

Host cell proteins bind to the cis-acting site required for virion-mediated induction of herpes simplex virus 1 α genes

(α -trans-induction factor/gel assay/DNase I/DNA-protein complexes)

THOMAS M. KRISTIE AND BERNARD ROIZMAN

The Marjorie B. Kovler Viral Oncology Laboratories, University of Chicago, 910 East 58th Street, Chicago, IL 60637

Contributed by Bernard Roizman, September 18, 1986

ABSTRACT The herpes simplex virus 1 genes form at least five groups (α , β_1 , β_2 , γ_1 , and γ_2) whose expression is coordinately regulated and sequentially ordered in a cascade fashion. In productively infected cells, the α genes are expressed first, and a virion protein, the α -trans-inducing factor (α -TIF), acts in trans to enhance their expression. Induction of the α genes by α -TIF requires the presence of a trans-induction cis-acting site (α -TIC), and one to three homologs of the α -TIC sequence are contained in the regulatory domains of all α genes. We report that small DNA fragments from regulatory domains of $\alpha 0$, $\alpha 4$, and $\alpha 27$ genes containing α -TIC homologs formed complexes with host but not viral proteins. DNase protection studies indicated that the major host protein complex α -H1 detected in DNA gel retardation assays bound asymmetrically across the α -TIC site. All DNA fragments containing α -TIC homologs, but not those lacking the homolog, competed for the binding of this complex. The location of the binding site of the other host proteins is not yet known. Simian virus 40 DNA fragments containing a homolog of the α -TIC sequence also competed with herpes simplex virus DNA fragments carrying authentic α -TIC homologs for the α -H1 protein complex.

The genes encoded in the herpes simplex virus type 1 (HSV-1) genome form several groups whose synthesis is coordinately regulated and sequentially ordered in a cascade fashion (ref. 1; reviewed in ref. 2). The five α genes, $\alpha 0$, $\alpha 4$, $\alpha 22$, $\alpha 27$, and $\alpha 47$, are the first to be transcribed, and at least one functional α protein, $\alpha 4$, is required for the expression of β and γ genes transcribed later in infection (2-5). Recently, we and others have shown that the $\alpha 4$ protein binds to specific sites in the promoter-regulatory domains of α , β , and γ genes (6-8, 38). The sites of binding of the $\alpha 4$ protein in the domains of α genes are consistent with studies showing that the expression of α genes is regulated, at least in part, by the products of the $\alpha 4$ gene and that this *trans*-acting protein regulates transcription both positively and negatively.

A remarkable property of HSV-1 gene regulation is that a structural protein located in the tegument (i.e., between the capsid and the envelope of the virion) induces the expression of α genes in trans (9, 10). The gene specifying the α -trans-inducing factor (α -TIF) has been mapped and sequenced (11-14). Studies on chimeric genes consisting of the promoter-regulatory domains of α genes fused to an indicator gene have shown that α genes contain a sequence that acts as a cis site for the induction of α genes by α -TIF (14-20). Homologs of this α -trans-induction cis site (α -TIC) are present in at least one ($\alpha 27$ gene) to three copies ($\alpha 0$ gene) in the regulatory domains of α genes (15).

In the course of studies on the binding of $\alpha 4$ protein to HSV-1 DNA, we noted that DNA fragments containing the

α -TIC sequence bind specifically cellular rather than virally encoded proteins (6, 7). In this paper we report that at least one cellular protein binds specifically to the α -TIC site. These data suggest that, in contrast to the $\alpha 4$ protein, the *trans*-acting function of α -TIF is mediated through its effect on one or more host proteins.

MATERIALS AND METHODS

Cells, Virus, and Protein Extracts. HeLa cells, grown to confluency in 850-cm² roller bottles, were mock-infected or infected with 5 plaque-forming units of HSV-1 strain F [HSV-1(F)] (21) and maintained for 2 hr or 12 hr. Nuclear extracts (8-9 mg of protein per ml) were prepared as described (22).

