Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 696–700, January 1990
Biochemistry

Searching sequence space by definably random mutagenesis: Improving the catalytic potency of an enzyme

(evolution/pseudorevertants/spiked oligonucleotides/triose-phosphate isomerase)

JEFFREY D. HERMES*, STEPHEN C. BLACKLOW, AND JEREMY R. KNOWLES†

Departments of Chemistry and Biochemistry, Harvard University, Cambridge, MA 02138

Contributed by Jeremy R. Knowles, October 20, 1989

ABSTRACT **How easy is it to improve the catalytic power of an enzyme? To address this question, the gene encoding a sluggish mutant triose-phosphate isomerase (D-glyceraldehyde-3-phosphate ketol-isomerase, EC 5.3.1.1) has been subjected to random mutagenesis over its whole length by using "spiked" oligonucleotide primers. Transformation of an isomerase-minus strain of *Escherichia coli* was followed by selection of those colonies harboring an enzyme of higher catalytic potency. Six amino acid changes in the Glu-165 → Asp mutant of triose-phosphate isomerase improve the specific catalytic activity of this enzyme (from 1.3-fold to 19-fold). The suppressor sites are scattered across the sequence (at positions 10, 96, 97, 167, and 233), but each of them is very close to the active site. These experiments show both that there are relatively few single amino acid changes that increase the catalytic potency of this enzyme and that all of these improvements derive from alterations that are in, or very close to, the active site.**

There are currently two ways that the manipulation and mutagenesis of genes that encode interesting proteins can give new insights into structure–function relationships. In the first of these, the presumed importance of a particular amino acid or of a group of amino acids is probed by changing or deleting it and examining the functional consequences (1–4). In its more ambitious forms, this line of attack has led to efforts to use site-specific mutagenesis in a rational way—for example, to produce enzymes of altered specificity (5–7) or repressors that recognize a different operator sequence (8, 9). These efforts have produced mixed results, however, and it is clear that our present understanding of protein structure and function does not yet guarantee that rationally designed changes will yield the predicted outcomes.

The second and complementary approach does not test our current views of the functional consequence of a particular structural change but rather aims for a deeper understanding of the problem by identifying all of those structural changes that produce a particular functional effect. Random mutagenesis allows a defined region of mutant sequence "space" (10) to be searched by creating a library of amino acid changes in the protein. Selection from this library then identifies mutants that confer the desired phenotype. Since this approach avoids all preconceived ideas about what is important, and since (at least in principle) all possible single amino acid changes and some fraction of double and higher order changes can be made, the potential exists for the unbiased discovery of alternative structural responses to a given functional challenge. In a sense, of course, the random mutagenesis approach mimics the evolutionary refinement of biological function at the molecular level.

In the work reported here, we have identified the single amino acid changes that improve the catalytic potency of a mutant triose-phosphate isomerase (D-glyceraldehyde-3-phosphate ketol-isomerase, EC 5.3.1.1) that has 1/1000th the activity of the wild type. Three questions about the relationship between enzyme structure and catalytic power have been answered. These questions are: (*i*) can the catalytic activity of a sluggish enzyme be improved other than by mere reversion to the wild type (this question is equivalent to asking whether there are other accessible "solutions" to the catalytic "problem"); (*ii*) if such improvement were possible, would it be rare (that is, are there only a few ways in which a protein can accommodate to and compensate for a particular catalytic lesion); and (*iii*) how are such second-site suppressor changes distributed, and can we learn from their distribution?

The recovery of function of a mutant protein by pseudoreversion is, of course, a well-established phenomenon. The early work of Benzer (11), of Crick *et al.* (12), and of Yanofsky (13) and later studies [such as those of Sauer and his colleagues on phage λ repressor (14)] have shown the feasibility of rescuing protein function (15). Yet such investigations have used methods either where the generation of mutations is not random and the search has been limited by mutational "hot spots," or where a random search is made only over a limited portion of the gene of interest. In contrast, we report here the results from a complete and unbiased search across the whole gene that allow us to identify those alterations that produce catalytic improvement and to recognize those tracts of sequence within which no changes of amino acid residues give a more efficient catalyst.

