[intensity $I > 1.5\sigma(I)$]. The merging R was 0.068 and the data were complete to 2.4 Å, while more than 60% of the data in the last shell were also measured.

14. S. B. H. Kent, *Annu. Rev. Biochem.* **57**, 957 (1988); J. Schneider and S. B. H. Kent, *Cell* **54**, 363 (1988).
15. The starting model was a crystallographic dimer of HIV-1 PR with perfect twofold symmetry. Rotation function calculations with either MERLOT [P. M. D. Fitzgerald, *J. Appl. Crystallogr.* **21**, 53 (1988)] or PROTEIN [Steigemann, thesis, Technical University, Munich (1974)] were unambiguous. The final Crowther rotation angles were (−2.5°, 85.5°, 99.5°), and this peak was found in the Patterson superposition at 4.3σ level for the data in the 10 to 3 Å shell. Several translation function algorithms also yielded consistent results, with the highest peak (6.8σ using the program RVAMAP in MERLOT) corresponding to fractional unit cell translations of (0.14, 0.405, 0.48). The preliminary orientation was further optimized with rigid-body refinement in X-PLOR [A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987)]. The molecular dynamics refinement with X-PLOR, followed by restrained least-squares refinement using PROLSQ [W. A. Hendrickson, *Methods Enzymol.* **115**, 252 (1985)], lowered the R factor from 0.49 to 0.221 for the data in the 10 to 3 Å shell. This process, which did not require manual intervention, introduced shifts as large as 7 Å in the flap regions (residues 45 to 52). Although the actual positions of the flaps resulting from this refinement were incorrect, the resulting electron density was unambiguous and allowed us to retrace the polypeptide chain without any difficulty. Further refinement with PROLSQ included the inhibitor as well as a limited number of water molecules (70 at present). The tetrahedral configuration of C and N atoms of the reduced amide in the inhibitor was specifically restrained. The solvent model is not yet complete and requires further rebuilding, but that procedure should not affect the description of the enzyme-inhibitor interactions, which are well defined. The present model is characterized by an R factor of 0.176 for 7813 reflections between 10 and 2.25 Å, with the deviations from ideality of 0.019 Å for bonds, 0.014 Å for the planes, and 0.22 Å^3 for the chiral volumes.
16. M. V. Toth and G. R. Marshall, unpublished results.
17. Abbreviations: A, Ala; D, Asp; G, Gly; I, Ile; L, Leu; P, Pro; R, Arg; T, Thr; and V, Val; W is water and X is Nle.
18. We thank M. Jaskólski for advice in the initial stages of this project, D. Davies for providing us with standard groups for least-squares refinement, and L. Palmer for critical reading of the manuscript and checking of some data. The Advanced Scientific Computing Laboratory, FCRF, provided a substantial allocation of time on their CRAY X-MP supercomputer. Research sponsored in part by the National Cancer Institute, DHHS, under contract N01-C0-74101 with BRI, in part by funds from the NSF Biological Instrumentation Division to S.B.H.K., by NIH grants SM-24483 and A-127302, and by Monsanto grant 44353K to G.R.M. The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government.

27 September 1989; accepted 20 October 1989

Structural Basis for Misaminoacylation by Mutant *E. coli* Glutaminyl-tRNA Synthetase Enzymes

JOHN J. PERONA, ROBERT N. SWANSON,* MARK A. ROULD, THOMAS A. STEITZ, DIETER SÖLL

A single-site mutant of *Escherichia coli* glutaminyl-synthetase (D235N, GlnRS7) that incorrectly acylates in vivo the amber suppressor *supF* tyrosine transfer RNA ($tRNA^{Tyr}$) with glutamine has been described. Two additional mutant forms of the enzyme showing this misacylation property have now been isolated in vivo (D235G, GlnRS10; I129T, GlnRS15). All three mischarging mutant enzymes still retain a certain degree of tRNA specificity; in vivo they acylate *supE* glutaminyl tRNA ($tRNA^{Gln}$) and *supF* $tRNA^{Tyr}$ but not a number of other suppressor tRNA's. These genetic experiments define two positions in GlnRS where amino acid substitution results in a relaxed specificity of tRNA discrimination. The crystal structure of the GlnRS:$tRNA^{Gln}$ complex provides a structural basis for interpreting these data. In the wild-type enzyme Asp^{235} makes sequence-specific hydrogen bonds through its side chain carboxylate group with base pair G3 · C70 in the minor groove of the acceptor stem of the tRNA. This observation implicates base pair 3 · 70 as one of the identity determinants of $tRNA^{Gln}$. Isoleucine 129 is positioned adjacent to the phosphate of nucleotide C74, which forms part of a hairpin structure adopted by the acceptor end of the complexed tRNA molecule. These results identify specific areas in the structure of the complex that are critical to accurate tRNA discrimination by GlnRS.

