Identification of Important Residues within the Putative Nucleoside Binding Site of HSV-1 Thymidine Kinase by Random Sequence Selection: Analysis of Selected Mutants *in Vitro*†

Margaret E. Black and Lawrence A. Loeb*

Joseph Gottstein Memorial Cancer Research Laboratory, Department of Pathology, SM-30, School of Medicine, University of Washington, Seattle, Washington 98195

Received July 6, 1993; Revised Manuscript Received August 23, 1993®

ABSTRACT: Random sequence mutagenesis in conjunction with genetic complementation was used to map the function of amino acid residues within the putative nucleoside binding site of the herpes simplex virus type 1 (HSV-1) thymidine kinase (TK). Six codons of the putative nucleoside binding site of the HSV-1 tk were substituted by a duplex of extended oligonucleotides containing 20% random sequences. Approximately 260 mutants were screened for the ability to genetically complement a TK-deficient *Escherichia coli*. Of those screened, 32% conferred TK activity. Approximately 60% of the TK positive clones contained single amino acid changes, 23% contained double changes, and 13.4% encoded the wild-type TK amino acid sequence. A small percentage of clones, 2.4% and 1.2%, contained triple or quadruple alterations, respectively. Three residues (D162, H163, and R164) appeared to be highly conserved especially with regard to the type of residues able to substitute. Secondary screening results indicated that several of the mutants had higher affinities for acyclovir and/or 3′-azido-3′-deoxythymidine than thymidine in complementation assays. In addition, a number of clones were unable to form colonies on selection medium at elevated temperatures (42 °C). Eight selected mutants were subcloned into an *in vitro* transcription vector and the derived transcripts used to program a rabbit reticulocyte lysate cell-free translation system. Biologically active translation products were then analyzed *in vitro* for thymidine kinase activity, for thermal stability, and for the ability to phosphorylate selected nucleoside analogues. Two of the eight mutants had an elevated thymidine kinase activity, two were significantly thermolabile, and three exhibited enhanced efficiency in phosphorylation of nucleoside analogues.

Thymidine kinase (TK, EC 2.7.1.21)[1] is a key enzyme in the salvage pathway of nucleoside metabolism, catalyzing the transfer of the γ-phosphate from ATP to thymidine to produce dTMP. The exploitation of HSV-1 thymidine kinase as an effective target for antiherpetic drugs is largely due to the differences in substrate specificity between the viral and the host (human) TK enzymes. The development of such drugs utilizes the ability of HSV-1 TK to phosphorylate guanosine analogues such as acyclovir and gancyclovir that the cellular enzyme is incapable of phosphorylating. A knowledge of the molecular basis of this difference in substrate specificity, *i.e.*, which residues or structures are involved in substrate binding and/or phosphorylation, could be key to rational drug design. This is an especially important question in light of the recent use of the HSV-1 thymidine kinase gene in gene therapy studies. In one such study, the HSV-1 tk gene was introduced into artificially induced tumor cells via retroviral vectors, and these cells were selectively killed in animals by administration of high levels of gancyclovir (Culver et al., 1992).

In an alignment of 12 *Herpesviridae* thymidine kinases, Balasubramaniam et al. (1990) identified 6 highly conserved regions and designated them sites 1–6. Two of these domains, sites 3 and 4 (HSV-1 TK residues 162–164 and 171–173, respectively), were suggested as comprising the putative nucleoside binding site either alone or in a combined fashion. One of these domains (residues 165–177 in HSV-1 TK, inclusive of site 4) was the focus of previous random sequence selection studies by Munir et al. (1992) that identified the three residues of site 4 as essential for TK activity. In the present study, we have identified a consensus sequence based on the type of residues at positions 155–165, determined the essentiality of site 3 and flanking residues using mutants obtained by random sequence selection, and assessed the role of individual mutant thymidine kinases produced *in vitro* with regard to substrate specificity, thermal stability, and thymidine phosphorylation.

MATERIALS AND METHODS

Bacterial Strains. *Escherichia coli* strain KY895 (F⁻, *tdk*⁻, 1-*ilv*), originally described by Igarashi et al. (1967), was used in the genetic complementation assays for thymidine kinase activity. *E. coli* strain NM522 (F′ *lacI*^{q} Δ (*lacZ*)M15 *pro*AB/ *sup*E *thi* Δ (*lac* proAB)Δ(*hsd*MS-*mcr*B)5(r_k^- McrB⁻)) (NEB, Beverly, MA) was used as a recipient in all subcloning experiments. Helper phage VCM13 (Stratagene, La Jolla, CA) was used in the production of single-stranded phage for sequencing.

Materials. L-[^{35}S]Methionine/cysteine (specific activity, 1140 Ci/mmol) for protein synthesis determination and [*methyl*-^{3}H]thymidine (specific activity, 87 Ci/mmol) were purchased from Amersham. Other radioisotopes [[*side chain*-2-^{3}H]acyclovir (specific activity, 28.6 Ci/mmol) and [5-^{3}H]-deoxycytidine (specific activity, 29 Ci/mmol)] were purchased from Du Pont—New England Nuclear (Boston, MA), and

† This work was supported by grants from the National Institutes of Health, an OIG grant (R35-CA3990), a postdoctoral research fellowship (1F32CA09384-01), and a postdoctoral training grant (T32CA09437).

* To whom correspondence should be addressed. Telephone: (206) 543-0557. FAX: (206) 543-3967.

® Abstract published in *Advance ACS Abstracts,* October 1, 1993.

[1] Abbreviations: HSV-1, herpes simplex virus type 1; TK, thymidine kinase; tk, thymidine kinase gene; ACV, acyclovir; GCV, gancyclovir; AZT, 3′-azido-3′-deoxythymidine.

0006-2960/93/0432-11618$04.00/0

[8-^{3}H]gancyclovir (specific activity, 22 Ci/mmol) and [*methyl*-^{3}H]-3′-azido-3′-deoxythymidine (specific activity, 14 Ci/mmol) were from Moravek (Brea, CA). Restriction endonucleases and T4 DNA ligase were purchased from New England Biolabs (NEB). Promega (Madison, WI) was the source of the *in vitro* transcription and translation reagents except for the cap analogue, ^{7}m(5′)Gppp(5′)G, which was purchased fron NEB. Oligonucleotides used for sequencing and polymerase chain reaction amplifications were obtained from Operon (Alameda, CA). Other chemicals were purchased from Sigma (St. Louis, MO) except where designated.

