Pollard, T. D., & Weeds, A. G. (1984) *FEBS Lett. 170*, 94–98.
Schulz, G. E., & Schirmer, R. H. (1979) *Principles of Protein Structure*, Springer-Verlag, New York.
Tanswell, P., Thornton, J. M., Korda, A. V., & Williams, R. J. P. (1975) *Eur. J. Biochem. 57*, 135–145.
Vandekerckhove, J., Deboben, A., Nassal, M., & Wieland, T. (1985) *EMBO J. 4*, 2815–2818.
Waechter, F., & Engel, J. (1977) *Eur. J. Biochem. 74*, 227–232.
Walker, J. E., Saraste, M., Runswick, M. J., & Gay, N. J. (1982) *EMBO J. 1*, 945–951.
Wanger, M., Keiser, T., Neuhaus, J.-M., & Wegner, A. (1985) *Can. J. Biochem. Cell Biol. 63*, 414–421.
Wegner, A. (1976) *J. Mol. Biol. 108*, 139–150.

Structure–Activity Relationships in Engineered Proteins: Analysis of Use of Binding Energy by Linear Free Energy Relationships†

Alan R. Fersht,* Robin J. Leatherbarrow, and Tim N. C. Wells
Department of Chemistry, Imperial College of Science and Technology, London SW7 2AY, U.K.
Received January 21, 1987; Revised Manuscript Received April 27, 1987

ABSTRACT: The activity of mutant enzymes can be analyzed quantitatively by structure–activity relationships in a manner analogous to Brønsted or Hammett plots for simple organic reactions. The slopes of such plots, the β values, indicate for the enzymatic reactions the fraction of the overall binding energy used in stabilizing particular complexes. In particular, information can be derived about the interactions between the enzyme and the transition state. The activities of many mutant tyrosyl-tRNA synthetases fit well simple linear free energy relationships. The formation of enzyme-bound tyrosyl adenylate (E·Tyr-AMP) from enzyme-bound tyrosine and ATP (E·Tyr·ATP) results in an increase in binding energy between the enzyme and the side chain of tyrosine and the ribose ring of ATP. Linear free energy plots of enzymes mutated in these positions give the fraction of the binding energy change that occurs on formation of the transition state for the chemical reaction and the various complexes. It is shown that groups that specifically stabilize the transition state of the reaction are characterized by β values $\gg 1$. This is found for residues that bind the γ-phosphate of ATP (Thr-40 and His-45) and have previously been postulated to be involved in transition-state stabilization. The importance of linear free energy plots is that (i) they bring order to the analysis of structure–activity relationships since they allow a large amount of data to be simplified and systematized, (ii) they allow transition-state structure to be inferred from ground-state structure, and (iii), perhaps the most important point at this stage of our knowledge in understanding enzyme structure, they collectively show trends, and exceptions are readily apparent. The observation that the activities of a large number of mutants of the tyrosyl-tRNA synthetase conform to linear free energy equations is the best evidence yet that mutation of the enzyme is probing general properties and trends in the relationship between structure and activity.

The quintessential feature of enzyme catalysis is the use of binding energy to give rate enhancement. It may be shown quite simply by theory (see Appendix) that the increase in rate between an enzyme-catalyzed reaction and its uncatalyzed counterpart in aqueous solution proceeding by the same transition states and intermediates is directly related to the dissociation constants of the substrates, transition states, intermediates, and products from the enzyme. The free energy of transfer of the reagents from water to enzyme is responsible for lowering the free energy of activation and for altering equilibrium constants between enzyme-bound reagents. Thus, to understand enzyme catalysis, one must determine the interaction energies between the enzyme and reagents throughout the complete reaction profile and the energetic changes within the enzyme itself.

