

The A+T-rich genome of *Herpesvirus saimiri* contains a highly conserved gene for thymidylate synthase

(oncogenic γ herpesvirus/DNA sequencing/biased mutations/enzyme-substrate complex)

R. W. HONESS*, W. BODEMER†, K. R. CAMERON*, H.-H. NILLER†, B. FLECKENSTEIN†,
AND R. E. RANDALL*‡

*Division of Virology, National Institute for Medical Research, Mill Hill, London NW7 1AA, United Kingdom; and †Institut für Klinische Virologie, Universität Erlangen-Nürnberg, Löschgestrasse 8, 8520 Erlangen, Federal Republic of Germany

Communicated by Sydney Brenner, January 10, 1986

ABSTRACT *Herpesvirus saimiri* (HVS) is the prototype member of a distinctive subset of lymphotropic herpesviruses (the γ_2 subgroup) with A+T-rich coding sequences. In this paper, we show that cells productively infected with HVS contain high concentrations of a virus-specified thymidylate synthase (5,10-methylenetetrahydrofolate:dUMP C-methyltransferase, EC 2.1.1.45); we identify the active polypeptide and present the sequence of the virus gene. The predicted amino acid sequence of the 294-residue subunit of the virus enzyme is 70% homologous with the sequence of the human enzyme and about 50% homologous with prokaryotic thymidylate synthases, illustrating the remarkable structural constraints imposed by the thymidylate synthase function. However, the presence of the enzyme is not a conserved property of herpesviruses. We find no evidence for a virus-encoded thymidylate synthase activity (or a homology to a thymidylate synthase sequence) in G+C-rich representatives of α_1 (e.g., herpes simplex viruses, 66–68% G+C), β (i.e., human cytomegalovirus, 58–59% G+C), and γ_1 (i.e., Epstein-Barr virus, 60% G+C) herpesvirus subgroups. The production of excess thymidylate by a virus thymidylate synthase in cells infected with an A+T-rich herpesvirus would provide one plausible source of biased mutations by the virus-encoded replicative enzymes, which we have previously suggested as the likely general cause of differences in the mean nucleotide compositions of herpesvirus genomes.

Herpesviruses all have double-stranded, linear, DNA genomes of >100 kilobase pairs (kbp), but the mean nucleotide compositions of genomes from different members of the group range from 36% to 75% G+C. Herpesviruses also differ markedly in their biological properties, and a subclassification into α , β , and γ herpesvirus subgroups on the basis of gross differences in these properties seems useful and is consistent with more objective measures of relatedness (1, 2). Thus, the α herpesviruses are neurotropic viruses with either G+C-rich [e.g., herpes simplex viruses (HSV), 66–68% G+C; of the α_1 subgroup] or A+T-rich [e.g., varicella-zoster virus (VZV), 46% G+C; of the α_2 subgroup] coding sequences. The β herpesviruses are the salivary gland inclusion viruses or cytomegaloviruses (55–60% G+C), and the γ herpesviruses include the G+C-rich lymphotropic viruses of man and Old World monkeys [e.g., Epstein-Barr virus (EBV), 60% G+C; of the γ_1 subgroup] and the lymphotropic herpesviruses of New World monkeys and lower vertebrates, which have A+T-rich coding sequences [e.g., *Herpesvirus saimiri* (HVS), 36% G+C; of the γ_2 subgroup (3)]. Even within a subgroup, the sequences of homologous proteins can be widely divergent, and comparisons of the sequences of

HSV, VZV, and EBV, while revealing some conserved functions, also provide examples of coding sequences in one herpesvirus that have no recognizable counterpart in another (2, 4, 5). However, no correlation has yet been established between the presence or absence of a given gene and any subgroup-specific molecular or biological properties.

The virion DNA of HVS consists of a 110- to 112-kbp segment of unique sequences with a mean composition of 36% G+C, terminated by multiple tandem arrays of a 1.44-kbp, noncoding, repeat unit of 71% G+C (3). The virus naturally persists as an inapparent infection involving T lymphocytes of the squirrel monkey (*Saimiri sciureus*) but produces malignant lymphoproliferative diseases in other New World primates. Tumor tissues and cell lines derived from diseased animals or from *in vitro* immortalization experiments with HVS, bear T-cell markers and contain multiple copies of the virus genome as circular episomes (3). Transcripts arising from the rightmost 3.5 kbp of the *Kpn* I D fragment of virion DNA (3, 6) have been detected in such cell lines. In this paper we show that this region of the HVS genome contains the coding sequences for a highly conserved thymidylate (dTMP) synthase (5,10-methylenetetrahydrofolate:dUMP C-methyltransferase, EC 2.1.1.45) and that this enzyme is expressed at a high level in cells infected with HVS. We further show that a virus-specific dTMP synthase is not produced in cells infected with representatives of G+C-rich herpesvirus subgroups.

