGDTVHTAVSIV.PFAHDDDPYVMEVVWVRFD.VPE
       *               *            *******       *              *      *
 PRV  RPPAPLAP.....VGPARHEPRFHALGFHSQLFSPGDTFDLMPRVVSDMGDSRENFTATLDWY.YARAPP
       **    *       *      *   * **     *  * *     *         * *   **
EHV1  TPPPVTVP.....QVPVKTHTDFVVHGYHSRVYRDGESFELSVNLESHIVEP..SFSAEIQWYYMNTSSS
            *       *      *    *** *   *  * *   *   * *    *     *
 VZV  EKTRNPTP....AVTPQPRGAEFHMWNYHSHVFSVGDTFSLAMHLQYKIHEA..PFDLLLEWL.YVPIDP
```

```
            280        290         300        310          320        330
HSV2   SCADMRIYEA
       *** *****
HSV1   SCAEMRIYESCLYHPQLPECLSPADAPCA..ASTWTSRLAVRSYAGCSRTNP...PPRCSAEAHMEPVPG
        * ****** ***** **** ********  ** ** ***** *  ***  *   ****   * **  **
 SA8   ECGEMRIYEPCLYHPRLPECRSPADAPCA..ASVWTERLAVRRYGPCSRHVP...PPRCPTDAAMEARPG
        *     **** **** ***  * *  *            * * *  **          ** **  *
 PRV   RCLLYYVYEPCIYHPRAPECLRPVDPACSFTSPARAALVARRAYASCSPLLGDRWLTACPFDAFGEE...
        * *  * * ** ** *  ** *    ****** **     * *  **   *  *   *     *
EHV1   SCDLFRVFETCIFHPTAMACLHPEQHTCSFTSPIRATKILHRVYGNCSDH.GNSWPSRCHSTLLGNRLYF
                                                      *  ********** *  *
EHV4                                                  GSTWPSRCHSTLLGDRPHF
        *   *   **  ** *  **      * ****  *      ** **           *
 VZV   TCQPMRLYSTCLYHPNAPQCLSHMNSGCTFTSPHLAQRVASTVYQNCEHADN..YTAYCLGISHMEPSFG
```



```
          340        350        360        370        380        390        400
HSV1   LAWQAASVNLEFRDASPQHSGLYLCVVYVNDHIHAWGHITISTAAQYRNAVVEQPLPQRGADL.AEPTHP
       * *    ***  ****    *** ****** * *****   **** *******  **         *  *
 SA8   LGWYGPTVNLQLRDASEASGGLYVCVVYVNGHVHAWGHVVVSTAARYRNAVVERSLPRYRPP....PAAP
                   *     **** *   ****  *    * *** *     *   *   *     *
 PRV   ..........VHTNATADESGLYVLVMTHNGHVATWDYTLVATAAEYVTVIKELTAPARAPGTPWGPGGG
                      *   **** *   ***   * ***  **     *    * *          *
EHV1   IQPAQNRVDLLFKDTPASATGLYVFVLLYNGHPEAWTYTLLSTANHFMNVLTDVTRPRLGEHFYTDLGHK
       **** ********* * ************************************ ************ **
EHV4   IQPAPNRVDLLFKDIPESATGLYVFVLLYNGHPEAWTYTLLSTANHFMNVLTDRTRPRLGEHFYTDHGHQ