The target enzyme was triose-phosphate isomerase, which catalyzes the interconversion of dihydroxyacetone phosphate and D-glyceraldehyde 3-phosphate (Fig. 1). The enzyme is an α_2-chain dimer of subunit M_r 26,500, and the gene encoding the chicken muscle enzyme has been cloned and expressed (16) in an *Escherichia coli* strain (DF502) from which the endogenous isomerase gene (*tpi*) has been excised (17). Crystal structures of both the chicken (18) and yeast (19) enzymes alone and in combination with substrate (20) or inhibitors (21) have been solved to high resolution. Mechanistic studies (22) have shown that the isomerization proceeds through an enediol(ate) intermediate: Glu-165 acts as the base (23, 24), and His-95 (and possibly Lys-13) acts as the electrophile. The rate constants for every step of the catalyzed reaction have been determined both for the wild-type enzyme (25) and for a mutant in which Glu-165 has been changed to aspartate (26); for simplicity, this Glu-165 → Asp mutant enzyme is designated "E165D" (single-letter amino acid code). The wild-type enzyme is an extremely powerful catalyst whose rate of reaction is encounter-controlled (27); the E165D mutant, in contrast, has a $k_{catalysis}$ value for glyceraldehyde phosphate that is smaller by a factor of about 1000 (28). Difference Fourier maps of crystals of these two

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.

*Present address: Merck and Co., Inc., P.O. Box 2000, Rahway, NJ 07065.
†To whom reprint requests should be addressed.

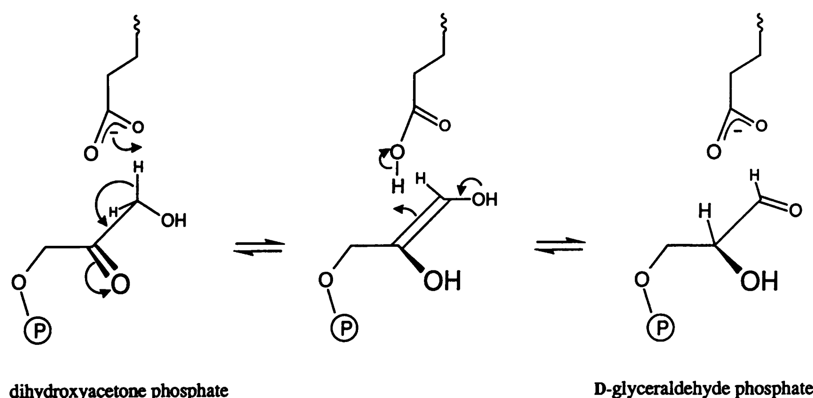

FIG. 1. The interconversion of dihydroxyacetone phosphate and D-glyceraldehyde phosphate, catalyzed by triose-phosphate isomerase.

enzymes show that the only significant change in the structure of the E165D enzyme is a movement of the catalytic carboxylate group by about 1 Å (E. Lolis and G. Petsko, personal communication). These findings emphasize the sensitive dependence of the reaction rate on the precise positioning of catalytic groups. The E165D isomerase mutant, fully characterized in both structural and energetic terms, provides an excellent starting point for catalytic refinement by random mutagenesis and selection.

MATERIALS AND METHODS

Reagents and Strains. Reagents for DNA synthesis were obtained from commercial sources for use on a Milligen/Biosearch model 7500 automated DNA synthesizer. Primer-directed mutagenesis was performed by using the Amersham oligonucleotide-directed *in vitro* mutagenesis system. *E. coli* strain DF502 was provided by D. Fraenkel. *E. coli* strains DH5 (29) and TG1 (30), used in the amplification and sequencing steps, were obtained from commercial sources.

Plasmid Constructs. *pBSRP*. A *Pvu* II fragment containing the multiple cloning site and part of the *lacI* and *lacZ* genes (genes encoding respectively the repressor for the β-galactosidase gene and β-galactosidase) was excised from pBS(+) (Stratagene Cloning Systems). Oligonucleotide-directed mutagenesis (Amersham) was used with a synthetic 84-base oligonucleotide to introduce unique *Eco*RI and *Pst* I sites.

pBSTIM. The triose-phosphate isomerase gene (*tpi*), behind the *trc* promoter (a *tac*-related promoter containing an additional nucleotide between the −35 and −10 regions), was introduced into the pBSRP vector from pX1c (16, 31) on an *Eco*RI–*Pst* I fragment. The *tpi* gene is expressed in the same orientation as the ampicillin-resistance gene in pBSTIM. This vector expresses triose-phosphate isomerase at roughly 100 μg/ml of cell culture and was used in the isolation of purified proteins.

ptsTIM. The ptsTIM[E165D] plasmid was used in the mutagenesis and selection steps. The plasmid was produced by introducing the *Sca* I–*Pst* I fragment of pX1c[E165D], containing the E165D isomerase gene and its promoter, into the pBSRP vector, which also contains unique *Sca* I and *Pst* I sites. This plasmid expresses triose-phosphate isomerase at 7 μg/ml of cell culture.

