Permissible Amino Acid Substitutions within the Putative Nucleoside-binding Site of Herpes Simplex Virus Type 1 Established by Random Sequence Mutagenesis*

Khan M. Munir, David C. French, Dipak K. Dubeć, and Lawrence A. Loeb

From the Joseph Gottstein Memorial Cancer Research Laboratory, Department of Pathology, University of Washington, Seattle, Washington 98195

We determined the essentiality of all amino acid replacements within an 11-codon sequence in the putative nucleoside-binding site of thymidine kinase encoded by herpes simplex virus type 1. This involved partial randomization of 11 codons in the gene to create a degenerate library, followed by genetic complementation using a tk* Escherichia coli strain and selection of unnatural active enzymes. We produced and tested 53,000 variants; of which 190 were found to be biologically active. Sequence analyses of functional variants revealed a high degree of flexibility in accommodating different types of amino acid substitutions in this region. However, no replacement was tolerated at proline-173, whereas tyrosine-172 could be replaced by only phenylalanine. To further define permissible substitutions at specified positions, we constructed a library with randomization at only four test codons. We produced and tested 600,000 variants; of which only 5 were active. Again proline-173 was conserved, and only tyrosine and phenylalanine were found at position 172. The identification of these conserved amino acids should provide important insights into the understanding of the structural basis of catalysis by this enzyme.

Site-directed mutagenesis has been an important method to study the roles of individual amino acids in the function of proteins (1, 2). Even though the wide use of this technique has provided insight into the understanding of protein folding (3, 4), structure-function relationships of many enzymes (5-7), and receptor binding (8), it has important limitations. The most frequently cited concern is the absence of the rules governing the interchangeability of different amino acids. In absence of these rules there are 19 choices to replace each amino acid residue at any given position in a protein. Therefore, to analyze the active site of an enzyme requires the construction and testing of an unmanageably large number of variants.

Random mutagenesis presents an alternative to site-directed mutagenesis because of its potential to test all possible substitutions starting from a population of random nucleotide replacements in DNA (9). If combined with biological selection, this approach provides a powerful method for identifying unnatural biologically active molecules from a large pool of random sequences (9, 10). The effective generation of functional mutant enzyme variants by this technique is based on the hypothesis that multiple amino acid substitutions can be tolerated within the active site, and that these substitutions give rise to molecules with altered properties. These altered molecules should be useful tools in the study of protein structure to function relationships. We have applied this technique to define Escherichia coli promoter sequences (9, 11), and to remodel the active site of RTEM-1 β-lactamase (12). Other investigators have utilized a similar approach to probe signal sequences of secreted proteins (13), binding sites of transcription factors (14, 15), and a ribosomal binding site (16), to improve the catalytic capability of an enzyme (17), ribozymes, and new RNA molecules (18, 19), and to construct random gene libraries of unnatural peptides (20-22).

In this article we report the analysis of the spectrum of mutations produced in the herpes simplex virus 1 (HSV-1)

1 thymidine kinase gene (tk) generated by replacing a portion of the putative nucleoside-binding site with random nucleotide sequences. HSV-1-specified thymidine kinase (TK) is a multifunctional enzyme that can catalyze the phosphorylation of thymidine, thymidylic acid, and cytidine in the presence of ATP (23, 24). TK has also been implicated to play a role in the uptake of thymidine in E. coli (25, 26). Even though the gene that encodes TK has been cloned and sequenced (27, 28), little is known about the structural basis of its catalytic activity. Kinetic studies suggest that TK has two binding sites, one for ATP and the other for thymidine and related nucleosides (29). Liu and Summers (29) have mapped the ATP-binding site. The thymidine-binding site or a portion thereof has been putatively assigned to the sequence between amino acid residues 168 and 176 (30). The designation of nucleoside-binding site at these positions rests on the characterization of drug-resistant mutants of HSV-1. Mutants with alterations at this site exhibit higher \( K_a \) for thymidine due to substitutions of alanine-168 to threonine and arginine-176 to glutamine, respectively, and also show differences in rates of phosphorylation of nucleoside analogues (31). Previously, we reported the effect of randomization of two codons, 166 and 167, of the tk gene (32). Here, we present a new approach with limiting randomization of codons spanning 165-175 to analyze the essentiality of each amino acid within this domain.