Site-directed mutagenesis is eminently suited to studying such structure–activity relationships since the structure of an enzyme may be systematically varied. Unfortunately, protein engineering is an invasive technique since mutagenesis perturbs the very system it studies. The structures of mutant enzymes may be studied by X-ray crystallography or spectroscopy, which may detect changes in structure. But, these techniques can be applied only to stable complexes and cannot be used for examining the important structure for determination of rate, the transition state. Further, X-ray crystallography may not have sufficient resolution for detecting miniscule changes in the structure of an enzyme or its solvation shell that cause significant changes in energetics of catalysis. The only means at present of probing transition-state structure experimentally is the use of kinetics. Accordingly, we are attempting to devise kinetic methods that may be applied in conjunction with X-ray crystallographic data to analyze structure–activity relationships of enzymes and, in particular, how binding energy is utilized. In this paper, we outline the theory of how the use of binding energy may be analyzed by protein engineering, propose criteria for the design of experiments that are open to interpretation by such theory, and apply the theory to reactions of a particular set of mutants of the tyrosyl-tRNA synthetase from *Bacillus stearothermophilus*.

Application of Linear Free Energy Relationships. A traditional way of probing transition-state structure in the reactions of simple organic molecules is the construction of linear

† This work was funded by the Medical Research Council of the U.K.

0006-2960/87/0426-6030$01.50/0

free energy plots such as the Brønsted and Hammett plots. In these, plots of rate constant against equilibrium constant are made for a reaction in which the structures of the reagents are systematically varied. The plots can give information on how much the transition state resembles starting materials or products. The plots assume that there is the relationship between the equilibrium constant K of a reaction and the rate constant k for the formation of products of the general form

$$k = AK^{\beta} \quad (1)$$

where A and β are constants. The plot of log k against log K when eq 1 holds is a straight line of slope β (eq 2). Since

$$\log k = \text{constant} + \beta \log K \quad (2)$$

log k is proportional to the free energy of activation of the reaction, ΔG^*, and log K is proportional to the free energy change for the equilibrium, ΔG_e, eq 2 is equivalent to

$$\Delta G^* = \text{constant} + \beta \Delta G_e \quad (3)$$

A qualitative interpretation of eq 1–3 is that, all things being equal, a value of β that is close to zero means that the transition state resembles starting materials and a value of β close to 1 means that the transition state resembles products. We have recently shown that linear free energy relationships can be applied to the changes in *binding energy* that occur when enzyme structure is varied in a site-directed mutagenesis experiment (Fersht et al., 1986a).

Linear free energy relationships must not be applied blindly to the kinetics of mutant enzymes since the area is rife with potential artifacts. We thus first classify the effects of a single amino acid substitution to decide which experiments are worthy of extensive analysis.

Classification of Effects of Mutation on Structure. The study of enzyme structure and activity by systematic mutagenesis of the enzyme involves at the very simplest the substitution of just one amino acid for another. More complex experiments may involve multiple substitutions, deletions, or even changing whole domains. But even the substitution of the side chain of an amino acid by another can lead to a variety of effects, some of which may be too complicated for simple analysis. The substitutions with which we are primarily concerned are those that effect the interactions between enzyme and ligand and those that perturb the structure of the enzyme. These are classified as follows.

(*1*) *Nondisruptive Deletion.* A side chain is replaced by another that lacks a group involved in a specific interaction. This is done in such a manner that there are no effects on the structure of the enzyme or any of its complexes with the substrates other than the loss of this interaction. This could be, for example, the removal of an interaction between the enzyme and substrate that does not affect any other interaction between the enzyme and substrate or any interaction within the enzyme itself. These involve almost always the replacement of one side chain by another that is smaller and lacks the functional group under discussion. A typical example is the mutation of Tyr → Phe-34 in the tyrosyl-tRNA synthetase. The mutation just removes a group which, in the free enzyme, makes interactions with solvent only. Nondisruptive deletion may, however, perturb solvation.

(*2*) *Disruptive Deletion.* Replacement of a side chain may lead to a perturbation of structure elsewhere in the protein. One example that we have found in our studies is the removal of the carboxylate of an aspartate residue, which forms part of a hydrogen-bonded network (Lowe et al., 1987).

(*3*) *Conservative Substitution.* A side chain is replaced by one that can substitute in the same interactions. For example, on mutation of His → Asn-48 in the tyrosyl-tRNA synthetase, the amide nitrogen of $-CONH_2$ can on some, but not all, occasions substitute for the δ-NH of histidine in forming a hydrogen bond with the substrate (Lowe et al., 1985).

