physiological role of GAP is to down-regulate *ras* p21 and that GAP, therefore, acts upstream from *ras* p21 in a signal transduction pathway (*4*). However, the results presented here, which suggest that GAP and the *ras* effector bind to the same site on p21, raise the possibility that GAP is the effector protein itself. If so, it would be expected that the interaction between GAP and p21 occurs on the plasma membrane, since membrane association is known to be required for *ras* function (*12*). We therefore propose that GAP, which has been shown previously to be a cytosolic protein (*4*), translocates to the membrane through association with the GTP-bound form of p21. We speculate that this translocation is necessary to bring GAP in close proximity to its own target (or targets) and that the formation of this complex generates a signal for cellular proliferation. Thus, GAP activation may be analogous to that of protein kinase C, which has been shown to involve translocation to the plasma membrane in response to mitogenic stimulation (*13*). An interesting aspect of the model proposed here is that it provides a means for down-regulating the active complex through the stimulation of GTP hydrolysis by GAP, resulting in the dissociation of GAP from the complex. Since the GTPase activity of oncogenic p21 mutants cannot be stimulated by GAP, constitutive association between GAP and its putative target (or targets) takes place, which in turn results in loss of growth control.

REFERENCES AND NOTES

1. M. Barbacid, *Annu. Rev. Biochem.* **56**, 779 (1987); D. R. Lowy and B. M. Willumsen, *Cancer Surveys* **5**, 275 (1986).
2. T. Toda *et al.*, *Cell* **40**, 27 (1985); D. Broek, *ibid.* **41**, 763 (1985); M. S. Marshall, J. B. Gibbs, E. M. Scolnick, I. S. Sigal, *Mol. Cell. Biol.* **8**, 52 (1988).
3. J. P. McGrath, D. J. Capon, D. V. Goeddel, A. D. Levinson, *Nature (London)* **310**, 644 (1984); R. W. Sweet *et al.*, *ibid.* **311**, 273 (1984); J. B. Gibbs, I. S. Sigal, M. Poe, E. M. Scolnick, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5704 (1984); V. Manne, E. Bekesi, H. F. Kung, *ibid.* **82**, 373 (1985).
4. M. Trahey and F. McCormick, *Science* **238**, 542 (1987).
5. J. Field, D. Broek, T. Kataoka, M. Wigler, *Mol. Cell. Biol.* **7**, 2128 (1987); T. Satoh, S. Nakamura, Y. Kaziro, *ibid.*, p. 4553.
6. H. Adari, D. R. Lowy, B. M. Willumsen, C. J. Der, F. McCormick, unpublished data.
7. M. E. Furth, L. J. David, B. Fleurdelys, E. M. Scolnick, *J. Virol.* **43**, 294 (1982).
8. B. M. Willumsen *et al.*, *Mol. Cell. Biol.* 6, 2646 (1986).
9. M. Trahey and F. McCormick, unpublished data.
10. A. M. de Vos *et al.*, *Science*, in press.
11. I. S. Sigal, J. B. Gibbs, J. S. D'Alanzo, E. M. Scolnick, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4725 (1986).
12. B. M. Willumsen, A. Christensen, N. L. Hubbert, A. G. Papageorge, D. R. Lowy, *Nature (London)* **310**, 583 (1984).
13. A. S. Kraft, W. B. Anderson, H. L. Cooper, J. J. Sando, *J. Biol. Chem.* **257**, 13193 (1982); A. S. Kraft and W. B. Anderson, *Nature (London)* **301**, 621 (1983).
14. R. Dahr *et al.*, *Science* **217**, 934 (1982); C. J. Tabin *et al.*, *Nature (London)* **300**, 143 (1982); E. P. Reddy, R. K. Reynolds, E. Santos, M. Barbacid, *ibid.*, p. 149; C. J. Der, T. Finkel, G. M. Cooper, *Cell* **44**, 167 (1986).
15. The following abbreviations were used for amino acids: A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.
16. C. J. Der and McDonald, in preparation.

17. We would like to thank M. Trahey, B. Rubinfeld, R. Clark and G. Wong for useful discussions and assistance; C. Levenson for oligonucleotide synthesis; M. Innis for DNA sequencing; and S. Nilson and D. Jackson for help in preparing the manuscript.

