

Nucleotide Sequence of the Herpes Simplex Virus Type 2 Thymidine Kinase Gene

MARGARET A. SWAIN AND DENISE A. GALLOWAY*

Fred Hutchinson Cancer Research Center, Seattle, Washington 98104

Received 20 January 1983/Accepted 16 March 1983

We have determined the complete nucleotide sequence of the thymidine kinase gene of herpes simplex virus (HSV) type 2 strain 333. The sequence of the thymidine kinase gene exhibits an open translational reading frame of 1,128 nucleotides encoding a protein of 376 amino acids. The DNA sequence was compared with that of the HSV type 1 thymidine kinase gene from strain MP (S. L. McKnight, *Nucleic Acids Res.* 8:5949-5964, 1980) and from strain CL 101 (M. J. Wagner, J. A. Sharp, and W. C. Summers, *Proc. Natl. Acad. Sci. U.S.A.* 78:1441-1445, 1981) to assess the extent of intra- and intertypic variation for one viral gene. The nucleotides encoding the structural gene varied 1.7% between the two HSV type 1 strains and 19% between HSV type 1 and HSV type 2, which translated to differences in the amino acid sequence of the two proteins of 1.9 and 27%, respectively. The DNA encoding the 5' regulatory sequences appeared to be more conserved than the DNA coding for the structural gene, and the DNA at the 3' end of the gene was the least homologous.

A great deal of interest has centered on the thymidine kinase (TK) gene encoded by herpes simplex virus (HSV) for several reasons. First, the gene is readily amenable to biochemical and genetic analyses because of the powerful selection procedures for and against the gene (15) and because TK is not an essential viral gene. Second, it is a model of gene regulation, in that in virally infected cells, TK transcription is regulated both positively and negatively by other viral products (11, 13, 26). However, when the gene is introduced into other cells it can be expressed in the absence of other viral products (16, 40) although the regulatory sequences recognized by superinfection remain closely associated with the TK gene (13, 41). Third, the ease with which the TK gene can be transferred to cells has facilitated the study of integration and retention of viral sequences into chromosomal DNA (14, 25) and has provided a versatile method for introducing other foreign genes into eucaryotic cells (12, 17). And finally, because the viral TK has a much broader range of phosphorylation substrates than does the cellular enzyme, it has become a target for antiviral therapy (5).

Both genetic (9) and biochemical transfer experiments (20, 28, 29, 40) have located the TK gene of HSV types 1 and 2 (HSV-1 and HSV-2, respectively) at a colinear position on the HSV genome around map position 0.30. The exact nature of the sequences coding for the HSV-1 TK gene has been determined for two strains, MP (21) and CL 101 (39). These results indicate

that (i) the mRNA specifying TK is slightly more than 1,300 nucleotides, and the gene apparently contains no intervening sequences; (ii) there is an untranslated stretch of 107 nucleotides at the 5' end of the RNA, and, beginning with the ATG closest to the start of the message, there is an open reading frame of 1,128 nucleotides; and (iii) signals which have been implicated in the control of transcription of several eucaryotic genes were found at appropriate locations at the 5' and 3' ends of the TK gene.

To study the HSV-2 TK gene more fully, we have determined the nucleotide sequence of this region of the genome from a plasmid clone containing a *Bgl*II-*Hind*III fragment of HSV-2 strain 333 DNA inserted into pBR322 (7). We have previously shown that the cloned DNA is able to biochemically transform Ltk⁻ cells (20). A restriction map of the plasmid is shown in Fig. 1, as is the strategy used to sequence the HSV-2 TK gene. Fragments of viral DNA were purified from the recombinant plasmid, cleaved with restriction enzymes, and cloned into appropriate sites made in the replicative form of the single-stranded phage vectors M13 mp8, mp9, and mp10 (24). Initially, a fragment extending from the *Sac*I (position 0.298) to the *Sal*I site in pBR322 was purified and was treated as follows: (i) the fragment was cleaved with *Pst*I and cloned into the *Pst*I site of M13 mp8 to yield four phage, i.e., the two *Pst*I fragments in both orientations; (ii) the fragment was cleaved with three combinations of enzymes *Pst*I and *Eco*RI,