Random Mutagenesis of the E165D Isomerase. The protocol for mutagenesis using spiked oligonucleotides has been described (32). The mutant E165D isomerase was subjected to mutagenesis by using 10 "spiked" oligonucleotides that had been synthesized with a total contamination level of the three "wrong" phosphoramidites of 2–2.5%. These 10 oligonucleotides ranged from 69 to 92 bases in length and spanned the entire isomerase gene. At least 150,000 transformants of *E. coli* DH5 were obtained from each oligonucleotide window (and at least 300,000 transformants from those oligonucleotides that encode the known active site functionalities Asp-165 and His-95). Several sets of transformants were checked to confirm the efficiency of mutagenesis and the randomness of base misincorporation. The transformants from each oligomer were separately pooled, and the plasmid DNA from each pool was isolated. Transformation of the isomerase-minus *E. coli* strain (DF502) then allowed the selection of colonies that harbor isomerases having a specific catalytic activity greater than that of the starting E165D enzyme.

Selection of Pseudorevertants. Selection for isomerases of increased specific catalytic activity was performed on M63 minimal agar plates (33) containing glycerol as the sole carbon source as described (32). Plasmid DNA from each colony that survived the selection step was then isolated by alkaline lysis (34) after growth of an overnight culture, and the isomerase gene was sequenced across the region of the mutagenic oligonucleotide. Each mutant *tpi* gene, together with the *trc* promoter used for its expression, was excised on an *Eco*RI–*Pst* I fragment, and this fragment was introduced into the pBSRP vector [a modified version of pBS(+), described above]. Each mutant-bearing plasmid was reintroduced into the isomerase-minus strain DF502, and overnight cultures of each mutant were assayed for isomerase activity as described (32). The isomerases from clones that grew on minimal plates containing glycerol, that gave a higher isomerase activity in crude lysates, and that showed an amino acid change in the DNA sequence were purified as described (32).

Steady-State Kinetics. Kinetic measurements were performed on a Hewlett–Packard 8452A diode array spectrophotometer with a Brinkmann RMS 6 temperature controller. Enzyme assays, based on the method of Putman *et al.* (35), were performed at 30°C in 0.1 M triethanolamine hydrochloride buffer (pH 7.6) containing 10 mM EDTA.

RESULTS AND DISCUSSION

Random Mutagenesis. The random mutagenesis protocol that we adopt must be definably random, and the amino acid sequences generated must not be biased or limited by features such as hot spots (11, 36) or unreactive regions in the DNA substrate. Though there are many ways of introducing undirected changes into particular genes (37), the criterion of assured randomness rules out all *in vivo* methods such as the use of mutator strains (38), the application of mutagenic agents to whole cells, or the induction of error-prone repair systems (39). The use of enzyme-dependent methods (40), however ingenious (41), with the isolated gene can also suffer from sequence selectivity. Even the chemical mutagenesis procedure developed by Myers *et al.* (42) gave an uneven distribution of mutations in our system. A second requirement of the mutagenesis protocol is that the most extensive region of sequence space be searched. The protocol must

maximize the access to the other 19 amino acids that can be encoded at each position. When one base in a codon is changed, an average of only 5.7 (of the 19) other amino acids can be encoded (and these alterations tend to be conservative), but when two bases in a codon are changed, the average number of other amino acids that can be accessed rises to 15.7. This number obviously includes amino acids having a variety of side-chain characteristics.

We developed the spiked oligonucleotide primer approach (32) to generate our library of mutants. Long oligonucleotides were synthesized that contain a small fixed percentage of the three wrong bases at every position (43), and these spiked primers (32, 43) were used in a high-yield mutagenesis protocol (44) to produce a collection of mutants that span the length of the primer. This procedure is illustrated in Fig. 2. By using several primers, the whole of the structural gene was subjected to mutagenic variation, the severity of which was determined by the percentage of wrong amidite reagents used in the synthesis of the primer.

The probability P of finding n errors in an m-long oligonucleotide that is synthesized with a fraction α of the three wrong nucleotides at each position is:

$$P(n, m, \alpha) = [m!/(m - n)!n!][\alpha]^n[1 - \alpha]^{m-n}. \quad [1]$$

Specifically, by screening 150,000 transformants with mutagenic 75-base primers at a total contamination level of 2.5% of the three wrong phosphoramidites, we have covered >99% of the one-base changes and ≈75% of the two-base changes that are possible in that oligomer window. Two other features of the method are worth noting. First, we can specify "immutable" positions, where for particular codons, clean (unspiked) reagents are used. This possibility was important in the work described here, where we wished to minimize the background of true revertants in the search for pseudorevertant clones. Second, the incorporation of a silent base change within each primer as a marker provides an early check on the efficiency of the mutagenesis procedure and confirms that a high percentage of the resulting clones are derived from the primer.

