VIROLOGY 211, 307–311 (1995)

SHORT COMMUNICATION

Herpes Simplex Virus Type 1 (HSV-1) Uracil–DNA Glycosylase: Functional Expression in *Escherichia coli*, Biochemical Characterization, and Selective Inhibition by 6-(*p*-*n*-Octylanilino)Uracil

R. ARGNANI,* F. FOCHER,† S. ZUCCHINI,* A. VERRI,† G. E. WRIGHT,‡ S. SPADARI,† and R. MANSERVIGI§[1]

*Institute of Microbiology and §Interdepartmental Center for Biotechnology, University of Ferrara, I-44100 Ferrara, Italy; †Istituto di Genetica Biochimica ed Evoluzionistica, CNR, I-27100 Pavia, Italy; and ‡Department of Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 10655.

Received February 13, 1995; accepted May 5, 1995

The Herpes simplex virus type 1 (HSV-1) uracil–DNA glycosylase (UDG) is encoded by the UL2 gene. The translation from the first putative start codon of UL2 predicts a polypeptide of 334 residues, while the translation from the second start codon predicts a polypeptide of 244 residues. We have cloned and expressed the two forms of UDG, by means of the prokaryotic expression vector pMAL-c2, and both of them were enzymatically active. Furthermore, the enzymatic properties of the recombinant UDGs and of the enzyme purified from HSV-1-infected cells were similar. The two UDG polypeptides have molecular weights of 27 and 37 kDa, respectively. The 37-kDa form of recombinant UDG is consistent with the reported molecular mass of 37 kDa for the enzyme purified from HSV-1-infected cells. Both recombinant UDGs were as sensitive as UDG purified from HSV-1-infected cells to 6-(*p*-*n*-octylanilino)uracil, the most potent of a series of uracil analogs that inhibit the viral enzyme.

Uracil in DNA arises either from cytosine deamination (*1*) or from dUTP misincorporation in place of dTTP during DNA synthesis (*2–4*). If not removed, in both cases uracil alters specific DNA/protein interactions (*4, 5*), and, if present in regulatory sequences (such as promoters, enhancers, origins of replication, etc.), it could affect gene expression, regulation, and DNA replication. Furthermore, uracil deriving from cytosine deamination is also highly mutagenic by resulting in GC → AT transition during DNA replication. Uracil is rapidly removed from DNA by the action of uracil–DNA glycosylase (UDG) which cleaves the N-glycosylic bond linking the uracil to the deoxyribose phosphate backbone (*6, 7*). This enzyme is highly represented in proliferating mammalian cells, and it is cell-cycle regulated, showing an increase during cell proliferation (*8, 9*). UDG is, thus, an essential component of the cellular DNA replication and repair machineries (*3*), and it is present in a wide range of organisms, including viruses of the Herpes (*10*) and Pox (*11*) families.

The sequences of UDGs show a high degree of conservation over broad evolutionary distances, with almost 50% identity at the amino acid level existing between *Escherichia coli* (*12*) and the dominant human form (*13*). Structural mutations in human UDG are associated with Bloom's syndrome (*14, 15*), a rare genetic inherited defect of DNA repair. Induction of UDG activity has been observed in cells infected with Herpes simplex virus type 1 (HSV-1) and type 2 (HSV-2) (*16*), and the UL2 gene encoding for HSV-1 and HSV-2 UDG has been identified and characterized (*10, 17*). Like its cellular counterpart, the viral enzyme acts to reduce the damage that is associated with the presence of uracil in DNA (*6*) and may also help to maintain regulatory sequences present in the HSV genome. In this regard, the presence of uracil in the HSV unique short origin of replication (OriS) has been demonstrated to alter the interaction of the origin-binding protein with its target sequence (*4*).

As it has been reported, there are two possible translation start codons for both HSV-1 and HSV-2 UL2 (*18*). The first ATG occurs at base 9886, and it is conserved between the two HSV serotypes, whereas the second ATG is at bases 10,123 in HSV-2 and 10,156 in HSV-1. The conserved ATG at base 9886 might serve as an initiator in both viruses (*19*). Translation beginning at the first start codon corresponds to a polypeptide of 334 amino acids with a predicted M_r of 36,300, whereas the second start codon would give a protein of 244 residues in HSV-1 and of 256 residues in HSV-2 of predicted M_r of 27,300 and 28,600, respectively. The DNA sequence beginning from the second start codon has a high degree

[1] To whom correspondence and reprint requests should be addressed at Center for Biotechnology, University of Ferrara, Via Fossato di Mortara 64, I-44100 Ferrara, Italy. Fax: (39)-532-247618.