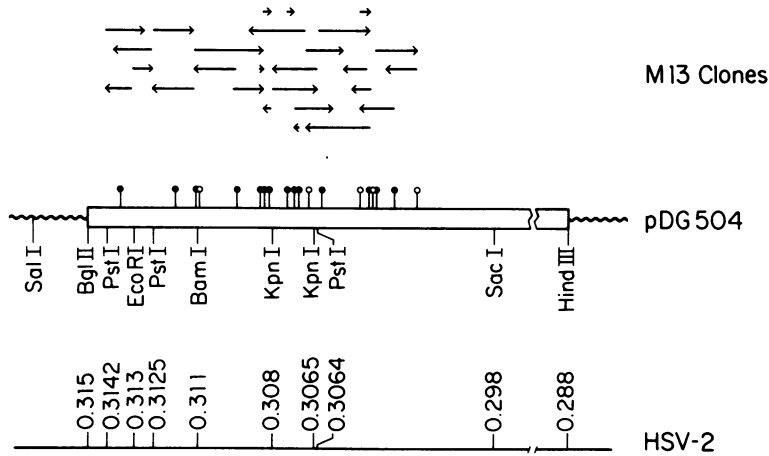


FIG. 1. Strategy used to sequence the HSV-2 TK gene. The bottom line indicates the map coordinates of restriction enzyme sites around the TK gene along the HSV-2 genome. Plasmid pDG504 contains a *Bgl*II-*Hind*III fragment of HSV-2 DNA inserted into pBR322. The positions of cleavage by *Hpa*II (♣) and *Sau*3A (♠) are shown above the fragment. The M13 clones are denoted by arrows which indicate which strand of DNA was contained within the clone. The length of the arrow indicates the amount of sequence that was obtained, not necessarily the size of the fragment.

*Pst*I and *Bam*HI, and *Eco*RI and *Bam*HI, and the resulting fragments were cloned into M13 mp8 and M13 mp9 that had been cleaved with the same combination of enzymes. Three smaller fragments of viral DNA were purified from the TK plasmid, including one extending from the *Bam*HI site (position 0.311) to the *Sac*I site (0.298); a second fragment extending from the *Pst*I site (0.313) to the *Kpn*I site (0.308); and a third fragment extending from the *Bam*HI site (0.311) to the *Kpn*I site (0.308). The first fragment was treated as follows: (i) it was cleaved with *Sau*3A and cloned into *Bam*HI-cleaved M13 mp8; (ii) it was cleaved with *Hpa*II and cloned into *Acc*I-cleaved M13 mp8; or (iii) it was cleaved with *Pst*I and *Sau*3A and inserted into M13 mp8 which had been cleaved with *Pst*I and *Bam*HI. The two smaller fragments of viral DNA were cleaved with *Hpa*II and cloned into *Acc*I-cleaved M13 mp10. This library of recombinant phage (Fig. 1) assured that sequence data could be obtained from both strands of the entire HSV-2 TK gene, usually from more than one template.

The mature single-stranded phage were purified from the culture media, and the DNA which was obtained by deproteinization (33) was used as a template for sequence analysis by the dideoxynucleoside triphosphate chain termination method (32). A 17-residue fragment (a gift from R. Gelinas, Fred Hutchinson Cancer Research Center, Seattle, Wash.) adjacent to the cloning site was used to prime the DNA synthesis reaction which was extended by Klenow polymerase (New England Nuclear Corp.) in the presence of [³²P]dCTP and [³²P]TTP, unlabeled dATP and

dGTP, one dideoxynucleotide in each of four reactions, 6.6 mM NaCl, 6.6 mM Tris-hydrochloride (pH 7.5), 6.6 mM MgCl₂, and 6.6 mM dithiothreitol for 15 min at room temperature. The products of elongation were displayed on thin (0.35-mm) 6 or 8% urea polyacrylamide gels (31). To resolve compressed bases, we separated the products on 15% formamide polyacrylamide gels. The nucleotide sequence data from each template was entered into a VAX/VMS computer, and the sequence was compiled and analyzed using previously established programs (8, 35, 36).