8 February 1988; accepted 18 March 1988

Evidence from Cassette Mutagenesis for a Structure-Function Motif in a Protein of Unknown Structure

NEIL D. CLARKE,* DONALD C. LIEN, PAUL SCHIMMEL

The three-dimensional structure of most enzymes is unknown; however, many enzymes may have structural motifs similar to those in the known structures of functionally related enzymes. Evidence is presented that an enzyme of unknown structure [Ile-transfer RNA (tRNA) synthetase] may share a functionally important structural motif with an enzyme of related function (Tyr-tRNA synthetase). This approach involves (i) identifying segments of Ile-tRNA synthetase that have been unusually conserved during evolution, (ii) predicting the function of one such segment by assuming a structural relation between Ile-tRNA synthetase and Tyr-tRNA synthetase, and (iii) testing the predicted function by mutagenesis and subsequent biochemical analysis. Random mutations were introduced by cassette mutagenesis into a ten-amino-acid segment of Ile-tRNA synthetase that was predicted to be involved in the formation of the binding site for isoleucine. Few amino acid substitutions appear to be tolerated in this region. However, one substitution (independently isolated twice) increased the Michaelis constant K_m for isoleucine in the adenylate synthesis reaction by greater than 6000-fold, but had little effect on the K_m for adenosine triphosphate, the apparent K_m for tRNA, or the rate constant k_{cat}.

ALTHOUGH FUNCTIONALLY RELATED, the members of the aminoacyl–transfer RNA (tRNA) synthetase class of enzymes are heterogeneous in quaternary structure and subunit size and in general show little sequence similarity. However, a subset of these enzymes (including Ile-tRNA synthetase) share a similar 11-amino-acid segment known as the signature sequence (*1*). In the two aminoacyl-tRNA synthetase structures that have been partially solved (Tyr- and Met-tRNA synthetase), this sequence occupies the same position in a mononucleotide binding fold (*2*) (Fig. 1). In the case of *Bacillus stearothermophilus* Tyr-tRNA synthetase, residues in this sequence are known to be involved in the binding of adenosine triphosphate (ATP) (*3*) and in the catalytic formation of the aminoacyl-adenylate intermediate (*4*). We present evi-

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

*Present address: Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205.

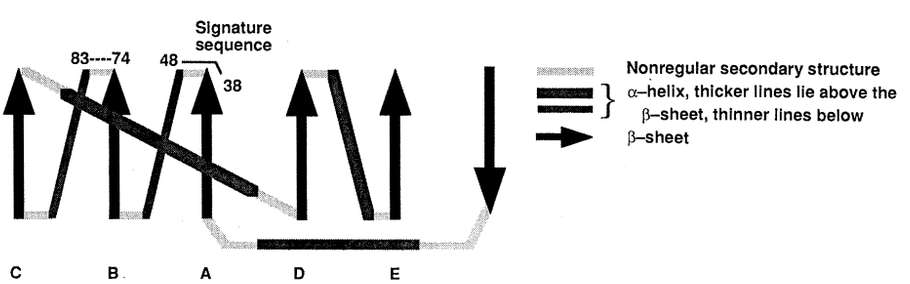


Fig. 1. Schematic of the dinucleotide binding fold type structure of *B. stearothermophilus* Tyr-tRNA synthetase. This drawing (not to scale) represents one selected part of the solved structure of Tyr-tRNA synthetase based on published drawings and descriptions (*2, 8, 9*). The letters below the schematic indicate the β-strands according to the convention of Blow *et al.* (*2*). The approximate location of the signature sequence found in a subset of aminoacyl-tRNA synthetase is shown, as is the location of a segment that starts 26 amino acids beyond the end of the signature sequence. This segment may be analogous to the segment of Ile-tRNA synthetase that we have subjected to mutagenesis in this work.

dence that Ile-tRNA synthetase may share with Tyr-tRNA synthetase a functionally important beta-alpha-beta motif that includes and extends beyond this sequence.

