Elsevier has created a Monkeypox Information Center in response to the declared public health emergency of international concern, with free information in English on the monkeypox virus. The Monkeypox Information Center is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its monkeypox related research that is available on the Monkeypox Information Center - including this research content - immediately available in publicly funded repositories, with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the Monkeypox Information Center remains active.

VIROLOGY 135, 561-567 (1984)

Nucleotide Sequence of the Thymidine Kinase Gene Region of Monkeypox and Variola Viruses

JOSEPH J. ESPOSITO[1] AND JANICE C. KNIGHT

Molecular Virology Branch, Division of Viral Diseases, Centers for Disease Control, Department of Health and Human Services, Atlanta, Georgia 30333

Received February 27, 1984; accepted March 20, 1984

Among the orthopoxviruses variola virus induces in cells a characteristic thymidine kinase (TK) activity that can be feedback inhibited in reactions with thymidine triphosphate. Northern blot analyses of variola and monkeypox virus-infected cell extracts showed RNAs of the same molecular weight as the major (590-base) and minor (2380-base) TK transcripts described for vaccinia virus. The nucleotide sequences of 1275 bp in the TK gene region of variola and monkeypox viruses have been determined. When these sequences were compared with such sequences reported for vaccinia virus, differences were observed at 41 nucleotide positions. Examination of the putative encoded TK polypeptide for the three viruses revealed variation at eight amino acid positions. Two major differences in the amino acid composition of the variola virus TK were identified that might play a role in alteration of its kinetic properties.

Increased induction of thymidine kinase (TK) activity has been observed in cells infected with Shope fibroma virus (*2*) and with most orthopoxviruses (*3, 20, 21*). TK gene mapping (*7, 16, 30, 32*) with a more precise localization (*1*) of the TK transcript within the vaccinia conserved central DNA region (*Hin*dIII—"J"fragment), and the synthesis in rabbit reticulocyte lysates (*1, 14*) and *Xenopus laevis* oocytes (*18*) of enzymatically active TK from early mRNA from vaccinia virus (WR)-infected TK^- cells confirmed earlier observations (*25*) that indicated the TK activity is virus encoded.

McAuslan (*20–22*) originally suggested that the vaccinia TK was an allosteric-type enzyme because thymidine, ATP, and HeLa cell-soluble extract stabilized and reactivated TK that had been partially purified through Sephadex. The reactivated TK under appropriate levels of ATP could be feedback inhibited by thymidine triphosphate (TTP), however, TTP did not feedback inhibit the crude active TK. Bedson explored the possibility of distinguishing orthopoxviruses by resolving serologic and/or physico-chemical differences in the TK (*3*). With infected cell lysates in a standardized assay he observed that only the TK activities of variola viruses and the so-called "whitepox" viruses (which are identical to variola viruses by endonuclease analyses *8, 9, 12*, Esposito *et al.*, in preparation) were sensitive to feedback inhibition by TTP (4–8 μmol/liter). The 51 orthopoxviruses that Bedson compared included 11 variola viruses, 5 "whitepox" viruses, 10 variola–cowpox hybrids, 6 alastrim–rabbitpox hybrids (*4, 5*), 5 monkeypox strains, plus 1 white variant monkeypox virus, and 13 other members of the genus. Tests with the hybrid viruses indicated that the TK character segregated as an independent biologic marker.

Recently, Weir and Moss (*33*) and Hruby *et al.* (*17*) determined nucleotide sequences in the TK gene region of WR vaccinia virus. Weir and Moss (*33*) also sequenced the TK region of three vaccinia spontaneous TK^- mutants. Using primer extension and nuclease DNA–RNA hybrid digestion, they located precisely on the vaccinia genome the 5′ and 3′ ends of a 570-base TK transcript. Two possible TK transcriptional

[1] Author to whom requests for reprints should be addressed.

0042-6822/84 $3.00
Copyright © 1984 by Academic Press, Inc.
All rights of reproduction in any form reserved.

regulatory (A + T rich) signals (approximately at −10 and at −35 from the mRNA 5′ end) were recognized in the terminal sequences of an adjacent late gene. Identification of the 177 codons (531 bases) for the TK protein (20,077 Da) and sequencing the mutants disclosed the probable translation initiation and termination sites and revealed that the TK^- character was due to an apparent single-base reiteration which produces a reading frameshift and consequent nonsense codon downstream.