THE ACCURACY OF PROTEIN BIOSYNthesis in living cells depends critically on the acylation of transfer RNA (tRNA) molecules with the correct amino acid. Aminoacyl-tRNA synthetases, the enzymes catalyzing this reaction, have a high degree of specificity in discriminating among structurally similar tRNA molecules. The identification of the specific chemical groups responsible for the selectivity of interactions between tRNA and protein has been the focus of many studies in which various physical, biochemical, and genetic techniques were used (*1–5*). This has led (for some systems) to the partial identification of the set of nucleotides, known as identity elements, which serve to distinguish that set of isoacceptor RNA's specific to a given amino acid in vivo (*6*). The development of the technology required for the in vitro synthesis of any desired tRNA has allowed the identification of the nucleotides required for tRNA recognition by yeast phenylalanyl-tRNA synthetase (*7*). Genetic and biochemical studies have shown that some identity elements of glutaminyl tRNA ($tRNA^{Gln}$) are located at the acceptor end (nucleotides G73 and U1 · A72) (*3*) and in the anticodon (U35) (*2, 4*).

Department of Molecular Biophysics and Biochemistry, and the Howard Hughes Medical Institute, Yale University, New Haven, CT 06511.

*Present address: Service de Biochimie, CNRS de Saclay, Saclay, France.

Table 1. Suppression of the $lacZ_{1000}$ gene by different amber suppressor tRNAs (see text). Lysates of λ phages carrying the *glnS*$^+$, *glnS7*, *glnS10*, and *glnS15* alleles were spotted onto lactose minimal plates seeded with $lacZ_{1000}$ strains carrying the various amber suppressors. The plates were examined for growth after 2 to 3 days of incubation at 30°C. Bacterial growth (+) indicates mischarging of the suppressor tRNA with glutamine. The nomenclature for the in vitro derived Cys, Phe, and Ala suppressors is according to (*6*).

Suppressor	*glnS* allele			
	glnS$^+$	*glnS7*	*glnS10*	*glnS15*
supD (Ser)	−	−	−	−
supE (Gln)	+	+	+	+
supF (Tyr)	−	+	+	+
supP (Leu)	−	−	−	−
pGFIB:Cys	−	−	−	−
pGFIB:Phe	−	−	−	−
pGFIB:Ala2	−	−	−	−

There has been considerably less progress, however, in discovering which amino acids in synthetases are crucial for recognition of their respective tRNA's (*8*). In order to address this question we have used an in vivo genetic approach to generate mutations in *glnS* (the gene for *E. coli* GlnRS), which cause misacylation of noncognate tRNA species with glutamine. Selection for mutants of GlnRS that can mischarge noncognate tRNA's with glutamine is based on the suppression spectrum of an amber mutation in the gene for *E. coli* β-galactosidase (the $lacZ_{1000}$ mutation). If not suppressed by a suppressor tRNA this mutant gene gives rise to a truncated β-galactosidase protein. Insertion of an amino acid in response to the amber codon gives rise to full-length protein. Thus, suppression with the glutamine-inserting *supE* tRNA allows cells to grow on minimal lactose plates. However, the serine-

inserting *supD* and tyrosine-inserting *supF* tRNA cannot produce a functional β-galactosidase enzyme by suppression of the $lacZ_{1000}$ gene (*9*) (Table 1). Presumably insertion of Ser and Tyr at the site of the amber mutation gives an inactive protein. Thus a strain that is $lacZ_{1000}$, *supF* is phenotypically Lac^-. This feature can then be used to select for mutants of *glnS* that mischarge *supF* $tRNA^{Tyr}$ with glutamine and confer a Lac^+ phenotype on the strain.