The sequences of the genes for Ile-tRNA synthetase from *Escherichia coli* (*5*) and from the yeast *Saccharomyces cerevisiae* (*6*) show that the enzymes have diverged significantly since the prokaryote-eukaryote evolutionary split. Even if small gaps are allowed in the alignment of the sequences, there is only 27% amino acid identity between the two enzymes (*6*). However, the divergence is not uniform throughout the sequence, but rather is punctuated by blocks of strong sequence identity (Fig. 2). We surmise that these segments have been conserved during evolution because they are functionally important.

One of the most striking blocks of sequence identity lies 26 amino acids beyond the end of the signature sequence. This block of ten amino acids includes eight identities and two conservative substitutions. Although the Ile-tRNA synthetase polypeptide (a monomer of 939 amino acids) is more than twice the length of that of the tyrosine enzyme (an α_2 dimer with subunits of 419 amino acids), there is evidence for a nucleotide fold in the amino terminal part of Ile-tRNA synthetase that may be structurally related to that of Tyr-tRNA synthetase (*7*). Because of this possible relation, the Tyr-tRNA synthetase-tyrosyl-adenylate co-crystal structure (*8, 9*) was examined in order to identify a possible function for the segment of Tyr-tRNA synthetase that lies 26 amino acids beyond the signature sequence. The segment is in relativelty close proximity to the bound tyrosyl-adenylate, and near the center of the segment is Asp^{78}, which is believed to hydrogen bond to the amino group of tyrosine (*10*).

In view of this, the aforementioned region that is conserved between two Ile-specific synthetases may contribute to the formation of the binding site for Ile. To investigate this possibility, we introduced random mutations into the region by cassette mutagenesis (Fig. 2) (*11*). The restriction sites that were used for cloning the mutagenic cassettes were introduced into the *E. coli ileS* gene by oligonucleotide-directed mutagenesis (*12*). The cassettes were cloned into *ileS* in such a way that greater than 99% of the transformants are attributable to the insertion of a cassette; 93 independent transformations were performed. Plasmid DNA was prepared from a single transformant from each of the 93 transformations and used to transform a *recA* derivative of the strain MI1. [MI1 (*13*) has a mutation in the chromosomal allele of *ileS* that renders the Ile-tRNA synthetase activity undetectable under standard assay conditions. When transformed with the *ileS*-bearing plasmid pMT521, the plasmid-encoded aminoacylation activity under standard conditions is more than 100,000 times greater than that of the chromosomal-encoded MI1 activity, and the Ile-tRNA synthetase activity in crude extracts can be attributed to the plasmid allele.] The transformants were assayed under four sets of substrate concentrations to screen for changes in the Michaelis constant K_m values of each of the three substrates (Ile, ATP, and tRNA).

Of the 93 transformants examined, 62 were phenotypically wild type with respect to the K_m for all three substrates, whereas 28 produced neither detectable activity nor detectable levels of protein. One mutant (IN102), which encodes an Ile to Asn substitution at position 9 of the identity block, had only minor effects (less than tenfold) on the kinetic parameters of Ile-tRNA synthetase and was not studied in detail. The remaining two mutants were found to have severely altered activity. These two mutants each encode a Gly^{94} to Arg substitution (codon changes of GGC to CGA and CGG,

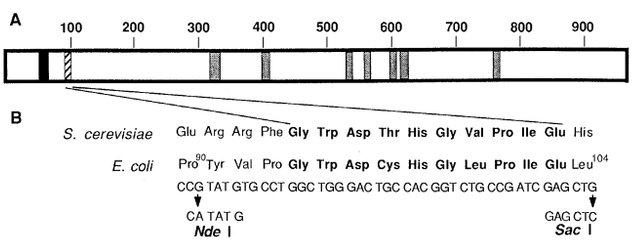
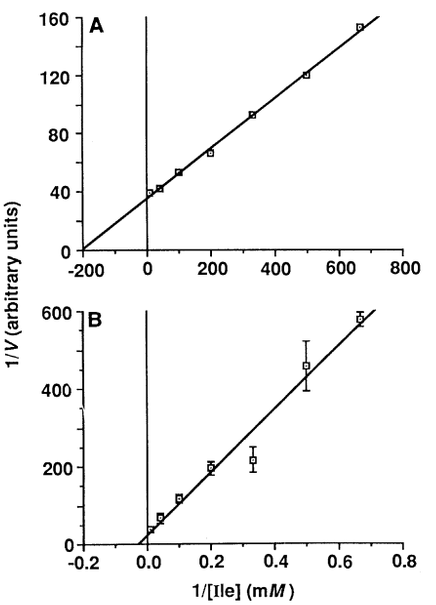