**EXPERIMENTAL PROCEDURES**

**Materials**—Oligonucleotides were synthesized by Operon Technologies (San Pablo, CA). [methyl-\(^{3}H\)]Thymidine (28 Ci/mmol) was

---

*This work was supported by Grants OIG R35-CA-39903 (to L. A. L.) and AG-00067 (to K. M. M.) from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

1 The abbreviations used are: HSV-1, herpes simplex virus type 1; TK, thymidine kinase; tk, thymidine kinase gene.
purchased from Amersham Corp. Restriction enzymes were obtained commercially, and used according to the suppliers’ instructions. Standard molecular cloning techniques were employed (33).

Plasmids and Bacteria—The *E. coli* tk strain KY895 (K12, F′,ydy-, ilv 276) (34) and plasmids pHETK1 and pHETK2 (35) were generous gifts of Dr. William Summers (School of Medicine, Yale University, New Haven). Plasmids pHETK1 and pHETK2 are expression vectors that contain the HSV-1 tk structural gene and are derivatives of pBR322 (35). Restriction maps of pHETK1 and pHETK2 can be found in Waldman et al. (35). Plasmid pHETK2 contains XbaI and XhoI promoters, ampR, and a cl857 temperature-sensitive repressor, whereas pHETK1 contains all the above except the XhoI promoter.

Construction of a Dummy Vector—A plasmid that expresses a totally inactive TK hetero referred to as “dummy vector” was used for the insertion of random nucleotide sequences. Previously we reported that a plasmid vector pKTPD (32), pHETK2 was converted to a dummy vector, pKTPD, by replacing the nucleotide sequence between the unique KpnI and SacI sites within the tk gene by a smaller unrelated nucleotide sequence. Since the presence of two promoters in PHETK2 leads to overproduction of the tk gene product that may be lethal to cells, we constructed from pHETK1 and pKTPD a new chimeric dummy vector, designated pMDC, that contains only one promoter, XbaI. Upon digestion with SpeI and PoulI pHETK1 is cut into two fragments. The larger fragment contains ampR, cl857, XbaI sequences, and part of the tk gene spanning from the BamHI to the SpeI site. The smaller fragment contains the remainder of the tk gene from SpeI to PoulI. Similarly, pKTPD upon digestion with the same two enzymes is cut into one larger and one smaller fragment. The smaller SpeI/PoulI fragment of pKTPD contains the dummy sequence within the KpnI and SacI sites of the tk gene. Ligation of the larger fragment from pHETK1 with the smaller fragment of pKTPD results in a chimeric vector, pMDC, that produces an inactive tk gene product. Another chimeric vector, pMCC, containing the wild-type tk gene was similarly constructed by ligating the larger fragment from PHETK1 with the smaller fragment of pHETK2.

Construction of a Library with Oligonucleotides Containing Partial Random Nucleotide Substitutions—A library containing 20% random nucleotide sequences was constructed as described by Dube et al. (32) with modifications. A 52-mer oligonucleotide (oligo I) with the wild-type tk sequence was hybridized to a 56-mer that contained degenerate nucleotides spanning from codon 165 through 175 of the tk gene (oligo II), where n = 80% wild-type nucleotides and a 20% mixture of the other three at each position. Both oligomers were complementary to each other along 12 bases at their 3′-ends.