(*4*) *Semiconservative Substitution.* Some of the function is conserved on replacement. For example, replacement of a histidine by an asparagine when the imidazole ring of the histidine is charged and in a salt bridge. The $-CONH_2$ of the Asn may still form a hydrogen bond but lacks the formal positive charge.

(*5*) *Disruptive Substitution.* Substitution of a large side chain for a small one in a buried close-packed region of a protein may clearly cause severe structural changes because of steric repulsions. There are also substitutions that appear innocuous at first sight but cause severe problems because they add alternative functions. For example, the mutations Asp → Asn and Asn → Asp appear to be extremely conservative since $-CO_2^-$ and $-CONH_2$ are isosteric and both polar. But, whereas both oxygen atoms of $-CO_2^-$ are hydrogen-bond acceptors, as is the oxygen of $-CONH_2$, the $-NH_2$ group is a hydrogen-bond donor.

(*6*) *Nondisruptive Addition.* Bulky groups may be added to the surface of proteins without necessarily causing perturbation of structure. Similarly, the binding cavity of an enzyme may be partly filled so that substrates smaller than the specific substrate can still bind. For example: indole binds to the active site of chymotrypsin and enhances the reaction rate with *p*-nitrophenyl acetate (Foster, 1961; Landis & Berliner, 1980); mutation of residue Gly-166 in subtilisin, which is part of the S_1 subsite for the side chain of a substrate, increases the reactivity toward substrates with small amino acids in the P_1 position (Estell et al., 1986).

The disruptive mutations are likely to be too complicated for simple analysis. The most promising class to analyze is that of nondisruptive removal since, in the ideal case, there is just the removal of a simple interaction with no complications from disruption of structure or the substitution of alternative interactions. Conservative and semiconservative substitutions may also be amenable to analysis. Mutants with nondisruptive additions may perhaps be analyzed where alternative substrates are available. This study concerns primarily nondisruptive deletants, but the principles may be extended to other cases.

Experimental Procedures

The mutant TyrTS(His→Asn-48)[1] was prepared according to Lowe et al. (1985). Kinetic measurements were all performed at 25 °C at pH 7.78 in a buffer containing 144 mM Tris-HCl, 0.14 mM 2-mercaptoethanol, and 10 mM $MgCl_2$ (free). The pre-steady-state kinetic measurements of the rate constants for the activation of tyrosine were determined as before (Wells & Fersht, 1986).

Results

The dissociation constant of tyrosine from the E·Tyr complex with TyrTS(His→Asn-48) (K_t) was found to be 23 μM; the dissociation constant of ATP from the E·Tyr·ATP complex (K'_a), 3.8 mM; the rate constant for the formation of E·Tyr-AMP·PP_i from the E·Tyr·ATP complex (k_3), 27 s^{-1}; the rate constant for the formation of E·Tyr·ATP from E·Tyr-AMP·PP_i (k_{-3}), 23.4 s^{-1}; and the dissociation constant of py-

[1] Abbreviations: TyrTS, tyrosyl-tRNA synthetase; T, Tyr; A, ATP; T-A, Tyr-AMP (tyrosyl adenylate); [T-A]*, transition state for formation of tyrosyl adenylate; ET, enzyme-bound tyrosine; E[T-A]*, enzyme-bound transition state for formation of tyrosyl adenylate; Tris, tris(hydroxymethyl)aminomethane.

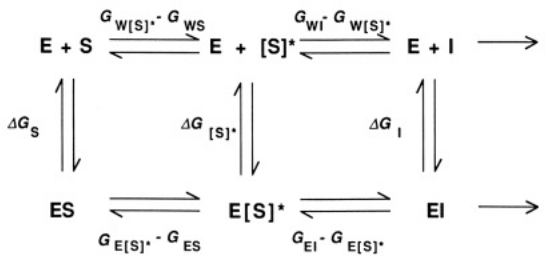

FIGURE 1: Thermodynamic cycles for comparing Gibbs free energy changes for uncatalyzed reaction in solution with its enzyme-catalyzed counterpart. G denotes free energy of reagent (subscript S = substrate, I = intermediate, and [S]* = transition state) in water (subscript W) or bound to enzyme (subscript E). ΔG denotes free energy of transfer of reagent from enzyme to water.

rophosphate from the last complex (K_{pp}), 0.57 mM.