Fig. 2. Blocks of strong sequence identity and the strategy for mutagenesis. (**A**) A representation of *E. coli* Ile-tRNA synthetase; numbering is from the NH_2-terminus. Regions of strong sequence identity (a stretch of at least ten amino acids with 60% or more colinear amino acid identity) with the *S. cerevisiae* enzyme are represented as shaded, crosshatched, or solid blocks. The total amino acid identity in the shaded blocks is 71%. The total amino acid identity in the rest of the sequence is 21%, provided that small gaps in the sequence are allowed. The black block represents the signature sequence. The cross-hatched region begins 26 amino acids beyond the end of the signature sequence. (**B**) The *E. coli* sequence targeted by the mutagenic cassettes (*Nde* I to *Sac* I) is shown (*11*). The bold type in the *E. coli* and yeast sequences indicates the conserved residues.



C

	$K_{m(Ile)}$ (μM)	$K_{m(ATP)}$ (mM)	$K_{m(tRNA)}$ (apparent) (μM)	Relative K_{cat}
Wild type	5	0.4	0.7	1.00
GR94	33,000	1.5	0.2	0.85

Fig. 3. Kinetic parameters of the Gly^{94} to Arg mutant. Lineweaver-Burk plots were used to determine the $K_{m(Ile)}$ of (**A**) the wild-type and (**B**) the mutant Ile-tRNA synthetase. Extracts were prepared essentially as described in (*18*), and dialyzed against 50 mM tris-HCl (pH 7.5), 100 mM KCl, 1 mM EDTA, and 10% (v/v) glycerol. Activity was measured with the pyrophosphate exchange assay as described (*19*) except that the ATP concentration was 1 mM. This is the optimal concentration for both the wild-type and the GR94 mutant enzymes (*14*) and is near to the $K_{m(ATP)}$ for each enzyme. Higher concentrations of ATP are inhibitory (*20*). The Ile concentrations used for the mutant are 1000-fold higher than for the wild type. The error bars for the mutant data represent the standard deviation of measurements made on the two independent isolates. At high substrate concentrations, greater than 98% of the activity is attributable to the GR94 mutant enzyme rather than to the MI1 chromosomal mutant. (**C**) Kinetic parameters for wild-type and mutant Ile-tRNA synthetase. The estimated accuracy of these parameters is ±20%. The K_m values for ATP and Ile were measured with the pyrophosphate exchange assay. The Ile concentration for each $K_{m(ATP)}$ determination was at least threefold above the $K_{m(Ile)}$ for that enzyme. Relative k_{cat} values were determined by normalizing the pyrophosphate exchange values for V_{max} to the amount of Ile-tRNA synthetase in crude extracts as determined by densitometry of Western blots with antibodies to Ile-tRNA synthetase. The elevation in $K_{m(Ile)}$ in the mutant enzyme makes the measurement of these values by the aminoacylation assay (*14*) unfeasible (the higher concentrations of unlabeled Ile that are required necessarily reduce the specific activity of [^{3}H]Ile to too large an extent). Thus, the apparent K_m for tRNA was arbitrarily defined and determined at 1 mM ATP and 25 μM Ile with purified $tRNA^{Ile}$ (Subriden, Inc., Rolling Bay, Washington).

		Tyr	Val	Pro	Gly	Trp	Asp	Cys	His	Gly	Leu	Pro	Ile	Glu
Wild type		--T	GTG	CCT	GGC	TGG	GAC	TGC	CAC	GGT	CTG	CCG	ATC	G--
Phenotypic mutants														
GR94 (a)	(1)	-	-	-	Arg CGA	-	-	-	-	-	-	-	-	-
GR94 (b)	(1)	-	-	-	Arg CGG	-	-	-	-	-	-	-	-	-
IN102	(1)	-	-	-	-	-	-	-	-	-	-	-	Asn AAC	-
Phenotypic wild types														
	(2)	-	Val GTA	-	-	-	-	-	-	-	-	-	-	-
	(1)	-	-	-	Gly GGG	-	-	-	-	-	-	-	-	-
	(2)	-	-	-	-	-	Asp GAT	-	-	Gly GGA	-	-	-	-
	(3)	-	-	-	-	-	-	Ser TCC	-	-	-	Pro CCC	-	-
	(3)	-	-	-	-	-	-	-	-	Gly GGG	-	-	-	-
	(3)	-	-	-	-	-	-	-	-	Gly GGG	-	Pro CCT	-	-
	(1)	-	-	-	-	-	-	-	-	-	Leu CTA	-	-	-
	(1)	-	-	-	-	-	-	-	-	-	-	-	Ile ATT	-