Preparation of DNA Probes and Competitor DNAs. The cloning of the DNA fragments shown in Fig. 1B has been described (6, 7, 19). For DNA probes, the appropriate plasmid DNA was digested with *EcoRI* and *HindIII*, dephosphorylated with calf intestinal alkaline phosphatase (Boehringer Mannheim), purified from polyacrylamide gels, and 5'-end labeled with [γ -³²P]ATP (>7000 Ci/mmol, New England Nuclear; 1 Ci = 37 GBq) by phage T4 polynucleotide kinase (United States Biochemical, Cleveland, OH) to an activity of 25,000 to 50,000 cpm/ng of DNA fragment (23, 24). All DNA fragments used in the competition assays were purified from restriction enzyme digests of CsCl gradient-purified plasmid DNAs separated by polyacrylamide gel electrophoresis as described. The concentrations of all competitor DNA fragments were determined by comparison with known standards in agarose gels.

RESULTS

One or More Host Proteins Bind to DNA Fragments Containing the α -TIC of α Genes 0, 4, and 27. Fig. 1B shows the genomic locations of the four DNA fragments tested for their ability to bind host cell and viral-encoded proteins in the presence of excess competitor nucleic acids using the gel retardation assay originally described by Fried and Crothers (28) and Garner and Revzin (29). 150 α 4R and 70 α 4R represent the promoter proximal 150-base-pair (bp) and promoter distal 70-bp portions of the $\alpha 4$ gene regulatory domain, respectively. 48 α 27R and 29 α 0R designate a 48-bp and a 29-bp sequence cloned from the regulatory domains of the $\alpha 27$ and $\alpha 0$ genes, respectively. Each of these cloned DNA fragments contains a homolog of the α -TIC sequence required for induction of α indicator genes by α -TIF (15). These fragments were labeled with [γ -³²P]ATP and incubated, in the presence of excess competitor nucleic acids, with protein

Abbreviations: HSV-1, herpes simplex virus 1; HSV-1(F), HSV-1 strain F; bp, base pair(s); α -TIF, α -trans-inducing factor; α -TIC, α -trans-induction cis-acting site; α -H_x, host protein(s)_x binding to α -TIC; SV40, simian virus 40; CAT, chloramphenicol acetyltransferase.

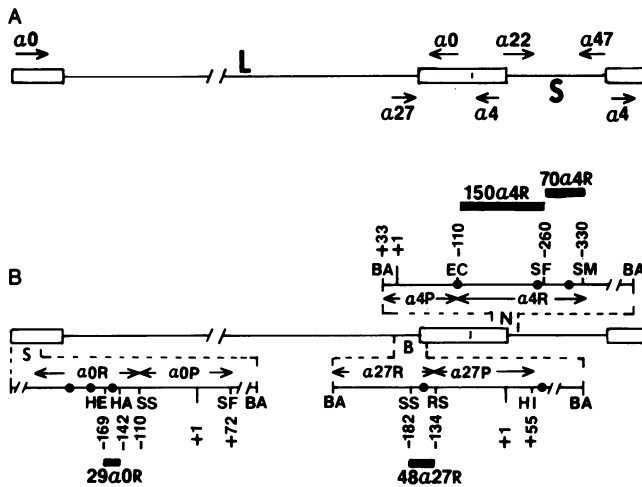


FIG. 1. (A) Sequence arrangement of the HSV-1 genome and location of the α genes. The genome is shown in the prototype arrangement (25). The boxed regions represent the inverted repeat sequences *ba* and *b'a'* flanking the L (long) component and *a'c'* and *ac* flanking the S (short) component. Arrows indicate the location and orientation of the genes encoding the α proteins 0, 4, 27, 22, and 47 (26, 27). (B) Sequence arrangement of the promoter-regulatory domains of α genes 0, 27, and 4. The expanded scales show the restriction patterns and the locations of the regulatory (R) and promoter (P) domains of $\alpha 0$, $\alpha 27$, and $\alpha 4$ located within the viral *Bam*HI DNA fragments S, B, and N, respectively. The filled circles indicate the location of the α -TIC homologs (15). Bold lines delineate the DNA fragments derived from the α -gene regulatory domains tested in these studies. All numbers refer to the site of transcription initiation. HE, *Hae* III; HA, *Hae* II; SS, *Sac* II; SF, *Sfa*NI; BA, *Bam*HI; RS, *Rsa* I; HI, *Hin*I; EC, *Eco*RI; SM, *Sma* I.

extracts of cells mock-infected, infected with HSV-1(F) for 2 hr, or infected with HSV-1(F) for 12 hr. The resulting

complexes were electrophoretically separated in a nondenaturing gel and visualized by autoradiography (Fig. 2).