The nucleotide sequence of the noncoding strand of the HSV-2 TK gene is shown in Fig. 2. Beginning at the *Pvu*II site, 1,600 nucleotides which overlap the previously reported sequence of the HSV-1 TK gene (21, 39) are presented. We have not located the exact start of the HSV-2 TK mRNA, but have compared the DNA sequence of the region with that of HSV-1. Sequences located at the 5' end of the TK gene, i.e., between the *Pvu*II site and the ATG initiation codon (nucleotides 310 to 313), are very similar between HSV-1 and -2. The numbers of nucleotides are identical, and the sequences are 89% homologous. Comparison of the DNA sequences of a number of eucaryotic and viral genes have implicated two classes of conserved sequences as important in transcriptional control (1, 3). Presumably, the sequences CATATTA and GGCGAATTC which are identical in both HSV-1 and HSV-2 and are located the correct distance outside the transcribed region have a regulatory function. McKnight et al. have studied the transcriptional control signals

```

CAGCTGCTTC ATGCCCGTGG TCCGCTGTTC GCGTTGCTG GCCGTGTCCC CGGAAGAAAT CGATTTCAT      70
GTCTTTAGCT CCAGGATGAC GCACACACCT CCCAAGCTTT TGTCAATTGGC GAATTGGAAC ACGCAGATGC      140
AGTCTGGGCG GCGCGGCCG AGGTCCAATT CGCATATTAAGGTGACGCG GTGGCCTCGA ACAGCGAGCG      210
ACCTGCAGC GACCCGCTCA TCAGCGTCAG CAGCGTTCCA CAAATCCTGG TGGCGTTGAA CTCCCGCACC      280
TCTCGGGCGA ACGCCTTGTA GAAGCGGGTATGCGCTTCTCA GCGCGGCCAA CAGCACGCGC CTGCGTTCGG      350
TCAGGCTGCT CGTGCAGAGC GCCTACCGA GCGCCGCGCG GCGTCCCGTC CTAGCCATCG CCAGGGGGCC      420
TCCGAAGCCC GCGGGGATCC GGAGCTGCCC ACGCTGCTGC GGGTTTATAT AGACGGACC CACGGGGTGG      490
GGAAGACCAC CACCTCCGCG CAGCTGATG AGGCCCTGGG GCCGCGCGAC AATATCGTCT ACGTCCCGA      560
GCCGATGACT TACTGGCAGG TGCTGGGGGC CTCCGAGACC CTGACGAACA TCTACAACAC GCAGCACCGT      630
CTGACCGCG GCGAGATATC GGCCGGGGAG GCGGCGGTGG TAATGACCAG CGCCAGATA ACAATGAGCA      700
CGCCTTATGC GCGCAGCGAC GCCGTTTTGG CTCCTCATAT CGGGGGGGAG GCTGTGGGCC CGCAAGCCCC      770
GCCCCCGGCC CTCACCCTG TTTTCGACCG GCACCCTATC GCCTCCCTGC TGTGCTACCC GGCCGCGCGG      840
TACCTCATGG GAAGCATGAC CCCCCAGGCC GTGTGGCGGT TCGTGCCCT CATGCCCCCG ACCGGCCCCG      910
GCACGAACCT GGTCTGGGT GTCCTTCCGG AGGCCGAACA GCGCGACCGC CTGGCCAGAC GCCAAGCCCC      980
GGCGGAGCGG CTGACCTGG CCATGCTGTC GCGCATTCGC CGTGTCTACG ATCTACTCGC CAACACGGTG      1050
CGGTACCTGC AGCGCGGCGG GAGGTGGCGG GAGGACTGGG GCCGGCTGAC GGGGGTCGCC GCGGCGACCC      1120
CGCGCCCCGA CCCCAGGAC GCGCGGGGGT CTCGTCCCCG CATCGAGGAC ACGCTGTTG CCCTGTTCGG      1190
CGTTCCCGAG CTGCTGGCCC CCAACGGGGA CTGTACCAC ATTTTGCCT GGGTCTTGA CGTCTGGCC      1260
GACCGCTCC TCCGATGCA TCTATTGTC CTGGATTACG ATCAGTCGCC CGTCCGGTGT CGAGACGCC      1330
TGTTCGCGCT CACCGCCGGG ATGATCCCAA CCCGCGTCAC AACCGCCGGG TCCATCGCCG AGATACCGA      1400
CCTGGCGCGC ACGTTTGCC GCGAGGTGGG GGGAGTTTAG TCAAACACG GAAGCCCGAA CGGAAGCCCT      1470
CCCGGCGATG ACGGCAATAA AAGAACAGAA TAAAAGGCAT TGTGTGCTG TGGTGTGTC ATAAGCGCGG      1540
GGGTTCCGGG CCAGGGCTGG CACCGTATCA GCACCCACC GAAAACCGGA GCGGGCCGAT C      1610
    
```