In order to begin understanding the molecular aspects of evolutionary divergence among orthopoxviruses, we have compared the DNA sequences reported for the vaccinia virus TK gene region with DNA sequences derived from the analogous regions in variola and monkeypox viruses. Further, we compared the amino acid sequences of the putative encoded TK polypeptide of the three viruses. Two single-base substitutions were observed in the variola TK gene that resulted in the replacement of an asparagine codon by an aspartic acid codon and of a glutamic acid codon by a lysine codon. These changes might be responsible for the characteristic TTP feedback inhibition of the variola virus TK activity that was reported by Bedson (*3*).

Figure 1A shows restriction maps (derived in another study, Esposito and Knight, in preparation) that indicate the positions of the analogous *Hin*dIII fragments containing the TK gene sequences—vaccinia "J," variola "L," and monkeypox "K." With standard methods (*23*) we have inserted in bacterial plasmids (pBR-322 or -328) most of the *Hin*dIII DNA fragments of the Bangledesh-1975 strain of variola virus (fragments C, E, F, H-P, Fig. 1A) and all fragments except the terminal "J" of Copenhagen monkeypox virus and progeny transposition-deletion variants CpCR, CpCW-N1, and CpCW-N2 (Fig. 1A) (11). Large-scale preparations (*23*) were made of the pBR-322 recombinants pBSH-L and pN2-K that contain respectively *Hin*dIII fragments "L" of variola virus (Bangledesh-1975) and "K" of monkeypox virus (Copenhagen) white variant N2. Plasmid insert restriction fragments were used for blot hybridizations and for sequencing the viral TK gene region. A description is presented in Fig. 1B of pBSH-L and pN2-K together with a diagram showing our DNA sequencing strategy.

Expression of TK gene activity in BSH and N2 virus-infected and mock-infected human amnion (FL) cells was determined by TK assays (*15*) performed with cells harvested at 5 hr postinfection. Cytoplasmic NP-40 cell extracts (*15*) showed a twofold or greater increase (not shown) in the level of TK activity in BSH and N2 virus-infected FL cells compared with the mock-infected cells (TK^+). The viral origin of the observed increase in TK activity was established with northern-blot assays (*28*) performed with RNAs purified from cells by treatment of the NP-40 cytoplasmic extracts (prepared with RNAsin) with proteinase-K, and SDS, followed by phenol/chloroform extraction, and ethanol precipitation. The RNAs were separated by electrophoresis in a horizontal submerged 1.4% agarose gel in pH 3.3 citrate-urea buffer (*13*). Gels were washed repeatedly with a dilution series (100 to 10 m*M*) of neutral sodium phosphate buffer, and then RNAs were blotted onto nitrocellulose sheets (*28*). TK-specific RNA transcripts were identified by hybridizations with an *Rsa*I-*Eco*RI fragment from pBSH-L and pN2-K (Fig. 1B). These fragments are analogous to the 220-bp vaccinia TK gene region fragment that was used by Weir and Moss (*33*) to locate the vaccinia TK mRNA 5′ end by primer extension ("G" at position 494, Fig. 3). As shown in Fig. 2, BSH and N2 virus-infected cells clearly contain viral TK gene sequences as is revealed by the hybridization of the *Rsa*I-*Eco*RI fragment to specific cytoplasmic RNAs. The hybridizing monkeypox and variola virus RNAs, which could be enriched if cyclohexamide was present during infection, contained 600–700 nucleotides relative to the migration of the marker 890 nucleotide RNA, the influenza virus NS gene (provided by Nancy Cox, CDC Influenza Branch). After a prolonged exposure of the northern-blot radioautographs (not shown), we observed a second RNA of about 2400 nucleotides, present at an es-