Oligo I: 5′-TGG GGA GCT CAC ATG CCC GGC CCC GGC CCC GCG TCT ATG TAC GCC ATC-3′
Oligo II: 5′-ATG TAG TAC CGN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN NNN TGG CGA TCG AA-3′

The hybrid was extended with the Klenow fragment of *E. coli* DNA polymerase I to produce a complete double-stranded DNA product. The above strategy was taken in order to avoid synthesizing a long random nucleotide containing oligonucleotide (oligo II) since the locations of KpnI and SacI sites (insertion sites) in the vector require a long cassette. The Klenow fragment generated double-stranded DNA was then subjected to polymerase chain reaction amplification by using two synthetic primers: the first primer, a 5′-TGG GAG CTG ACA TCC GGC GCC CCC-3′ corresponds to the 21-base sequence of 5′ terminus of oligo I. The second primer, b 5′-ATG AGG TAC CGC-3′ corresponds to the 11-base sequence of 5′ terminus of oligomer II. The polymerase chain reaction amplification reactions contained 20 mM Tris-HCl (pH 8.3), 25 mM KCl, 1.5 mM MgCl2, and 0.05% Tween 20, 0.1 mg/ml BSA, 50 μM each of the four deoxynucleoside triphosphate, 20 pmol of primer a, 40 pmol of primer b, approximately 1 pmol of the extended double-stranded oligonucleotide as template, and 2 units of Taq polymerase (Cetus) in 100-μl final reaction volume. Each reaction was overlaid with mineral oil and subjected to 30 rounds of temperature cycling: 94 °C for 1 min, 34 °C for 2 min, and 72 °C for 7 min. Low molecular weight components and excess primers were removed from the polymerase chain reaction-amplified product by centrifugation through a Centricon 30 ultrafiltration unit, and the amplified DNA was digested with KpnI and SacI. The digested double-stranded oligonucleotide containing the random sequence was again purified by a Centricon 30 unit, and ligated in place of the dummy segment between the KpnI and SacI site of pMDC (Fig. 1).
Sequence Analysis—Double-stranded DNA templates were prepared from 5 ml of bacterial culture by a combination of alkaline lysis and polyethylene glycol precipitation methods (33). The DNA sequencing primer was 5'-CAT GTC TTA TGC CGT GA-3', and the nucleotide sequence was determined by the dideoxynucleotide chain termination method (37).

E. coli Extracts and Assay for TK Activity—Crude extracts were obtained from infected E. coli by a modification of the method of Garapin et al. (38). Thirty ml of cultures were grown at 30 °C up to an optical density of 0.1 at 540 nm, shifted to 37 °C, and allowed to grow to 1.0. The cells were collected by centrifugation at 4 °C for 15 min, washed with 5 ml of 25% (w/v) sucrose, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, and resuspended in 0.1 ml of the same buffer which contained 50 μg of chicken egg white lysozyme (Sigma). After 10 min at 0 °C the cell suspension was mixed with 0.5 ml of chilled TK buffer (50 mM Tris-HCl, pH 7.5, 4 mM MgCl₂, 50 mM KCl, 5 mM dithiothreitol, 0.1% Triton X-100, and 2.5 μg/ml each of leupeptin, pepstatin, 50 μg/ml of aprotinin, and 1 mM phenylmethylsulfonyl fluoride). The suspension was maintained on ice until the bacteria lysed. After centrifugation at 120,000 × g for 30 min, the resulting supernatant contained 2–4 mg/ml protein, and was stored in aliquots at −70 °C.

The activity of TK in vitro was determined by the conversion of labeled thymidine to phosphorylated products that absorb onto DEAE-cellulose discs (24, 38). The reaction mixture in a volume of 50 μl contained 50 mM Tris-HCl (pH 7.5), 4 mM MgCl₂, 2.5 mM dithiothreitol, 5 mM ATP, 12 mM KCI, and 1 μl [³H]thymidine (1 μCi), 2.5% glycerol, 0.1 mg/ml bovine serum albumin, and 10 μl of crude extract. After incubation for 10 min at 37 °C, the reaction was stopped by chilling the mixture to 0 °C. A 10-μl aliquot was pipetted onto a DEAE-cellulose disc (25 mm in diameter), nonphosphorylated [³H] thymidine was removed by washing (39), and the absorbed products were counted in a scintillation counter.