Fig. 4. Sequences of phenotypically wild-type and kinetically altered cassette mutants. The cassettes include 35 base pairs of double-stranded DNA. The wild-type coding strand for these 35 base pairs is shown along with the amino acids that potentially can be mutated by nucleotide substitutions in this region. Kinetically altered cassette mutants were cloned into an M13 vector and sequenced by the Sanger method according to Biggin *et al.* (*20*). In the case of the phenotypic wild-type clones, the 62 plasmids were pooled, and a fragment containing the cassette region was cloned into an M13 vector. Seventy-two members of this pool were then sequenced, of which 16 showed differences with the wild-type sequence. Maintenance of the wild-type codon is indicated by a dash. The sequences of mutant codons appear beneath the corresponding wild-type codon. Note that, with one exception, none of the other nucleotide substitutions among the phenotypic wild types results in a change in amino acid sequence. Amino acid changes are indicated in bold type. The numbers in parentheses to the left of the mutant sequences represent the number of times the sequence was obtained. For the phenotypic wild types, these do not necessarily represent independent isolates.

respectively) at the first position of the identity block. The mutants have a K_m for Ile in the adenylate synthesis reaction of 33 m*M*, which is more than 6000 times greater than the $K_{m(Ile)}$ of the wild-type enzyme (Fig. 3, A and B). The K_m value for ATP is increased only fourfold and the rate constant k_{cat} is essentially unchanged (Fig. 3C). To our knowledge, the magnitude and specificity of the change in kinetic parameters caused by this mutation is unusual for a multisubstrate enzyme. That only one kinetic parameter is significantly altered suggests that the structural change caused by the mutation is highly localized and does not induce a global distortion of the enzyme. Although other possibilities exist, these data suggest that Gly^{94} contributes to formation of the Ile binding site.

We examined further the 62 transformants that are phenotypically wild type. Plasmid DNA from these transformants was pooled and 72 isolates from the pool were sequenced. At least 11 independent nucleotide substitutions were found among these phenotypic wild-type transformants. With one exception, all of these substitutions preserve the wild-type amino acid sequence (Fig. 4). The exception is a Cys^{97} to Ser substitution at a position at which the yeast enzyme has a Thr.

A statistical analysis based on the distribution of nucleotide changes among the phenotypic wild-type mutants suggests that more than 50 amino acid substitutions are expected to be represented by the transformants that we examined (*14*). Because only four amino acid substitutions were found among either the active or the partially active transformants, many of the remaining amino acid substitutions that occurred apparently gave rise to inactive enzyme. Most of the plasmids derived from strains that fail to produce plasmid-encoded activity are smaller than the wild-type plasmid. Although we have not examined these clones in detail, it seems likely that deletions have occurred among the transformants with amino acid substitutions in this region. This suggests that some substitutions are deleterious to the cell. Among other possibilities, this deleterious effect could be due to misacylation of $tRNA^{Ile}$ caused by a relaxed specificity of the mutant enzyme for its amino acid substrate.

The pronounced and specific effect of the Gly^{94} to Arg mutation suggests that a structural element for binding Ile to the Ile-tRNA synthetase occurs in approximately the same location, relative to the signature sequence, as an element that contributes to binding Tyr to Tyr-tRNA synthetase. This is further evidence for a structural relatedness between aminoacyl-tRNA synthetases that are superficially quite different [compare with (*7*)].