THEORY

Consider a reaction that involves the reversible binding of a substrate S to an enzyme to form a Michaelis complex, ES, followed by the formation of an enzyme-bound transition state, E[S]*, and then an enzyme-bound intermediate, EI (eq 4).

$$E + S \overset{K_S}{\rightleftharpoons} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EI \rightarrow \quad (4)$$

The reaction of the enzyme with its substrates, their transition states, and intermediates may be described by the formal series of thermodynamic cycles in Figure 1. These link the equilibria between the enzyme-bound species and their counterparts in aqueous solution by dissociation constants. From standard transition-state theory and equilibrium thermodynamics

$$RT \ln (k_2/k_{-2}) = G_{ES} - G_{EI} \quad (5)$$

$$RT \ln k_2 = kT/h - (G_{E[S]^*} - G_{ES}) \quad (6)$$

$$RT \ln (k_2/K_S) = kT/h - (G_{E[S]^*} - G_{WE} - G_{WS}) \quad (7)$$

where G is the Gibbs free energy and the subscripts WS and WE refer to the substrate and enzyme in water. It is shown in the Appendix that these equations may be cast entirely in terms of binding energies for analyzing the changes in rate and equilibrium constants that occur when a series of mutated enzymes reacts with the same substrate. That is

$$RT \ln (k_2/k_{-2}) = \Delta G_S - \Delta G_I + \text{constant} \quad (8)$$

$$RT \ln k_2 = \text{constant} - \Delta G_{[S]^*} + \Delta G_S \quad (9)$$

$$RT \ln (k_2/K_S) = \text{constant} - \Delta G_{[S]^*} \quad (10)$$

where ΔG_S, $\Delta G_{[S]^*}$, and ΔG_I are the free energies of binding of the substrate, transition state, and intermediate to the enzymes. The standard equations (e.g., 5, 6, and 7) can be written as usual in terms of the free energies of the relevant states, but it can be considered that the only relevant changes are those in binding energy. Hence, any linear free energy relationships will apply to changes in binding energies. The only assumption is that the structure of the transition state of the substrate does not change on mutation—see Appendix.

Application to Reactions of Tyrosyl-tRNA Synthetase. Mutants of the tyrosyl-tRNA synthetase have been generated that have been altered in side chains that make hydrogen bonds with the substrates (Figure 2) (Winter et al., 1982; Fersht et al., 1984, 1985, 1986b). The first step of the enzymatic reaction is the formation of tyrosyl adenylate (eq 11). Model

$$E + Tyr + ATP \rightleftharpoons E\text{-}Tyr\text{-}AMP + PP_i \quad (11)$$

building followed by mutagenesis has revealed that there is a major catalytic factor caused by two side chains that contribute binding energy to the γ-phosphate of ATP only when it is in the transition state (Leatherbarrow et al., 1985). The binding energies of groups far removed from the seat of reaction are also used to increase the chemical rate constants for the reaction (Wells & Fersht, 1985). Complete free energy profiles for activation of tyrosine by the mutant enzymes have been measured (Fersht et al., 1986a; Wells & Fersht, 1986; Ho & Fersht, 1986). The comparison of these with the profile for wild-type enzyme gives the *apparent* binding energy (Wells & Fersht, 1986) of the relevant side chain with the substrates, transition states, and intermediates throughout the reaction. These data have been analyzed in each individual case to show how the changes in binding energy affect the rate and equilibrium constants for the interconversion of free and enzyme-bound species. Several of the side chains that bind the side chain of tyrosine and the ribose ring of ATP appear to stabilize the enzyme-bound tyrosyl adenylate complex more than the transition state for its formation and more still than the initial enzyme–substrate ternary complex.