              * * *** *  ******   *** **  **  **  ** *        *             *
 VZV   LILHDGGTTLKFVDTPESLSGLYVFVVYFNGHVEAVAYTVVSTVDHFVNAIEERGFPPTAGQ..PPATTK
```


410 420 430 440 450
HSV1 .....HVGAPPHAP.PTHGAL..RLGAVMGAALLLS....ALGLSVWACMTCWRRRAWRAVKS
SA8 .....TPSARPQGPRPALRSP..RLVGVFGAALGLA....AAGLSVWACVTCRRARAWRAVKK
PRV DDAIYVDGVTTPAPPARPWNPYGRTTPG..RLFVLALGSFVMTC...VVGGAVWLCVLCSRRRAASRPFR
EHV1 .....IITPHPSVA....TTEELGAWTRHYLAFLLVIICTCAALLVALVVWGCILYIRSNRKPYEVL
EHV4 .....LFTPHPSEA....TTQELGAWTRHYLAFLLIIICTCAALLIALVVWGCILYIRSNRKPYEVL
VZV .....PKEITP...VNPGTSPLLRYAAWTGG.LAAVVLL.CLVIFLICTA
460 470 480 490 500 510 520
HSV1 RASGKGPTYIRVAYSELYADWSSDS..EGERDQVPWLAPPERPDSPSTNGSGFEILSPTAPSVYPRSDGH
SA8 RDPGT.QTYIRLADDELYADLSSDGGWEDSEDDDSDDDRLPGTDRPPKRGSGFQILSGTKAD..PWSPEA
PRV VPTRAGTRMLSPVYTSLPTHEDYYDGDDDDEEAGDARRRPSSPGGDSGYEGPYVSLDAEDEFS.SDEDDG
EHV1 NPFETVYTSVPSNDPSDEVLVFERLASDSDDSFDSDSDEELEYPPPPKPAPQ.LPPYQFVDGGDAPSGRS
EHV4 NPFETVYTSVPSNDPTDEVLVFERLASDSDDSFDSSSDEELELPQPP.PAAQ.LQPYSSLESADASRGRS
VZV KRMRVKAYRVDKSPYNQSMYYAGLPV...DDFEDSESTDTEEEFGNAIGGSHGGSSYTVYIDKTR

```
          530         540         550
HSV1  QSRRQLTTFGSGRPDR.RYSQASDSSVFW
         * * **       * *     *
 SA8  RRGRDLVTFRVDDAARYRDASPPDPPHRR
         *     *        **   *
 PRV  LYVRPEEAPRSGFDVWFRDPEKPEVTNGPNYGVTASRLLNARPA
                     *    *        ** *   ** *
EHV1  GFKVWFRDTPEASPVPLHKPTL....QGPDYSRVASKLKSILK
      ************** *** **       ***** ***** ****
EHV4  GFKVWFRDTPEASPEPLHRPTP...PVGPDYSKVASKLRSILK
```

these viruses was more completely assessed by using the UWGCG program GAP to generate pairwise alignment scores for SA8, PRV, EHV1 and VZV gE sequences with the HSV1 sequence. Table 2 summarize the results. The predicted amino acid sequences of gE polypeptides show 49.1% identity between the SA8 and HSV1, 27.8% between PRV and HSV1, 23.4% between EHV1 and HSV1 and 24.2% between VZV and HSV1. In addition, the results of pairwise similarity scores were used to hierarchically cluster gE sequences with other sequences, with the most closely related sequences being clustered together. The SA8 gE sequence was more closely related to the HSV1 and HSV2 gEs than to the gEs of the other 4 viruses (Figure 8).