The analysis reported here may be extended to other aminoacyl-tRNA synthetases. This approach is likely to be most fruitful for those enzymes for which sequences are known from a prokaryotic and a eukaryotic organism, as this will maximize the significance of conserved sequence elements. In addition, those aminoacyl-tRNA synthetases that contain the conserved signature sequence may be most amenable to this analysis because they may contain a mononucleotide fold similar to those found in the Tyr and Met enzymes. The Val- and Gln-tRNA synthetases meet these criteria (*15–17*). The introduction of mutations at appropriate residues in these enzymes may provide further evidence for structural similarities within this seemingly heterogeneous family of enzymes.

REFERENCES AND NOTES

1. P. Schimmel, *Annu. Rev. Biochem.* **56**, 125 (1987).
2. D. M. Blow, T. N. Bhat, A. Metcalfe, J.-L. Risler, C. Zelwer, *J. Mol. Biol.* **171**, 571 (1983).
3. A. R. Fersht *et al.*, *Nature (London)* **314**, 235 (1985).
4. R. J. Leatherbarrow, A. R. Fersht, G. Winter, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7840 (1985).
5. T. Webster, H. Tsai, M. Kula, G. A. Mackie, P. Schimmel, *Science* **226**, 1315 (1984).
6. U. Englisch *et al.*, *Biol. Chem. Hoppe-Seyler* **368**, 971 (1987).
7. R. M. Starzyk, T. A. Webster, P. Schimmel, *Science* **237**, 1614 (1987).
8. T. N. Bhat, D. M. Blow, P. Brick, J. Nyborg, *J. Mol. Biol.* **158**, 699 (1982).
9. D. M. Blow and P. Brick, in *Biological Macromolecules and Assemblies*, vol. 2 of *Nucleic Acids and Interactive Proteins*, F. Jurnak and A. McPherson, Eds. (Wiley, New York, 1985), p. 442.
10. A. R. Fersht *et al.*, *Angew. Chem.* **94**, 455 (1984).
11. Two single nucleotide mutations were introduced into the *E. coli ileS* gene by oligonucleotide-directed mutagenesis (*12*). Oligonucleotides were synthesized on a Systec 1450A automated DNA synthesizer and gel-purified. Mutagenic cassettes were synthesized with phosphoramidite solutions that had been spiked with each of the other three phosphoramidites. The oligonucleotides were synthesized twice by using two different concentrations of spiked phosphoramidites, such that the mean number of nucleotide substitutions per strand is expected to be 0.7 or 2.1. Both sets of syntheses were used in this experiment. Cassettes (which lacked 5′ phosphates) were ligated to an *Nde* I– and *Sac* I–digested *ileS* backbone. The backbone was derived from a plasmid in which a spacer fragment of several kilobase pairs had been inserted between the *Nde* I and *Sac* I sites. This procedure assures that purified backbone fragment has been digested at both sites, thereby maximizing the number of transformants that arise from introduction of a mutagenic cassette. The partial distribution of expected mutations among cassette-dependent transformants is: 0: 25%; 1: 27%; 2: 19%; 3: 11%.
12. M. J. Zoller and M. Smith, *Methods Enzymol.* **100**, 468 (1983).
13. M. Iaccarino and P. Berg, *J. Bact.* **105**, 527 (1970).
14. N. D. Clarke, thesis, Massachusetts Institute of Technology, Cambridge (1987).
15. T. J. Borgford, N. J. Brand, T. E. Gray, A. R. Fersht, *Biochemistry* **26**, 2480 (1987).
16. X. Jordana, *J. Biol. Chem.* **262**, 7189 (1987).
17. S. W. Ludmerer and P. Schimmel, *ibid.*, p. 10801.
18. E. Seeberg, J. Nissen-Meyer, P. Strike, *Nature (London)* **263**, 524 (1976).
19. S. D. Putney, R. T. Sauer, P. R. Schimmel, *J. Biol. Chem.* **256**, 198 (1981).
20. F. X. Cole and P. R. Schimmel, *Biochemistry* **9**, 3143 (1970).
21. M. D. Biggin, T. J. Gibson, G. F. Hong, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3963 (1983).

22. We thank D. Blow for instructive comments on structural features of the tyrosine tRNA synthetase, and D. Blow and A. Fersht for comments on the manuscript. Supported by NIH grant GM15539. N.C. was an NIH predoctoral trainee.

11 December 1987; accepted 4 March 1988