RESULTS

Strategies for Partial Random Mutagenesis and the Identification of Functional Mutants—Eleven codons (33 nucleotides) spanning codons 165–175 of the HSV-1 tk gene were replaced by random nucleotide sequences. The strategy used in the construction and generation of unnatural functional mutants is depicted in Fig. 1. A central feature of this strategy is the utilization of a dummy vector, pMDC, that produces a complete random sequence containing mutants. Target codons in the tk gene (Fig. 1) were replaced by random nucleotide sequences. The strategy used in the construction and generation of unnatural functional mutants is reported by Oliphant and Struhl (40), and Dube and Loeb (12).

Characterization of the 20% Random Library—From a total of 53,000 transformants that grew on carbenicillin medium, 190 were able to complement E. coli KY 895 for TK function. Plasmid DNA from the 190 TK positive clones along with 10 others from the carbenicillin medium were sequenced. Among these 90 TK positive mutants, there were a total 233 nucleotides that differed from the wild-type sequence, of which 114 were changed to thymine, 63 to adenine, 47 to guanine, and 9 to cytosine. The lower representation of cytosine substitutions could be due to a bias against cytosine addition during chemical synthesis of 20% random oligonucleotide or could be indicative of selection for active molecules. A similar under-representation of cytosine in random sequence mutants was reported by Oliphant and Struhl (40), and Dube and Loeb (12).

Spectrum of Amino Acid Substitutions—Fig. 3 is a tabulation of the different amino acid substitutions at the target region of the 90 functional TK mutants that were sequenced. There is a high degree of plasticity in the substitutions of wild-type amino acids by other amino acids with a variety of physical properties (Table I). Nine of 11 wild-type amino acid residues in this region are neutral and hydrophobic. Each of these were substituted by 1 or more amino acids with different functional groups. Wild-type polar residues, cysteine-171 and tyrosine-172, were changed to leucine and phenylalanine, respectively. Proline-173 was conserved. Codons 171, 172, and 173 were not conserved among the mutants that were obtained from carbenicilllin medium. Mutants with substitutions at these positions were devoid of detectable TK activity as judged by both in vitro and in vivo assays (results not shown).

Mutants with Multiple Amino Acid Substitutions—Many functional mutants contained amino acid substitutions at two or more positions within the target sequence. As shown in Fig. 4, two active mutants contained 4 substitutions, two contained 3 substitutions, and 16 had 2 substitutions. The remaining 70 showed only single amino acid substitutions. An average of 3 amino acid substitutions were observed in mutants that were obtained from nonselective carbenicilllin medium (data not shown).

Fig. 3. Functionally acceptable amino acid substitutions at the target region. Wild-type amino acid residues are shown in boldface.
**TABLE 1**

**Percentage of amino acid substitutions at each position in active TK mutants**

| Wild-type residues | Neutral and hydrophilic | Neutral and polar | Acidic | Basic |
|--------------------|-------------------------|------------------|--------|-------|
| Proline-165         | 55 (3/9)                | 33 (3/9)         | None   | 11 (1/9) |
| Isoleucine-186      | 86 (6/7)                | 14 (1/7)         | None   | None   |
| Alanine-167         | None                    | 80 (12/15)       | 61 (1/15) | 13 (2/15) |
| Alanine-168         | 17 (4/23)               | 75 (18/23)       | None   | 4 (1/23) |
| Leucine-169         | 80 (4/5)                | None             | 20 (1/5) |
| Leucine-170         | 83 (5/6)                | 17 (1/6)         | None   | None   |
| Cysteine-171        | 100 (1/1)               | None             | None   | None   |
| Tyrosine-172        | 100 (3/5)               | None             | None   | None   |
| Proline-173         | None                    | None             | None   | None   |
| Alanine-174         | 33 (4/12)               | 58 (7/12)        | 8 (1/12) |
| Alanine-175         | None                    | 100 (1/1)        | None   | None   |

* Number indicates the sequence position in wild-type TK.