For each fragment and extract tested, the amounts of protein extract and competitor nucleic acids were 1 μ g and 1.7 μ g or 5.0 μ g and 5.2 μ g. The following results were obtained. (i) The pattern of DNA-mock-infected protein complexes did not differ from that of complexes formed with proteins from 2- or 12-hr HSV-1-infected cells. (ii) Each of the DNA fragments formed a major comigrating DNA-protein complex (designated α -H1) at the higher protein/competitor nucleic acid ratio with all of the extracts tested. A comigrating DNA-protein complex was also formed in the lower protein/competitor ratio reaction with the 70 α 4R probe. Several minor 48 α 27R DNA-protein complexes were also evident (designated α -H2 and α -H3).

Specificity of the Protein-DNA Binding in the Major α -H1 Protein-DNA Complexes. Fig. 3 shows the results of a series of competition studies designed to determine the specificity of the protein-DNA sequence interactions in the major, comigrating DNA-protein complexes (designated α -H1 in Fig. 2). In these studies, labeled DNA fragments 48 α 27R (Fig. 3A) and 29 α 0R (Fig. 3B) were incubated with mock-infected cell extract in the presence of increasing quantities of various unlabeled competitor DNA fragments. The unlabeled homologous 48 α 27R competitor DNA, but not the control fragments (PUC and CAT), efficiently competed with the labeled 48 α 27R probe for the DNA-binding proteins in the α -H1 complex (Fig. 3A). The 29 α 0R, the 70 α 4R, and the 150 α 4R fragments also competed for the proteins in this complex. It is of interest to note that the SV40 enhancer-early promoter region also competed efficiently, in contrast to the α -gene promoter and the metallothionein-gene promoter DNA fragments. Therefore, these results suggest that the proteins involved in the α -H1 complex are sequence-specific binding proteins, recognizing a motif common to the four α DNA fragments and also common to the SV40 early promoter