Construction of the Random Sequence Insert. The two oligonucleotides used in random sequence selection were synthesized by American Synthesis, Inc. (Pleasanton, CA): MB110 (70mer) is 5′-TGGGAGCTCACATGCCCCGCCC-[CCG]GCCCTCACCCTCATC[TTCGACCGCCATCCC]-ATCGCCGCCCTCCTG-3′; MB111 (38mer) is 5′-ATGAG-GTACCGCGCAGCTGGGTAGCACAGGAGGGCGGC-3′, where the nucleotides in brackets were synthesized as 80% wild-type nucleotide, 20% the other three nucleotides. At the 5′ end of MB110 is an *Sac*I restriction site, and at the 5′ end of MB111, a *Kpn*I site is present. These restriction sites were utilized at a later step after second-strand synthesis occurred. Furthermore, as an internal control, a *Pvu*II site was introduced (silent change) in MB111 in order to allow confirmation of random sequence insertion prior to sequencing. Twelve nucleotides at the 3′ ends of each oligonucleotide are complementary to allow for hybridization of the two strands to each other. Each oligonucleotide was subjected to electrophoresis on a 20% acrylamide–urea gel and visualized by UV-shadowing on a PEI–cellulose TLC plate (Baker, Phillipsburg, NJ), the portion of the gel containing the correct sized oligonucleotide was excised, and the oligonucleotide was eluted from the gel in 0.5 M NH_4OAc/10 mM $MgOAc_2$ overnight at 37 °C. The eluted oligonucleotide was then ethanol-precipitated and resuspended in H_2O. An OD_{260} measurement was taken, and the extinction coefficient for each oligo was used to determine the concentration.

Equimolar amounts of MB110 and MB111 (25 pmol) were annealed in a small volume (20 μL) in 1 × annealing buffer [10 × annealing buffer = 70 mM Tris (pH 7.5)/60 mM $MgCl_2$/200 mM NaCl] for 5 min at 95 °C, then moved to 65 °C for 20 min, and slowly cooled to room temperature. To the annealed oligonucleotides (20 μL) were added 2 μL of 10 × annealing buffer, 2.8 μL of 10 mM dNTPs, 0.8 μL of 0.1 M dithiothreitol (DTT), 2.4 μL of DNA polymerase I Klenow fragment (5 units/μL), and H_2O to bring the volume to 40 μL. The mixture was placed at 37 °C for 30 min, at 65 °C for 10 min, and finally at room temperature for 10 min. Verification of fully extended radioactive oligonucleotides was done by subjecting samples to denaturing acrylamide gel electrophoresis and autoradiography. Amplification of the extended products was performed using the polymerase chain reaction with Taq polymerase (Stratagene). The 100-μL reactions contained 20 mM Tris (pH 8.3)/25 mM KCl/1.5 mM $MgCl_2$/0.05% Tween 20/0.1 mg/mL BSA/50 μM of each of the four deoxynucleoside triphosphates (dNTPs)/22 pmol of PCR primer 1/20 pmol of PCR primer 2/2 units of Taq polymerase and 6 pmol of the extended random oligonucleotide; primer 1 = 5′-TGGGAGCTCACATGCCCC-GCC-3′ and primer 2 = 5′-ATGAGGTACCG-3′. One drop of mineral oil was added to each tube which was then placed in a Perkins Elmer—Cetus thermal cycler (Norwalk, CT) and programmed for 30 cycles of 95 °C for 1 min and 34 °C for 2 min. At the end of the 30 cycles, the reactions were left at 72 °C for 7 min, and then the cycler was maintained at 4 C. After confirmation of amplification by 2% agarose gel electrophoresis, the product-containing reactions were pooled, precipitated, and digested with *Kpn*I and *Sac*I. Doubly restricted fragments were distinguished from single cut or uncut fragments on nondenaturing acrylamide gels, and the appropriate fragment was excised and isolated as described above.

Library Vectors. pMCC contains the wild-type HSV-1 tk open reading frame and was the parent plasmid for the construction of pMDC ("dummy" vector) that contains a stuffer fragment in the *Kpn*I–*Sac*I sites. pMCC and pMDC were a gift from Khan Munir and, the construction of which, are described in Munir et al. (1992). The vector was derived from pBR322 and contains, in addition to HSV-1 TK sequences, a λPr promoter, sequences for the temperature-sensitive λ repressor (cI857), and β-lactamase [to confer ampicillin (carbenicillin) resistance]. The original plasmid, pHETK2, is described by Waldman et al. (1983).

Library Construction and Selection. Cesium chloride gradient purified pMDC ("dummy" vector) (Munir et al., 1992) was digested with *Kpn*I and *Sac*I restriction endonucleases and gel-isolated from a 1% agarose/1 × TBE gel using GenClean II (Bio101, La Jolla, CA). This vector was ligated with the gel-isolated PCR-amplified random fragment overnight at 16 °C with 1 unit of T4 DNA ligase. The ligated mixture was then used to transform KY895 by electroporation (BioRad gene pulser, 2 kV, 25 μF, 400 Ω). Cells were prepared for electroportion according to BioRad (Richmond, CA). After each pulse, 1 mL of SOC (2% Bactotryptone/0.5% yeast extract/10 mM NaCl/2.5 mM KCl/10 mM $MgCl_2$/10 mM $MgSO_4$/20 mM glucose) was added to the cuvette and the electroporation mixture transferred to a 25-mL snap-cap Falcon tube. After the tubes were shaken for 1 h at 37 °C, the cells were plated onto LB [per liter: 10 g of tryptone/5 g of yeast extract/10 g of NaCl (pH 7)] containing carbenicillin (50 μg/mL) plates (LB + carb^{50}) and the plates incubated at 37 °C overnight. The number of colonies was counted, and they were picked with a toothpick and streaked on TK selection [2% BBL Trypticase peptone (Becton Dickenson, Cockeysville, MD)/0.5% NaCl/0.8% Gel-Rite (Scott Laboratories, Carson, CA)/0.2% glucose/50 μg/mL carbenicillin/10 μg/mL 5′-fluorodeoxyuridine/2 μg/mL thymidine/12.5 μg/mL uridine] and LB + carb^{50} plates. The basis of this selection is that 5′-fluorodeoxyuridine (FUdR) is phosphorylated by thymidine kinase to form FdUMP, an inhibitor of the *de novo* pathway enzyme, thymidylate synthase. The requirement for dTMP can then be fulfilled only by an active thymidine kinase. Uridine is supplied to inhibit thymidine phosphorylase. After 16–24 h, the TK selection plates were scored for growth and any positives picked and restreaked on TK selection plates and LB + carb^{50} plates to confirm the phenotype. The lower limit of TK activity that allowed colonies to be scored as positives is described in Dube et al. (1991).

DNA Isolation and Sequencing. Mutant DNA was isolated from overnight cultures grown in 2 × YT (per liter: 16 g of tryptone/10 g of yeast extract/5 g of NaCl) + carb^{50} using the Promega Magic miniprep kit according to the manufacturer except that 3 mL of culture was used per isolation because of the low copy number of the plasmid. Ten microliters of each dsDNA was alkaline-denatured, precipitated, and resuspended in Sequenase reaction buffer, H_2O, and sequencing primer (5′-CATGCCTTATGCCGTGA-3′). The primer was then annealed and the DNA subjected to dideoxy sequencing (Sanger et al., 1977) using Sequenase according to the

manufacturer's instructions (USB, Cleveland, OH). The region introduced (*Kpn*I to *Sac*I site) and limited flanking regions were sequenced and any mutations noted.