TABLE II

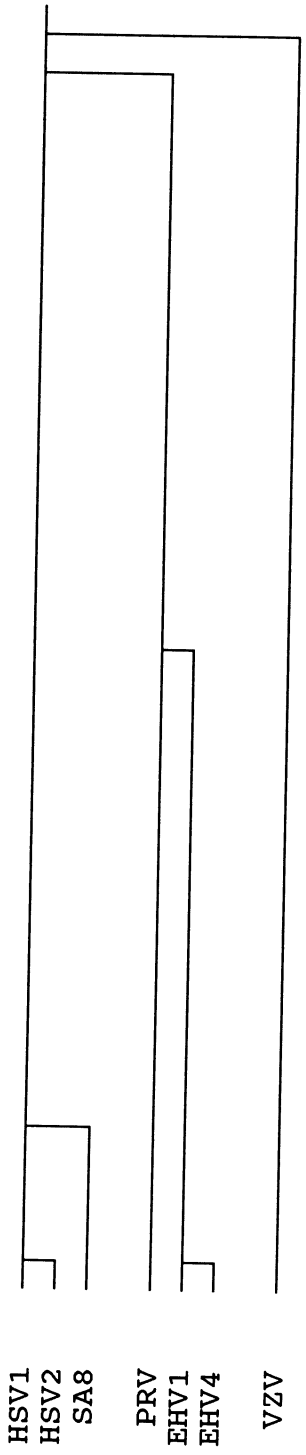
RELATEDNESS OF gE AMINO ACID SEQUENCES TO HSV1 gE

	HSV2[1]	SA8	PRV	EHV1	VZV
% Identity[2]	(66.1)	49.1	27.8	23.4	24.2
% Similarity[3]	(78.8)	64.7	49.1	42.5	45.0

1) Only partial sequence is available for the HSV2 gE gene; these values are based on the first 280 amino acids of the HSV1 & HSV2 gE sequences.

2) Percent of positions in the aligned sequences where an identical amino acid occurs in both sequences.

3) Percent of positions in the aligned sequences where similar amino acids (based on the Dayhoff mutational data matrix) occur.


HSV1
HSV2
SA8
PRV
EHV1
EHV4
VZV

CHAPTER IV

DISCUSSION

Comparison of the partial DNA sequences of several SA8 subclones to the DNA sequence of the HSV1 Us region (37) identified regions of homology with the US3, US6(gD), US7(gI), US8(gE) and US10 genes of HSV1. Based on comparison of the distances between these regions of homology and the relative orientation of the SA8 and HSV1 sequences (Figure 3), it is apparent that both the order and orientation of these five genes are the same in SA8 as in HSV1. Although this needs to be confirmed with additional studies, the results do suggest that the Us region of SA8 is co-linear with the HSV1 Us region. At the least we can say that no major insertions or deletions exist in the SA8 Us region relative to the HSV1 Us region.

Comparison of the DNA sequence and the predicted amino acid sequence of SA8 gE with the HSV1 gE show significant homology at both the DNA and amino acid levels. The SA8 gE gene ORF is 1602 bp and 58.8% identical to the HSV1 gE gene ORF (1650bp). At the level of the amino acid sequence, the HSV1 and SA8 gE polypepetides are 49.1% identical and 64.7% similar to each other.

Both pairwise comparison and cluster analysis of

HSV1, SA8, PRV, EHV1, VZV and a portion of the HSV2 and EHV4 gE amino acid sequences further demonstrate that the SA8 gE is more similar to the HSV1 (and HSV2) gE than it is to the PRV, VZV, EHV1 or EHV4 gEs. Although only 280 amino acids of the HSV2 gE sequence were available for analysis, it was still apperent that the HSV2 gE is more closely related to HSV1 gE than to the SA8 gE. This result parallels that obtained for the gB polypeptides of these same viruses (81). This may indicate that some evolutionary pressure is applied equally to all herpesvirus glycoproteins. Adaptation of the viral surface antigens to surface determinants of cells of their particular host species is one such possible pressure.

Multiple alignments demonstrated that in the C-terminal half of the extracellular hydrophilic region there is significant sequence homology among the HSV1, SA8, PRV, VZV, EHV1 and EHV4 gEs. The amino acid sequences in this region are quite similar and six of the seven Cys residues are conserved. The strongly conserved nature of this region implies a structural similarity in this region for the gEs of all these viruses. Since this region is located in the extracellular domain, it is possible that it serves an important function. One obvious functional possibility is binding the Fc region of anti-viral IgG.