MATERIALS AND METHODS

DNA Sequencing, Sequence Assembly, and Analysis. The DNA sequence was determined independently both by directed DNA-sequencing of *Hind*III subclones of the *Kpn* I D fragment of HVS DNA (7) by the method of Maxam and Gilbert (8) and, as part of the determination of the complete sequence of the *Kpn* I D and E fragments (7) of HVS DNA, by the application of the Sanger (9) dideoxy chain-terminating sequencing method to random subfragments cloned into M13 mP18 using the methods developed and described in detail by Barrell and his colleagues (10). The sequence was assembled with the Staden (11) DB programs and analyzed with the Molecular Genetics and Sequencing (MGS) suite of programs developed and supported by B. Greer and P. Gillette (Computing Laboratory, National Institute for Medical Research).

Cells and Viruses. Conditions for the growth of uninfected Owl monkey kidney OMK-210 cells and for infections with HVS strain 11 and with HSV type 1 (HSV-1) have been

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HVS, *Herpesvirus saimiri*; HSV, herpes simplex virus; EBV, Epstein-Barr virus; VZV, varicella-zoster virus; kbp, kilobase pair(s); dTMP synthase, thymidylate synthase.

‡Present address: Department of Biochemistry, University of St. Andrews, St. Andrews, Fife KY16 9AL, Scotland, U.K.

described in detail elsewhere (12). For the experiment illustrated in Fig. 2A, cells were infected with 20 plaque-forming units per cell of HVS and harvested 40 hr later and with 10 plaque-forming units per cell of HSV-1 strain KOS and harvested 12 hr later. For the experiment illustrated in Fig. 2B, cultures were labeled with 10 μ Ci (1 Ci = 37 GBq) of [³⁵S]methionine per ml (>800 Ci/mmol; Amersham International, Bucks, U.K.) from 10 to 30 hr and were harvested at 30 hr.

Preparation of Cell Extracts and Active-Site Labeling of dTMP Synthase. Extracts of infected or uninfected cells were prepared by sonically disrupting washed (phosphate-buffered saline) cell suspensions (2×10^7 cells per ml) in a buffer containing 50 mM Tris-HCl (pH 7.5), 80 mM KCl, 10 mM 2-mercaptoethanol, 1 mM EDTA, and 0.1% Triton X-100 and then taking the fraction that remained soluble after sedimentation for 30 min at 30,000 rpm in a Beckman SW 41 rotor. [³²P]FdUMP (>5000 Ci/mmol) was synthesized from 5'-fluoro-2-deoxyuridine (Sigma) and [γ -³²P]ATP (>5000 Ci/mmol, Amersham International) using a partially purified preparation of thymidine kinase from HSV-1-infected cells and purified by preparative paper chromatography on Whatman 3 MM paper with 1-butanol/acetic acid/water, 2:1:1 (vol/vol), as the developing solvent. Reaction mixtures (0.1 ml) for the formation of the binary complexes (see text) contained the soluble extract from 10^4 to 10^6 cells (Fig. 2) with 1–2 μ Ci of [³²P]FdUMP in a solution containing 50 mM Tris-HCl (pH 7.4), 10 mM KF, and 20 mM 2-mercaptoethanol. Reaction mixtures for the formation of the ternary complexes (see text) also contained 0.23 mM 5,10-methylenetetrahydrofolate formed by the addition of excess (1.25 mM) formaldehyde to a 0.23 mM solution of tetrahydrofolic acid (*dl*-5,6,7,8-tetrahydropteroyl-L-glutamic acid; Sigma). Mixtures were incubated at 25°C in the dark for 1 hr and then either were denatured and analyzed by NaDodSO₄ gel electrophoresis or used as antigens in immunoprecipitation reactions.

In Vitro Translations, Immunoprecipitation, and Gel Electrophoresis. *In vitro* translations used 0.04 ml of a commercial nuclease-treated rabbit reticulocyte cell lysate (Anglian Biotechnology, Colchester, U.K.) programed with 2–4 μ g of poly(A)⁺ RNA prepared from uninfected or HVS-infected OMK cells by guanidinium isothiocyanate extraction and cesium chloride sedimentation (13), followed by oligo(dT)-cellulose chromatography (14). The preparation and properties of the antiserum against viral 28- to 30-kDa polypeptides, which we have called "anti-(28–30K)," and the monoclonal antibodies and their use in immunoprecipitation have been reported previously, as have the identification and properties of HVS-specified polypeptides (12, 15, 16). All samples were made to 2% NaDodSO₄, 5% 2-mercaptoethanol, 0.05 M Tris (pH 7.0), and 5% glycerol and were heated to 80°C for 10 min prior to electrophoretic separation on 15% polyacrylamide gels crosslinked with *N,N'*-diallyltartardiamide. Gels were fixed, stained, dried, and exposed to x-ray film for the periods indicated (Fig. 2).