Secondary Screening Selection. Log-phase cultures of TK positive clones were serially diluted in 0.9% NaCl and spread onto acyclovir or AZT plates (TK selection plates except 1 μg/mL thymidine + 1 μg/mL acyclovir or 0.05 μg/mL AZT). Mutant cultures were also spread onto duplicate TK selection and LB + carb^{50} plates. One set of TK selection plates and LB + carb^{50} plates were incubated at 42 °C. All other plates were incubated at 37 °C. After 16–24 h, the plates were scored.

Subcloning of Selected Mutants. DNAs of selected clones were restricted with *Mlu*I and *Bss*HII to release a 1.07 kbp fragment [nucleotide numbers ~335 and 1400 on the McKnight sequence (1980)]. The fragments were gel-isolated from 1% agarose gels using GenCleanII and ligated to pT7: HSVTKII vector DNA which had been restricted with *Mlu*I and *Bss*HII, treated with calf intestinal alkaline phosphatase, and gel-isolated. pT7:HSVTKII was derived from pT7: HSVTK transcription vector described by Black and Hruby (1992). pT7:HSVTKII differs from pT7:HSVTK only by the loss of an *Nco*I–*Bam*HI fragment 3′ to the end of the HSV-1 tk gene which was originally used to aid in the initial cloning of the tk gene. No difference in activity was observed from products derived from the two vectors. The entire open reading frame of each mutant clone subcloned into the *in vitro* vector was sequenced.

In Vitro Transcription and Translation. Production of mutant and wild-type mRNAs by *in vitro* transcription was described previously (Wilson et al., 1989). The amount of protein synthesis was determined by the amount of TCA-precipitable counts. Cell-free translation was according to Promega using nuclease-treated rabbit reticulocyte lysates. Analysis of protein synthesis levels was according to Black and Hruby (1990). Thymidine kinase activity assays were performed as described by Hruby and Ball (1981) except that 0.044 μCi of [*methyl*-^{3}H]thymidine was used per reaction, and all washes were done at room temperature. The TK assays are a filter binding assay which monitors the conversion of [^{3}H]thymidine to labeled dTMP. For the enzyme assays using tritiated thymidine, deoxycytidine, gancyclovir, acyclovir, and azidothymidine, 48 μmol of each substrate was used. All assays were performed in triplicate for the designated period of time. Unless otherwise stated, all assays were performed at 30 °C. Normalization of enzyme assays to protein synthesis levels was as according to Black and Hruby (1990, 1992).

RESULTS

Alignment of Herpes TK Site 3 Residues. The amino acid sequences from 13 *Herpesviridae* family thymidine kinases were aligned in the region spanning residues 155–165 of HSV-1 TK. A consensus sequence based on the HSV-1 TK sequence (McKnight, 1980) was derived from this aligment and is shown in Figure 1. The DRH (amino acids 162–164 where D is aspartic acid, R is arginine, and H is histidine) motif (site 3; Balasubramaniam et al., 1990) is completely conserved in all 13 sequences. The proline at position 165 is highly conserved in 11 out of the 13 sequences while the phenylalanine residue at position 161, just N-terminal to the core conserved motif, appears conserved primarily with respect to hydrophobicity. This appears to also be the case for L159 and I160. (To specify a particular residue, the single amino acid designation and the position within the HSV-1 TK polypeptide will be the nomenclature used, *e.g.*, L159 for leucine at residue 159.)

								Site 3			
HSV-1 TK	$P_{(9)}$	$A_{(3)}$	$L_{(4)}$	$T_{(6)}$	$L_{(7)}$	$I_{(4)}$	$F_{(6)}$	D	R	H	$P_{(11)}$
(residue position)	155	156	157	158	159	160	161	162	163	164	165
Type of residue	O	(-)	O	I	O	O	O	(-)	(+)	(+)	O
Number of occurrences (out of 13)	10	7	10	8	13	13	11	13	13	13	12
Other types of residues	2 I 1 (+)	4 O 2 I	2 I 1 (+) 1 (-)	5 O			1 I 1 (+)				1 I

FIGURE 1: Amino acid type consensus sequence derived from an alignment of 13 *Herpesviridae* family thymidine kinases spanning site 3. The amino acid sequence shown in the top row is that of HSV-1 TK with the number of times the particular residue shown is present in other members of this family (in subscript). Site 3 (Balasubramaniam et al., 1990) residues are in the shaded box and comprise residues 162–164: D, R, and H (D, aspartic acid; R, arginine; H, histidine). The second row indicates the type of residue found in the majority of sequences where O = hydrophobic, I = hydrophilic, (+) = positively charged, or (–) = negatively charged residues. In the third row, the number of times the type of residue shown in the second row was found is given. Below the third row are the number and types of other residues found at each position. Sequences used include herpes simplex virus types 1 (McKnight, 1980) and 2 (Swain & Galloway, 1983), marmoset herpesvirus (Otsuka & Kit, 1984), varicella zoster virus (Davison & Scott, 1986), feline herpes virus (Nunberg et al., 1989), pseudorabies virus (Kit & Kit, 1985), equine herpesvirus type 1 (Roberson & Whalley, 1988), bovine herpesvirus type 1 (Mittal & Field, 1989), turkey herpesvirus (Martin et al., 1989), Marek's disease virus (Scott et al., 1989), herpes saimiri (Honess et al., 1989) Epstein-Barr virus (Baer et al., 1984), and thymidine kinase from infectious laryngotracheitis virus (Griffin & Boursnell, 1990).

Except for P155, the other residues (A156, L157, T158) are less conserved since different types of residues are found at these positions. At the N-terminus of this alignment, a proline is present in 9 out of the 13 sequences compared and may display a structural function.

Random Sequence Selection Library. To address the importance of particular residues used in the alignment, random sequence mutagenesis and genetic complementation were employed as outlined in Figure 2. The codons for P155, F161, D162, R163, H164, and P165 in the HSV-1 tk gene were targeted for mutagenesis. At each of the 18 nucleotides targeted for substitution, the wild-type nucleotide was introduced in the oligonucleotide synthesis at an 80% frequency, and the other 3 nucleotides were introduced 20% of the time. Approximately 260 random transformants were screened for their ability to complement KY895, a TK-deficient *E. coli* on TK selection media. Of these, 82 were scored as positives and sequenced. Therefore, approximately 32% of all transformants encoded functional enzymes.

Single and Multiple Amino Acid Mutations. Eleven of the clones encoded the wild-type amino acid sequence (13.4%) with seven of these containing the wild-type nucleotide sequence. Three clones with wild-type residues contained single nucleotide changes (all different), and one contained three nucleotide changes. A total of 49 TK positive clones containing single amino acid changes (59.8%) were identified (Figure 3A). Nineteen double amino acid mutations (23.2%), two triple (2.4%), and one clone containing four amino acid changes (1.2%) were identified. The amino acid sequences of clones with multiple alterations are shown in Figure 3B.