**Fig. 4.** TK mutants with 2 or more amino acid substitutions. Wild-type sequence is shown in **boldface** at the bottom.

**TK Mutants from Codon-specific Library**—In order to challenge the stringency for substitutions in functional TK mutants at cysteine-171, tyrosine-172, proline-173, and alanine-175, we constructed a library that contained 100% random nucleotides at these codons but maintained the wild-type sequences at all other positions within the 11-amino acid target sequence. This library has the potential to code for 20 or 160,000 different amino acid sequences. To verify the randomness of the codon-specific library, we first sequenced 25 mutants that grew on carbamoylaminicillin medium. Nucleotide substitutions were only observed at the target codons. The distribution of substitutions for all the 4 nucleotides at the target codons was approximately equal, that is, 20% G, 34% T, 25% A, and 21% C. The sequences of some of these mutants as well as their ability to grow on TK selection medium by complementation are listed in **Fig. 5**. Interestingly, none of them could grow on the TK selection medium, and only one of the mutants, LCC3, exhibited detectable TK activity in *vitro* (1 pmol/mg/30 min), which is less than that required for growth on TK selection plates.

Of 600,000 total transformants from the codon-specific library that were tested for TK complementation, only five were found to complement *E. coli* KY 895. The sequences of the 5 mutants are shown in **Fig. 6**. As expected each of the mutants had an alanine at position 174. With respect to positions 172 and 173, the stringency of substitution was the same as that obtained in the 20% library. No amino acid substitution was observed for proline-173, whereas tyrosine-172 changed to only phenylalanine. However, in this challenge experiment, positions 171 and 175 were less stringent. Cysteine-171 and alanine-175 could be changed to other amino acid residues and the transfectants could grow on TK selection medium. In each of the five mutants, 19 other amino acids could have substituted for proline and none were found. In contrast, substitutions for cysteine-171 and alanine-175 were not as stringent as indicated in the analysis of the 20% library (**Fig. 3**).

**Catalytic Activity of Various TK Mutants**—The ability of some of the TK mutants to phosphorylate thymidine in *vitro* relative to that of wild-type extract is shown in **Fig. 7**. These mutants were sampled to highlight the importance of some of the amino acids in the target domain for the catalytic activity of TK. The replacement of even 1 amino acid with another...
can alter the catalytic effectiveness of TK. While substitution at alanine-168 to serine did not alter the activity significantly of mutant TKF36 compared to the wild-type TK, tyrosine substitution at the same position in mutant TKF102 rendered it 50% less active. In contrast, TKF2 with 3 amino acid substitutions exhibited 70% of the activity of the wild-type extracts.

**DISCUSSION**

The catalytic function of an enzyme is specified by its amino acid sequence which directs the folding into a correct three-dimensional structure. However, predicting protein structure from primary sequence is currently an insurmountable task as we do not yet know how to infer the structure or function of a protein from its amino acid sequence (41). It is presently thought that only a small fraction of the amino acid residues in a given enzyme contributes to its activity (10, 42), and it should be possible to assign specific amino acid residues for specific functional properties. Site-directed mutagenesis, in combination with x-ray crystallography, and other physical methods, are providing significant insights in this direction (2). There are, however, limitations to the site-directed approach for defining involvement of different amino acids. One requires preconceived ideas as to which of 19 other amino acid replacement at each position should yield functional consequences. The crystal structure of a protein frequently permits one to rationalize the effects of amino acid replacements; however, in many cases this is not available. In this report we describe a random sequence mutagenesis approach to probe the amino acid residues at a site that is thought to be the nucleoside-binding domain of TK specified by HSV-1. This technique in combination with a selection assay offers a powerful alternative method to identify residues that are required for enzyme function.