It is of interest that the identity of gE gene DNA sequences (58.7%) between SA8 and HSV1 is higher than

their amino acid identity (49.1%). This is the reverse of what is commonly observed for related genes. One possible reason that could explain this is the high G+C content of these two genes. The high G+C content in the gE genes of both SA8 (76.2%) and HSV1 (66.6%) limit the frequency for use of A or T in the coding sequences. The high G+C content also results in almost exclusive use of only G or C in the 3rd position of codons (Table 3). Because of the strong bias in HSV1 and especially SA8 for codons with G or C in the 3rd position, the probability of having a non-identical base in the 3rd position is considerably reduced. For example, where Gly could usually be coded for by any one of four possible codons (GGA, GGC, GGG, or GGT), due to the strong GC bias in HSV1 and SA8 the only practical possibilities for Gly are GGC and GGG. This usually increases the probability of DNA homology in the 3rd position and, thus, in the gene overall.

Summary

The purpose of this study was to determine if SA8 has a homolog of the HSV1 gE gene within the Us region and, if so, to clone and sequence this gene. Since the gE gene is located within the Us region of HSV1, the BamHI D fragment (about 10 Kbp) containing most Us region of SA8 was cloned. DNA sequencing of the terminal sequences of a number of subclones of this 10 Kbp BamHI D fragment of SA8

Table III

COMPARISON OF CODON USAGE AND PREDICTED AMINO ACID COMPOSITION OF SA8 AND HSV1 gEs

	No. Residues		% Codons with G/C in:					
			3rd Position			1st Position		
AA	SA8	HSV1	Avg[1]	SA8	HSV1	Avg	SA8	HSV1
Non-Polar:								
Ala	74	60	50	93.2	88.3			
Ile	7	11	33.3	71.4	45.4			
Leu	35	41	50	97.1	80.5	66.7	100	80.5
Met	6	10	100	-	-			
Phe	10	9	50	90.0	55.6			
Pro	64	65	50	92.2	78.5			
Trp	13	15	-	-	-			
Val	56	50	50	92.9	82.0			
Polar:								
Asn	6	8	50	100	62.5			
Cys	11	13	50	100	30.8			
Gln	7	16	50	100	75.0			
Gly	47	38	50	91.5	78.9			
Ser	30	52	50	96.7	75.0			
Thr	26	32	50	96.2	87.5			
Tyr	14	15	50	92.9	80.0			
Acidic:								
Asp	32	29	50	93.8	69.0			
Glu	22	26	50	100	69.2			
Basic:								
Arg	57	38	50	87.7	71.1	66.7	94.7	78.9
His	16	16	50	93.8	75.0			
Lys	5	6	50	80.0	50.0			

1) Values represent the expected frequency of occurrence of G/C vs A/T if all codon choices were utilized with equal frequency.

indicated that the Us region of SA8 is co-linear with the Us region of HSV1, since the sequenced regions of this 10 Kbp fragment were found to have good homology with the US3, US6(gD), US7(gI), US8(gE), and US10 genes of HSV1 and distances between these homologous sequences in HSV1 and SA8 were also similar. Both strands of subclones spanning the SA8 gE gene homolog were sequenced completely. The entire sequence of the SA8 gE gene was assembled based on overlapping sequences. The SA8 gE gene ORF is 1602 base pairs long and has 58.8% identity with the HSV gE ORF at the DNA level. The predicted amino acid sequence of the SA8 gE gene product is 534 residues in length and showed 49.1% identity and 64.7% similarity to the HSV1 gE. The predicted secondary structural properties of the SA8 gE polypeptide were very similar to those predicted for the HSV1 gE polypeptide. They included a small N-terminal hydrophobic signal peptide (about 21-25 AAs), a conserved N-linked glycosylation site, an long hydrophilic region representing the extracellular domain, a hydrophobic transmembrane domain, and a highly charged and divergent cytoplasmic tail domain. A highly conserved region in the extracellular domain was apparent which extended from SA8 gE amino acid 205-401. This region was located adjacent to the transmembrane domain. This conserved region of the SA8 gE polypeptide had 63% amino acid sequence identity with the corresponding region of HSV1 gE polypeptide. All

seven Cys residues located in this region were conserved between HSV1 and SA8. Other predicted secondary structure features also indicated that the structure of this region is very similar in both HSV1 and SA8 gE polypeptides. Multiple sequence alignment of seven gE amino acid sequences and hierarchical cluster analysis of these seven gE sequences indicated that the SA8 gE is more similar to the HSV1 gE than the PRV, VZV, or EHV1 gE homologs.