Secondary Screening. Thirty-five representative clones were further characterized. The ability of pMCC (KY895) and the 35 log-phase mutant pMDC (KY895) cultures to produce colonies on acyclovir or AZT plates was determined as described under Materials and Methods. Colony formation on these plates suggests that *E. coli* harboring these plasmids

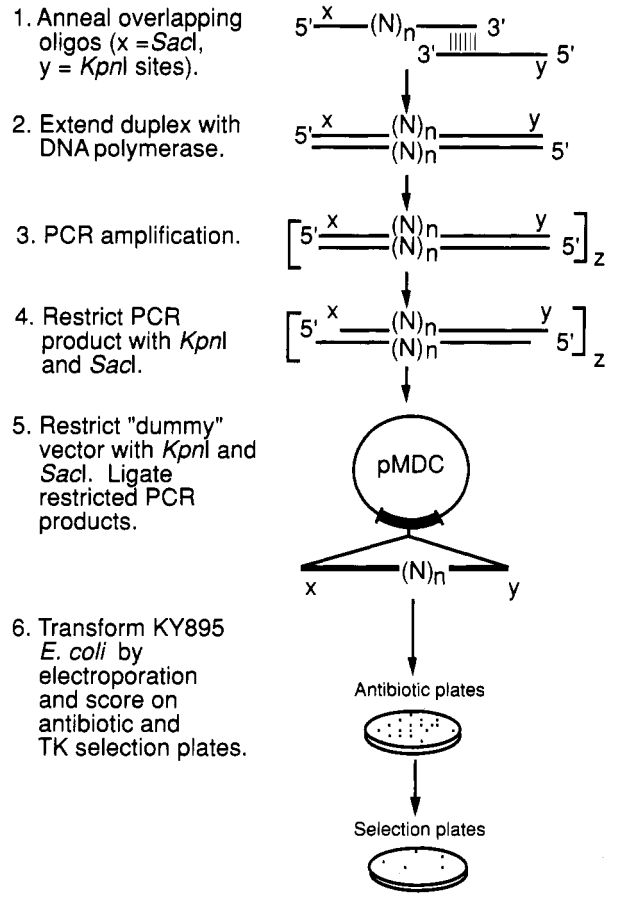

FIGURE 2: Outline of the random sequence selection procedure. The restriction sites *Sac*I and *Kpn*I are designated by x and y, respectively. N represents random nucleotides used in the oligonucleotide synthesis where 80% of the wild-type nucleotides were used and 20% of the other three nucleotides. The small, subscripted n designates the number of positions in which random nucleotides were introduced and, in these experiments, was 18.

A

Wild Type Sequence	O **P** 155	O **F** 161	(-)I **D** 162	(+)I **R** 163	(+)I **H** 164	O **P** 165
Substitutions at Each Position	3 L 2 A 2 T 1 Q 1 R	4 I 4 Y 3 C 2 L 1 S	5 E 1 G	5 C 1 S	3 N 1 T	3 L 2 T 2 S 1 N 1 A
Types of Substitutions	11% (+) 33% I 56% O	57% I 43% O	83% (-)I 17% I	100% I	100% I	10% (+) 50% I 40% O

B

Number of changes	P 155	F 161	D 162	R 163	H 164	P 165
Doubles	A	V				
	Q	I				
	Q		E			
	R		E			
(4)	R		G			
	T		E			
(2)		I		H		
		I				R
		N				S
			Y	C		
			N		K	
(2)			E		N	
				P	Q	
					Q	L
Triples	Q		E			L
	A			P		T
Quadruple			N	S	N	A

FIGURE 3: (A) Single amino acid changes in the putative nucleoside binding site of TK positive mutants obtained by random sequence selection. The wild-type HSV-1 TK amino acids mutated are given in the boldface box with the residue number and the type of residue found in the majority of sequences [O = hydrophobic; I = hydrophilic; (+) = positively charged; (-) = negatively charged residues]. Below the wild-type residue are the number of times a particular amino acid substitution was found. In the bottom section, the percentages of each type of residue found are listed. (B) Occurrence of multiple mutations within the putative nucleoside binding site of HSV-1 TK. The wild-type amino acids and their positions in the HSV-1 TK polypeptide are indicated at the top of the figure. Double, triple, and quadruple amino acid substitutions are shown in the respective categories. If a set of mutations was identified more than once, the number of occurrences is noted on the left in parentheses.

are able to phosphorylate thymidine in preference to phosphorylating acyclovir and/or AZT. However, mutant cells able to more efficiently phosphorylate the nucleoside analogues than thymidine will be unable to form colonies since incorporation of phosphorylated ACV or AZT would inhibit DNA synthesis. All cultures formed colonies on control TK selection and LB + carb[50] plates. The ability to form colonies on acyclovir or AZT plates is shown in Figure 4. In comparison to the wild-type, several mutants appeared to preferentially utilize one or both nucleoside analogues over thymidine (P155A/F161V, F161I, F161C, and R163P/H164Q).

All cultures were also tested for their ability to complement the TK-minus KY895 on TK selection plates at the elevated temperature of 42 °C. Several mutants were unable to form colonies on TK selection plates at 42 °C (F161L and R163P/H164Q), and one (F161I/R163H) showed a severely reduced ability to form colonies at 42 °C (Figure 4).

Expression of Mutant Enzymes in a Cell-Free Translation System. In order to study the properties of the mutant TKs, the 1.07 kbp *Mlu*I–*Bss*HII fragment of each of the eight mutants was subcloned into the *in vitro* vector pT7:HSVTKII as described under Materials and Methods. The derived transcripts were used in a rabbit reticulocyte lysate cell-free translation system to synthesize active enzymes. Expression of full-length proteins was analyzed by subjecting ^{35}S-radiolabeled cell-free translation products to SDS–PAGE and autoradiography (described under Materials and Methods). An autoradiograph of the mutants selected for further characterization is shown in Figure 5. The major radiolabeled translation product from each mutant transcript migrates during electrophoresis as a ~43-kDa protein with the same electrophoretic mobility as that observed with translation products from wild-type pT7:HSVTKII transcripts. To quantitate the level of protein synthesis for each translation, determination of trichloroacetic acid precipitable counts from each of the same samples was performed in triplicate (data not shown). The amount of acid-precipitable counts roughly parallels the band intensity of each mutant in Figure 5.

pMCC versus pT7:HSVTKII. In the final sequence analysis of the eight mutant fragments subcloned into the pT7:HSVTKII vector, two additional amino acid differences were identified between these tk genes. The sequence of pT7:HSVTKII is exactly the same as that published by McKnight (1980). pMCC, the parental plasmid of pMDC and hence the vector into which the random sequences were ligated, contains two amino acid aberrations from the McKnight

Clones	ACV	AZT	LB	37C	42C
pMCC (wild-type)	++	++	++	++	++
P155A/F161V	++	+	++	++	++
F161I	+	+	++	++	++
F161C	+	-	++	++	++
F161L	++	++	++	++	-
R163P/H164Q	+	+	++	++	-
F161I/R163H	++	++	++	++	+
pMDC	-	-	++	-	-

FIGURE 4: Ability of mutants to grow on ACV or AZT selection plates or at 42 °C on TK selection plates. Thirty-five cultures were grown to mid-log phase, diluted, and spread onto plates containing either ACV or AZT as described under Materials and Methods. An additional set of TK selection plates was also spread with each culture and scored for growth at 42 °C. Only mutants that gave results which differed from those observed with the wild-type pMCC (KY895) are shown. The mutants are designated with the wild-type residue and position number followed by the amino acid substitution deduced from the nucleotide sequence; *e.g.*, F161I indicates that isoleucine replaces phenylalanine at residue 161 in this particular mutant. ++ = same number of colonies observed as on control plates; + = fewer (<20% those observed with pMCC) and generally smaller (~50% smaller diameter) colonies; (–) = no colonies.