FIG. 2. Nucleotide sequence of the HSV-2 TK gene. The noncoding strand of the nucleotide sequence is shown. Features of interest are boxed. Symbols above a nucleotide indicate: *, the nucleotide differed from the one that was present in either of the two sequences of the HSV-1 TK gene; -, the nucleotide was deleted in the HSV-1 sequence; +n, some number (n) of nucleotides was present in the HSV-1 sequence that was not present in the HSV-2 sequence. The positions of the insertions or deletions were made to maximize homology. The arrow above nucleotide 201 indicates the start of the HSV-1 TK mRNA.

of the HSV-1 TK gene by constructing a series of deletion mutants (22) and "linker scanning" mutants (23) and assaying the transcriptional competence of the mutants by microinjection into frog oocytes. Their results identify three separate segments of DNA located upstream from the 5' terminus of the TK gene into which the introduction of mutations alters transcrip-

tion. The first segment, located between 16 and 32 nucleotides upstream from the start of the mRNA, contains the TATA homology and is identical between HSV-1 and HSV-2. The second element, located between 47 and 61 nucleotides upstream from the mRNA start, contains the guanosine-rich sequence 5'-GGGGCGGCGCG-3'. These sequences are

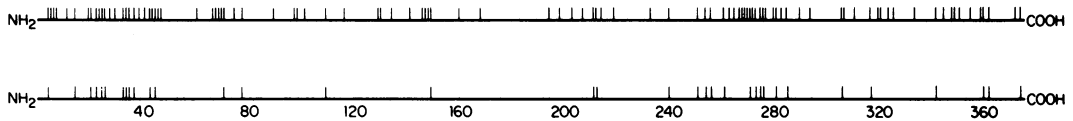


FIG. 3. Location of amino acid substitutions along the TK enzyme of HSV-2 compared with that of HSV-1. Along the top molecule, the vertical bars indicate the positions of any amino acid differences. The vertical bars along the bottom molecule indicate the positions of amino acids with different charges.

the same in HSV-2 except that the 5'-most guanosine is replaced by thymidine. The third element is located between 80 and 105 nucleotides from the start site and contains a cytosine-rich component, six base pairs of which pair exactly with sequences in the guanosine-rich segment CCGGCC. In HSV-2, however, the sequence reads CACCTCC, making an intrastand interaction between these two elements unlikely. Within the third component, two other nucleotide changes were present, a cytosine-thymidine transition at -92 and a guanosine-adenosine transition at -96. Our sequence of this region did not contain the guanosine-adenosine transition at -83 tentatively reported by Reyes et al. (28) and replaced their X at -92 with a thymidine. By simply comparing those sequences which are most conserved between HSV-1 and HSV-2 as an indication of which sequences may serve a regulatory function common to the two viruses, our data have located two nucleotide stretches of 34 and 28 which are perfectly identical. In Fig. 2 these regions fall between nucleotides 110 to 144 and 161 to 189. These regions encompass the so-called CAT and TATA homologies and are located 57 to 91 and 12 to 40 nucleotides upstream from the start of the mRNA (nucleotide 201 in Fig. 2).