Because there are no hot spots in a synthetic oligonucleotide, provided that the synthesis is performed with pure reagents of equal reactivity, there is an equal probability of mutation at every position along the primer. In principle, there is no bias in the percentage of spiking, in the nature of the improper base, or in the position of the change when this

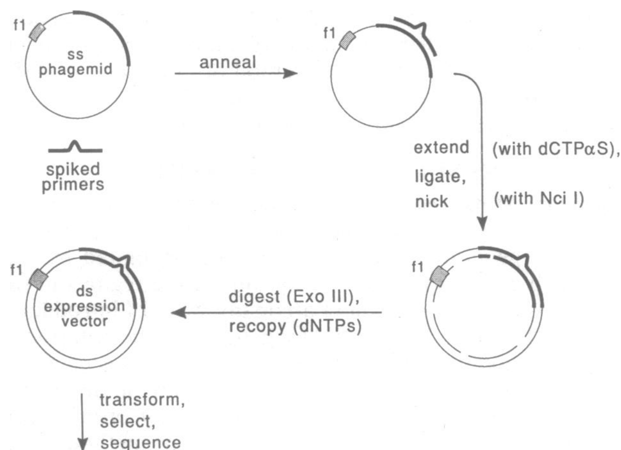

FIG. 2. The mutagenesis protocol. Single-stranded (ss) template was prepared from early logarithmic-phase *E. coli* TG1(ptsTIM-[E165D]) cells superinfected with VCS-M13 phage (Stratagene). Only one of a pool of spiked oligonucleotide primers is shown. ds, Double-stranded; dCTPαS, deoxycytidine 5′-[α-thio]triphosphate; Exo III, exonuclease III.

method is used; in practice, each of these concerns is checked experimentally (31, 32). Deviations from randomness deriving from differential phosphoramidite stability, from synthesizer malfunction, from differences in coupling kinetics, from inadequate amidite mixing, or from sequence-dependent effects on base incorporation rates, can be identified and eliminated. In the present case, up to 100 individual clones from a mutagenesis experiment were isolated without any phenotypic selection and were sequenced to verify that the efficiency and randomness criteria were satisfied (32).

Identification and Selection of Pseudorevertants. At least 10^6 DF502 transformants deriving from each of the 10 spiked primers were subjected to selection on glycerol plates. This selection is based on the following trick. While the isomerase-minus host strain will grow on glycerol (from which dihydroxyacetone phosphate is made) containing lactate (which provides D-glyceraldehyde phosphate), it cannot grow on either glycerol alone or lactate alone (16). Plasmid-mediated expression of wild-type isomerase from the *trc* promoter allows DF502 to grow on either glycerol or lactate as the sole source of carbon. When the mutant E165D isomerase is expressed from the same promoter, however, the transformants grow on lactate but not on glycerol (28). Herein lies the basis for the selection of mutants: colonies that appear on glycerol plates have more isomerase activity than the E165D mutant. [The reason why colonies expressing less than about 100 picounits per cell of isomerase activity grow on lactate but not on glycerol is not known, but it seems likely that cells grown on glycerol accumulate dihydroxyacetone phosphate that decomposes nonenzymatically to toxic methylglyoxal (45). Cells grown on lactate will not accumulate such high levels of D-glyceraldehyde phosphate (the path to which is longer and thermodynamically more unfavorable). Less methylglyoxal forms, and the cells survive.] Happily, the selection is not absolute, and even cells that only produce the E165D mutant may form small colonies on glycerol plates after 48 hr. As a result, colony size can be used on the borderline as an indicator of isomerase activity, and mutants that are only slightly better catalysts than the starting E165D enzyme can be scored and identified.

Overall, 436 clones were picked for further characterization. This number is very much higher than the number of true pseudorevertants for several reasons. First, we picked candidate colonies generously to be sure that true but only marginally improved pseudorevertants would not be missed. Second, we were aware of secondary structure near the initiation codon that could modulate both transcription and translation (31), and the selection could not distinguish between a more active enzyme and more of a less active one. Third, although the synthesizer misincorporation frequency is low when unspiked phosphoramidites are used, these errors, as well as polymerase infidelity during the mutagenesis protocol, lead to the appearance of some true revertants at position 165. From all of the generated mutants, which peppered the entire isomerase gene, six pseudorevertants were identified, sequenced, and characterized (see Table 1). The kinetic characteristics of each of the purified proteins, improved as catalysts by between 1.3- and 19-fold, are summarized in Fig. 3*A* and Table 2.