RESULTS

DNA Sequence and Predicted Amino Acid Sequence of the HVS dTMP Synthase Gene. Analyses of the DNA sequence of the *Kpn* I D fragment of HVS DNA revealed that the rightmost 3.5 kbp contained a single, large, "leftward"-directed, open reading frame of 294 amino acids (Fig. 1a). A global search for sequences homologous to this virus reading-frame in the National Biomedical Research Foundation amino acid sequence data bank (Georgetown University, Washington, DC; 1984 edition) located highly significant homologies to sequences of the dTMP synthases of *Lactobacillus casei* and *Escherichia coli*. The significance of

these homologies was reinforced after the publication of the sequence of the human dTMP synthase gene (18). Thus, the predicted polypeptide product of the HVS reading frame was found to be some 70% homologous to the sequence of the human dTMP synthase gene and to exhibit the high degree of homology to prokaryotic enzymes previously noted (18) between the human and prokaryotic sequences (Fig. 1b). We know of no proven precedents for this activity encoded by a herpesvirus or any other eukaryotic virus. Therefore, we undertook to identify and characterize the virus gene product and to determine if a dTMP synthase were encoded by representatives of other subgroups of herpesviruses.

Direct Demonstration of a Functional dTMP Synthase in Cells Infected with HVS. Extensive studies of the structure and function of the dTMP synthase of *L. casei* have revealed a number of properties of the reaction intermediates that appear to be common to all such enzymes and that provide an unusually direct, sensitive, and highly specific means of detecting and identifying these proteins. Most usefully, FdUMP acts as a "dead-end" inhibitor and forms a binary complex with the enzyme so that ≈ 1 mol of FdUMP is bound per enzyme dimer to the active site cysteine (arrow in Fig. 1b; Cys-198 of the *L. casei* and Cys-146 of the *E. coli* sequences) and, in the presence of 5,10-methylenetetrahydrofolate, this is converted to a covalent ternary complex with about 2 mol of FdUMP per mol of enzyme dimer (24, 25). These complexes are stable to boiling in NaDodSO₄, and FdUMP has previously been shown to label and permit the identification of the single cellular dTMP synthase polypeptide in crude extracts of uninfected mouse and human cells separated by NaDodSO₄ gel electrophoresis (26). Fig. 2 summarizes results exploiting this assay to identify and characterize the HVS-specific dTMP synthase polypeptide.

Evidence That Representatives of G+C-Rich Herpesvirus Subgroups Do Not Encode dTMP Synthase. With the high-specificity substrate used in these experiments, we can readily detect the labeled HVS-specific polypeptide in extracts of 10^4 infected cells; under these same conditions we detect no labeled virus-specific polypeptide in extracts from up to 10^6 OMK (or rabbit kidney) cells infected with HSV-1 (e.g., Fig. 2a, compare lanes 4, 5, and 6), 10^6 Vero cells infected with pseudorabies virus (strain Dekking), or 10^6 MRC-5 cells infected with human cytomegalovirus (strain AD169, not shown). The electrophoretic mobilities of the dTMP synthase polypeptides of uninfected OMK and rabbit cells differ in a characteristic and reproducible way (e.g., Fig. 2c, lanes 10, 11, and 14), and the failure to observe any virus-specific band after infections of either cell type with HSV is therefore unlikely to be attributable to the coincidental comigration of virus-induced and host polypeptides. No significant homology to any of the known dTMP synthase amino acid sequences (Fig. 1b) was found in the 83 major reading frames predicted from the complete sequence of the human γ herpesvirus EBV (4). Given the remarkable conservation of the sequence of this enzyme, the absence of a homology constitutes good evidence for the absence of a functional dTMP synthase from the genome of the B95-8 strain of EBV. In contrast, the sequence of the VZV genome (46% G+C; α_2 subgroup) contains a highly conserved homologue to the HVS dTMP synthase sequence (A. J. Davison, personal communication), and enzyme-substrate complexes with a virus-specific dTMP synthase polypeptide are readily detected in extracts of cells infected with VZV (A. Davison and R.W.H., unpublished results).

Identification of the HVS dTMP Synthase Polypeptide *in Vivo* and *in Vitro*. We have shown that there are at least four independent virus-specific polypeptides with molecular masses in the range of 28–32 kDa (12, 15, 16) and that two of these polypeptides have precursors that also migrate with apparent masses in this range (i.e., 30–28 kDa and 31–29

a

-100
TAC TTT ATT TGC AGA TAA AAA TGA CAA AAT TTA AGT TTT ATG TTT TAT TTG TAT TAT GGA CAC TTT ATA CAG TAT ATT AAC TAA AAC TTG ACA AAT AAT AAG ATG TCA
M S
+1
I

ACA CAC ACA GAA GAA CAG CAT GGA GAA CAC CAG TAC CTC TCA CAG GTA CAG 60 CAC ATT TTA AAT TAT GGG TCA TTT AAA AAT 90 AGA ACA GGG ACT GGA ACA CTG AGT
T H T E E Q H 10 G E H Q Y L S Q V Q 20 H I L N Y G S F K N 30 D R T G T G T L S