0042-6822/95 $12.00

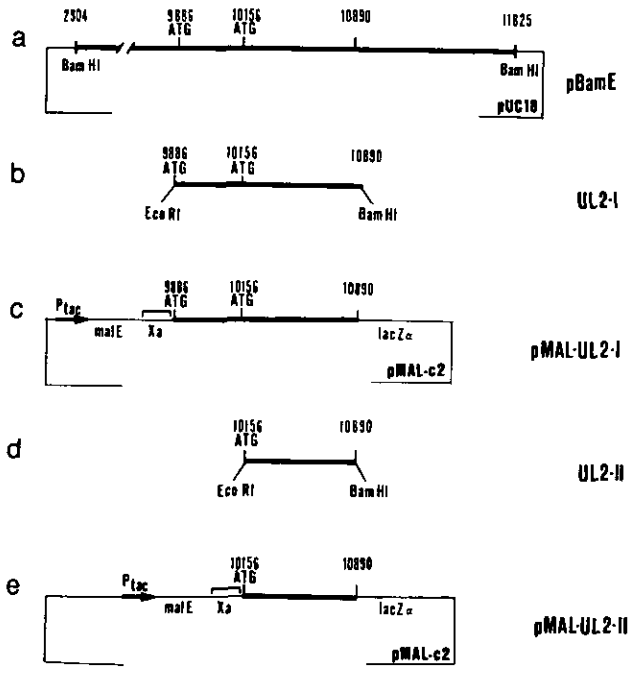

FIG. 1. Schematic diagrams of recombinant plasmids and PCR-amplified DNA fragments. Sequences deriving from HSV-1 are shown in heavy lines. (a) p*Bam*HI: pUC18 containing *Bam*HI E fragment of HSV-1 strain *F*. (b) UL2-I: PCR-amplified full coding sequence of UDG from p*Bam*E template, with *Eco*RI and *Bam*HI sites introduced, respectively, at 5′ and 3′ ends. (c) pMAL–UL2-I: UL2-I fragment inserted into pMAL-c2 cut with *Eco*RI–*Bam*HI. (d) UL2-II: PCR-amplified shorter coding sequence of UDG from p*Bam*E template, with *Eco*RI and *Bam*HI sites introduced by PCR primers. (e) pMAL–UL2-II: UL2-II fragment inserted into pMAL-c2 cut with *Eco*RI–*Bam*HI.

of homology with the UDG sequences from other evolutionary distant organisms (*13*).

In order to study the enzymatic function of HSV-1 UDG an expression system is needed yielding sufficiently large amount of the enzyme. The two DNA sequences corresponding to each start codon have recently been cloned using a prokaryotic expression vector (*20*). However, so far only the shorter sequence has been expressed and demonstrated to produce the active enzyme. We have therefore tried to generate a system capable of expressing both DNA sequences, and to demonstrate *in vitro* the activity of both the 334- and the 244-residue enzymes. To this purpose, the *Bam*HI *E* fragment (2907–11,825) cloned into pUC18 (Fig. 1a) was used as a PCR template to amplify the UL2 coding sequences.

The UDG coding sequences beginning from either of the two start codons were amplified by using the following three primers: (A) 5′ TCAGAATTCAAGCGGGCCTGCAGCCGA 3′; (B) 5′ GGGGGATCCCTTTCAAACCGACCAGTCGAT 3′; and (C) 5′ TCAGAATTCGATTTAACAAACGGGGGG 3′.

First start codon

The full coding sequence of UDG (UL2-I) (9885–10,890) was amplified with primers A and B which introduced *Eco*RI and *Bam*HI sites, respectively, at the 5′ and 3′ ends, and the UDG-I sequence was inserted into the pMAL-c2 vector which was cut with *Eco*RI/*Bam*HI. This yielded a construct (pMAL–UL2-I) in which the UL2-I sequence is located downstream from and in frame with the C-terminus of the maltose-binding protein (MBP)-encoding gene (*mal*E) as shown in Fig. 1c. The cloning of the UL2-I sequence results in the expression of the MBP–UDG-I fusion protein. A sequence coding for the recognition site of the specific protease Factor Xa and located at the 5′ end of the UL2-I coding sequence allows UDG-I to be cleaved from the MBP.