LITERATURE CITED

1. Roizman B. 1979. The structure and isomerization of herpes simplex virus genomes. Cell 16:481-494.

2. Dumas AM, Geelen JLMC, Maris W, et al. 1980. Infectivity and molecular weight of varicella-zoster virus DNA. J Gen Virol 47:233-235.

3. Henle W, Diehl V, Henle G, et al. 1967. Herpes type-virus and chromosone marker in normal leukocytes after growth with irradiated Burkett cells. Science 157:1064-1065.

4. Malherbe H, Harwin R. 1958. Neurotropic virus in African monkeys. Lancet ii 530.

5. Pastoret PP, Thiry E, Derboven G, et al. 1982. Bovine herpesvirus 1 infection of cattle: pathogenesis, latency, and consequences of latency. Ann Rech Vet 13:221-235.

6. Barahona H, Melendez LV, Melnick JL. 1974. A Compendium of herpesviruses isolated from non-human primates. Intervirol 3:175-192.

7. Roizman B. 1982. The Herpesviridae, vol 1. New York: Plenum Press 1-17.

8. Roizman B. 1985. The herpesviruses, vol 3. New York: Plenum Press 105-173.

9. Stevens JG. 1978. Latentcy characteristics of selected herpesvirus. Adv Cancer Res 26:227-256.

10. Plummer G, Goodheart CR, Studdert MJ. 1973. Equine herpesviruses: antigenic relationships and DNA densities. Infect Immun 8:621-627.

11. Smith MG. 1956. Propagation in tissue cultures of a cytopathogenic virus from human salivary gland virus (SGV) disease. Proc Soc Exp Biol Med 92:424-430.

12. Wittmann G. 1989. Herpesvirus diseases of cattle, horses, and pigs. Kluwer Academic Publishers. Boston. 176-229.

13. Wittmann G. 1989. Herpesvirus diseases of cattle, horses, and pigs. Kluwer Academic Publishers. Boston. 96-115.

14. Daniel MD, Melendez LV, King NW, et al. 1971. Herpesvirus aotus: a latent herpescirus from owl monkeys (Aotus trivirgatus)-isolation and characterization. Proc Soc Exp Biol Med 138:835-845.

15. Roizman B. 1985. The herpesvirus. Plenum Press. New York 12-89.

16. Barahona HH, Melendez LV, King NW, et al. 1973. Herpesvirus aotus type 2: a new viral agent from owl monkeys (Aotus trivirgatus). J Infec Dis 127:171-178.

17. Cebrian J, Bucchini D, Sheldrick P. 1983. Endless viral DNA in cells infected with channel catfish virus. J Virol 46:405-412.

18. Rouse BT, Lopez C. 1984. Immunology of herpes simplex virus infection. CRC Press, Florida 1-8.
19. Baringer JR. 1975. Herpes simplex virus infection of nervous tissue in animals and man. Prog Med Virol 20:1-26.
20. Corey L, Adams H, Holmes K, et al. 1983. Genital herpes simplex virus infection: clinical manifestation, course and complications. Ann Intern Med 98:958-960.
21. Baringer JR. 1976. The biology of herpes simplex virus infection in humans. Surv Ophthalmol 21(2):171-174.
22. Centifanto YM, Hildebrandt RJ, Held B, et al. 1971. Relationship od herpe simplex genital infection and carcinoma of the cervix: population studies. Am J Obstet Gynecol 110:690-692.
23. Furlong D, Swift H, Roizman B. 1972. Arrangement of herpesvirus deoxyribonucleic acid in the core. J Virol 10:1071-1074.