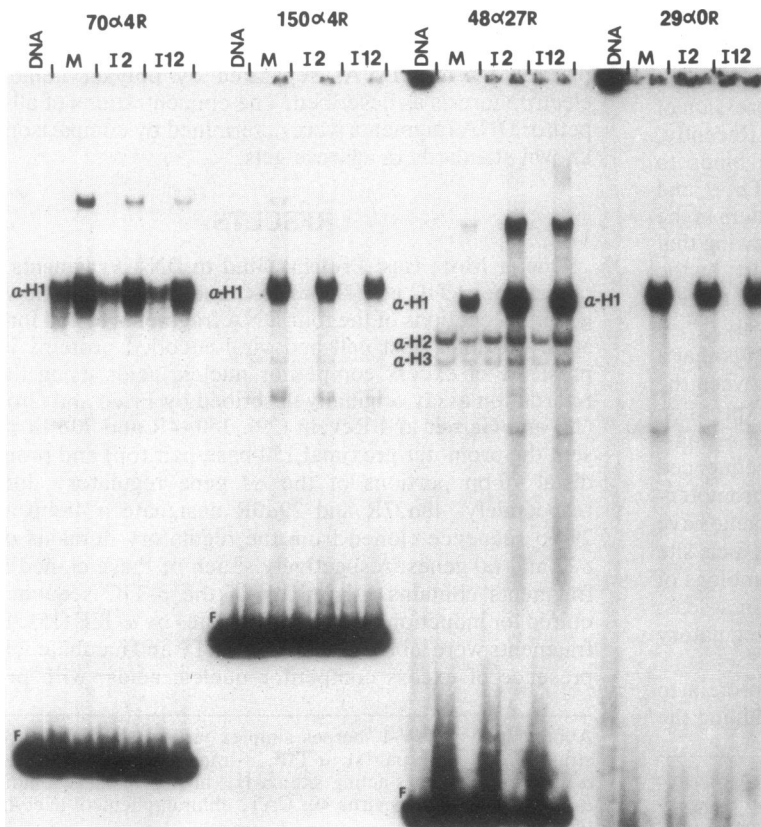


FIG. 2. Autoradiographic images of labeled DNA-protein complexes subjected to electrophoresis in nondenaturing gels. The DNA fragments from α -gene regulatory domains were complexed with proteins from mock- and HSV-1-infected cells as described (6, 7). Specifically, labeled DNA probes (0.5 ng) were incubated with protein extract in the presence of poly(dI)poly(dC) (Pharmacia P-L Biochemicals) in 20 mM Tris-HCl, pH 7.6/50 mM KCl/1 mM EDTA/0.05% Nonidet P-40/5% (vol/vol) glycerol/5 mM 2-mercaptoethanol/50 μ g of bovine serum albumin (Sigma) per ml in a reaction volume of 15 μ l for 45 min at 25°C. The reaction mixtures were electrophoretically separated in a 4% polyacrylamide gel, dried, and exposed to Kodak XS film for 12-16 hr. Lanes marked DNA contain the DNA probes only. The protein extract tested is indicated at the top of each pair of reactions as follows: M, extract from mock-infected cells; I2, extract from cells infected with HSV-1(F) for 2 hr; and I12, extract from cells infected with HSV-1(F) for 12 hr. The first lane of each pair represents reactions with 1 μ g of protein extract and 1.7 μ g of poly(dI)poly(dC), whereas the second lane represents reactions with 5 μ g of protein extract and 5.2 μ g of poly(dI)poly(dC). The protein-DNA complexes are designated α -H1, α -H2, and α -H3, whereas the position of the free DNA probe is indicated by F. The unbound DNA of 29 α 0R fragment migrates slightly faster than that of 48 α 27R and is not represented on this gel.