We have previously reported (32) on the effect of randomization of the two codons that specify isoleucine-166 and alanine-167 of HSV-1 TK. The above report, along with others (30, 32), indicate that a region of 11 amino acid residues spanning from codons 166 to 176 comprises the nucleoside-binding site. To further define the amino acid requirements for catalytic activity of this site, we have replaced the nucleotide sequences in codons 165–175 with oligonucleotides containing 20% random substitutions at each position. Since 100% randomization of 11 codons (33 nucleotides) would require screening of a vast number of clones (4^33 \approx 10^9), a significant portion of which would be inactive due to random insertions of stop codons, we chose to construct a library with 20% degeneracy. After screening only 53,000 transformants, we obtained 190 functional mutants (0.36%). Sequence data of 90 functional and 10 nonfunctional mutants allowed us to characterize the quality of this library. All the four bases were inserted in approximately equal proportion except for cytosine, which was reduced in frequency. Since we did not observe a bias in base distribution in our codon-specific 100% library (above), the low frequency of cytosine in mutants obtained from 20% random library could be indicative of the repertoire of substitutions that give rise to functional mutants.

Nine of 11 wild-type amino acid residues that constitute the target sequence are neutral and hydrophobic. Fig. 3 is the spectrum of acceptable functional substitutions at these positions. Substitutions were obtained for most of the amino acids in this region. From these, Table I was constructed which tabulates the percentage of tolerated amino acid residues at any given position in terms of their physical properties. A general pattern was observed for the substitutions; any hydrophobic residue could be replaced only by another hydrophobic residue but also by polar, acidic, or basic residues. However, no acidic residue substitution was observed at proline-165, isoleucine-166, alanine-168, leucine-169, leucine-170, and alanine-175. The simplest explanation is that we may not have screened a sufficient number of clones to sample functional replacement by acidic residues at these positions. This can be substantiated by two observations: (i) alanine at positions 167 and 174 did get replaced with acidic residues, and (ii) even though we did not find any hydrophobic residue substitution at alanine-167, functional mutants with hydrophobic residue replacements at this position were observed by Dube et al. (32).

In contrast to the wide spectrum of permissible mutations at positions 165–170, a functional requirement for cysteine-171, tyrosine-172, and proline-173 is suggested. The informational content of these residues can be judged by the percent of allowed substitutions that were obtained. The informational content of any amino acid residue in an enzyme, as defined by Reidhar-Olson and Sauer (10), is a value that decreases as the number of allowed substitutions increases at any position. Hence, the informational content of any particular amino acid is highest if no other substitution is allowed. Since random sequence mutagenesis can test all possible substitutions at any given position, it offers the widest approach to determine the informational content of that amino acid, and hence, its contribution to the structure-function relationship. Proline-173 of TK appears to have the highest informational content since no other residues were found to functionally replace it. To substantiate the essentiality of specific amino acids, we constructed a codon-specific library in which the nucleotides of codons 171, 172, 173, and 175 were replaced by 100% random nucleotides, and functional clones were scored on TK selection medium. Only five functional clones were obtained (5/600,000), and none of them contained any substitution for proline-173 (Fig. 6), even though nucleotide substitutions at this codon that still coded for proline were observed. This absolute stringency was not due to under-representation of proline codons in the library, since proline was substituted by other amino acids in unselected tk mutants that were obtained from carbenicillin medium. The proline-substituted clones, however, contained nonfunctional TK as judged by in vivo TK selection assay and in vitro biochemical assay. That proline-173 appears to
have a specifically high informational content can also be judged by the fact that another proline residue at position 165, which was inside our mutagenesis window, could be replaced by many other different residues (Fig. 3 and Table I). Unlike proline-173, tyrosine-172 is not 100% conserved; it can be replaced by phenylalanine (Fig. 6). Since phenylalanine and tyrosine are structurally similar, a change of tyrosine to phenylalanine may represent an isosteric replacement that does not lead to conformational instability. If this is the case, the hydroxyl side chain of tyrosine does not have a dominant role in the conformational stability or in the function of the enzyme.