24. Wildy P, Watson DH. 1963. Electron microscopic studies on the architecture of animal viruses. Cold Spring Harbor Symp Quant Biol 27:25-47.
25. Fong CKY, Tenser RB, Gross PA, et al. 1973. Ultrastructual studies of the envelopment and release of guinea pig herpeslike virus in cultured cells. Virol 52:468-477.
26. Stannard LM, Fuller AO, Spear PG. 1987. Herpes simplex virus glycoproteins associated with different morphological entities projecting from the virion en-

velope. J Gen Virol 68:715-725.

27. Epstein MA. 1962. Observation of the mode of release of herpes virus from infected Hela cells. J Cell Biol 12:589-597.

28. Goding CR, O'Hare P. 1989. Herpes simplex virus vmw65-octamer binding protein intraction:a paradigm for combinatorial control of transcription. Virol 173:363-367.

29. Yasuda J, Milgrom F. 1968. Haemadsorption by herpes simplex infected cell cultures. Inte Arch Allergy Appl Immun 33:151-170.

30. Westmoreland D, Watkins JF. 1974. The IgG receptor induced by herpes simplex virus: studies using Radioiodinated IgG. Virol 24:167-178.

31. Keller R, Peitchel R, Goldman M, et al. 1976. An IgG-Fc receptor induced in cytomegalovirus-infected human fibroblasts. J Immunol 116:772-777.

32. Ogata M, Shigeta S. 1979. Appearance of immunoglobulin G Fc receptors on cultured human cells infected with varicella-zoster virus. Infec Immun 26:770-774.

33. Spear PG, Baucke RB. 1979. Membrane proteins specified by herpes simplex viruses. V.identification of an Fc-binding glycoprotein. J Virol 32:779-789.

34. Spear PG, Para MF, Goldstein L. 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.

35. Para MF, Zezulac KM, Spear PG, et al. 1983. Use of monoclonal antibodies against two 75,000-molecular-weight glycoproteins specified by herpes simplex virus types 1 and 2 in glycoprotein identification and gene mapping. J Virol 45:1223-1227.

36. Lee GTY, Para MF, Spear PG. 1982. Location of the structural genes for glycoproteins gD and gE and for other polypeptides in the S component of herpes simplex virus type 1 DNA. J Virol 43:41-49.

37. McGeoch DJ, Dolan A., Donald S. et al. 1985. Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. J Mol Biol 181:1-13.

38. Friedman HM, Frank I. 1989. A novel function of the herpes simplex virus type 1 Fc receptor: participation in bipolar bridging of antiviral immunoglobulin G. J Virol 63:4479-4488.

39. Adler R, Glorioso JC, Cossman J, Levin M. 1978. Possible role of Fc receptors on cells infected and transformed by herpesvirus: escape from immune cytolysis. Infect Immun 21:442-447.

40. Johansson PJH, Schroder AK, Christensen P, et al. 1986. Interaction between herpes simplex virus type 1-induced Fc receptor and human and rabbit immunoglobulin

(IgG) domains. Immunology 58:251-255.

41. McKendall RR. 1985. IgG-mediated viral clearance in experimental infection with herpes simplex virus type 1: role for neutralization and Fc-dependent function but not C' cytolysis and C5 chemotaxis. J Infect Dis 151:464-470.

42. Chatterjee S, Koga J, Whitley RJ. 1989. A role for herpes simplex virus type 1 glycoprotein E in induction of cell fusion. J Gen Virol 2158-2163.

43. Cai W, Gu B, Person S. 1988. Role of glycoprotein B of herpes simplex virus type 1 in viral entry and cell fusion. J Gen Virol 62:2607-2612.

44. Johnson DC, McDermott MR. 1988. Herpes simplex viruses lacking glycoprotein D are unable to inhibit virus penetration:quantitative evidence for virus-specific cell surface receptor. J Virol 51:389-394.