120
ATT TTT GGA ACA CAG TCT AGG TTT AGT TTA GAA AAT 150 GAA TTT CCA CTT TTA ACT ACT AAG AGA 180 TTT TGG AGA GGT GTT GTT 210 TGG TTT ATC AGA
I F G T Q S R F S L E N E F P L L T T T K R A G V F W R G V E E L T W F I R
40 50 60 70 80 90 100 110

GGA TCT ACT GAC AGC AAA GAG CTA TCA GCA GCT GGC GTG CAC ATT 270 TGG GAT GCG AAT GGA TCT AGA TCA TTT TTA 300 GAT AAA CTT GGT TTT TAT GAC AGA GAT GAA GGA 330
G S T D S K E L S A A G V H I W D A N G S R S F L D K L G F Y D R D E G
80 90 100 110

GAT CTT GGA CCT GTG TAT GGA TTT CAG TGG AGG CAT TTC GGA GCA GAA TAT AAA GGT GTT 390 GGG CGC GAT TAT AAA GGA GAA GGA GTT 420 GAC CAA TTA AAA CAG TTG ATT
D L G P V Y G F 120 W R H F G A E Y K G V V G R D Y K G E G V D Q L K A Q L L I
130 140

450
GAT ACT ATA AAA ACA AAC CCT ACA GAT AGA AGG ATG TTA ATG 480 TGT GCT TGG AAT GTT TCA GAC ATT CCT AAA ATG GTG TTG CCG CCT TGT CAT GTA TTA AGT CAA TTT
D T I K T N P T D R R M L H C A W N V S D I P K M V L P P C H V L S Q F
150 160 170 180

TAT GTC TGT GAT GGA AAG CTT TCC TGT CAA CTG TAT CAA AGA TCA GCT GAT ATG GGG TTA GGA GTG CCA TTT AAC ATT GCT AGC TAT TCT CTT TTA ACT TGC ATG ATT
Y V C D G K L S C Q L Y Q R S A D H G L G V P F N I A S Y S L L T C M I
190 200 210

660
GCT CAT GTC ACT AAT CTA GTG CCT GGA GAG TTT ATC CAT ACT ATA GGA GAT GCT CAC ATC TAT GTA GAT CAC ATT GAT GCT CTT AAA ATG CAG TTG ACG CGA ACT CCT
A H V T N L V P G E F I H T I G D A H I Y V D H I D A L K M Q L T R T P
220 230 240 250

780
AGA CCG TTT CCA ACA CTT AGA TTT GCT AGA AAT GTT TCA TGC ATT GAT GAC TTT AAA GCA GAT GAC ATT ATA TTA GAA AAC TAT AAC CCA CAC CCT ATA ATT AAA ATG 870
R P F P T L R F A R N V S C I D D F K A D D I I L E N Y N P H P I I K M
260 270 280 290

900
CAT ATG GCT GTT TAA GAA AAA TAA CAT GTA AAA CGT ATC ATA AAC TAA AAC ACT TAC ATT GTA ACT ACA TTT TTC ATA GCT GTT TTA AAA AAT AAA AGT TTT AAA ACG
H M A V :
294

990 1000
TAT TTA CTT GTC TGT TGA ATG A

b

1 2 29 30 51
(A) M-----S T H T E E Q H G E L L Q Y L S Q V Q H I L L N Y G S F R K N-----D R T G T G T L S I F G T T Q S R F S L R E N E
(B) M P V A G S E L P R R P L P P A A Q E R D A E P R P P H G E L L Q Y L S Q V Q H I L L N Y G S F R K N-----D R T G T G T L S I F G T T Q S R F S L R E N E
(C) M-----L E L L Q Y L S Q V Q H I L L N Y G S F R K N-----D R T G T G T L S I F G T T Q S R F S L R E N E
(D) M-----K Q Y L S Q V Q H I L L N Y G S F R K N-----D R T G T G T L S I F G T T Q S R F S L R E N E
(E) M-----K Q Y L S Q V Q H I L L N Y G S F R K N-----D R T G T G T L S I F G T T Q S R F S L R E N E
(F) M-----K Q Y L S Q V Q H I L L N Y G S F R K N-----D R T G T G T L S I F G T T Q S R F S L R E N E

52 folate binding 82 83 109
(A) P L L T T K R V F V R G V V E E L L W F I R G S T D S K E L L-----S A A G V H W D A N G S R S F L D K L G F Y D R D E-----
(B) F P L L T T K R V F V R G V V E E L L W F I R G S T D S K E L L-----S A A G V H W D A N G S R S F L D K L G F Y D R D E-----
(C) F P L L T T K R V F V R G V V E E L L W F I R G S T D S K E L L-----S A A G V H W D A N G S R S F L D K L G F Y D R D E-----
(D) F P L L T T K R V F V R G V V E E L L W F I R G S T D S K E L L-----S A A G V H W D A N G S R S F L D K L G F Y D R D E-----
(E) F P L L T T K R V F V R G V V E E L L W F I R G S T D S K E L L-----S A A G V H W D A N G S R S F L D K L G F Y D R D E-----
(F) P L L T T K R V F V R G V V E E L L W F I R G S T D S K E L L-----S A A G V H W D A N G S R S F L D K L G F Y D R D E-----