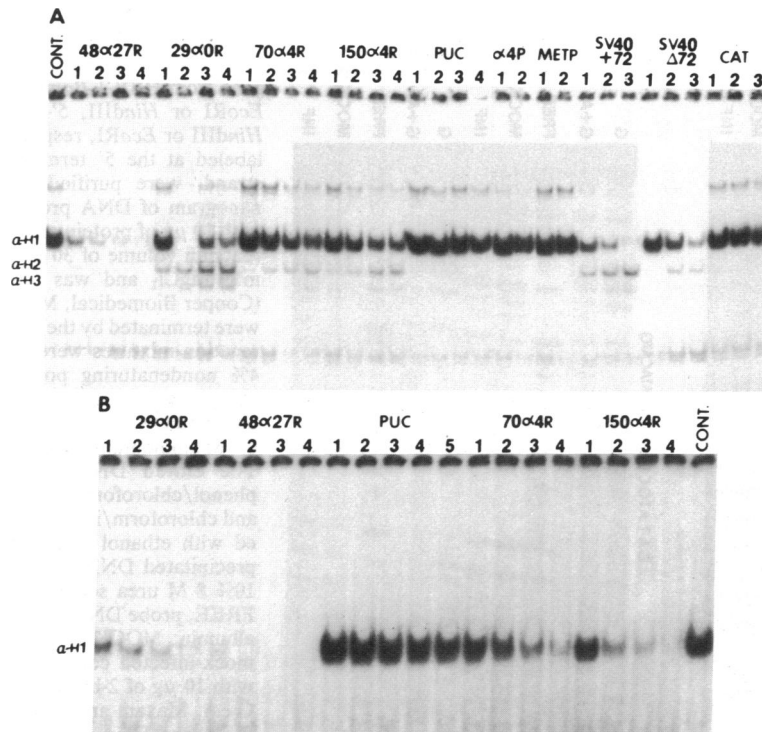


FIG. 3. Autoradiographic images of labeled DNA-protein complexes formed in the presence of various competitor DNAs and electrophoretically separated in nondenaturing gels. DNA-protein binding assays were done as described in the legend to Fig. 2 with 5 μ g of mock-infected cell protein and 5.2 μ g of poly(dI)-poly(dC) per reaction. In addition, increasing concentrations of the unlabeled DNA fragment indicated at the top of the lanes were added to the reaction mixture prior to the addition of the protein extract as follows: lanes 1, 0.15 pm; lanes 2, 0.38 pm; lanes 3, 0.78 pm; lanes 4, 1.17 pm; and lanes 5, 1.4 pm (representing approximately 15-, 38-, 78-, 117-, and 140-fold molar excess over the probe DNA concentration). The DNA fragments derived from α -gene regulatory domains and used as competitor DNAs are shown in Fig. 1. PUC, the 175-bp *Pvu* II-*Hind*III fragment of pUC9 (30); α 4P, the promoter domain of the α 4 gene extending from +33 to -110 relative to the transcription initiation site (16); METP, the 216-bp *Sac* I-*Bgl* II fragment derived from pMK (gift of R. Palmiter) (31) and containing the mouse metallothionein gene promoter domain; SV40 +72, the 340-bp *Pvu* II-*Hind*III fragment of pSV2cat (gift of G. Khoury), which contains the SV40 early promoter and both enhancer repeats (32); SV40 Δ 72, the 200-bp *Bgl* II-*Hind*III fragment of pA10cat (gift of G. Khoury), which contains the SV40 early promoter and \approx 0.25 of the origin-proximal enhancer element (32); CAT, the 160-bp *Pvu* II-*Hind*III fragment of pSV2cat and containing a portion of the chloramphenicol acetyltransferase (CAT) gene coding sequences; CONT., control reaction in the absence of competitor DNA fragments. Protein-DNA complexes are labeled α -H1, α -H2, and α -H3 as in Fig. 2. (A) Competition for proteins involved in the formation of the 48 α 27R DNA probe-protein complexes. (B) Competition for proteins involved in the formation of the 29 α 0R DNA probe-protein complexes.

region. A potentially significant observation was the enhancement of the formation of the 48 α 27R α -H2 and α -H3 complexes as the formation of the 48 α 27R α -H1 complex was reduced by competing DNA fragments.

In the reciprocal experiment, shown in Fig. 3B, the four unlabeled fragments from the α -gene regulatory domains, but not the control fragments, competed with the labeled 29 α 0R fragment for the proteins in the major α -H1 complex. As expected from the first series of competition analyses, the 48 α 27R fragment competed most efficiently for the proteins involved in the 29 α 0R major complex (15 \times molar excess). The homologous 29 α 0R, 70 α 4R, and 150 α 4R fragments also competed significantly relative to the control PUC competitor DNA. The competition series emphasizes the specificity of the binding of the proteins in the α -H1 complex to DNA sequences common to the four α -gene fragments. In contrast to the results obtained in Fig. 3A with the 48 α 27R labeled probe, no secondary protein-DNA complexes were formed by competition for the 29 α 0R major complex proteins, supporting the conclusion that the proteins involved in 48 α 27 α -H2 and α -H3 formation are unrelated to those involved in α 0-fragment α -H1 formation.

DNase I Cleavage-Inhibition Patterns of the Major α -TIC Protein-DNA Complex. To determine the binding site of the protein(s) involved in the formation of the α -H1 protein-DNA complex, the 48 α 27R fragment was labeled with γ - 32 P at the 5' end of the coding and noncoding strands, incubated

with mock-infected or 2-hr HSV-1(F)-infected extract, and partially digested with DNase I (33). The α -H1 complex and the unbound DNA were electrophoretically separated on a nondenaturing gel, eluted from the gel, and electrophoretically separated on a standard denaturing gel as described (24, 34). The autoradiographic images of the coding- and noncoding-strand domains protected from DNase I digestion are shown in Fig. 4. The protected domain in both cases was unaltered by increased concentrations of DNase I (Fig. 4, lanes B). Comparisons of the protected domains of the DNA fragment treated with extracts from mock-infected cells with those obtained with extracts of 2-hr HSV-1(F)-infected cells showed no significant differences, in accord with the conclusion that the proteins involved in the α -H1 complex are of host origin.