Second start codon

The shorter coding sequence of UDG (UL2-II), beginning at the second start codon (10,155–10,890), was amplified with primers C and B introducing *Eco*RI and *Bam*HI sites. This fragment was inserted into pMAL-c2 giving a construct named pMAL–UL2-II (Fig. 1e).

The full-length construct pMAL–UL2-I and the short construct pMAL–UL2-II were used to transform competent *E. coli* XL1 Blue cells, and several clones were screened for the presence of the insert. Two bacterial clones containing the pMAL–UL2-I and the pMAL–UL2-II recombinant DNAs were selected and grown overnight to saturation at 23 and 37°, respectively. The two cultures were then diluted 1/100 into LB broth modified with the addition of D-(+)-glucose, grown at 23 and 37°, respectively, to mid-logarithmic phase, and expression was induced by addition of IPTG at a final concentration of 1 m*M*. Cells were harvested after 4 hr of induction by centrifugation at 4000 *g* for 20 min at 4°. The pellet was resuspended in 50 ml of 20 m*M* Tris–HCl, pH 7.4, 200 m*M* NaCl, 1 m*M* EDTA (buffer A), lysed by sonication in short pulses of 15 sec for 3 min, and centrifuged at 9000 *g* for 30 min. The supernatants (crude extracts) containing the fusion proteins (MBP–UDG-I and MBP–UDG-II) were collected and the expression of the fusion proteins was tested on SDS–PAGE followed by staining with Coomassie blue. As shown in Fig. 2, two bands of molecular weight 79 (lane 3) and 69 (lane 7) kDa were clearly detected in lysates derived from IPTG-induced cells harboring pMAL–UL2-I and pMAL–UL2-II plasmids, respectively. These proteins were absent in cell lysates harboring the pMAL-c2 vector alone (lane 1). The size of these two bands was similar to that expected from the UDG-I (37 kDa) and the UDG-II (27 kDa) sequences linked to the MBP (42 kDa).

Because of the affinity of the MBP for amylose, the MBP–UDG-I and MBP–UDG-II fusion proteins were purified from the bacterial cell extracts by passage through an amylose resin column. After the adsorption of the fusion proteins to the amylose, the columns were washed with 10–12 vol of buffer A. Taking advantage of the stronger affinity of the MBP for maltose, the fusion

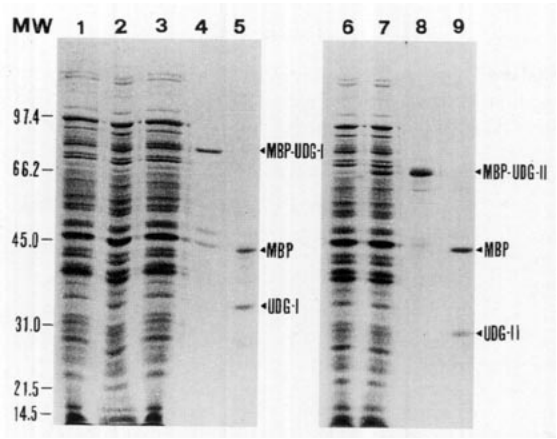

FIG. 2. SDS–polyacrylamide gel (15%) stained with Coomassie blue showing fractions from the purification stages of recombinant MBP–UDGs. Lane 1, control cells transformed with pMAL-c2 plasmid. Lane 2, uninduced pMAL–UL2-I cells. Lane 3, IPTG-induced pMAL–UL2-I cells. Lane 4, MBP–UDG-I fusion protein eluted from amylose affinity column with maltose. Lane 5, MBP and UDG-I proteins after cleavage with Factor Xa. Lane 6, uninduced pMAL–UL2-II cells. Lane 7, IPTG-induced pMAL–UL2-II cells. Lane 8, MBP–UDG-II fusion protein eluted from amylose affinity column with maltose. Lane 9, MBP and UDG-II proteins after cleavage with Factor Xa.

proteins were eluted with buffer A containing 10 m*M* D-(+)-maltose. The fractions were analyzed by SDS–PAGE after staining with Coomassie blue (Fig. 2, lanes 4 and 8) and those containing the fusion proteins were pooled and concentrated to about 1 mg/ml by centrifugation in Centricon-30 microconcentrators (Amicon Division).