The nucleotide sequence of the HSV-2 TK gene has an open reading frame of 376 codons extending from the ATG most proximal to the 5' end of the TK mRNA to a TAG termination signal 1,128 nucleotides downstream. The reading frame is the same as that used by HSV-1, and the number of amino acids specifying the two enzymes is constant, except that the HSV-2 sequence contains an additional three nucleotides, GAG (nucleotides 520 to 522), resulting in the insertion of a glutamic acid residue after amino acid 70. To compensate, the termination codon in the HSV-2 mRNA precedes that of HSV-1 by three nucleotides, having a sequence of UAGUUC in HSV-2 and AACUGA in HSV-1. The nucleic acid homology within the coding region of the TK gene is less well conserved than is the homology within the 5' flanking sequences. Approximately 20% of the nucleotides do not match, and these are scattered throughout the gene. Nearly half of the nucleotide differences result in amino acid substitutions. Figure 3 indicates the positions at which the 102

amino acid substitutions occur. The changes are spread throughout the protein, with particular clusters from amino acids 20 to 40 and 260 to 280. Although some of the substitutions are conservative, e.g., valine to isoleucine, many of the changes affect either the charge or the polarity of the residues. Without further experiments, it is impossible to predict which residues are involved in the catalytic activities of the enzyme or which residues confer properties which are type specific, e.g., thermostability (37), sensitivity to bromovinyl-2-deoxyuridine (6), etc. In comparison, there were many fewer intratypic changes between the two strains of HSV-1. Within the coding region, there were 19 nucleotide changes (1.7%) which resulted in seven different amino acids (1.9%), five of which were conservative changes. In addition, at five of those seven substitutions, the HSV-2 residue was identical to the amino acid of one of the two strains of HSV-1.

The degree of amino acid relatedness between the two viral enzymes is not very different from that seen for the enzyme dihydrofolate reductase among animal species. Amino acid sequence identities of 72 to 89% for this enzyme from chicken, bovine, and porcine liver proteins have been reported (38), in contrast to around 25% identity between animal and bacterial forms or even between different species of bacteria (18). Despite the marked differences in amino acid sequence, the crystal structures of the various dihydrofolate reductases were very similar (18, 38). A review of the DNA sequences of representatives of different classes of adenoviruses has shown that the relatedness of viral proteins can vary greatly (30). For instance, the tumor antigens encoded by early region 1a are only about 40% identical, whereas the IVa2 protein, which is involved in virion assembly, exceeds 90% amino acid identity.

The sequence of HSV-2 DNA located at the 3' end of the gene is the region with the least similarity to HSV-1, although significant features are identical. Both viruses contain two polyadenylation signals, AATAAAA (27). Interestingly, in both viral DNAs the first adenosine of the signal is located 46 bases downstream of the last nucleotide of the termination codon although the intervening sequences are only 75% homologous, including insertions and deletions

TABLE 1. Comparison of intra- and intertypic differences in the TK gene

Region ^a	No of:		No. of differences (%)					
	Nucleotides	Amino acids	Strain MP/strain 101		Strain MP/strain 333		Strain 101/strain 333	
			Nucleotides	Amino acids	Nucleotides	Amino acids	Nucleotides	Amino acids
5'	309		3 (1.0)		33 (11)		34 (11)	
Coding	1,128	376	19 (1.7)	7 (1.9)	215 (19)	101 (27)	219 (19)	102 (27)
3'	164 ^b		0 (0)		40 (24)		40 (24)	
Total	1,601		22 (1.4)		288 (18)		393 (18)	

^a The regions 5' and 3' refer to the DNA that was sequenced on either side of the coding region of the gene, and this does not imply that all of the sequences have any physiological relation to the TK gene.