Number and Nature of the Pseudorevertants. The search for pseudorevertant isomerases from the E165D mutant has produced a variety of amino acid changes, scattered across the entire sequence, that can partially compensate for the lesion represented by E165D. The wide distribution of mutations suggests that the spiked primer approach is gratifyingly thorough in producing random changes in a target gene. It is also clear from Table 1 that all of the effective changes are in or close to the active site. Indeed, none of those oligonucleotides that only encode regions of the protein remote (i.e., >5 Å) from the active site (e.g., numbers 2, 3, 5, 6, and 8) give rise

Table 1. Second-site suppressor mutations that improve the catalytic activity of the sluggish E165D mutant of chicken triose-phosphate isomerase

Spiked oligonucleotide	Amino acids encoded	Part of active site?*	Isolated pseudorevertants†
1	1–26	++	G10S
2	25–52	−	None‡
3	51–77	+	None‡
4	76–100	++	S96P, or S96T, or E97D
5	99–126	+	None‡
6	125–151	−	None‡
7	151–179	++	V167D
8	178–205	−	None‡
9	204–228	++	None‡
10	226–248	++	G233R

*++, At least one amino acid residue in the encoded region falls within 5 Å of enzyme-bound phosphoglycolohydroxamate; +, the nearest amino acid residues in the encoded region lie between 5 Å and 8 Å of the bound inhibitor; −, no amino acid residue in the encoded region lies nearer than 8 Å from the bound inhibitor. Distances were measured by using the molecular graphics program QUANTA 10 using the coordinates of the yeast isomerase complexed with phosphoglycolohydroxamate[21].
†Table designations are in single-letter amino acid code and are explained in the legend to Fig. 3.
‡Of the colonies picked for further characterization, none showed an increase (in lysates of cells transformed with the gene that had been subcloned into pBSRP) in the catalytic activity towards D-glyceraldehyde phosphate.

to any pseudorevertants. While it is obviously premature to generalize, it may turn out for enzymes such as triose-phosphate isomerase, where there is no evidence for information transfer either within or between protein subunits, that only local changes in the first or second shell of amino acid residues near the active site will affect the specific catalytic activity. Perhaps this finding should not surprise us. When one examines the conserved regions of particular enzymes from a range of species, these regions are usually localized in and near to the active site. Such is certainly the case for triose-phosphate isomerase, for which some 12 sequences are known

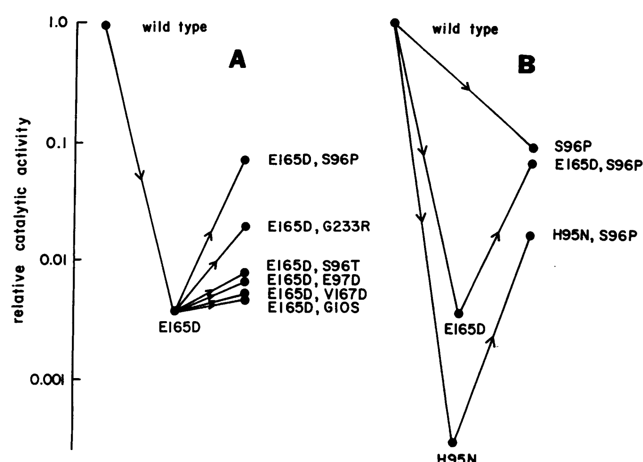

FIG. 3. (A) Catalytic activity of various mutant triose-phosphate isomerases obtained by spiked primer mutagenesis of the gene encoding the sluggish isomerase mutant E165D. (B) Catalytic activity of the pseudorevertants obtained by chemical mutagenesis of the gene encoding two sluggish triose-phosphate isomerase mutants: E165D and H95N. The relative catalytic activity of the S96P point mutant is also shown. Mutant designations: E165D, Glu-165 → Asp; S96P, Ser-96 → Pro; G233R, Gly-233 → Arg; S96T, Ser-96 → Thr; E97D, Glu-97 → Asp; V167D, Val-167 → Asp; G10S, Gly-10 → Ser; H95N, His-95 → Asn.