To separate the UDG-I (650 U/mg) and the UDG-II (2000 U/mg) from the bacterial MBP, Factor Xa was added at a w/w ratio of 1% of the amount of the fusion protein, and the reaction mixtures were incubated for 48 hr at 4°. Fractions at different stages of purification were analyzed on SDS–PAGE. As shown in Fig. 2 (lanes 5 and 9), the cleavage of each of the two fusion proteins yielded two bands corresponding to the MBP (42 kDa) and the UDG-I (37 kDa) or the UDG-II (27 kDa), respectively.

Subsequently, UDG-II (0.25 mg/ml; approximately 2000 units) was purified from the fusion protein cleavage mixture by FPLC chromatography using a Mono S column (Pharmacia) as follows. Four milliliters of the cleavage mixture was dialyzed against 20 m*M* KPO_4, pH 7.2, 1 m*M* DTT, 0.5 m*M* EDTA, and 0.5 m*M* PMSF (buffer B) and then loaded on a Mono S column (1 ml). UDG-II was eluted with a linear gradient between 20 and 600 m*M* KPO_4 in buffer B at the flow rate of 0.5 ml/min. Proteins were separated on 12% SDS–PAGE and bands were stained with Silver (Merck). The UDG-II activity, assayed as described in (*21*), elutes as a single sharp peak at 230 m*M* KPO_4 (Fig. 3A) which mirrors the peak of protein eluted from the column (Fig. 3B). MBP is not retained by the column and it is recovered in the flowthrough (Fig. 3B).

Contrary to UDG-II, UDG-I is highly unstable and could not be further purified by FPLC chromatography. Since purified UDG-II shares all biochemical properties tested with the enzyme present in the fusion protein cleavage mixture, the biochemical properties of UDG-I and UDG-II reported here were compared by using their fusion protein cleavage mixtures. The pH profile of the UDG activity shows optimal activity occurring between pH 8 and 8.3; monovalent ions (K^+, Fig. 4A) inhibit the UDG activities. All three enzymes (1) retained their activity after incubation at 45° up to 40 min, (2) lost over 90% of activity after 5 min incubation at 65°, and (3) showed similar heat inactivation profiles when incubated at 55° (Fig. 4B). When the initial velocity of the enzymes was measured using various concentrations of uracil-containing native DNA and the results were plotted according to Lineweaver and Burk, the K_m varied between 2 and 7 μM (Table 1). Finally, we have studied the effect of several 6-(*p*-alkylanilino) uracils that we have found to inhibit the HSV-1 UDG (*21*). Both cloned UDGs were as sensitive as the purified HSV-1 to these analogs. Table 1 shows the IC_{50} values obtained with 6-(*p*-octylanilino)uracil (octyl-AU), the most potent inhibitor.

Savva and Pearl have cloned in an *E. coli* expression system sequences allowing translation from the two pu-