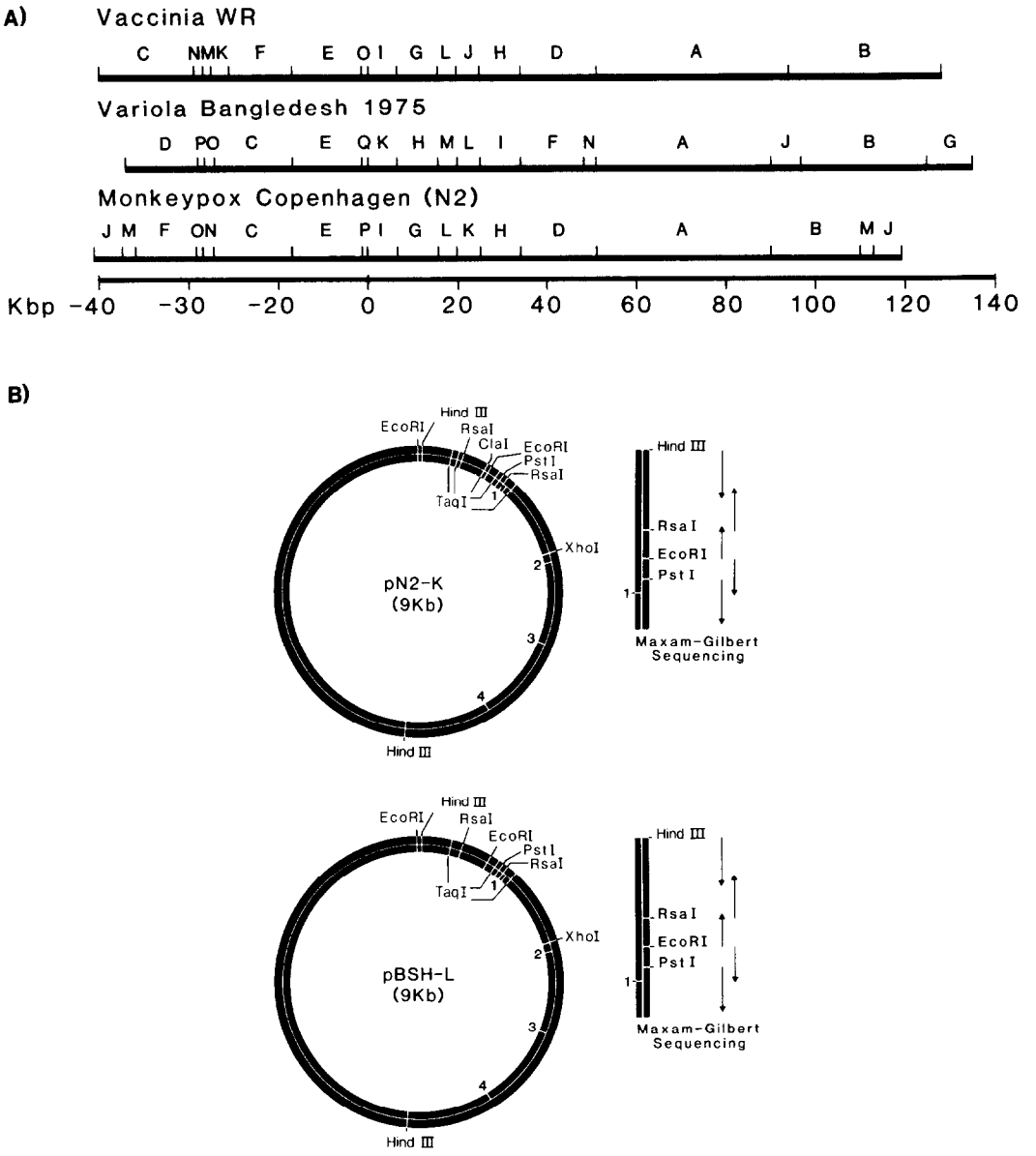

FIG. 1. (A) *Hind*III restriction sites mapped on the genome DNAs of vaccinia, variola, and monkeypox viruses (from Esposito and Knight, in preparation). (B) Endonuclease sites within *Hind*III fragment "L" of BSH variola virus and *Hind*III fragment "K" of N2 monkeypox virus. As described in the text, these *Hind*III fragments were cloned into pBR-322 by standard methods, mapped, and the sequences in the thymidine kinase gene region (Fig. 3) were determined for complementary DNA strands by the Maxam-Gilbert (*24*) method with the strategy indicated.

timated concentration one-twentieth the level of the major RNA. Both RNA types also have been identified in vaccinia-infected cells by Bajszar *et al.* (*1*) who concluded that the longer RNA resulted from a small percentage of the TK transcripts terminating at the end of a 3′ adjacent gene.