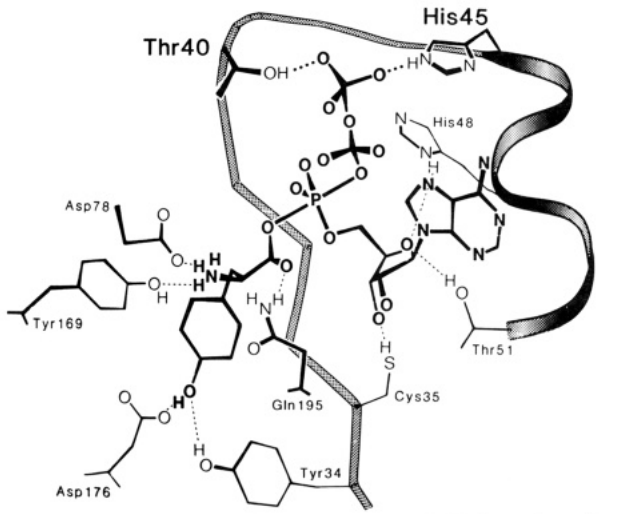

FIGURE 2: Transition state of tyrosine + ATP built into the active site of the tyrosyl-tRNA synthetase [from Leatherbarrow et al. (1985)].

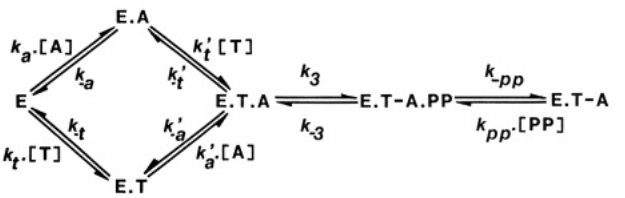

FIGURE 3: Reaction scheme for the activation of tyrosine that defines the relevant rate and dissociation constants.

We have measured the relevant equilibrium and rate constants for Figure 3, which describes the formation of tyrosyl adenylate by the tyrosyl-tRNA synthetase (Wells & Fersht, 1986; Ho & Fersht, 1986; Leatherbarrow et al., 1985; T. N. C. Wells, R. J. Leatherbarrow, and A. R. Fersht, unpublished results). The free energy terms are defined thus: G_{WE} = free energy of unligated enzyme; G_{ET} = free energy of enzyme–tyrosine complex; G_{ETA} = free energy of enzyme–tyrosine–ATP complex; $G_{E[T\text{-}A]^*}$ = free energy of transition state for the interconversion of the ternary complexes of enzyme-bound tyrosine and ATP and of enzyme-bound tyrosyl adenylate and pyrophosphate; $G_{ET\text{-}A\cdot PP}$ = free energy of the ternary complex with pyrophosphate and tyrosyl adenylate; $G_{ET\text{-}A}$ = free energy of enzyme-bound tyrosyl adenylate complex.

A series of linear free energy relationships were observed that relate various rate constants with corresponding equilibria. The equations were set up by first writing down the initial,