In the aforementioned way, we selected earlier a variant form of GlnRS, GlnRS7 (*9*). The molecular change responsible for the mischarging phenotype was an Asp→Asn substitution at position 235 (D235N) (*10*). The corresponding change in the *glnS7* gene is a G→A transition, which interrupts a *Bst* NI restriction site. This feature was used to isolate further variants that are not of the *glnS7* type. After screening of DNA blots we found another variant, *glnS10*, which did not interrupt this Bst NI site. Sequence analysis of the gene showed an A→G transition immediately adjacent (at the 3′ side) to the base pair altered in *glnS7*. The *glnS10* gene is altered at the same codon, resulting in an Asp^{235} to Gly^{235} change (D235G) in the gene product (Fig. 1). Since both the *glnS7* and *glnS10* mutations involve G residues, DNA sequencing of G tracks could identify additional mischarging mutants. An isolate (*glnS15*) was found with the wild-type codon at amino acid 235. Complete sequence analysis showed a T→C transition in the gene causing replacement of Ile^{129} with Thr^{129} (I129T) in GlnRS15 (Fig. 1).

Three tests were performed with the mutants. The ability of the mutant enzymes to acylate $tRNA^{Gln}$ with glutamine in vivo was verified by complementation of the temperature-sensitive *glnS1* gene (*11*) to allow a temperature-sensitive *E. coli* strain to survive at the nonpermissive temperature. What is the efficiency of in vivo mischarging by these enzymes? In order to assess the relative degree of mischarging by GlnRS10 and GlnRS15, we assayed the β-galactosidase levels in lysogens harboring the different mutant prophages and found that *glnS15* has a weaker mischarging phenotype than does *glnS10* (Table 2). In similar experiments *glnS10*, in turn, mischarges less well than does the original mischarging mutant *glnS7* (*12*). Both *glnS10* and *glnS7*, however, are better mischargers than a number of in vitro–derived *glnS* alleles coding for different amino acid replacements at position 235 (*13*). Thus, *glnS15* seems to cause the least mischarging of the three in vivo selected misacylation mutants. This corresponds well with the observed slower growth rate of the strain carrying the λgt*glnS15* prophage on lactose minimal medium relative to most other isolates obtained during selection of mischarging mutants.

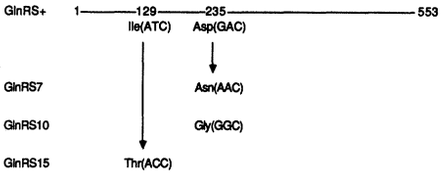

Fig. 1. In vivo derived mischarging mutants of GlnRS. Mutants of the *glnS* gene (carried as a duplicate copy on phage λ over the chromosomal $glnS^+$ gene) were obtained either by hydroxylamine treatment of isolated phage or via plate lysates with *E. coli* strain ES1578 (*mutD5*) as a host (*12*). Isolation of mischarging mutants by selection of phenotypically Lac^+ colonies of *E. coli* strain BT32 (relevant markers $lacZ_{1000}$, *supF*) was as described (*9, 12*). Screening of mutants to determine new isolates which differed from *glnS7* was by restriction analysis, genetic mapping, and DNA sequencing (*18*).

Table 2. Mischarging levels in *glnS10* and *glnS15* strains. Lysogens were constructed by infecting strain RS109 (relevant markers are *supF*, $lacZ_{1000}$, *recA*) with the phages λgt$glnS^+$, λgt*glnS10*, and λgt*glnS15* at a low multiplicity of infection (about 1:5) and then selecting for λ immunity. Lysogens were also made that carried the phage λpsu^+2 (*15*), which contains the gene for the *supE* suppressor $tRNA^{Gln}$. Six isolates of each type of lysogen were selected and β-galactosidase assays were performed to quantitate mischarging. The β-galactosidase content of the different isolates of a given type of lysogen were generally consistent with one another. If, however, the β-galactosidase activity of one isolate was about twice that of the others, the isolate was presumed to be a double lysogen and was discarded. RS109(λ*supE*) measures the efficiency of glutamine insertion into $lacZ_{1000}$ by *supE* (*17*).