110 140 165
(A)-----G D L G P V Y G F Q W R H F G A E Y K G V G R D Y K G E G V D Q L K Q L I D T I K T N P T D R R I M L M C A W N V
(B)-----G D L G P V Y G F Q W R H F G A E Y R D H E S D Y S G Q G V D Q L K Q R V I D T I K T N P T D R R I M L M C A W N V
(C) A V Y H E E M A K F D D R V L H D D A F A A K Y G D L G P V Y G F Q W R H F G A E Y R D H E S D Y S G Q G V D Q L K Q R V I D T I K T N P T D R R I M L M C A W N V
(D)-----G D L G P V Y G F Q W R H F G A E Y R D H E S D Y S G Q G V D Q L K Q R V I D T I K T N P T D R R I M L M C A W N V
(E)-----G E L L G P V Y G F Q W R H F G A E Y R D H E S D Y S G Q G V D Q L K Q R V I D T I K T N P T D R R I M L M C A W N V
(F)-----G T T L G H A Y G F Q L G K K N-----R S L N G E K V D Q I E V T D R T K K L N P T D R R I M L M C A W N V

166 FdUMP 190 200 220 245
(A) S D L P K M V L P P C H V L S Q F Y V C D G K L S C Q L Y Q R S A D M G L G V P F N I A S Y L L T C M I A H V T N L V P G E F I H T I G D A H I V D H I D A
(B) R D L P L M A L P P C H A L L C Q F Y V N S E L S C Q L Y Q R S A D M G L G V P F N I A S Y L L T Y M I A H I T G L K P G D F I H T L G D A H I Y L N H I E P
(C) E D V P T M A L P P C H T L Y Q F Y V N D G K L S C Q L Y Q R S A D I T F L G V P F N I A S Y L L T V H M A H E C L G L E V G E F I H T G G D A H T Y V N H H L D Q
(D) G E L D L P P C H A L F Y Q F Y V A D G K L S C Q L Y Q R S A D I T F L G V P F N I A S Y L L V H M A H A Q C C D L E V G D F V W T G G D T H T Y N H H V E Q
(E) A E L L K Y M A L P P C H M F Y Q F Y V R H G V L D L Q M V Y R S W D V F E L G L P F N I A S Y L L V H M A H A Q C C D L E V G D F V W T G G D T H T Y N H H V E Q
(F) D L D A M A L P P C V Y E T Q V Y V K Q G K L H L E V R A R S N D A L G N P F N V F Q V R V L V Q R T A Q V T G Y E L G E Y I F N I G D C H V Y T R H I D N

246 272 273 294 (294 residues)
(A) L K M Q L T R I P R P F P T L R F A R N V S C I D D F-----K A D G L I L E N Y N F P H P I K H H A V I. 313
(B) L K I Q L Q R E P R P P K L R L B K V E K I D D F-----K A E D Q L E G A Y P P H P T I K M E M A V I. 316
(C) I K E Q L S R P P P A P T L L Q L N P D K H D I F D F-----D M K D K L L N Y P P P A I K A P V A V I. 286
(D) T H L Q L S R E P R P P L K L I I K K P E S I F D Y-----R F D E I E G Y D P H P G I K A P V A V I. 286
(E) C K E I L R R E P K E L C E L V I S G L P Y K F R Y L S T K E Q L K Y V L K L R P K D F V I N N Y V S H P P L K G K H A V I. 286
(F) L K I Q M E R E Q F E A P E L W I N P E V K D F Y N F-----T V D D F K L I N Y K H G D K L L F E V A V I. 279

FIG. 1. (a) Nucleotide sequence and predicted amino acid sequence of the HVS dTMP synthase gene. The sequence shown is of the "message-sense" strand of a leftward reading frame located between 100.5 and 99.6 kbp on the 111.0-kbp unique sequence component of the HVS genome. The HindIII site indicated at nucleotides 562-567 is that separating the 1284-bp and 921-bp HindIII subfragments of Kpn I D fragment (7, 17). (b) Alignments of the predicted amino acid sequence of the HVS dTMP synthase (line A) with sequences of human (18) (line B) and of prokaryotic [Lactobacillus casei, thyA (19) (line C); Escherichia coli, thyA (20) (line D); coliphage T4, thyA (21) (line E); and Bacillus subtilis phage 3T, thyP (22) (line F)] dTMP synthase polypeptides. Residues that are identical in three or more sequences in the alignments shown are boxed and residues that are identical in all six sequences are also underlined. Gaps introduced to give these alignments are indicated by dashes. The conserved cysteine that forms a covalent bond with FdUMP in the binary complex (19, 20) and the region implicated as the binding site for folate and its analogues in the ternary complex (23) are indicated by arrows. The numbering is with respect to the predicted amino acid sequence of the HVS polypeptide.