Fig. 5 shows the location of the nucleotide sequences of the 48 α 27R fragment protected from DNase I digestion by the protein(s) in the α -H1 complex. The protected domain extends for 20 nucleotides on each strand and appears to be asymmetrically positioned on the coding and noncoding strands. Significantly, the protected regions coincide with the α -TIC, present in the regulatory domains of all α genes, and required for the *in vivo* regulation of an α gene by α -TIF. Preliminary data indicates that the α -TIC homolog in 29 α 0R is similarly protected by the proteins of the 29 α 0R α -H1 complex. As could be predicted, the SV40 DNA sequences sharing homology with α -TIC homologs successfully com-

The Binding of Secondary (α -H2, α -H3) Proteins to DNA Fragments Containing the α -TIC Site. Low concentrations of protein extracts with the 48 α 27R fragment yielded several minor bands migrating faster than the α -H1 DNA complex, in contrast to the reactions with the 29 α 0R fragment. The evidence that the protein in these complexes (α -H2 and α -H3 in Figs. 2 and 3) may not be related to α -H1 rests on competition studies. Pending the DNase I inhibition pattern of the binding site(s) of α -H2 and α -H3 in the 48 α 27R fragment, their concordance with the α -TIC homolog and their specificity remain in question. It should be noted that *in vivo* studies with chimeric α genes produced by fusion of the 48 α 27R or the 29 α 0R fragments to the promoter domains of the α 4 or the β thymidine kinase gene have suggested that, although the α -TIC homolog contained within both α donor fragments is necessary for α -gene regulation, additional sequence elements contained only within the larger 48 α 27R fragment were both necessary and sufficient to regulate the heterologous β -class promoter as an α gene (19). Therefore, it would be consistent that the 48 α 27R fragment contain additional cis-acting sites that would be lacking in the 29 α 0R fragment.

The Nature of the Host Proteins Binding to the α -TIC Homolog. The nature of the proteins with affinity for the α -TIC homolog is not known. Consistent with the results reported here, we noted earlier (20) that homologs of the consensus sequence for the α -TIC site are present in promoter-regulatory domains of several inducible eukaryotic genes. *In vivo*, the α -TIF protein will induce the expression of other viral α or immediate early genes such as the 95-kDa protein of the human cytomegalovirus in cells permissive for cytomegalovirus growth (35). These observations suggest that the proteins binding to the α -TIC homolog may represent a subset of transcriptional factors involved in the expression of inducible genes. HSV-1 has a wide experimental host range. It would follow that, if induction of α genes is required for viral multiplication, all of the proteins interacting with the DNA fragments containing the α -TIC site and capable of conferring α -gene regulation upon promoter elements must be well conserved among vertebrates.

How Does α -TIF Induce the Expression of the α Genes? In contrast to the α 4 protein, the major HSV regulatory protein required for the induction of HSV-1 β and γ genes, the α -TIF does not appear to bind to viral DNA. The absence of demonstrable binding is not *per se* conclusive, but the observation that the cis-acting site required for induction of α genes by α -TIF specifically binds host proteins suggests that α -TIF may, in fact, effect transcription of α genes by modifying host transcriptional factors rather than by binding directly to viral DNA. In this respect, α -TIF may resemble the E1A protein of adenoviruses (36) rather than the α 4 protein of HSV-1 (6–8, 38).

The size and nucleotide sequence of α -gene promoter-regulatory domains suggest that they contain cis-acting sites for a variety of transcriptional factors but that the distribution of these factors among the α genes is not uniform. No fewer than four SP1 transcriptional factor binding sites have been mapped within the α 4 promoter-regulatory domain (37), but its consensus sequence is not present in the domains of all of the α genes. It is conceivable that each α gene has a nonuniform set of cis-acting sites for a variety of transcriptional factors which, in concert, determine the level and other features of the expression of each gene, but that the initiation of high-level transcription would be determined by activation of a subset of these transcriptional factors. In contrast to the α 4 protein, which appears to be able to regulate gene expression both positively and negatively, the operational effect of α -TIF would be solely positive, but its mode of action is less clear at this time. The apparent multiplicity of

host proteins binding to small DNA fragments containing the α -TIC homolog does not allow us to determine at this time if α -TIF inactivates the major binding proteins (α -H1), enabling the minor ones (e.g., α -H2 and α -H3) to enhance transcription, or activates the existing binding protein, thereby changing the rate of transcription.

We thank G. Khoury for the gift of plasmids pSV2cat and pA10cat and R. Palmiter for the gift of plasmid pMK. These studies were aided by grants from the National Cancer Institute (CA08494 and CA19264) and the American Cancer Society (MV2T).