Strain	β-Galactosidase (Miller units)
RS109(λ$glnS^+$)	6
RS109(λ*glnS10*)	32
RS109(λ*glnS15*)	20
RS109(λ*supE*)	36

In a third experiment, the ability of the mischarging mutant enzymes to misacylate a wider range of suppressor tRNA's was evaluated. In vivo tests for the mischarging suppression of the $lacZ_{1000}$ gene were carried out in the presence of the *supD* $tRNA^{Ser}$ and the *supP* $tRNA^{Leu}$, as well as in the presence of plasmids carrying in vitro–constructed $tRNA^{Phe}$, $tRNA^{Cys}$, and $tRNA^{Ala}$ amber suppressors (*6*). Apart from *supE*, only *supF* was an active suppressor (Table 1). Thus, the mutant GlnRS enzymes still retain some limited specificity in tRNA recognition.

The structure of the GlnRS:$tRNA^{Gln}$ complex has been determined to a resolution of 2.8 Å by Rould *et al.* (*14*). A large interface comprising interactions of approximately 50 to 60 amino acid residues interacting with about 30 tRNA nucleotides is observed. The mischarging mutant proteins identify specific sites on the macromolecular contact surface that are involved in tRNA discrimination. We can thus begin to distinguish those intermolecular interactions responsible for specificity from those playing only a general role in complex formation. The aspartyl side chain at residue 235 of the wild-type GlnRS makes hydrogen bonding interactions with both G3 of base pair G3 · C70 of the tRNA as well as with a buried water molecule (*14*) (Fig. 2A). This observation strongly implicates base pair 3 · 70 as one of the identity determinants of $tRNA^{Gln}$. Position 235 is located at the amino end of an α helix within the dinucleotide fold of the enzyme. The carboxylate side chain forms part of a hydrogen bonding protein surface that is complementary to base pairs G2 · C71 and G3 · C70 in the minor groove of the acceptor stem. Mutation of the aspartate to asparagine in GlnRS7 or glycine in GlnRS10 must result in the disruption or alteration of at least part of the observed hydrogen bonding network between the protein and $tRNA^{Gln}$ in this region. This disruption might result in the observed decreased selectivity of the mutant enzymes for $tRNA^{Gln}$. Alternatively or in addition, the mischarging mutants could, conceivably, interact more readily than does $GlnRS^+$ with the *supF* $tRNA^{Tyr}$. Determination of the crystal structures of complexes of the mutant enzymes with $tRNA^{Gln}$ and *supF* $tRNA^{Tyr}$ will be required to understand the detailed basis for the mischarging shown by these enzymes.

The Cγ atom of Ile^{129} is positioned approximately 3.6 Å distant from the phosphate group of nucleotide C74 (Fig. 2B). This amino acid is found in the 110-residue acceptor-binding domain of the enzyme that is inserted between the two halves of the dinucleotide binding fold. The primary function of this domain appears to be the binding and stabilization of the hairpin structure at the acceptor end of the tRNA such that the terminal ribose of nucleotide A76 is precisely placed in the active site. Rould *et al.* (*14*) suggest that U1 · A72 and G73 in $tRNA^{Gln}$ may be identity elements for this tRNA because they can facilitate the adoption of this conformation at a lower free energy cost than other nucleotides. Mutation of residue 129 to threonine in GlnRS15 could allow the γ-hydroxyl of the mutant enzyme to make additional interactions with the phosphate of nucleotide C74 of *supF* $tRNA^{Tyr}$. This proposed interaction might help to stabilize the required hairpin structure at the acceptor end of this species even in the absence of the $tRNA^{Gln}$-specific nucleotides at positions 1 · 72 and 73.

On the basis of the combined genetic, biochemical, and crystallographic data Rould *et al.* (*14*) have proposed that some of the identity elements of $tRNA^{Gln}$ are U35, base pairs U1 · A72, G2 · C71, G3 · C70, and the discriminator base G73. A comparison of the nucleotides at these positions among the suppressors tested in Table 1, however, does not readily explain the specificity of the mischarging mutant enzymes for the charging of only *supF* $tRNA^{Tyr}$. Likewise, other bases in contact with the protein (*14*) are not preferentially present in *supF* $tRNA^{Tyr}$. Both the *supD* $tRNA^{Ser}$ as well as the in vitro–constructed $tRNA^{Ala}$ have a larger number of identical bases with $tRNA^{Gln}$ at these positions but are not recognized by GlnRS7, GlnRS10, or GlnRS15. Measurements of the suppressor efficiency of *supF* have shown it to be intermediate relative to other amber suppressors (*15*). This suggests that the observed mischarging phenotype of *supF* is not due to a superior capability of this tRNA to function in protein synthesis.