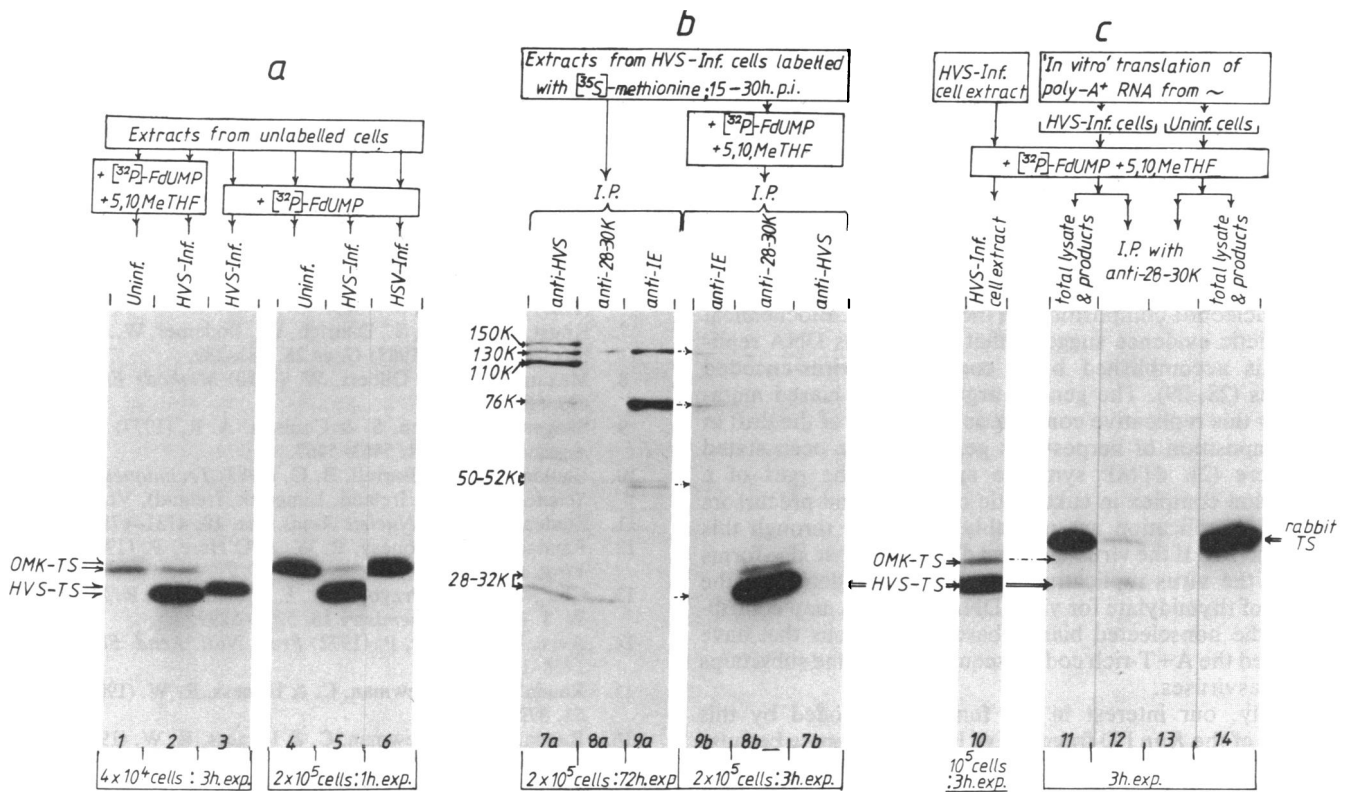


FIG. 2. Identification of the HVS dTMP synthase (abbreviated here TS) polypeptide in extracts of infected cells (a and b) and in the products of *in vitro* translation programmed by poly(A)⁺ RNA from infected cells (c). The figure shows autoradiograms of polypeptides (designated by their size or size range in kDa, abbreviated "K") separated by NaDodSO₄/polyacrylamide gel electrophoresis from extracts of infected (HSV-Inf.) and uninfected (Uninf.) OMK-210 cells labeled *in vitro* with [³²P]FdUMP to give the covalent binary complexes or in the presence of 5,10-methylenetetrahydrofolate (5,10-MeTHF) to give the ternary complexes. h, hour(s); exp., exposure; p.i., postinfection; I.P., immunoprecipitation. (a) Separations of binary and ternary complexes from HVS-infected cells readily detect novel virus-specific complexes (HVS-TS) with sizes of 28–30 kDa in addition to the host enzyme complexes (OMK-TS). Virus-induced and cellular polypeptides are converted to slightly faster migrating forms and are labeled at higher specific activity in the presence of 5,10-methylenetetrahydrofolate in addition to [³²P]FdUMP (lanes 2 and 3). No virus-specific band is detected in extracts from cells infected with HSV (lane 6). (b) The virus-specific TS is completely precipitated by a monoclonal antibody, polyclonal, rabbit antiserum to the HVS-specified 28-kDa and 30-kDa polypeptides ("anti-28–30K"; lanes 8a and 8b). The HVS-TS is not precipitated by a hyperimmune, polyprecipitin antiserum to HVS-specified polypeptides, which reacts with many virus-specified polypeptides, including others in the 28- to 32-kDa range ("anti-HVS"; lanes 7a and 7b), nor is it precipitated by a range of monoclonal antibodies, including one directed against the 52-kDa immediate-early protein of HVS ("anti-IE"; lanes 9a and 9b). Prominent virus-specified polypeptides represented in the immunoprecipitates from [³⁵S]methionine-labeled extracts are annotated with their molecular mass (150–28 kDa) as described previously (12). (c) Poly(A)⁺ RNA from infected cells specifically directs the synthesis of a polypeptide that reacts with the active site-directed label (lanes 11 and 14), is precipitated by the "anti-(28–30K)" antiserum (lane 12), and comigrates with the virus-specific complex formed with extracts of infected cells (compare lanes 10, 11, and 12).