Cysteine-171 and alanine-175 seemed to have moderate informational content (Fig. 3). When alterations of these 2 residues were forced using the codon-specific library, they were found to be limited to hydrophobic residue substitutions. Three residues, alanine, valine, and glycine replaced cysteine-171, as single amino acid substitutions, and only valine could replace alanine-175. This limited repertoire of replacement is not due to the under-representation of other codons since we have seen different types of changes in the sequences of nonfunctional mutants at the same site (Fig. 5). It has been predicted (30) that 2 cysteine residues at positions 171 and 336 play a role in the conformational stability of wild-type TK enzyme. We obtained a mutant with a cysteine-171 to leucine replacement, that had a nearly normal activity, but appeared to be an unstable enzyme as judged by the rate of loss of activity over time (data not shown). Inactivation of this mutant enzyme might indicate a critical structural role for cysteine-171 in TK. It is worth noting that while proline-173 and tyrosine or phenylalanine at 172 are absolutely required, substitution at many of other positions alters the catalytic effectiveness of TK (Fig. 7).

Darby et al. (30) and Larder et al. (31) described a mutant TK with an arginine-176 to glutamine replacement. This mutation decreased the affinity of TK for thymidine. Unfortunately, arginine-176 lies just outside our mutagenesis cassette. Still, our data taken together with the above suggest that residues 168–176 comprise a domain that is required to bind thymidine, or is necessary to form a conformation that accentuates the enzymatic reaction of TK. However, a three-dimensional crystal structure of key mutants should establish the exact nature and function of each of the amino acids within this domain.

One important advantage of random mutagenesis coupled with selection is that one can compare functional versus nonfunctional sequences. We find that functional mutants with two or more amino acid substitutions can fold to form a correct three-dimensional structure as evidenced by catalytic activity. To ascertain that nonfunctional phenotypes of inactive mutants were not due to lack of expression or degradation, we performed Western immunoblot studies on several functional and nonfunctional mutants; all of them showed equal amounts of protein expression, and proteolytic degradation was not evident (data not shown).

In summary, we have presented data that show the essentiality of 2 amino acid residues that are important for the function of thymidine kinase. Using a 20% random library we have established the probability of amino acid substitutions within the putative nucleoside-binding site of HSV-1 TK. By limiting the number of possibilities we have been able to evaluate the informational content of amino acids that are infrequently substituted in the random library. This approach to define the involvement of different amino acids in the catalysis of an enzyme can be utilized for any protein in which one can design a biological assay for function based on genetic selection.

Acknowledgements—We thank Drs. M. Fry, K. Williams, and A. Blank for comments on the manuscript.