The primary structure of the monkeypox and variola virus TK genes was examined by determining the nucleotide sequences

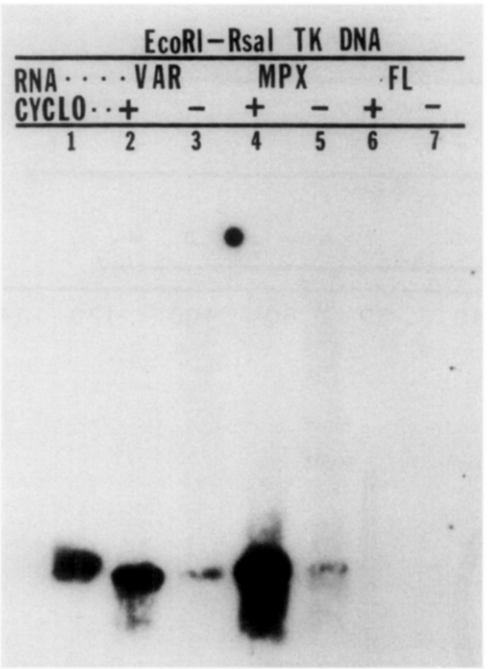

FIG. 2. Northern-blot analysis for viral thymidine kinase RNA transcripts in variola (lanes 2 and 3) and monkeypox (lanes 4 and 5) virus-infected human amnion (FL) cells. Cytoplasmic RNAs (5 hr postinfection) produced in infected and mock-infected FL cells (lanes 6 and 7) in the presence (+) or absence (–) of cyclohexamide (CYCLO) were separated by electrophoresis in citrate-urea agarose gels. RNAs were transferred to nitrocellulose and viral TK-specific transcripts were identified by hybridization with a *Rsa*I–*Eco*RI fragment from pBSH-L or pN2-K (Fig. 1B) as described in the text. The size of the transcripts were estimated by parallel electrophoresis with an influenza NS gene RNA probed with a DNA cloned copy (lane 1).

for 1275 bp of pN2-K and pBSH-L with the strategy shown in Fig. 1B. Complementary DNA strands were sequenced separately by the method of Maxam and Gilbert (*24*). Usually, about 10–100 pmol of plasmid DNA was restricted (Fig. 1B), and depending on the endonuclease site, fragments either were dephosphorylated and 5′ end labeled with [γ-^{32}P]ATP and polynucleotide kinase or were 3′ end labeled with [α-^{32}P]dNTPs and DNA polymerase I (Klenow fragment) or with 3′-[α-^{32}P]dATP and terminal transferase (*23, 29*). After cleavage with a second endonuclease, fragments labeled at one end for sequencing were purified by electroelution from polyacrylamide gels.

Figure 3 shows a comparison of the nucleotide sequences for the TK gene region in monkeypox and variola virus DNAs together with such sequences that have been determined for vaccinia virus by Weir and Moss (*33*). We also compared the computer-derived amino acid sequence of the putative encoded polypeptide. The illustration shows all the TK region nucleotide sequences (transcript sense DNA strand) and the TK polypeptide for monkeypox virus. For clarity, only the different TK region nucleotides (* = deletion) and TK amino acids have been shown for variola and vaccinia viruses. Nucleotide sequences were analyzed with the SEQ program (*6*) and managed with the NIH DEC-10 programs FUR, SOS, and SEQTRAN (J. V. Maizel, Jr., personal communication).

Although correlations of orthopoxvirus base sequences with eucaryotic regulatory signal sequences have not been fully established, the locations of seemingly analogous poxvirus sequences (Fig. 3) are summarized as follows: The locations of candidate vaccinia TK promotor sequences (*27, 33*) have been indicated (+). The variola TK gene showed an adenine deletion in the proximal candidate promotor region (position 485). Variola and vaccinia DNAs each showed a thymidine deletion (position 472) in sequences between the two possible promotors. Examination of computer-deduced open-reading frames revealed that the net result of these deletions in monkeypox, variola, and vaccinia would be the presence of a different carboxy-terminal amino acid sequence and a different translation termination codon position for the 5′ adjacent (late) gene in the respective viruses. In vaccinia the translation stop codon (TAA) for this 5′ adjacent gene occurs at position 483–485, in monkeypox the stop codon (TAA) occurs at position 479–482, and in variola the stop codon TGA (position 501–503) overlaps the open-reading frame of the TK gene. The effect on the regulation of TK by this crowding of genes is unknown; there might be no effect on the TK early gene since the adjacent gene is ex-