^b The number of nucleotides indicated was determined from the HSV-2 sequence. The HSV-1 3' region has 166 nucleotides, and the other two regions have the same number of nucleotides in both HSV-1 and HSV-2.

of nucleotides. Within this stretch is a recognition site (AGGCCT) for cleavage by *StuI*, which explains the reduced transforming efficiency seen by Reyes et al. (28) and corrects the position to within the 3' regulatory sequences. Within the 100 nucleotides after the polyadenylation signals, the homology between HSV-1 and HSV-2 was also around 75%, whereas there were no differences in the entire 3' region between the two strains of HSV-1. A recent report by Sharp et al. (34) has located the 5' end of a late mRNA encoded by HSV-1 23 nucleotides downstream from the first polyadenylation site for the TK mRNA and places the regulatory sequences for the message within the 3' untranslated region of the TK mRNA. Interestingly, the region, which appears to serve as a CAT homolog, has an entirely different sequence in HSV-2. Precisely at the start of the HSV-1 sequence GAGACAATAC there are four fewer nucleotides, followed by the sequence CCCGAA. The homology on either side of the CAT box is good, with perfect matches 11 bases upstream and 7 bases downstream. The other putative regulatory sites appear to be better conserved. A repeat of the sequence AATAAAA, which could serve as the TATA homolog, is present in HSV-2, as is the sequence at the start of the HSV-1 mRNA, CATAA. It will be important to determine whether this RNA is synthesized during an HSV-2 infection and, if it is, to determine how it is regulated and which sequences are important in its regulation. A summary of the intra- and intertypic nucleotide differences located throughout the TK region is shown in Table 1.

The information derived from comparative DNA sequence studies of related viruses is of central importance in determining the true degree of relatedness of the HSVs and in suggesting a mechanism of evolution. The analysis of this single gene only represents a first step, but may indicate that the homology is more substantial than was predicted by hybridization (10) or heteroduplex methods. If the nucleotide changes

are randomly scattered along the genome so that segments with perfect complementarity are limited to less than 25 nucleotides, it may be that the sequences of HSV-1 and HSV-2 are 70 to 90% homologous.

Although dramatic chromosomal rearrangements produced by transposable elements (2) may be the major means of affecting adaptation and speciation (19), there is no evidence for this sort of change within this region of the HSV genome. Clearly, a series of single base changes are the basis of variation between the HSV-1 and HSV-2 TK gene. The observed variation is very conservative. Transitions and transversions occur at approximately equal frequency and are much more common than insertions or deletions. When the latter occur they are more common within the noncoding region, in this case within the 3' untranslated sequences. In the one instance in which a deletion occurred within the coding region, three nucleotides were deleted, which maintained the previous reading frame. Whether this DNA sequence analysis provides any generalities about the evolution of the HSVs remains to be seen. Perhaps there are constraints on the variability of the TK gene that will not be found in other classes of viral proteins, e.g., virion components, or in the noncoding regions of the genome. Our studies on a 38,000-molecular-weight protein encoded around position 0.58 on the HSV-2 genome have revealed a similar pattern of variation compared with the homologous HSV-1 protein (4) as seen for the TK gene, except that significant differences occur within the 5' untranslated region (unpublished data). We hope that DNA sequence analysis will eventually shed some light upon the progenitors of this interesting class of viruses.

We thank Richard Gelinas for his patient instructions on M13 cloning, dideoxynucleotide sequencing, and computer analysis. We also thank Devin Harrison for technical assistance, Paul Su for artwork, and Ann Kritzer for help with the manuscript.

This work was supported by grant CA26001 from the National Cancer Institute.