45. Gompels U, Minson A. 1986. The properties and sequence of glycoprotein H of herpes simplex virus type 1. Virol 153:230-247.

46. Hanke T, Graham FL, Johnson DC, et al. 1990. Herpes simplex virus IgG Fc receptors induced using recombinant adenovirus vectors expressing glycoproteins E and I. Virol 177:437-444.

47. Kalter SS, Heberling RL. 1971. Comparative virology of primates. Bact Rev 35:310-364.

48. Malherbe H, Strickland-Cholmley M. 1969. Simian herpesvirus SA8 from a baboon. Lancet 2:1427

49. Hull RN. 1971. B virus vaccines. Lab Anim Sci 21:1608-1071.

50. Levin JL, Hilliard JK, Goodwin WJ, et al. 1988. A naturally occurring epizootic of simian agent 8 in the baboon. Lab Anim Sci 38:394-397.

51. Audonnet JC, Winslow A, Paoletti E, et al. 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.

52. Petrovskis EA, Timmins JG, Post LE. 1986. Use of lambda gtll to isolate genes for two pseudorabies virus glycoproteins with homology to herpes simplex virus and varicella-zoster virus glycoproteins. J Virol 60:185-193.

53. Davison A, Scott J. 1983. DNA sequence of the Us component of the varicella-zoster virus genome. EMBO J 2:2203-2209.

54. McGeoch DJ, Moss HWM, Frame MC., et al. 1987. DNA sequence and genetic content of the hindIII l region in the short unique component of the herpes simplex virus type 2 genome: identification of the gene encoding glycoprotein G, and evolutionary comparisons. J Gen Virol 68:19-38.

55. Stevens DA, Pincus T, Hampar B, et al. 1968. Serologic relationship of a simian herpes virus (SA8) and herpes simplex virus: heterogeneity in the degree of reciprocal cross-reactivity shown by rabbit 7S and 19S

antibodies. J Immunol 101:979-983.

56. Eberle R, Hilliard JK. 1984. Replication of simian herpesvirus SA8 and identification of viral polupeptides in infected cells. J Virol 50:316-324.

57. Booy FP, Newcomb WW, Steven AC, et al. 1991. Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. Cell 64:1007-1015.

58. Eberle R, Courtney RJ. 1980. Preparation and characterization of specific antisera to individual glycoprotein antigens comprising the major glycoprotein region of herpes simplex virus type 1. J Virol 35:902-917.

59. Wudunn D, Spear PG. 1989. Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. J Virol 63:52-58.

60. Heine JW, Honess RW, Roizman B, et al. 1974. Proteins specified by herpes simplex virus. XII. The virion polypeptides of type 1 strains. J Virol 14:640-651.

61. Mackem S, Roizman B. 1982. Structural features of the α gene 4, 0, and 27 promoter-regulatory sequences which confer αregulation on chomeric thymidine kinase genes. J Virol 44:939-949.

62. Stevens JG, Wagner EK, Devi-Rao GB, et al. 1987. RNA complementary to a herpesvirus α gene mRNA is prominent in latently infected neurons. Science 235:1056-1059.

63. Honess RW, Roizman B. 1974. Regulation of herpes macromolecular synthesis. I. cascade regulation of the

synthesis of three groups of viral proteins. J Virol 14:8-19.

64. Costa RH, Devi BG, Anderson KP, et al. 1981. Characteristic of a major late herpes simplex virus type 1 mRNA. J Virol 38:483-496.

65. Holland LE, Anderson KP, Wagner EK, et al. 1980. Viral DNA synthesis is required for the efficient expression of specific herpesvirus type 1 mRNA species. Virol 101:10-24.

66. Oakes JE. 1975. Role for cell-mediated immunity in the resistance of mice to subcutaneous herpes simplex infection. Infect Immun 12:166-172.

67. Oakes JE., Lausch RN. 1981. Role of Fc fragments in antibody-mediated recovery from ocular and subcutaneous herpes simplex virus infection. Infect Immun 33:109-114.