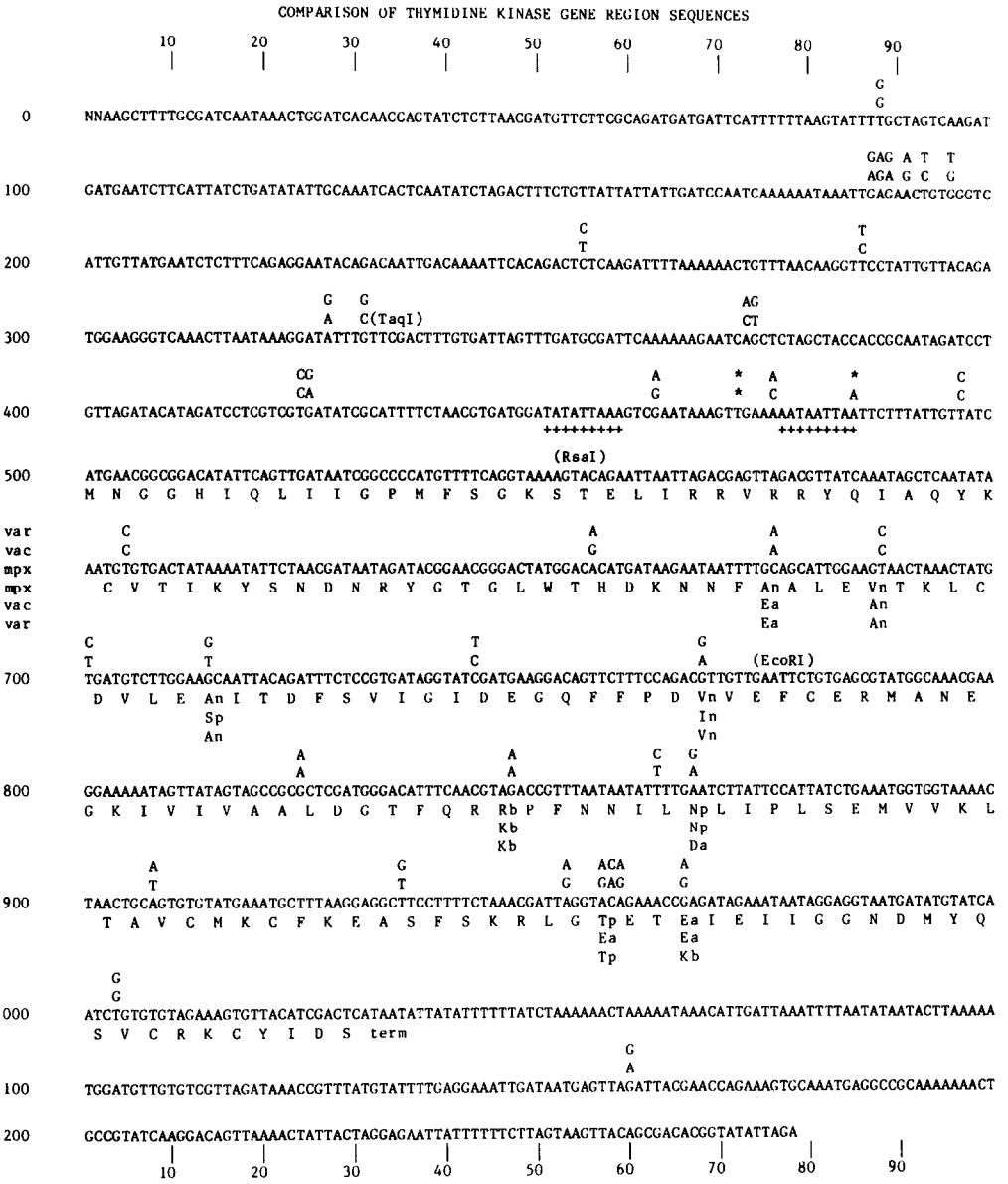

FIG. 3. Nucleotide sequences in the thymidine kinase gene region of variola, monkeypox, and vaccinia virus DNAs and the computer-derived amino acid sequence of the putative TK polypeptide. Variola and monkeypox virus nucleotides were sequenced by the Maxam-Gilbert method with the strategy shown in Fig. 1B as described in the text. The complete base sequence of the monkeypox TK gene region is shown with positions of nucleotide variation indicated for variola and vaccinia DNAs. The complete amino acid sequence of the monkeypox TK polypeptide is shown with positions of amino acid variation indicated for the variola and the vaccinia polypeptide. Amino acid type: a = acidic, b = basic, p = polar, n = nonpolar. Symbols: (*) deletion, (+) candidate promotor sequences. Vaccinia sequences were from Weir and Moss (*33*).