The isolation of mutant forms of the GlnRS enzyme with reduced discrimination for tRNA has enabled us to identify specific interactions in the three-dimensional structure of the complex that are important to tRNA selectivity. In addition, the location of the altered amino acids in the mischarging enzymes in proximity to the acceptor arm of the tRNA adds further to the evidence that nucleotides in this part of the $tRNA^{Gln}$ molecule are critical to its identity. The lack of correlation between the presence of proposed identity elements and the acylation of noncognate suppressors by the mischarging mutant enzymes suggests that aspects of the $tRNA^{Gln}$ structure in addition to those already identified are important for specific interaction with GlnRS. For example, sequence-specific distortions in the tRNA structure, perhaps in the acceptor stem, may be necessary for specific binding. In this case, precise tRNA recognition in the glutamine system may be a consequence of the small contribution of a larger number of nucleotides extending over a large area of the contact interface. Sequence-dependent distortability of the acceptor stem of the tRNA has been shown to be important to the specific recognition of *E. coli* $tRNA^{Ala}$ by AlaRS (*16*). In vitro analyses of modified $tRNA^{Gln}$ and GlnRS species are required to address these issues. Such data should provide us with a more complete understanding of the basis for tRNA identity and selectivity in this system.

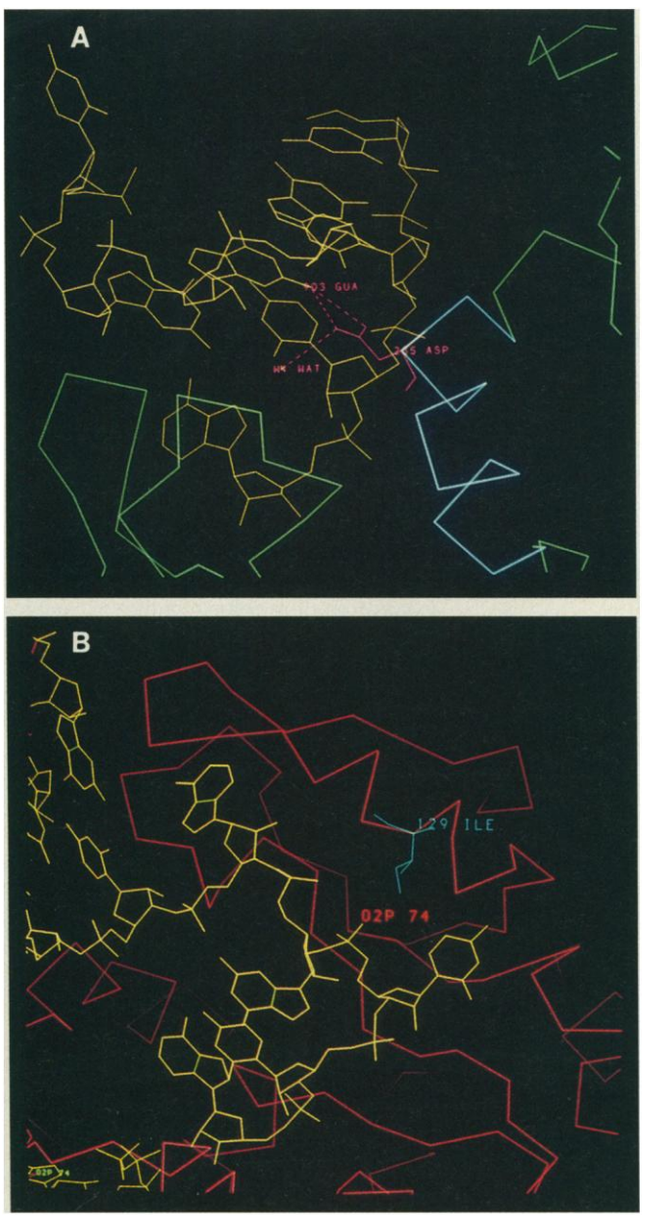

Fig. 2. (**A**) Location of the Asp^{235} side chain in proximity to the minor groove of the acceptor stem. The water molecule shown also interacts with base pair G2 · C71 and the backbone NH group of Ile^{183} to form complementary hydrogen bonding surfaces between the protein and tRNA. In this model one hydrogen atom bonded to the N2 group of G3 is bifurcated, bonding with both carboxylate oxygens of Asp^{235}, and the other is involved in the base pair with C70. (**B**) Location of the Ile^{129} side chain in proximity to the phosphate of nucleotide C74 of the tRNA. The γ-hydroxyl of threonine in the GlnRS15 enzyme is stereochemically equivalent to the ethyl side chain of isoleucine and could possibly interact with the analogous phosphate group in the *supF* $tRNA^{Tyr}$.

REFERENCES AND NOTES

1. M. Ruff *et al.*, *J. Mol. Biol.* **201**, 235 (1988); A. G. Redfield *et al.*, in *Structure and Dynamics of RNA*, P. H. van Knippenberg and C. W. Hilbers, Eds. (Plenum, New York, 1986), pp. 99–112; G. Zaccai *et al.*, *J. Mol. Biol.* **129**, 483 (1979).