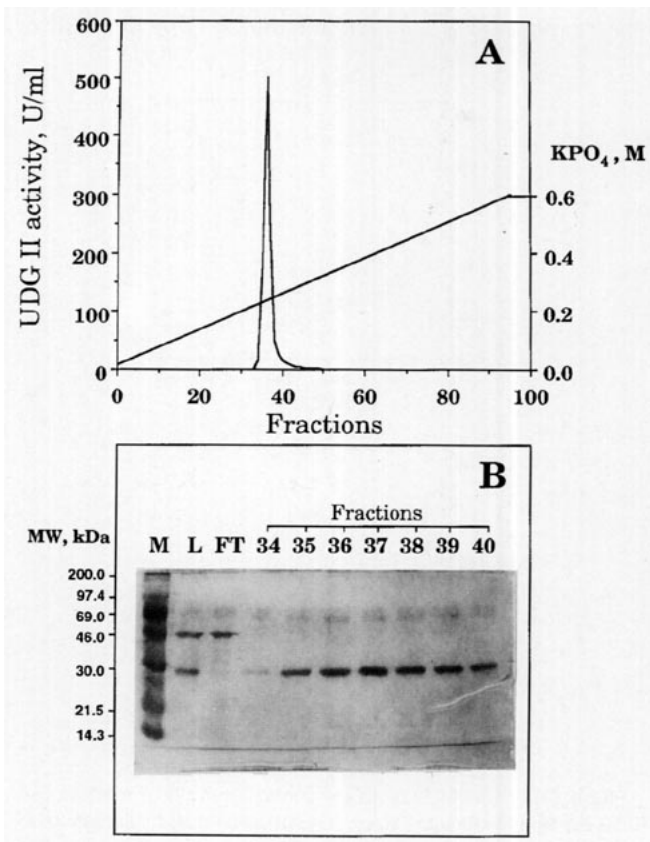

FIG. 3. (A) Elution profile of UDG-II from FPLC Mono S column. (B) SDS–PAGE (12%) of markers (M), loaded cleavage mixture (L), flowthrough (FT), and fractions 34–40. The bands were stained with silver stain (Merck).

tative start codons of HSV-1 UL2 (*20*), but were unable to express the full-length, presumably native form of UDG. Employing a different expression system based on the prokaryotic expression vector pMAL-c2 we have cloned and expressed both forms of UDG and both have been shown to be enzymatically active. The full-length form of the recombinant UDG (37 kDa) is consistent with the molecular mass of 37 kDa reported for the enzyme purified from HSV-1-infected cells (*21*). The same host–vector system is currently being employed for the expression of HSV-2 UDG necessary for comparison with HSV-1.

The availability of a large amount of recombinant UDG in both active and native forms will allow studies aimed to understand the role of the HSV-coded enzyme in DNA repair and replication as well as information concerning the three-dimensional structure of the protein.

Since insertion mutagenesis disrupting the UL2 gene does not prevent the growth of HSV in cell culture (*10*), it is unclear why the virus should have its own encoded UDG. Possibly, this is a requirement necessary for the *in vivo* virus–host interaction as suggested by recent research. We have in fact demonstrated that UDG is an enzyme no longer expressed in adult neurons (*3*) and we have hypothesized that viral UDG could play a role both in the "cleansing" of viral genome from uracils before the start of DNA replication and in the removal of misincorporated uracil residues during early steps of viral DNA replication that lead to the reactivation of the virus from latency (*3, 22*). This hypothesis received support by the Pyles and Thompson finding (*23*) that UDG$^-$ HSV-1 mutants display reduced neurovirulence and neuroinvasiveness and reactivate with reduced frequency *in vivo*. These data suggest that HSV-1 UDG activity is a valid target for antiviral chemotherapy. Agents targeting UDG activity should be able to reduce the frequency of recurrent disease. In this regard, we have recently reported the synthesis and characterization of several uracil analogues that specifically inhibit HSV-1 UDG (*21*), the first selective inhibitors of viral UDG that appear to act as substrate analogues. We have studied the effect of one of the most potent and selective of these analogues, 6-(*p-n*-octylanilino)uracil (octylAU), on the two forms of the recombinant UDGs. Our data show that both cloned UDGs were as sensitive as the enzyme purified from HSV-1-infected cells to this analogue. We are currently attempting to improve pharmacologically useful characteristics of the more promising of the above-mentioned compounds to facilitate both *in vivo* animal experimentation aimed at defining the role of UDG in latency reactivation and possible future administration as an antiviral drug.

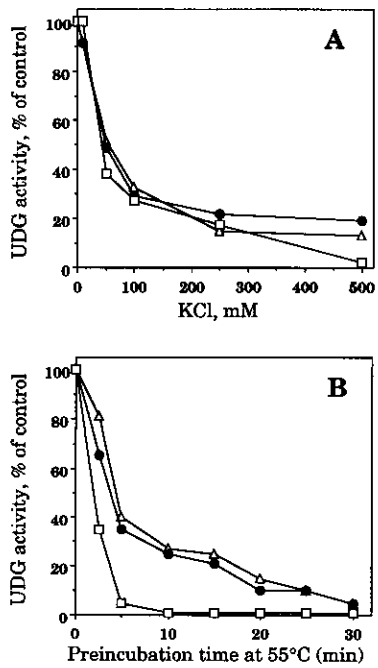

FIG. 4. (A) Effect of KCl on UDGs purified from HSV-1-infected cells (UDG-cp, ●) and cloned in *E. coli* and expressed as full-length (UDG-I, □) or short form (UDG-II, △). (B) Effect of temperature on UDGs purified from HSV-1-infected cells (UDG-cp, ●) and cloned in *E. coli* and expressed as full-length (UDG-I, □) or short form (UDG-II, △). Enzymes were preincubated at 55° for different times and then incubated in the assay mixture as described (*21*).