pressed late in infection (*33*). No characteristic eucaryotic mRNA cap signal (overlapping GCG triplets) was observed in the poxvirus sequences at the reported (*33*) vaccinia mRNA start site (position 494). At position 390 (−100 bp upstream)

the characteristic eucaryotic RNA polymerase II promotor sequence 5′-CAAT (*10*) was noticed, whether these function similarly for poxvirus RNA polymerase is yet unknown. A search for sequences similar to the eucaryotic polymerase activator-enhancer sequences (*31*) within the 1275 bp revealed such sequences possibly at positions 341–347 (5′-GTGATTA) and 891–898 (5′-GTGGTAA). At residues 1063–1068 the hexanucleotide, 5′-AATAAA, which typically precedes the polyadenylation site in eucaryotic genes was found immediately followed by CA, sequences presumed to specify the preferred site for poly-A addition (*26*). The ATG probable translation start codon (*19*) for the TK occurred at position 500–503 followed by an open-reading frame (177 codons) for a polypeptide of about 20,000 Da that terminates at (TAA) position 1032–1034 (*33*).

When we compared the nucleotide differences within the TK open-reading frame for the three viruses we found that there were 10 loci of silent (third base, "wobble" locations) mutational differences (positions 604, 656, 700, 743, 824, 863, 907, 935, 953, and 1033), that would produce no consequent changes in amino acid sequence of the putative TK polypeptide. However, sequence differences in eight codons were identified that would be expected to change the amino acid composition of the putative TK polypeptide (positions 676, 688, 714, 768, 847, 867, 957–959, and 966). Of the 1275-base TK gene region sequences compared, nucleotide variation was observed at 41 residue positions. The putative TK polypeptide that is encoded by each virus showed variation at eight amino acid positions. Two differences in the amino acid composition of the variola virus TK were identified that might play a role in alteration of its kinetic properties. Because of the nucleotide difference at residue 867, instead of the polar amino acid asparagine (N) there would be expected an aspartic acid (D) in variola; the nucleotide difference at residue 966 would change glutamic acid (E) for the basic amino acid lysine (K) in variola. The amino acid differences could affect the configuration of the variola TK, possibly exposing an allosteric-effector site, producing a variola TK activity sensitive to feedback inhibition in reactions with TTP.

The sequences of the three viruses obviously were highly conserved. Within the 1275 bp that we compared, 41 separate residue positions showed variation (Fig. 3). There was no single residue position at which a different nucleotide was present for each of the three viruses. Variola and vaccinia DNAs showed the same nucleotide, whereas monkeypox DNA contained a different base at 10 loci (residues 88, 424, 472, 497, 605, 676, 688, 824, 847, and 1004). Monkeypox and vaccinia DNAs showed the same nucleotide, whereas variola DNA contained a different base at 10 loci (residues 196, 327, 463, 485, 701, 743, 863, 867, 935, and 966). Vaccinia DNA contained a different base at 21 separate loci that showed the same nucleotide with variola and monkeypox DNA.

ACKNOWLEDGMENTS

We thank Dr. John Dunn for advice on sequencing and Dr. Bernard Moss for critical reading of the manuscript.