kDa). A panel of monoclonal antibodies and a number of polyclonal, monoclonal antisera have been useful in resolving some of this complexity (12, 15, 16); therefore, we tested our previously characterized antisera for their ability to precipitate the dTMP synthase polypeptide of HVS. A monoclonal rabbit antiserum against the 28- to 30-kDa virus polypeptide specifically and completely precipitated the virus dTMP synthase polypeptide (Fig. 2b, lanes 8a and 8b), whereas a polyprecipitin serum and monoclonal antibodies that react with other virus polypeptides in this mass range did not precipitate the virus dTMP synthase (e.g., anti-HVS; Fig. 2b, lanes 7a and 7b). The antiserum against the viral 28- to 30-kDa polypeptide did not precipitate the dTMP synthase activity from lysates of uninfected cells but it did precipitate the host enzyme from infected cells (Fig. 2b; lane 8b; see also Fig. 2c, lanes 12 and 13). We currently assume that this is due to the formation of a proportion of dimers of virus- and host-specified polypeptides. The active-site-directed label also permits detection of the *de novo* synthesis of the virus enzyme in a reticulocyte lysate programmed with poly(A)⁺ RNA from infected cells (Fig. 2c; lanes 11 and 12), although no synthesis of the OMK cell enzyme was detected in experiments with RNA from infected or

uninfected cells (lanes 11 and 14). The antiserum against the viral 28- to 30-kDa polypeptide precipitates the virus enzyme and the endogenous reticulocyte enzyme in approximately equimolar amounts from lysates programmed with infected cell RNA but does not precipitate the reticulocyte enzyme in the absence of the virus enzyme (lanes 12 and 13), consistent with the formation of mixed dimers *in vitro* as well as *in vivo*.

DISCUSSION

The presence of an active, highly conserved, gene for dTMP synthase in representatives of A+T-rich herpesviruses and the absence of this function from representatives of G+C-rich herpesviruses has a number of implications. Firstly, the sequence of the HVS enzyme is clearly more closely related to the human sequence than it is to prokaryotic sequences. However, the HVS enzyme displays the same high degree of homology to the prokaryotic sequences as does the human sequence, and the HVS and human sequences are at least as similar to the bacteriophage sequences as the bacteriophage sequences are to each other or to the sequences of the genes from *E. coli* and *L. casei*. Despite this unusually high degree of structural conservation, differential inhibition of the *E. coli*

and coliphage T2 enzymes has been observed with some folate analogues (27). Therefore, virus-encoded dTMP synthase genes may represent a useful additional target for selective chemotherapy of infections with A+T-rich herpesviruses. In particular, the use of folate analogues to inhibit dTMP synthase may increase the effectiveness of those thymidine analogues that enter nucleotide pools after phosphorylation by virus-encoded thymidine kinases.

Secondly, the present results provide a specific example of a difference in the genetic repertoire of herpesviruses that may correlate with some of the marked differences in the mean nucleotide compositions of their genomes. Biochemical and genetic evidence suggests that herpesvirus DNA replication is accomplished by a complex of virus-encoded enzymes (28, 29). The general arguments for biased mutations by this replicative complex as the source of the drift in the composition of herpesvirus genomes have been stated elsewhere (2). dTMP synthase appears to be part of a replication complex in eukaryotic cells (30), and precursors for DNA replication are probably channeled through this complex (31). If the virus-encoded dTMP synthase also forms part of the virus replicative complex, its influence on the supply of thymidylate for virus DNA synthesis may contribute to the nonselected bias in base substitutions that have produced the A+T-rich coding sequences of some subgroups of herpesviruses.