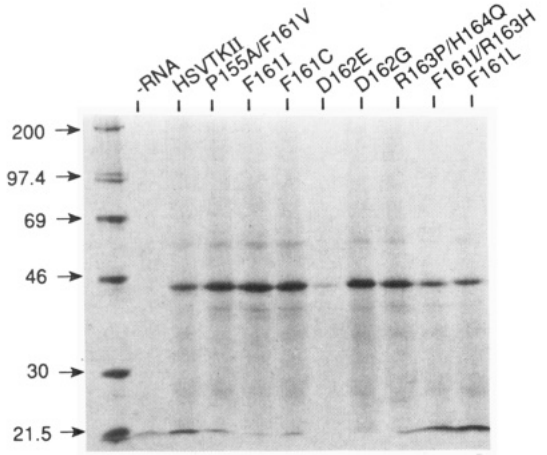

FIGURE 5: Autoradiograph of ^{35}S-radiolabeled cell-free translation products subjected to SDS–PAGE and TCA-precipitable counts. One microliter of each radiolabeled cell-free translation of *in vitro*-derived mutant mRNAs was subjected to SDS-containing polyacrylamide (12%) gel electrophoresis. An autoradiograph of this gel is shown. The first lane contained ^{14}C-labeled rainbow molecular weight markers (Amersham) with the apparent molecular weight (× 10^{-3}) given on the left. The second lane corresponds to a cell-free translation performed in the absence of any added mRNA. The third lane corresponds to the wild-type pT7:HSVTKII mRNA translation product. All other lanes contained translation products of the mutant mRNAs produced *in vitro* as described under Materials and Methods.

sequence. These are at position 434 (C→T) and 575 (G→A) and result in a proline-49 to leucine and an arginine-89 to glutamine change. Therefore, all mutants contain these two mutations in addition to those described. In addition, a single nucleotide difference at position 480 (C→T) was also identified but does not result in an amino acid change. Because all *in vitro* analyses were compared against pT7:HSVTKII as the wild-type, the *Mlu*I–*Bss*HII fragment from pMCC was subcloned into the corresponding sites of pT7:HSVTKII (now designated pT7:MCC), and the subsequent cell-free translation products were compared to those derived from pT7:HSVTKII. Time course and thermal stability analyses showed no significant difference between pT7:HSVTKII- and pT7:MCC-derived translation products (data not shown). No significant difference in phosphorylation efficiency was observed between pT7:MCC and pT7:HSVTKII when thymidine (1.3-fold), deoxycytidine (1.3-fold), GCV (0.8-fold), ACV (0.95-fold), or AZT (1.1-fold) were used as substrate (data not shown). Furthermore, Sanderson et al. (1988) reported that the K_m for thymidine and ATP and the V_{max} of TK purified from *E. coli* harboring pHETK2 (the parent plasmid of pMCC) and HSV-1-infected cells were indistinguishable. Therefore, the alterations observed in the properties of the mutant TKs can be attributed to the nucleotide substitutions within the target region and that any differences between the vectors (pT7:MCC and pT7:HSVTKII) exerted only minor changes in catalytic properties.

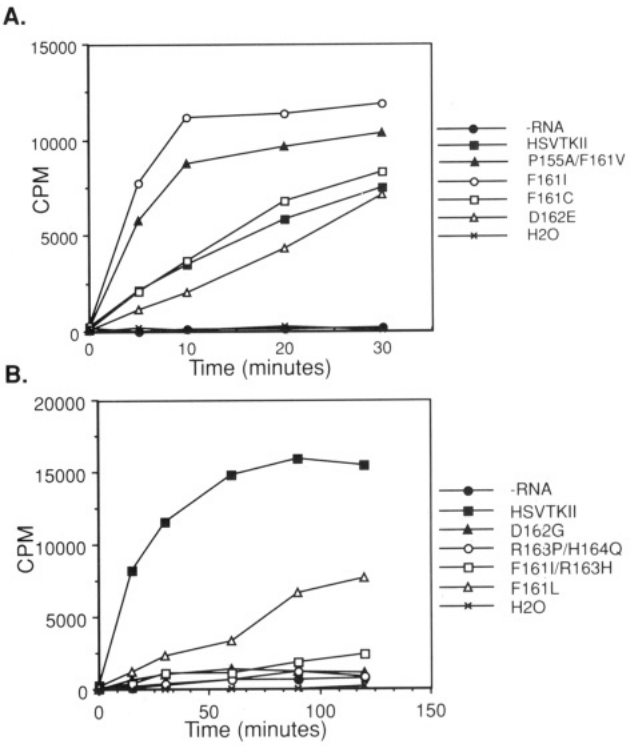

FIGURE 6: Time course analysis of high (A) or low (B) activity mutants produced in the rabbit reticulocyte lysate cell-free translation system. (A) TK assays of cell-free translation products of the high-activity mutants diluted 1/9 in distilled H_2O were performed in triplicate for 0, 5, 10, 20, and 30 min. The average of each time point was adjusted to reflect protein synthesis equivalent to wild-type protein synthesis levels. (B) In a similar fashion to (A), TK assays were performed in triplicate on 1/5 dilutions of the low-activity mutants with time points taken at 0, 15, 30, 60, 90, and 120 min intervals. Adjusted values are plotted.

Time Course Analysis of Mutant Enzymes. On the basis of TK activities, mutant TKs were classified into two subsets: (1) high-activity mutants (F155A/F161V, F161I, F161C, and D162E); (2) low-activity mutants (F161I/R163H, F161L, D162G, and R163P/H164Q). For the high-activity mutant enzymes, unlabeled translation products were diluted 1/9 and incubated for 0, 5, 10, 20, or 30 min at 30 °C (Figure 6A). The TK activity results (counts per minute) were adjusted to reflect equivalent protein synthesis levels using the corresponding TCA-precipitable counts (^{35}S cpm). Standard deviations for all time points were calculated (data not shown). Two of the mutants (F161I and P155A/F161V) demonstrated a statistically higher affinity for thymidine than the wild-type TK. Standard deviations of F161C and D162E activities (data not shown) indicate no difference in activities when compared to the wild-type TK enzyme activity. The low-activity mutants were diluted 1/5, and the rate of phosphorylation as a function