REFERENCES

1. BAJSZAR, G., WITTEK, R., WEIR, J., and MOSS, B., *J. Virol.* **45**, 62-72 (1983).
2. BARBANTI-BRODANO, G., PORTOLANI, M., BERNARDINI, A., STIRPE, F., MANNINI-PALENZONA, A., and LA PLACA, M., *J. Gen. Virol.* **3**, 471-474.
3. BEDSON, H. S., *Bull. WHO* **60**, 377-380 (1982).
4. BEDSON, H. S., and DUMBELL, K. R., *J. Hyg.* **62**, 141-146 (1964).
5. BEDSON, H. S., and DUMBELL, K. R., *J. Hyg.* **62**, 147-158 (1964).
6. BRUTLAG, D. L., CLAYTON, J., FRIEDLAND, P., and KEDES, L. H., *Nucleic Acids Res.* **10**, 279-294 (1982).
7. DUBBS, D. R., OTSUKA, H., QAVI, H., and KIT, S., *Virology* **126**, 408-411 (1983).
8. DUMBELL, K. R., and ARCHARD, L. C., *Nature (London)* **286**, 29-32 (1980).
9. DUMBELL, K. R., and KAPSENBERG, J. G., *Bull. WHO* **60**, 381-387 (1982).
10. EFSTRATIADIS, A., POSAKONY, J. W., MANIATIS, T., LAWN, R. M., O'CONNELL, C., SPRITZ, A. R., DERIEL, J. K., FORGET, B. G., WEISSMAN, S. M., SLIGHTOM, J. L., BLECHL, A. E., SMITHIES, O., BARALLE, F. E., SHOULDERS, C. C., and PROUDFOOT, N. J., *Cell* **21**, 653-668 (1980).

11. ESPOSITO, J. J., CABRADILLA, C. D., NAKANO, J. H., and OBIJESKI, J. F., *Virology* **109,** 231–243 (1981).

12. ESPOSITO, J. J., NAKANO, J. H., and OBIJESKI, J. F., *Virology* **89,** 53–66 (1978).

13. HOLLOWAY, B., and OBIJESKI, J. F., *J. Gen. Virol.* **49,** 181–195 (1980).

14. HRUBY, D. E., and BALL, L. A., *Virology* **113,** 594–601 (1981).

15. HRUBY, D. E., and BALL, L. A., *J. Virol.* **40,** 456–464 (1981).

16. HRUBY, D. E., and BALL, L. A., *J. Virol.* **43,** 403–409 (1982).

17. HRUBY, D. E., MAKI, R. A., MILLER, D. B., and BALL, L. A., *Proc. Nat. Acad. Sci. USA* **80,** 3411–3415 (1983).

18. HRUBY, D. E., MILLER, D. B., and BALL, L. A., *Virology* **123,** 470–473 (1982).

19. KOZAK, M., *Cell* **15,** 1109–1123 (1978).

20. MCAUSLAN, B. R., *Virology* **20,** 162–168 (1963).

21. MCAUSLAN, B. R., *Virology* **21,** 383–389 (1963).

22. MCAUSLAN, B. R., *In* "International Symposium on Enzymatic Aspects of Metabolic Regulation, Mexico City, 1966" (National Cancer Institute Monograph 27), pp. 211–219. DHEW, USPHS, NCI, Bethesda, Md., 1967.

23. MANIATIS, T., FRITSCH, E. F., and SAMBROOK, J., "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y., 1982.

24. MAXAM, A. M., and GILBERT, W., *In* "Methods in Enzymology" (L. Grossman and K. Moldave, eds.), Vol. **65,** pp. 499–560. Academic Press, New York, 1980.

25. MOSS, B., *In* "Comprehensive Virology" (H. Fraenkel-Conrat and R. R. Wagner, eds.), Vol. 3, pp. 405–474. Plenum, New York, 1974.

26. PROUDFOOT, N. J., and BROWNLEE, G. G., *Nature (London)* **252,** 359–362 (1974).

27. PUCKETT, C., and MOSS, B., *Cell* **35,** 441–448 (1983).

28. THOMAS, P. S., *Proc. Nat. Acad. Sci. USA* **77,** 5201–5205 (1980).

29. TU, C.-P. D., and COHEN, S. N., *Gene* **10,** 177–183 (1980).

30. VASSEF, A., BEN-HAMIDA, F., and BEAUD, G., *Ann. Virol. (Inst. Pasteur)* **134E,** 375–385 (1983).

31. WEIHER, H., KONIG, M., and GRUSS, P., *Science* **219,** 626–631 (1983).

32. WEIR, J. P., BAJSZAR, G., and MOSS, B., *Proc. Nat. Acad. Sci. USA* **79,** 1210–1214 (1982).

33. WEIR, J. P., and MOSS, B., *J. Virol.* **46,** 530–537 (1983).