Table 2. Kinetic characteristics* of the starting mutant triose-phosphate isomerase and of the six pseudorevertant enzymes of increased catalytic potency

Enzyme†	$\overrightarrow{k}_{cat}$, s^{-1}	$\overrightarrow{K}_m$, mM	$\overleftarrow{k}_{cat}$, s^{-1}	$\overleftarrow{K}_m$, mM	Relative $\overrightarrow{k}_{cat}/\overrightarrow{K}_m$‡
Wild type	600	0.65	8300	0.42	270
E165D	4.1	1.2	4.2	0.078	1.0
E165D, G10S	4.7	1.1	14	0.18	1.3
E165D, S96P	3.4	0.053	68	0.066	19
E165D, S96T	9.5	1.3	17	0.10	2.1
E165D, E97D	5.8	0.93	10	0.079	1.8
E165D, V167D	8.4	1.7	15	0.17	1.4
E165D, G233R	6.0	0.33	8.4	0.029	5.2

*An arrow pointing to the right signifies that dihydroxyacetone phosphate was substrate, and an arrow pointing to the left signifies that glyceraldehyde phosphate was substrate.
†Designations are in amino-acid single letter code and are explained in the legend to Fig. 3.
‡Relative to the E165D mutant enzyme.

from species as diverse as *Bacillus*, the coelacanth, and man (46–56). This view of enzyme catalysts (at least, of those that catalyze simple chemical reactions) would lead one to believe that enzymes are generated by the appropriate disposition of catalytic functionality upon a stable three-dimensional scaffolding. The arrangement must be precise [we must not forget that in the E165D mutant of isomerase, less than 1 Å is worth nearly a factor of 1000 in k_{cat} (26)], but the experiments described here suggest that there will be a number of ways to achieve that precision. Further, the idea of a range of pendant catalytic groups from a stable scaffolding is made attractive by the growing list of enzymes [now more than 14 (57, 58)] that catalyze a variety of different reactions but that all have their active sites at the C-terminal end of an eightfold β-barrel, the archetype of which is triose-phosphate isomerase (18).

The above speculations notwithstanding, all of the catalytically improving changes that we have found are in or close to the active site. These positions are highlighted in Fig. 4. Gly-10 lies three residues away from active-site Lys-13,

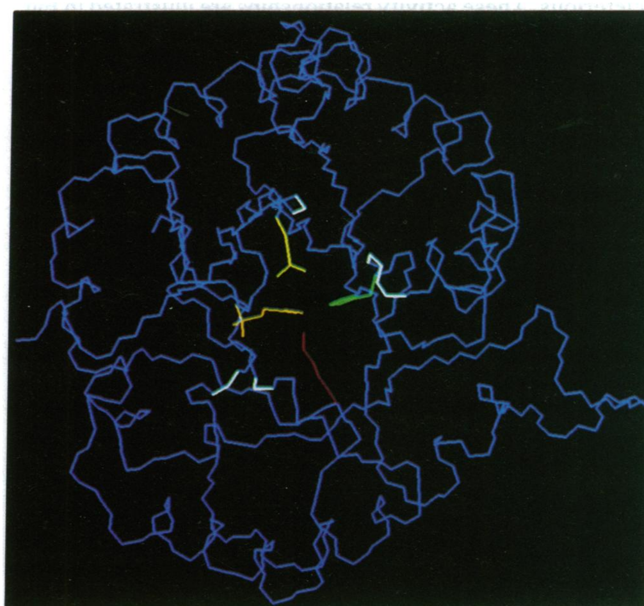

FIG. 4. Pseudorevertant triose-phosphate isomerases. The positions of amino acid changes that result in improvement of the catalytic potency of the E165D mutant isomerase are shown in white. The main chain is shown in blue, and three active site residues of the wild-type enzyme (Lys-13, His-95, and Glu-165) are shown in red, green, and yellow, respectively. The bound substrate is shown in orange.

Ser-96 and Glu-97 are adjacent to the catalytic electrophilic His-95, Val-167 is two residues from the essential catalytic base Glu-165 (which is aspartate in the sluggish mutant), and Gly-233 provides a hydrogen bond to substrate [as deduced from the crystal structure of the yeast enzyme complexed with phosphoglycolohydroxamate (21)]. While the molecular details of these changes and the correlation with their functional consequences must await the appropriate structural results, it is clear that each of these local alterations nudges the active site back towards higher catalytic effectiveness. Interestingly, all of the second-site suppressor mutations occur at positions where the sequence is highly conserved among species. In fact, only two of the altered residues show any variability at all, and each of these changes is conservative: Val-167 is a leucine in the *Bacillus stearothermophilus* enzyme (48), and Gly-10 is an alanine in the isomerase from *Trypanosoma bruceii* (56).