Finally, our interest in the functions encoded by this portion of the *Kpn* I D fragment of HVS DNA arose because of the detection of transcripts complementary to this region of the virus genome in a lymphoid tumor cell line (6) and in preparations of RNA from cells infected in the presence of cycloheximide (17). However, subsequent analyses have mapped the major immediate-early transcription unit elsewhere (within the *Msp* I H/*Hind*III G fragment, 25–29 kbp from the “left” end of the coding sequences of HVS DNA) and have shown that the major transcript crossing the dTMP synthase reading frame is the product of an abundantly transcribed late promoter in productively infected cells (E. Smith and R.W.H., unpublished data). We have shown here that the HVS dTMP synthase polypeptide is specifically immunoprecipitated by the monoprecipitin antiserum against the viral 28- to 30-kDa polypeptide, and we previously reported that both the 30-kDa nuclear precursor and the 28-kDa cytoplasmic product which are precipitated by this serum are late gene products (12). The significance of the detection of transcripts from this region of the HVS genome in lymphoblastoid cell lines is therefore uncertain. The availability of the sequence of this region of the virus genome should facilitate the construction of viruses with alterations to the dTMP synthase-encoding sequences in order to evaluate the role of this function in infection and transformation.

R.W.H. thanks B. G. Barrell and K. Weston for an introduction to their methods of “shotgun” cloning and DNA sequencing and Drs. R. Staden, J. Green, B. Greer, and P. Gillett for the necessary computer programs and support for the compilation and analysis of DNA sequences.

1. Roizman, B. (1982) in *The Herpesviruses*, ed. Roizman, B. (Plenum, New York), Vol. 1, pp. 1–23.
2. Honess, R. W. (1984) *J. Gen. Virol.* **65**, 2077–2107.
3. Fleckenstein, B. & Desrosiers, R. C. (1982) in *The Herpesviruses*, ed. Roizman, B. (Plenum, New York), Vol. 1, pp. 253–332.
4. Baer, R., Bankier, A. T., Biggin, M. D., Deininger, P. L., Farrell, P. J., Gibson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Seguin, C., Tuffnell, P. S. & Barrell, B. G. (1984) *Nature (London)* **310**, 207–211.
5. Davison, A. J. (1983) *EMBO J.* **2**, 2203–2209.
6. Knust, E., Dietrich, W., Fleckenstein, B. & Bodemer, W. (1983) *J. Virol.* **48**, 377–383.
7. Knust, E., Schirm, S., Dietrich, W., Bodemer, W., Kolb, E. & Fleckenstein, B. (1983) *Gene* **25**, 281–289.
8. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
9. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
10. Bankier, A. T. & Barrell, B. G. (1983) *Techniques in the Life Sciences* (Elsevier Ireland, Limerick, Ireland), Vol. B5.
11. Staden, R. (1982) *Nucleic Acids Res.* **10**, 4731–4751.
12. Randall, R. E., Honess, R. W. & O’Hare, P. (1983) *J. Gen. Virol.* **64**, 19–35.
13. Chirgwin, J. M., Przybyla, A. E., Macdonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
14. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
15. Randall, R. E., Newman, C. & Honess, R. W. (1984) *J. Virol.* **52**, 872–883.
16. Randall, R. E., Newman, C. & Honess, R. W. (1984) *J. Gen. Virol.* **65**, 1215–1219.
17. Bodemer, W., Knust, E., Angermuller, S. & Fleckenstein, B. (1984) *J. Virol.* **51**, 452–457.
18. Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. & Seno, T. (1985) *Nucleic Acids Res.* **13**, 2035–2043.
19. Maley, G. F., Bellisario, R. L., Guarino, D. U. & Maley, F. (1979) *J. Biol. Chem.* **254**, 1301–1304.
20. Belfort, M., Maley, G., Pedersen-Lane, J. & Maley, F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 4914–4918.
21. Chu, F. K., Maley, G. F., Maley, F. & Belfort, M. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3049–3053.
22. Kenny, E., Atkinson, T. & Hartley, B. S. (1985) *Gene* **34**, 335–342.
23. Maley, G. F., Maley, F. & Baugh, C. M. (1982) *Arch. Biochem. Biophys.* **216**, 551–558.
24. Santi, D. V., McHenry, C. S. & Perriard, E. R. (1974) *Biochemistry* **13**, 467–470.
25. Moore, M. A., Ahmed, F. & Dunlap, R. B. (1984) *Biochem. Biophys. Res. Commun.* **124**, 37–43.
26. Ayusawa, D., Iwata, K., Seno, T. & Koyama, H. (1981) *J. Biol. Chem.* **256**, 12005–12012.
27. Maley, G. F., Maley, F. & Baugh, C. M. (1979) *J. Biol. Chem.* **254**, 7485–7487.
28. Vaughan, P. J., Banks, L. M., Purifoy, D. J. M. & Powell, K. L. (1984) *J. Gen. Virol.* **65**, 2033–2041.
29. Chiou, H. C., Weller, S. K. & Coen, D. M. (1985) *Virology* **145**, 213–226.
30. Reddy, G. P. & Pardee, A. B. (1983) *Nature (London)* **304**, 86–88.
31. Reddy, G. P. & Pardee, A. B. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3312–3316.