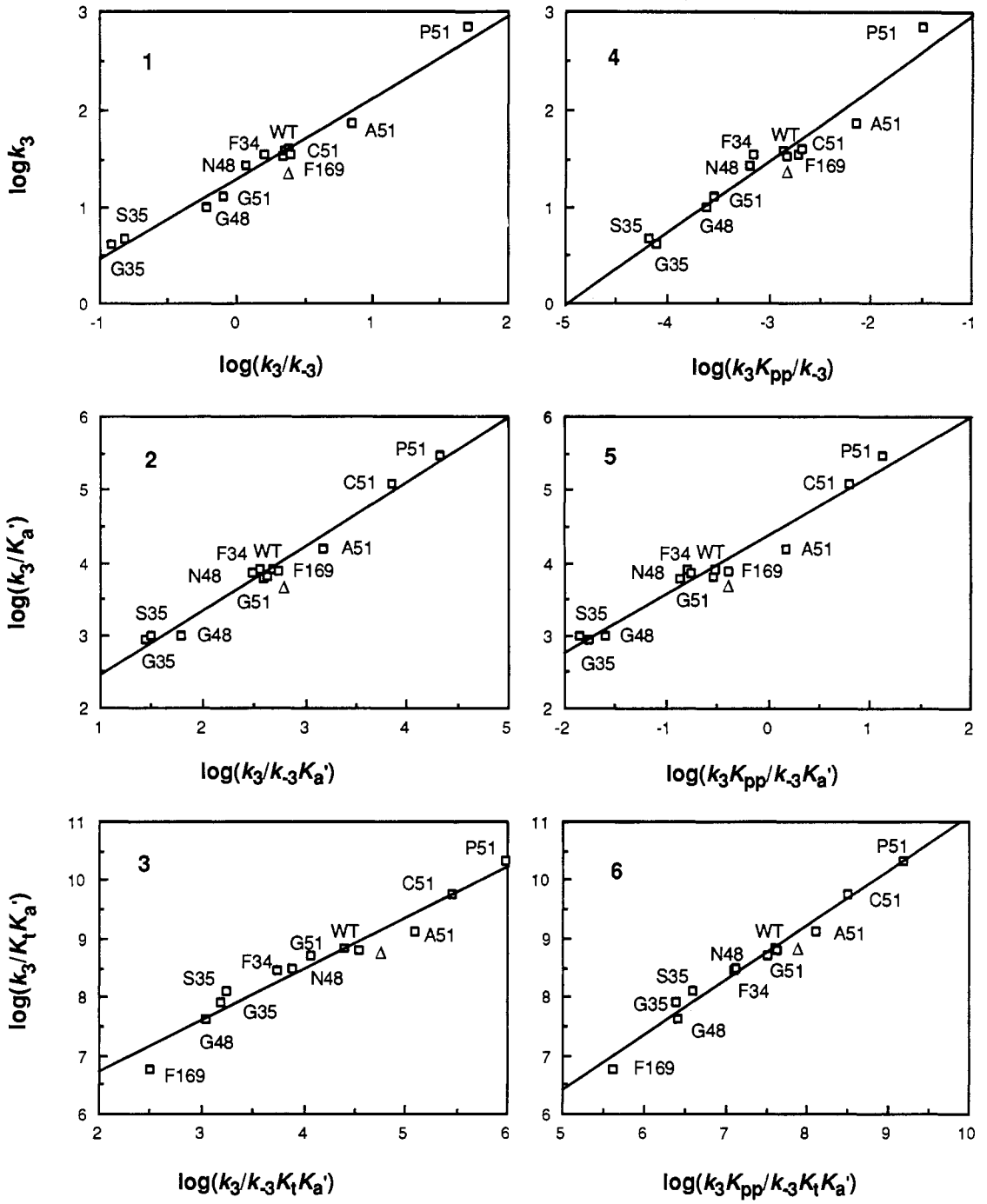

FIGURE 4: Linear free energy plots for different stages of the activation of tyrosine and the equilibria between enzyme-bound complexes.

transition (or intermediate), and final states to be compared. The equilibrium constant relating the initial and final states and the rate and equilibrium constant relating the transition state to the initial state (or, for an intermediate, just the equilibrium constant) were derived. The rate and equilibrium constants were transposed into the free energy terms just by considering the free energies of each state.

(A) Mutations in the Binding Sites for Tyrosine (Tyr → Phe-34, Tyr → Phe-169) and the Ribose of ATP (Cys → Gly-35, Cys → Ser-35, His → Gly-48, His → Asn-48, Thr → Gly-51, Thr → Ala-51, Thr → Cys-51, Thr → Pro-51). (Case 1) Relationship between k_3 and k_3/k_{-3}. A relationship of the form

$$k_3 = A(k_3/k_{-3})^{\beta} \quad (12)$$

is equivalent to

$$G_{E[T\text{-}A]^*} - G_{ETA} = \text{constant} + \beta(G_{ET\text{-}A\cdot PP} - G_{ETA}) \quad (13)$$

Using the manipulations described in the appendix, eq 13 transforms into

$$\Delta G_{[T\text{-}A]^*} - \Delta G_{TA} = \text{constant} + \beta(\Delta G_{T\text{-}A\cdot PP} - \Delta G_{TA}) \quad (14)$$

where $\Delta G_{[T\text{-}A]^*}$ represents the free energy of binding of the transition state [T-A]* to the enzyme, ΔG_{TA} represents the sum of the binding energies of tyrosine and ATP, etc. Equation 14 relates the change in binding energy on going from the ETA complex to E[T-A]* (transition state) compared with the total change in binding energy between ET-A·PP (tyrosyl adenylate + PP bound) and ETA. All the states are enzyme-bound and so no terms in K_t, K'_a, or K_{pp} are involved.