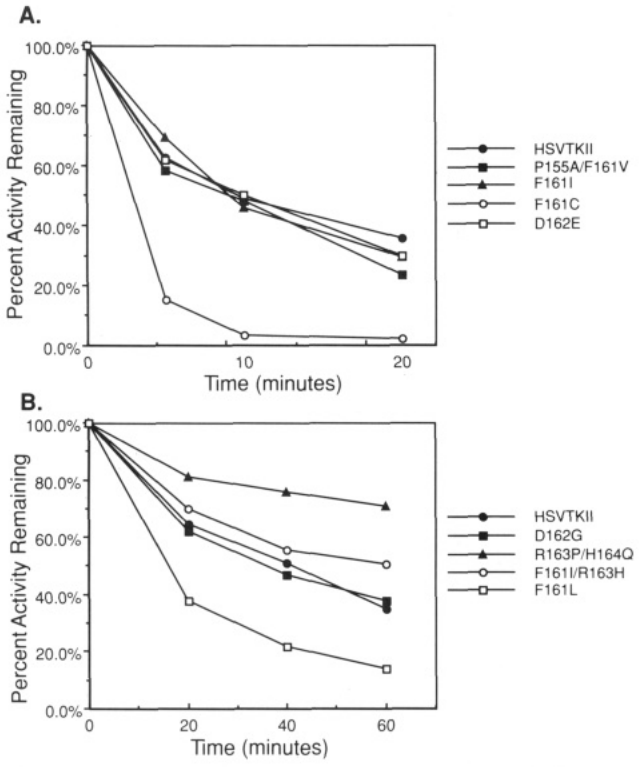

FIGURE 7: Thermal stability of high (A) or low (B) activity mutants. (A) CFT products of each high-activity mutant, –RNA, and HSVTKII samples were diluted 1/9 and incubated for 0, 5, 10, and 20 min at 42 °C. The preincubated samples were then assayed for 5 min (P155A/F161V and F161I) or 20 min (–RNA, HSVTKII, F161C, and D162E). The percent of activity remaining was determined with the untreated samples set at 100%. (B) Low-activity mutant CFT products were diluted 1/5 and incubated for 0, 20, 40, or 60 min at 42 °C. The preincubated samples were then assayed in triplicate for thymidine phosphorylation for 60 min. The percent of activity remaining was determined using the untreated (time 0) sample as 100%.

of time was determined (Figure 6B). The time course analysis indicates that most of the mutants had less than 10% wild-type activity. One, F161L, however, demonstrated a moderate ability to phosphorylate thymidine albeit at a much reduced rate from HSVTKII.

Thermal Stability Assays. In the assays for colony formation on TK selection plates, several mutants were unable to complement KY895 at 42 °C, suggesting that these mutant TKs were temperature-sensitive. To substantiate this observation, cell-free translation products were incubated at 42 °C for increasing times prior to being assayed for enzyme activity. Standard deviations for all time points were calculated (data not shown). Except for F161C, all high-activity mutants displayed thermal stabilities similar to HSVTKII after 42 °C preincubation periods as long as 60 min (data not shown). Because F161C lost greater than 90% of enzyme activity within the first 20 min at 42 °C, shorter incubation periods at 42 °C were performed (0, 5, 10, and 20 min). F161C was exceptionally thermolabile, demonstrating a ~85% activity loss after only 5 min at 42 °C (Figure 7A). Among the low-activity mutant subset, one translation product (F161L) was more thermolabile that HSVTKII (Figure 7B). Others in this set (R163P, F161I/R163H, H164Q, and D162G) were equivalent to HSVTKII.

Substrate Specificity Assays. Three of the mutants (P155A/F161V, F161I, and F161C) were assayed for their ability to phosphorylate equivalent concentrations of tritiated thymidine, deoxycytidine, ACV, GCV, or AZT. These

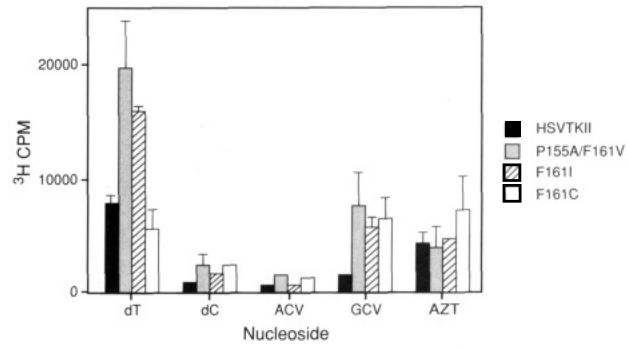

FIGURE 8: Phosphorylation of nucleosides and nucleoside analogues by mutant and wild-type HSV-1 thymidine kinases. Cell-free translation products of HSVTKII, MCC, P155A/F161V, F161I, and F161C were assayed in triplicate for the relative levels of phosphorylation using thymidine, deoxycytidine, ACV, GCV, or AZT as substrates. Forty-eight micromoles of each tritiated substrate was used in each assay reaction. Translation products were diluted for each nucleoside assay as follows (translation/H_2O): 1/100, thymidine; 2/3, deoxycytidine, GCV, and AZT; 4/1, ACV. Each set of assays was incubated for 2 h at 30 °C and the amount of phosphorylated product determined. The counts per minute were adjusted as described in Figure 6 and plotted. Standard deviations are shown as error bars.

mutants indicated a preference for acyclovir and/or AZT over thymidine in the secondary screening assays (Figure 4). R163P/H164Q was not tested due to a very low level of activity when thymidine was used as a radiolabeled substrate. The relative ability of each of these mutants, with respect to HSVTKII, to phosphorylate these compounds is shown in Figure 8. Both P155A/F161V and F161I displayed an elevated capacity to phosphorylate thymidine relative to HSVTKII, 2.6- and 2.2-fold, respectively. Phosphorylation of deoxycytidine by the mutant enzymes ranged from 1.9- to 2.8-fold over the wild-type enzyme (F161I, 1.9-fold; F161C, 2.8-fold; P155A/F161V, 2.8-fold). Two mutants appeared to share an increased ability to phosphorylate ACV (2.4- and 2-fold over HSVTKII by F155A/F161V and F161C, respectively). All mutants demonstrated approximately wild-type levels of AZT phosphorylation. All mutants assayed appeared to share a large increase in GCV phosphorylation at 3.9–5.2-fold compared to wild-type phosphorylation levels.

DISCUSSION

Although the HSV-1 thymidine kinase is the target for a number of antiherpetic drugs and, most recently, is being developed for use in gene therapy studies, the nucleoside binding site of this enzyme has not been definitively established. X-ray crystallographic studies of the purified protein have, to date, yielded very little detailed structural information (Sanderson et al., 1988). Amino acid alignments and sequence analysis of several thymidine kinases isolated on the basis of resistance to nucleoside analogues, most notably, acyclovir, have lead a number of investigators to focus research on a region corresponding to residues surrounding, and inclusive of, sites 3 and 4 described by Balasubramaniam et al. (1990). While others have investigated the role of residues surrounding site 4 (Munir et al., 1992), the work reported here focused on those residues that comprise site 3, the two adjacent residues, and an N-terminal proline residue possibly involved in structural integrity of the protein. In order to gain an understanding of the differences in substrate specificities between the various thymidine kinases that is the basis for the development of a number of antiviral drugs, identification of the nucleoside binding site is key.