REFERENCES

1. Smith, M. (1985) Annu. Rev. Genet. 19, 423–462
2. Knowles, J. R. (1987) Science 236, 1252–1258
3. Luger, K., Hommel, U., Herold, M., Hofsteenge, J., and Kirschner, K. (1986) Science 234, 296–310
4. Lin, L. and Brandts, J. F. (1984) Biochemistry 23, 5713–5723
5. Howell, E. E., Villafranca, J. E., Warren, S. M., Ostley, S. J., and Kraut, J. (1986) Science 231, 1125–1128
6. Craik, C. S., Largman, C., Fletcher, T., Rozenek, S., Barr, P. J., Pettersson, R., and Rutter, W. (1985) Science 228, 291–297
7. Sowek, J. A., Singer, B. L., Bringer, S., Mailey, M. F., Dougherty, T. J., Gougotas, J. Z., and Bush, K. (1991) Biochemistry 30, 3179–3188
8. Brodbelt, P. J., Hubler, L., Chen, W. S., Carpenter, C. D., Rosenfeld, M. G., and Gill, G. N. (1986) in "Growth Regulation of Cancer: UCLA Symposium on Molecular and Cellular Biology," Vol. 74, pp. 157–167, Alan R. Liss, New York
9. Horwitz, M. S. Z., and Loeb, L. A. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 7405–7409
10. Reishevar-Olson, J. F., and Sauer, R. T. (1988) Science 241, 53–57
11. Horwitz, M. S. Z., and Loeb, L. A. (1988) J. Biol. Chem. 263, 14724–14731
12. Dube, D. K., and Loeb, L. A. (1989) Biochemistry 28, 5703–5707
13. Kiser, C. A., Preuss, D., Grissau, P., and Botstein, D. (1987) Science 235, 312–317
14. Blackwell, T. K., and Weintraub, H. (1990) Science 250, 1104–1110
15. Oliphant, A., Brandl, C. J., and Struhl, K. (1989) Mol. Cell. Biol. 9, 2944-2949
16. Turets, C. and Gold, L. (1990) Science 249, 505–510
17. Herms, J. D., Blacklow, S. C., and Knowles, J. R. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 696–700
18. Robertson, D. L., and Joyce, G. F. (1990) Nature 344, 467–468
19. Ellington, A. D., and Szostak, J. W. (1990) Nature 346, 818–822
20. Masdeau, W. (1990) Protein Eng. 3, 221–226
21. Scott, J. K., and Smith, G. P. (1990) Science 249, 386–390
22. Devlin, J. D., Panganiban, L. C., and Devlin, P. E. (1990) Science 249, 409-415
23. Cheng, Y., and Ostrander, M. (1976) J. Biol. Chem. 251, 2636–2610
24. Chen, M. S., Walker, J., and Frutos, W. H. (1979) J. Biol. Chem. 254, 10747–10751
25. McKeown, M., Kahn, M., and Hanswalt, P. (1976) J. Bacteriol. 126, 814–822
26. Dube, D. K., Horwitz, M. S. Z., and Loeb, L. A. (1991) Gene (Amst.) 99, 25–29
27. Wagen, M. J., Sharp, J. A., and Summers, W. C. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 1441–1445
28. McKnight, S. L. (1986) Nucleic Acids Res. 14, 5049–5064
29. Liu, Q., and Summers, W. C. (1988) Virology 163, 638–642
30. Darby, G., Larder, B. A., and Inglis, M. M. (1986) J. Gen. Virol. 67, 753–769
31. Larder, B. A., Derse, D., Cheng, Y-C., and Darby, G. (1983) J. Biol. Chem. 258, 2027–2033
32. Dube, D. K., Parkcr, J. D., French, D. C., Cahill, D. S., Dube, S., Horwitz, M. S. Z., Munir, K. M., and Loeb, L. A. (1991) Biochemistry 30, 11760–11769
33. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
34. Igarashi, K., Hiraga, S., and Yura, T. (1967) Genetics 57, 643–654
35. Waldman, A. S., Huesselein, E., and Milman, G. (1985) J. Biol. Chem. 258, 11571–11576
36. Bio-Red (1985) Bacterial Electrotransformation, Version 2.89, pp. 14–16
37. Sonenberg, N., Fleischer, S., and Scolnick, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463–5467
38. Garagiz, M. C., Cesare, P., and Rodriguez, E., Horodincau, F., and Kounky, P. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 818–821
39. Elion, G. B., Furman, R. P., Hye, F. A., Miranda, P. D., Beaudon, J., and Strassman, E. J. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5718–5720
40. Oliphant, A. R., and Struhl, K. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 9094–9098
41. Griston, T. E. (1993) in Proteins: Structures and Molecular Interactions, Freeman, New York
42. Hecht, M. H., Nelson, H. C. M., and Sauer, R. T. (1983) Proc. Natl. Acad. Sci. U.S.A. 80, 2676–2680