LITERATURE CITED

1. Benoist, C., K. O'Hare, R. Breathnach, and P. Chambon. 1980. The ovalbumin gene-sequence of putative control regions. *Nucleic Acids Res.* 8:127-142.
2. Calos, M. P., and J. H. Miller. 1980. Transposable elements. *Cell* 20:579-595.
3. Corden, J., B. Wasyluk, A. Buchwalder, P. Sassone-Corsi, C. Kedinger, and P. Chambon. 1980. Promoter sequences of the eukaryotic protein coding genes. *Science* 209:1406-1415.
4. Draper, K. G., R. J. Frink, and E. K. Wagner. 1982. Detailed characterization of an apparently unspliced β herpes simplex virus type 1 gene mapping in the interior of another. *J. Virol.* 43:1123-1128.
5. Ellison, G. B., P. A. Furnam, J. A. Fyfe, P. DeMiranda, L. Beauchamp, and H. J. Schaeffer. 1977. Selectivity of action of an anti-herpetic agent 9-(2-hydroxyethoxymethyl) guanine. *Proc. Natl. Acad. Sci. U.S.A.* 74:5716-5720.
6. 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.* 2:432-437.
7. Galloway, D. A., and M. Swain. 1980. Cloning of herpes simplex type 2 DNA fragments in a plasmid vector. *Gene* 11:253-257.
8. Gingeras, T. R., J. P. Milazzo, D. Sciaky, and R. J. Roberts. 1979. Computer programs for the assembly of DNA sequences. *Nucleic Acids Res.* 7:529-545.
9. Halliburton, I. W., L. S. Morse, B. Roizman, and K. E. Quinn. 1980. Mapping of the thymidine kinase gene of type 1 and type 2 herpes simplex viruses using intertypic recombinants. *J. Gen. Virol.* 49:235-253.
10. Kieff, E., B. Hoyer, S. Bachenheimer, and B. Roizman. 1972. Genetic relatedness of type 1 and type 2 herpes simplex viruses. *J. Virol.* 9:738-745.
11. Kit, S., and D. R. Dubbs. 1977. Regulation of herpesvirus thymidine kinase activity in Lmtk⁻ cells transformed by ultraviolet light irradiated herpes simplex virus. *Virology* 76:331-340.
12. Lai, E. G., S. L. Woo, M. E. Bordelon-Riser, T. H. Fraser, and B. W. O'Malley. 1980. Ovalbumin is synthesized in mouse cells transformed with the natural chicken ovalbumin gene. *Proc. Natl. Acad. Sci. U.S.A.* 77:244-248.
13. Leiden, J. M., R. Buttyan, and P. G. Spear. 1976. Herpes simplex virus gene expression in transformed cells. I. Regulation of the viral thymidine kinase gene in transformed L cells by products of superinfecting virus. *J. Virol.* 20:413-424.
14. Leiden, J. M., N. Frenkel, and F. Rapp. 1980. Identification of the herpes simplex virus DNA sequences present in six herpes simplex virus thymidine kinase-transformed mouse cell lines. *J. Virol.* 33:272-285.
15. Littlefield, J. 1964. Selection of hybrids from matings of fibroblasts in vitro and their presumed recombinants. *Science* 145:709-710.
16. Maitland, N. J., and J. K. McDougall. 1977. Biochemical transformation of mouse cells by fragments of herpes simplex virus DNA. *Cell* 11:233-241.
17. Mantel, N., W. Boll, and C. Weissman. 1979. Rabbit β -globin mRNA production in L cells transformed with cloned rabbit β -globin chromosomal DNA. *Nature (London)* 251:40-46.
18. Matthews, D. A. 1979. Interpretation of nuclear magnetic resonance spectra for *Lactobacillus casei* dihydrofolate reductase based on X-ray structure. *Biochemistry* 18:1602-1610.
19. McClintock, B. 1956. Controlling elements and the gene. *Cold Spring Harbor Symp. Quant. Biol.* 21:197-216.
20. McDougall, J. K., T. H. Masse, and D. A. Galloway. 1980. Location and cloning of the herpes simplex virus type 2 thymidine kinase gene. *J. Virol.* 33:1221-1224.