TABLE 1

K_m for the Substrate [^{3}H]uracil–DNA and IC_{50} of octylAU for UDGs Purified from HSV-1-Infected Cells (UDG-cp) and Those Cloned in *E. coli* and Expressed As Full-Length Form (UDG-I) or Short Form (UDG-II)

Enzyme	[^{3}H]uracil–DNA K_m (μM)	octylAU IC_{50} (μM)
UDG-cp	2.0	18 (±5)
UDG-I	5.6	37 (±6)
UDG-II	7.1	21 (±4)

ACKNOWLEDGMENTS

This work was supported by grants from National Research Council (C.N.R. — Target Projects "Genetic Engineering" and "A.C.R.O."), AIDS-ISS (to S.S.), and from Telethon, Grant A. 24. The authors gratefully acknowledge the support of a collaborative research grant from NATO.

REFERENCES

1. Lindahl, T., and Nyberg, B., *Biochemistry* **13**, 3405–3410 (1974).
2. Friedberg, E. C., "DNA Repair," pp. 141–212. Freeman, New York, 1985.
3. Focher, F., Mazzarello, P., Verri, A., Hübscher, U., and Spadari, S., *Mutat. Res.* **237**, 65–73 (1990).
4. Focher, F., Verri, A., Verzeletti, S., Mazzarello, P., and Spadari, S., *Chromosoma* **102**, S67–S71 (1992).
5. Verri, A., Mazzarello, P., Biamonti, G., Spadari, S., and Focher, F., *Nucleic Acids Res.* **18**, 5775–5780 (1990).
6. Lindahl, T., *Prog. Nucleic Acid Res. Mol. Biol.* **22**, 135–195 (1979).
7. Caradonna, S. J., and Cheng, Y. C., *Mol. Cell. Biochem.* **46**, 49–63 (1982).
8. Gupta, P. K., and Sirover, M. A., *Cancer Res.* **41**, 3133–3136 (1981).

9. Weng, Y., and Sirover, M. A., *Mutat. Res.* **293,** 133–141 (1993).
10. Mullaney, J., Moss, H.-W., and McGeoch, D.-J., *J. Gen. Virol.* **70,** 449–454 (1989).
11. Upton, C., Stuart, D. T., and McFadden, G., *Proc. Natl. Acad. Sci. USA* **90,** 4518–4522 (1993).
12. Varshney, V., Hutcheon, T., and van de Sande, J. H., *J. Biol. Chem.* **263,** 7776–7784 (1988).
13. Olsen, L. C., Aasland, R., Wittwer, C. U., Krokan, H. E., and Helland, D. E., *EMBO J.* **8,** 3121–3125 (1989).
14. Vollberg, T. M., Seal, G., and Sirover, M. A., *Carcinogenesis* **18,** 1725–1729 (1987).
15. Seal, G., Tallarida, R. J., and Sirover, M. A., *Biochim. Biophys. Acta* **1097,** 299–308 (1991).
16. Caradonna, S. J., and Cheng, Y. C., *J. Biol. Chem.* **256,** 9834–9837 (1981).
17. Caradonna, S. J., Worrad, D., and Lirette, R., *J. Virol.* **61,** 3040–3047 (1987).
18. McGeoch, D. J., Cunningham, C., McIntyre, G., and Dolan, A., *J. Gen. Virol.* **72,** 3057–3075 (1991).
19. Singh, J., and Wagner, E. K., *Virology* **196,** 220–231 (1993).
20. Savva, R., and Pearl, L. H., *J. Mol. Biol.* **234,** 910–912 (1993).
21. Focher, F., Verri, A., Spadari, S., Manservigi, R., Gambino, J., and Wright, G., *Biochem. J.* **292,** 883–889 (1993).
22. Mazzarello, P., Focher, F., Verri, A., and Spadari, S., *Int. J. Neurosci.* **50,** 169–174 (1990).
23. Pyles, R. B., and Thompson, R. L., *J. Virol.* **68,** 4963–4972 (1994).