- Honess, R. W. & Roizman, B. (1974) *J. Virol.* **14**, 8–19.
- Roizman, B. & Batterson, W. (1985) in *Virology*, ed. Fields, B. (Raven, New York), pp. 497–526.
- Watson, R. J. & Clements, J. B. (1980) *Nature (London)* **285**, 329–330.
- Preston, C. M. (1979) *J. Virol.* **29**, 275–284.
- Dixon, R. A. F. & Schaffer, P. A. (1980) *J. Virol.* **36**, 189–203.
- Kristie, T. M. & Roizman, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3218–3222.
- Kristie, T. M. & Roizman, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4700–4704.
- Beard, P., Faber, S., Wilcox, K. W. & Pizer, L. I. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4016–4020.
- Post, L. E., Mackem, S. & Roizman, B. (1981) *Cell* **24**, 555–565.
- Batterson, W. & Roizman, B. (1983) *J. Virol.* **46**, 371–377.
- Campbell, M. E. M., Palfreyman, J. W. & Preston, C. M. (1984) *J. Mol. Biol.* **180**, 1–19.
- Pellett, P., McKnight, J. L. C., Jenkins, F. & Roizman, B. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5870–5874.
- Dalrymple, M. A., McGeoch, D. J., Davison, A. J. & Preston, C. M. (1985) *Nucleic Acids Res.* **13**, 7865–7879.
- McKnight, J. L. C., Kristie, T. M., Silver, S., Pellett, P. E., Mavromara-Nazos, P., Campadelli-Fiume, G., Arsenakis, M. & Roizman, B. (1986) in *Cancer Cell: Control of Gene Expression and Replication*, eds. Botchan, M., Grodzicker, T. & Sharp, P. (Cold Spring Harbor Laboratories, Cold Spring Harbor, NY), Vol. 4, pp. 163–173.
- Mackem, S. & Roizman, B. (1982) *J. Virol.* **44**, 939–949.
- Mackem, S. & Roizman, B. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4917–4921.
- Mackem, S. & Roizman, B. (1982) *J. Virol.* **43**, 1015–1023.
- Post, L. E., Mackem, S. & Roizman, B. (1981) *Cell* **24**, 555–565.
- Kristie, T. M. & Roizman, B. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4065–4069.
- Roizman, B., Kristie, T., Batterson, W. & Mackem, S. (1984) in *Transfer and Expression of Eukaryotic Genes*, eds. Ginsberg, H. S. & Vogel, H. J. (Academic, New York), pp. 227–238.
- Ejercito, P. M., Kieff, E. D. & Roizman, B. (1968) *J. Gen. Virol.* **2**, 357–364.
- Dignam, J. D., Lebovitz, R. M. & Roeder, R. (1983) *Nucleic Acids Res.* **11**, 1475–1489.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
- Roizman, B. (1979) *Annu. Rev. Genet.* **13**, 25–57.
- Mackem, S. & Roizman, B. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 7122–7126.
- Rixon, F. & Clements, J. (1982) *Nucleic Acids Res.* **10**, 2241–2256.
- Fried, M. & Crothers, D. (1981) *Nucleic Acids Res.* **9**, 6505–6525.
- Garner, M. M. & Revzin, A. (1981) *Nucleic Acids Res.* **9**, 3047–3060.
- Vieria, J. & Messing, J. (1982) *Gene* **19**, 259–268.
- Brinster, R. L., Chen, H. Y., Warren, R., Sarthy, A. & Palmiter, R. D. (1982) *Nature (London)* **296**, 39–42.
- Laimins, A. L., Khoury, G., Gorman, C., Howard, B. & Gruss, P. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 6453–6457.
- Galas, D. & Schmitz, A. (1978) *Nucleic Acids Res.* **5**, 3127–3170.
- Singh, H., Sen, R., Baltimore, D. & Sharp, P. A. (1986) *Nature (London)* **319**, 154–158.
- Spaete, R. R. & Mocarski, E. S. (1985) *J. Virol.* **56**, 135–143.
- Kovesdi, I., Reichel, R. & Nevins, J. R. (1986) *Science* **231**, 719–722.
- Jones, K. A. & Tjian, R. (1985) *Nature (London)* **317**, 179–185.
- Faber, S. W. & Wilcox, K. W. (1986) *Nucleic Acids Res.* **14**, 6067–6083.