Our finding of six pseudorevertant isomerases of improved catalytic potency by using the spiked primer approach sharply contrasts with the fact that chemical mutagenesis yielded only one. In three discrete experiments in which the gene for the E165D isomerase was subjected to the chemical procedures of Myers *et al.* (42), only one pseudorevertant [E165D, S96P (designation for the Ser-96 → Pro mutant)] was ever found (31). From this result alone, we could not be certain whether the Ser-96 → Pro mutation was the only possible way that the E165D mutant could be improved, or whether biases in the method (that is, the nature of the chemical mutagens used and of the particular gene being subjected to them) were simply finding a mutagenic hot spot at position 96. This concern that chemical reagents do not generate truly random changes was heightened by our finding that chemical mutagenesis of a quite different sluggish isomerase with a His-95 → Asn mutation (designated H95N, the specific catalytic activity of which is about 10^{-4} that of the wild type) also produced (59) a single pseudorevertant: H95N, S96P! The possibility that the S96P mutation is somehow a "global" suppressor [compare the global suppression of structural stability mutants in staphylococcal nuclease (60)] was ruled out by the finding that this change in the wild type is catalytically deleterious. These activity relationships are illustrated in Fig. 3*B* and suggest that the chemical procedure may indeed be biased. The spiked primer method is clearly superior in the exploration of sequence space.

These experiments not only demonstrate the feasibility of improving the catalytic effectiveness of an enzyme but also show that a range of different structural modifications can produce similar functional outcomes. Those who aim to manipulate the specificity or catalytic activity of enzymes can be encouraged by these results, for the notion that enzymes are the extremely rare products of an evolutionary process that cannot be simulated is unnecessarily defeatist. We have shown that six different changes will improve the catalytic activity of one sluggish isomerase mutant. Changes in the catalytic activity of enzymes by random mutagenesis and selection should become routine, and as these changes are scrutinized, our understanding of the structure–function relationship will deepen.

This work was supported by the National Institutes of Health.