Choice of Targeted Codons. Toward identifying important domains within the type I thymidine kinases, Balasubramaniam et al. (1990) aligned 12 different members of the *Herpesviridae* family thymidine kinases. Six highly conserved domains were identified and designated as sites 1 through 6. Two of these domains (sites 3 and 4, residues 162–164 and 171–173 in HSV-1 TK, respectively) were suggested to comprise the putative nucleoside binding site either alone or in a combined fashion. Because acyclovir, a guanosine analogue, is directly phosphorylated by HSV-1 TK, the location of amino acid alterations within acyclovir-resistant thymidine kinases is suggestive of an interaction of that residue with the nucleoside. One acyclovir-resistant HSV-1 TK was found to contain a mutation in a region adjacent to site 4 at arginine-176 (arginine to glutamine mutation) and led Darby and Larder (1987) to suggest a functional role for neighboring residues in nucleoside binding. Further evidence of a functional role for this region was provided by Sawyer et al. (1988), who isolated and identified a mutation in a site 3 residue (R130Q mutation in varicella zoster virus TK– corresponds to R163 in HSV-1 TK) that resulted in an acyclovir-resistant varicella zoster virus. To reveal potentially important residues in the region around site 3, 13 thymidine kinases were aligned, and a consensus sequence of the types of amino acids found was compiled (Figure 1). The highly conserved nature of the DRH motif of site 3 suggests that these residues play an important role in structural or enzymatic function. Two flanking residues, F161 and P165, are generally conserved on the basis of the type of residue (hydrophobic). A proline at the corresponding 165 position in HSV-1 TK is found in 11 of the 13 enzymes, suggesting a possible structural role. N-Terminal to site 3 at F161 in HSV-1 TK, predominantly hydrophobic residues are found with the exception of another aromatic residue (a histidine in Epstein Barr virus) or a serine (hydrophilic). Indeed, the overall hydrophobic nature of this region is consistent with these residues occupying positions within a catalytic pocket. Also of note is the proline at position 155 where a proline is found in 9 of the 13 proteins compared, again suggestive of a structural role. Despite the fact that natural selection continues to maintain these residues, especially those comprising site 3, it is not unlikely that other residues might also function in these position. To ascertain the role of selected residues in basic enzymatic function, random mutagenesis was performed on nucleotides comprising the site 3 residue codons (D162, R163, and H164) because of their complete conservation in all TKs aligned. In addition, nucleotides encoding both flanking amino acids, F161 and P165, were selected on the basis of their proximity to site 3 and their relatively high level of conservation. While somewhat distal to site 3, P155 may function indirectly (*e.g.*, structural integrity) with the putative catalytic site and therefore was chosen for random mutagenesis.

Random Sequence Selection. Because it was of interest to determine not only which residues were conserved but also the type of residue that is capable of producing an active enzyme, random sequence selection was employed. This technology allows one to create a library of clones containing random sequences at specific codons and to select for active recombinant molecules by genetic complementation. In this fashion, only functional clones are isolated and sequenced, and the amino acid alteration is deduced. A comparison of mutant sequences allows one to determine which residues are of little or no importance or are of significant value to enzymatic function. For example, if a number of different types of residues can substitute, that position is not likely to be essential. However, if no other residue or a similarly charged residue can replace a particular amino acid, then the importance of that residue is implied.

In this study, the target codons within the HSV-1 tk gene were replaced with a duplex of extended oligonucleotides that contained 80% of the wild-type nucleotides at each position and 20% of the other three nucleotides. The bias toward the wild-type sequence increases the likelihood of obtaining functional enzymes that have predominantly single nucleotide substitutions at each of the target codons. DNA sequence analysis of 10 nonselected mutants indicated that the 3 alternate nucleotides were present in equal amounts at each position. One question raised by the random sequence selection procedure is what is the likelihood that all amino acid substitutions will be represented in the population of transformants screened for activity. In this mutagenesis, the 18 nucleotides contained 20% random nucleotides (non-wild-type sequences) and thus represent 3.6 nucleotides changes per DNA molecule. In order to convert the possible nucleotide changes to amino acid changes, the average number of possible nucleotide changes per molecule which could yield a single amino acid change was calculated. For the 6 codons mutagenized, 72% should theoretically contain a single amino acid change with an average of 2.6 amino acid changes per molecule. At any 1 particular codon, 36 amino acid changes should be observed in the population of clones sequenced. This number of different amino acid changes at each position, however, was not observed in the population of mutants sequenced because the selection pressure imposed by screening the total population of 260 transformants for TK activity culls out the inactive mutants, i.e., the observed changes that occur at one particular residue directly reflect an amino acid change that confers activity. The absence of a particular amino acid substitution in the pool of sequenced mutants suggests that the substitution renders the enzyme nonfunctional. From the number and types of residues able to substitute for each mutagenized position obtained in this study, a number of inferences can be drawn.

Identification of Targeted Residues Important for TK Activity. A large number of different residues were able to substitute for P155, F161, and P165, such that it appears that these residues are not directly involved in nucleoside binding or structural integrity. In contrast, due to the few and/or similar types of residues obtained at D162, R163, and H164, it appears that they play some functional role. D162 appears to be the most conserved, only allowing glutamic acid, a similarly charged residue, or a small glycine residue to substitute. Interpretation of the multiple substitution mutations is much more complex and, in the absence of structural information, highly speculative. A number of the substitutions can be found as single mutations in Figure 3A, and thus the occurrence as multiple mutations is much more difficult to interpret. However, several of the mutations were not identified as single mutations such as D162N/H164K, and their occurrence implies some kind of compensatory affect between the two residues.

Nucleoside Complementation and Temperature Sensitivity. Toward identifying the nature of the mutations introduced, either structural or catalytic (involvement with nucleoside binding), the 35 representative mutants were screened for temperature sensitivity by complementation of the TK-deficient *E. coli* or for higher phosphorylation efficiency of nucleoside analogs (ACV or AZT) over thymidine. These experiments revealed several mutants that appeared to be temperature-sensitive as demonstrated by the lack of colony

formation on TK selection plates at 42 °C. The ability of a large number of different types of residues to substitute for both proline residues mutagenized in this study, and their apparent temperature insensitivity, argues against a significant role in the structural integrity of the enzyme. In addition, four mutants with apparent increased specificities for phosphorylating the analogues relative to thymidine when compared to that exhibited by the wild-type pMCC (KY895) were identified (Figure 4). Eight clones were selected for further analysis and subcloned into an *in vitro* transcription vector to allow for protein production in a cell-free translation system and *in vitro* biochemical analyses.

Putative Role of D162. While the D162E and D162G mutations did not display any phenotypic differences from pMCC in any of the secondary screening assays, the apparent essentiality of a negatively charged residue at position 162 was striking. An exception to this prediction was the isolation of D161G. In all *in vitro* assays, D162E behaved similarly to HSVTKII. However, D162G demonstrated a level of enzyme activity that was only a fraction of the wild-type activity. As with the other very low activity mutants (R163P/H164Q and F161I/R163H), D162G is likely to not be able to function sufficiently within the infected cell. In a comparison of the type of amino acid residues spanning site 3, Folkers and Trumpp (1987) and Black and Hruby (1992) showed that the major difference between the type I and type II thymidine kinases within this region is an aspartic acid (162) versus a glutamine (114), a negatively charged residue versus a hydrophilic residue. Site-directed mutagenesis of Q114 in VVTK to either an aspartic acid or a histidine residue resulted in enzymes that lost sensitivity to feedback inhibition by dTTP which corresponds to the lack of sensitivity to dTTP observed with HSV-1TK (Black & Hruby, 1992).