68. Kohl S, Adam E, Kaufman RH, et al. 1982. Kinetics of human antibody response to primary genital herpes simplex infection. Intervirology 18:164-168.

69. Kahlson J, Chatterjee S, Lakeman FD, et al. 1987. Detection of antibodies to herpes simplex virus in the cerebrospinal fluid of patients with herpes simplex encephalitis. J infect Dis 155:38-44.

70. Kohl S, Frazier JP, Loo LS, et al. 1981. Normal function of neonatal polymorphonuclear leukocytes in antibody-dependent cellular cytotoxicity to herpes simplex virus infected cells. J Pediatr 98:783-785.

71. Morahan PS, Kohl S, Murray BK, et al. 1981. Immune response to labial infection of Balb/c mice with herpes simplex virus type 1. Infect Immun 32:180-187.

72. Maniatis T, Fritsch EF, Sambrook J. 1982. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory. N.Y.

73. Sugino A, Goodman HM, Heyneker HL, et al. 1977. Interaction of bacteriophage T4 RNA and DNA ligases in joining of duplex DNA at base paired ends. J Biol Chem 252:3987-3994.

74. Birnboin HC, Doly J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl Acids Res 7:1513-1523.

75. Messing J, Vieira J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119.

76. Hattori M, Sakaki Y. 1986. Dideoxy sequencing method using denatured plasmid templates. Analytical Biochem 152:232-238.

77. Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463-5467.

78. Howard LA. 1989. Construction of cloning/sequencing vectors with an alternative polylinker. BioTechniques 7:940-941.

79. Promega Co. 1989. TaqTrack sequencing systems. Madison, WI.

80. Pustell J. Kafatos FC. 1982. A high speed, high capacith homology matrix: zooming through SV40 and polyoma. Nucleic Acids research 10:4765-4782.

81. Eberle R, Black D. 1991. The simian herpesvirus SA8 homologus of the herpes simplex virus gB gene: mapping, sequencing and comparison to HSV1 gB. Arch Virol (In press)

82. Perlman D, Halvorson HO. 1983. A putative signal peptidase recognition site and sequence in eukarytic and prokarytic signal peptides. J Mol Biol 167:391-409.

83. Hillard JK, Eberle R, Black D. 1989. Simian alphaherpesviruses and their relations to the human herpes simplex viruses. Arch Virol 109:83-102.

84. Rziha HJ, Mettenleiter TC, Schreurs C, et al. 1986. Location of the structural gene of pseudorabies virus glycoprotein complex gII. Virol 152:66-75.

85. Babiuk LA, Misra V, Gilchrist JE, et al. 1984. Interaction of monoclonal antibodies and bovine herpesvirus type 1 (BHV-1) glycoproteins: characterization of their biochemical and immunological properties. Virol 135:466-479.

VITA

MIN ZHANG

Candidate for the Degree of

Master of Science

Thesis: IDENTIFICATION AND SEQUENCING OF THE SIMIAN HERPESVIRUS SA8 HOMOLOG OF THE HUMAN HERPES SIMPLEX VIRUS TYPE 1 gE GENE

Major Field: Veterinary Parasitology

Biographical:

Personal Date: Born in Shanghai, China, October 22, 1962, the son of Mr. Nianguan Zhang and Mrs. Yipin Shan. Married Hua Yu on October 6, 1986.

Education: Graduated from New China High School, Shanghai, China in July, 1980; received Bachelor of Medicine from Shanghai Second Medical University, Shanghai, China in July, 1985; completed requirements for the Master of Science degree at Oklahoma State University in July, 1991.

Professional Experience: Pediatrician, Shanghai, China, August 1985-July 1989; Graduate Research Assistant, Department of Veterinary Parasitology and Microbiology, Oklahoma State University, August 1989-July 1991.

Professional Associations: Member of American Society for Microbiology.