1. Leatherbarrow, R. J. & Fersht, A. R. (1986) *Protein Eng.* **1**, 7–16.
2. Knowles, J. R. (1987) *Science* **236**, 1252–1258.
3. Shaw, W. V. (1987) *Biochem. J.* **246**, 1–17.
4. Gerlt, J. A. (1987) *Chem. Rev.* **87**, 1079–1105.
5. Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R. & Rutter, W. J. (1985) *Science* **228**, 291–297.
6. Cronin, C. S. & Kirsch, J. F. (1988) *Biochemistry* **27**, 4572–4579.
7. Wilks, H. M., Hart, K. W., Feeney, R., Dunn, C. R., Muirhead, H., Chia, W. H., Barstow, D. A., Atkinson, T. A., Clarke, A. R. & Holbrook, J. J. (1988) *Science* **242**, 1541–1544.
8. Wharton, R. P. & Ptashne, M. (1985) *Nature (London)* **316**, 601–605.
9. Knight, K. L. & Sauer, R. T. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 797–801.
10. Maynard-Smith, J. (1970) *Nature (London)* **225**, 563–564.
11. Benzer, S. (1961) *Proc. Natl. Acad. Sci. USA* **43**, 403–415.
12. Crick, F. H. C., Barnett, L., Brenner, S. & Watts-Tobin, R. J. (1961) *Nature (London)* **192**, 1227–1232.
13. Yanofsky, C. (1971) *J. Am. Med. Assoc.* **218**, 1026–1035.
14. Hecht, M. H. & Sauer, R. T. (1985) *J. Mol. Biol.* **186**, 53–63.
15. Dunn, I. S., Cowan, R. & Jennings, P. A. (1988) *Protein Eng.* **2**, 283–291.
16. Straus, D. & Gilbert, W. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2014–2018.
17. Frankel, D. (1986) *Annu. Rev. Biochem.* **55**, 317–337.
18. Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. D., Priddle, J. D. & Waley, S. C. (1975) *Nature (London)* **255**, 609–614.
19. Alber, T., Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D., Rivers, P. S. & Wilson, I. A. (1981) *Philos. Trans. R. Soc. London B* **293**, 159–171.
20. Alber, T., Davenport, R. C., Jr., Giammona, D. A., Lolis, E., Petsko, G. A. & Ringe, D. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52**, 603–613.
21. Davenport, R. C., Jr. (1986) Ph.D. Thesis (Massachussets Institute of Technology, Cambridge, MA).
22. Rieder, S. V. & Rose, I. A. (1959) *J. Biol. Chem.* **234**, 1007–1010.
23. Hartman, F. C. (1968) *Biochem. Biophys. Res. Commun.* **33**, 888–894.
24. de la Mare, S., Coulson, A. F. W., Knowles, J. R., Priddle, J. D. & Offord, R. E. (1972) *Biochem. J.* **129**, 321–331.
25. Albery, W. J. & Knowles, J. R. (1976) *Biochemistry* **15**, 5627–5631.
26. Raines, R. T., Sutton, E. L., Straus, D. R., Gilbert, W. & Knowles, J. R. (1986) *Biochemistry* **25**, 7142–7154.
27. Blacklow, S. C., Raines, R. T., Lim, W. A., Zamore, P. D. & Knowles, J. R. (1988) *Biochemistry* **27**, 1158–1167.
28. Straus, D., Raines, R. T., Kawashima, E., Knowles, J. R. & Gilbert, W. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 2272–2276.
29. Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557–580.
30. Carter, P., Bedouelle, H. & Winter, G. P. (1985) *Nucleic Acids Res.* **13**, 4431–4443.
31. Hermes, J. D. & Knowles, J. R. (1987) *Proceedings of the Robert A. Welch Foundation XXXI, Design of Enzymes and Enzyme Models* (Robert A. Welch Foundation, Houston, TX), pp. 207–226.
32. Hermes, J. D., Parekh, S. M., Blacklow, S. C., Köster, H. & Knowles, J. R. (1990) *Gene*, in press.
33. Miller, J. H. (1972) *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
34. Birnboim, H. C. & Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523.
35. Putman, S. J., Coulson, A. F. W., Farley, I. R. T., Riddleston, B. & Knowles, J. R. (1972) *Biochem. J.* **129**, 301–310.
36. Coulondre, C., Miller, J. H., Farabaugh, P. J. & Gilbert, W. (1978) *Nature (London)* **274**, 775–780.
37. Botstein, D. & Shortle, D. (1985) *Science* **229**, 1193–1201.
38. Cox, E. C. (1976) *Annu. Rev. Genet.* **10**, 135–156.
39. Muller, W., Weber, H., Meyer, F. & Weissmann, C. (1978) *J. Mol. Biol.* **124**, 343–358.
40. Shortle, D., Grisafi, P., Benkovic, S. J. & Botstein, D. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 1588–1592.
41. Lehtovaara, P. M., Koivula, A. K., Bamford, J. & Knowles, J. K. C. (1988) *Protein Eng.* **2**, 63–68.
42. Myers, R. M., Lerman, L. S. & Maniatis, T. (1985) *Science* **229**, 242–247.
43. Ner, S. S., Goodin, D. B. & Smith, M. (1988) *DNA* **7**, 127–134.
44. Taylor, J. W., Ott, J. & Eckstein, F. (1985) *Nucleic Acids Res.* **13**, 8765–8785.
45. Cooper, R. A. (1984) *Annu. Rev. Microbiol.* **38**, 49–68.
46. Artavanis-Tsakonis, S. & Harris, J. I. (1980) *Eur. J. Biochem.* **108**, 599–611.
47. Pichersky, E., Gottlieb, L. D. & Hess, J. F. (1984) *Mol. Gen. Genet.* **195**, 314–320.
48. Kolb, E., Harris, J. I. & Bridger, J. (1974) *Biochem. J.* **137**, 185–197.
49. Corran, P. H. & Waley, S. G. (1975) *Biochem. J.* **137**, 335–344.
50. Alber, T. & Kawasaki, G. (1982) *J. Mol. Appl. Genet.* **1**, 419–434.
51. Lu, H. S., Yuan, P. M. & Gracy, R. W. (1984) *J. Biol. Chem.* **259**, 11958–11968.
52. Maquat, L. E., Chilote, R. & Ryan, P. M. (1985) *J. Biol. Chem.* **260**, 3748–3753.
53. Russell, P. R. (1985) *Gene* **40**, 125–130.
54. McKnight, G. L., O'Hara, P. J. & Parker, M. L. (1986) *Cell* **46**, 143–147.
55. Marchionni, M. & Gilbert, W. (1986) *Cell* **46**, 133–141.
56. Swinkels, B. W., Gibson, W. C., Osingas, K. A., Kramer, R., Veeneman, G. H., vanBoon, J. H. & Borst, P. (1986) *EMBO J.* **5**, 1291–1298.
57. Brändèn, C. (1980) *Q. Rev. Biophys.* **13**, 317–338.
58. Chothia, C. (1988) *Nature (London)* **333**, 598–599.
59. Hermes, J. D., Blacklow, S. C., Gallo, K. A., Bauer, A. J. & Knowles, J. R. (1987) in *Protein Structure, Folding, and Design 2* (Liss, New York), pp. 257–264.
60. Shortle, D. & Lin, B. (1985) *Genetics* **110**, 539–555.