Increased Activity of F161 Mutants. Among the substitutions at phenylalanine at position 161, the majority of mutations (P155A/F161V, F161I) result in an increase in enzyme activity over wild-type levels or an equivalent level of wild-type activity (F161C) (Figure 6A). The other two mutations, F161L and F161I/R163H, result in a greatly reduced level of TK activity. If F161 is located within the nucleoside binding site, a replacement by the larger leucine side chain may place additional steric hindrance on the binding pocket and lead to a reduced level of catalytic activity. All the high-activity mutants contain equal or smaller sized residues than phenylalanine at position 161. In contrast, F161I/R163H is practically inactive. While the single amino acid change, F161I, leads to a very high activity mutant, the R163H alteration apparently suppresses the F161I change in the double mutant such that the improvement in activity from F161I is masked by the R163H mutation. This is further supported by the absence of the R163H substitution in the random sequence library.

Phosphorylation of Nucleoside Analogues. As a means toward identifying residues that interact with the substrate, the ability of mutants to differentially phosphorylate various nucleoside analogues may be indicative of substrate contact or the ability to subtly alter the binding pocket. The ability of different HSV-1 TK mutants (P155A/F161V, F161I, and F161C) to phosphorylate several nucleoside analogues (acyclovir, gancyclovir, and AZT) in addition to the natural substrates for HSV-1 TK, thymidine, and deoxyctidine was determined. As shown in Figure 8, the ability of these mutants to phosphorylate the nucleosides assayed varied from mutant to mutant and from nucleoside to nucleoside with respect to the wild-type enzyme. The *in vitro* results did not correlate completely with the *in vivo* complementation assays. However, it should be noted that the ability to complement a TK-deficient *E. coli* may not reflect the complexity of interactions within the milieu of a virus-infected mammalian cell and the requirements of a supposedly fully functional enzyme during viral replication and reactivation from the latent state. In addition, the TK proteins produced *in vivo* from the pMDC or pMCC vectors used are fusion proteins or proteolytic products of this fusion protein, none of which are apparently the authentic polypeptide (Waldman et al., 1983; Sanderson et al., 1988). Interestingly, P155A/F161V was able to phosphorylate gancyclovir approximately 5.2-fold better than HSVTKII, making it a good candidate for use in gene therapy studies. From these results, it appears that F161 plays a significant role in nucleoside binding. The capability of the different side chains in the mutants, especially at F161, to discern nucleosides and nucleoside analogues from one another, as evidenced by the differential ability of the mutants to phosphorylate the various analogues, is suggestive of a direct interaction with the substrate.

ACKNOWLEDGMENT

We thank Drs. J. Sweasy and T. Reid for critical comments on the manuscript and A. Herr and R. Duncan for their excellent technical assistance with sequencing.

REFERENCES

Baer, R., Bankier, A. T., Biggins, 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.

Balasubramaniam, N. K., Veerisetty, V., & Gentry, G. A. (1990) *J. Gen. Virol. 71*, 2979–2987.

Black, M. E., & Hruby, D. E. (1990) *J. Biol. Chem. 265*, 17584–17594.

Black, M. E., & Hruby, D. E. (1992) *J. Biol. Chem. 267*, 9743–9748.

Culver, K. W., Ram, Z., Wallbridge, S., Ishii, H., Oldfield, E. H., & Blaese, R. M. (1992) *Science 256*, 1550–1552.

Darby, G., Larder, B. A., & Inglis, M. M. (1986) *J. Gen. Virol. 67*, 753–758.

Davison, A. J., & Scott, J. E. (1986) *J. Gen. Virol. 67*, 1759–1816.

Dube, D. K., Parker, J. D., French, D. C., Cahill, D. S., Dube, S., Horwitz, M. S. Z., Munir, K. M., & Loeb, L. A. (1991) *Biochemistry 30*, 11760–11767.

Folkers, G., & Trumpp, S. (1987) *Med. Sci. Res. 15*, 1495–1496.

Griffin, A. M., & Boursnell, M. E. G. (1990) *J. Gen. Virol. 71*, 841–850.

Gruidl, M. E., Hall, R. L., & Moyer, R. W. (1992) *Virology 186*, 507–516.

Honess, R. W., Craxton, M. A., Williams, L., & Gompels, U. A. (1989) *J. Gen. Virol. 70*, 3003–3013.

Hruby, D. E., & Ball, L. A. (1981) *Virology 113*, 594–601.

Igarashi, K., Hiraga, S., & Yura, T. (1967) *Genetics 57*, 643–654.

Kit, S., & Kit, M. (1985) U.S. Patent 4514497.

Kit, S., Kit, M., Qavi, H., Trkula, D., & Otsuka, H. (1983) *Biochim. Biophys. Acta 741*, 159–170.

Littler, E., & Arrand, J. R. (1988) *J. Virol. 62*, 3892–3895.

Martin, S. L., Apariso, D. I., & Bayopadhyay, P. K. (1989) *J. Virol. 63*, 2847–2852.

Mittal, S. K., & Field, H. J. (1989) *J. Gen. Virol. 70*, 2901–2918.

Munir, K. M., French, D. C., Dube, D. K., & Loeb, L. A. (1992) *J. Biol. Chem. 267*, 6584–6589.

Nunberg, J. H., Wright, D. K., Cole, G. E., Petrovskis, E. A., Post, L. E., Compton, T., & Gilbert, J. H. (1989) *J. Virol. 63*, 3240–3249.

Ostuka, H., & Kit, S. (1984) *Virology 135*, 316–330.

Sanderson, M. R., Freemont, P. S., Murthy, H. M., Krane, J. F., Summers, W. C., & Steitz, T. A. (1988) *J. Mol. Biol. 202*, 917–919.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Sawyer, M. H., Inchaupse, G., Biron, K. K., Waters, D. J., Straus, S. E., & Ostrove, J. M. (1988) *J. Gen. Virol. 69*, 2585–2593.

Scott, S. D., Ross, N. L. J., & Binns, M. M. (1989) *J. Gen. Virol. 70*, 3055–3065.

Swain, M. A., & Galloway, D. A. (1983) *J. Virol. 46*, 1045–1050.

Waldman, A. S., Haeusslein, E., & Milman, G. (1983) *J. Biol. Chem. 258*, 11571–11575.

Wilson, E. M., Franke, C. F., Black, M. E., & Hruby, D. E. (1989) *Gene 77*, 69–78.

Zimmermann, N., Beck-Sickinger, A. G., Folkers, G., Krickl, S., Mueller, I., & Jung, G. (1991) *Eur. J. Biochem. 200*, 519–528.