2. L. H. Shulman and H. Pelka, *Biochemistry* **24**, 7309 (1985).
3. Y. Shimura *et al.*, *FEBS Lett.* **22**, 144 (1972); M. L. Hooper, R. L. Russell, J. D. Smith, *ibid.*, p. 149.
4. M. Yaniv *et al.*, *J. Mol. Biol.* **86**, 245 (1974).
5. N. E. Prather, E. J. Murgola, B. H. Mims, *ibid.* **172**, 177 (1984); J. Normanly, R. C. Ogden, S. J. Horvath, J. Abelson, *Nature* **321**, 213 (1986); W. H. McClain and K. Foss, *Science* **240**, 793 (1988); Y.-M. Hou and P. R. Schimmel, *Nature* **333**, 140 (1988); W. H. McClain and K. Foss, *Science* **241**, 1804 (1988); W. H. McClain and K. Foss, *J. Mol. Biol.* **202**, 697 (1988); L. H. Schulman and H. Pelka, *Science* **242**, 765 (1988).
6. J. Normanly and J. Abelson, *Annu. Rev. Biochem.* **58**, 1029 (1989).
7. J. R. Sampson *et al.*, *Science* **243**, 1363 (1989).
8. O. Leon and L. H. Schulman, *Biochemistry* **26**, 7113 (1987); H. Bedouelle and G. Winter, *Nature* **320**, 371 (1986).
9. H. Inokuchi, P. Hoben, F. Yamao, D. Söll, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5076 (1984).
10. J. Conley, H. Uemura, F. Yamao, J. Rogers, D. Söll, *Protein Sequence Data Anal.* **1**, 479 (1988).
11. A. Körner *et al.*, *J. Bacteriol.* **120**, 154 (1974).
12. R. Swanson, thesis, Yale University (1988); F. Hutchinson and J. Stein, *Mol. Gen. Genet.* **152**, 29 (1977); T. Maniatis, F. F. Fritsch, J. Sambrook, *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratories, Cold Spring Harbor, NY 1982).
13. H. Uemura *et al.*, *Protein Eng.* **2**, 293 (1988).
14. M. A. Rould *et al.*, *Science* **246**, 1135 (1989).
15. J. H. Miller and A. M. Albertini, *J. Mol. Biol.* **164**, 59 (1983); J. Normanly, J. M. Mason, L. G. Kleina, J. Abelson, J. H. Miller, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6548 (1986).
16. W. H. McClain *et al.*, *Science* **242**, 1681 (1988).
17. J. H. Miller, Ed., *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972).
18. Abbreviations: The mutant enzymes (GlnRS10) are denoted after the allele designation (*glnS10*). The mutational change from wild-type enzyme is given by the amino acid abbreviation before and after the position number. For instance, D235G indicates that in position 235 Asp in the wild-type enzyme is replaced by Gly in the mutant.

19. We thank M. J. Rogers for encouragement and many discussions. Supported by grants from NIH (T.A.S. and D.S.) and NSF (D.S.)

4 August 1989; accepted 20 October 1989