21. McKnight, S. L. 1980. The nucleotide sequence and transcript map of the herpes simplex virus thymidine kinase gene. *Nucleic Acids Res.* 8:5949-5964.
22. McKnight, S. L., and E. R. Gavis. 1980. Expression of the herpes thymidine kinase gene in *Xenopus laevis* oocytes: an assay for the study of deletion mutants constructed in vitro. *Nucleic Acids Res.* 8:5931-5948.
23. McKnight, S. L., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* 217:316-324.
24. Messing, J., and J. Vierira. 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. *Gene* 19:269-276.
25. Pellicer, A., M. Wigler, R. Axel, and S. Silverstein. 1978. The transfer and stable integration of the HSV thymidine kinase gene into mouse cells. *Cell* 14:133-141.
26. Preston, C. M. 1979. Control of herpes simplex virus type 1 mRNA synthesis in cells infected with wild-type virus or the temperature-sensitive mutant tsK. *J. Virol.* 29:275-284.
27. Proudfoot, N. G., and G. G. Brownlee. 1976. Non-coding region sequences in eukaryotic mRNA. *Nature (London)* 263:211-214.
28. Reyes, G. R., K.-T. Jeang, and G. S. Hayward. 1982. Transfection with the isolated herpes simplex virus thymidine kinase genes. I. Minimal size of the active fragments from HSV-1 and HSV-2. *J. Gen. Virol.* 62:191-206.
29. Reyes, G. R., R. LaFemina, S. D. Hayward, and G. S. Hayward. 1979. Morphological transformation by DNA fragments of human herpesviruses: evidence for two distinct transforming regions in herpes simplex virus types 1 and 2 and lack of correlation with biochemical transfer of the thymidine kinase gene. *Cold Spring Harbor Symp. Quant. Biol.* 44:629-641.
30. Sambrook, J., M. Sleigh, J. A. Engler, and T. R. Broker. 1980. The evolution of the adenoviral genome. *Ann. N.Y. Acad. Sci.* 354:426-452.
31. Sanger, F., and A. R. Coulson. 1975. A rapid method for determining sequences in DNA primed by synthesis with DNA polymerase. *J. Mol. Biol.* 94:441-448.
32. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74:5463-5467.
33. Schriber, P. H., and R. Cortese. 1979. A fast and simple method for sequencing DNA cloned in the single-stranded bacteriophage M13. *J. Mol. Biol.* 129:169-172.
34. Sharp, J. A., M. J. Wagner, and W. C. Summers. 1983. Transcription of herpes simplex virus genes in vivo: overlap of a late promoter with the 3' end of the early thymidine kinase gene. *J. Virol.* 45:10-17.
35. Staden, R. 1977. Sequence data handling by computer. *Nucleic Acids Res.* 4:4037-4051.
36. Staden, R. 1978. Further procedures for sequence analysis by computer. *Nucleic Acids Res.* 5:1013-1015.
37. Thouless, M., and G. R. G. B. Skinner. 1971. Differences in the properties of thymidine kinase produced in cells infected with type 1 and type 2 herpesvirus. *J. Gen. Virol.* 12:195-197.
38. Volz, K. W., D. A. Matthews, R. A. Alden, S. T. Freer, C. Hansch, B. T. Kaufman, and J. Kraut. 1982. Crystal structure of avian dihydrofolate reductase containing phenyltriazaeine and NADPH. *J. Biol. Chem.* 257:2528-2536.
39. Wagner, M. J., J. A. Sharp, and W. C. Summers. 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. *Proc. Natl. Acad. Sci. U.S.A.* 78:1441-1445.
40. Wigler, M., S. Silverstein, L. Lee, A. Pellicer, Y. C. Cheng, and R. Axel. 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. *Cell* 11:223-232.
41. Wilkie, N. M., J. B. Clements, W. Boll, N. Mantel, D. Lonsdale, and C. Weissman. 1979. Hybrid plasmids containing an active thymidine kinase gene of herpes simplex virus 1. *Nucleic Acids Res.* 7:859-877.