We find that a plot of log k_3 against log (k_3/k_{-3}) (Figure 4) gives a good linear free energy relationship of slope 0.79. This suggests that, on average, 79% of the binding energy change on going from ETA to ET-A·PP is realized in the E[T-A] transition state.

(*Case 2*) *Relationship between* k_3/K'_a *and* $k_3/k_{-3}K'_a$. The free energy changes are

$$G_{E[T\text{-}A]^*} - G_{ET} = \text{constant} + \beta(G_{ET\text{-}A\cdot PP} - G_{ET}) \quad (15)$$

which is equivalent to

$$\Delta G_{[T\text{-}A]^*} - \Delta G_T = \text{constant} + \beta(\Delta G_{T\text{-}A\cdot PP} - \Delta G_T) \quad (16)$$

Here, the initial state is the enzyme–tyrosine complex (ET). Thus, on formation of ET-A·PP and E[T-A]*, the binding energy of ATP to the ET complex is important and so K'_a enters the equations. β will contain in addition to case 1 the fraction of the total binding energy of ATP to the enzyme in the various states. The logarithmic plot (Figure 4) gives again a good line, but the slope β is now 0.92. That is, 92% of the binding energy change on going from ET to ET-A·PP is realized in E[T-A]*.

(*Case 3*) *Relationship between* $k_3/K_tK'_a$ *and* $k_3/k_{-3}K_tK'_a$. The free energy changes are

$$G_{E[T\text{-}A]^*} - G_{WE} = \text{constant} + \beta(G_{ET\text{-}A\cdot PP} - G_{WE}) \quad (17)$$

which is equivalent to

$$\Delta G_{[T\text{-}A]^*} = \text{constant} + \beta(\Delta G_{T\text{-}A\cdot PP}) \quad (18)$$

Here, the initial state is the free enzyme. Thus, formation of E[T-A]* and ET-A·PP from E involves the total binding energies of tyrosine and ATP at the different stages and so K_t and K'_a enter the equations. There is again a satisfactory logarithmic plot (Figure 4). The slope is about 0.95; i.e., 95% of the binding energy change on going from E to ET-A·PP is realized in the E[T-A]* complex.

(*Case 4*) *Relationship between* k_3 *and* k_3K_{pp}/k_{-3}. The energy changes are

$$G_{E[T\text{-}A]^*} - G_{ETA} = \text{constant} + \beta(G_{ET\text{-}A} - G_{ETA}) \quad (19)$$

which is equivalent to

$$\Delta G_{[T\text{-}A]^*} - \Delta G_{TA} = \text{constant} + \beta(\Delta G_{T\text{-}A} - \Delta G_{TA}) \quad (20)$$

This is the same as case 1, but the final state is the ET-A complex. The slope of the linear logarithmic plot (Figure 4) is 0.71, slightly less than that for case 1. This indicates that there is a slight increase in binding energy on going from ET-A·PP to ET-A.

(*Case 5*) *Relationship between* k_3/K'_a *and* $k_3K_{pp}/k_{-3}K'_a$. The energy changes are

$$G_{E[T\text{-}A]^*} - G_{ET} = \text{constant} + \beta(G_{ET\text{-}A} - G_{ET}) \quad (21)$$

which is equivalent to

$$\Delta G_{[T\text{-}A]^*} - \Delta G_T = \text{constant} + \beta(\Delta G_{T\text{-}A} - \Delta G_T) \quad (22)$$

This is the same as for case 2 except that the final state is now the ET-A complex because the term K_{pp} is involved. Again, β at 0.84 (Figure 4) is slightly lower than that for case 2.

(*Case 6*) *Relationship between* $k_3/K_tK'_a$ *and* $k_3K_{pp}/k_{-3}K_tK'_a$. The energy changes are

$$G_{E[T\text{-}A]^*} - G_{WE} = \text{constant} + \beta(G_{ET\text{-}A} - G_{WE}) \quad (23)$$

which is equivalent to

$$\Delta G_{[T\text{-}A]^*} = \text{constant} + \beta(\Delta G_{T\text{-}A}) \quad (24)$$

This is the same as for case 3 except that the final state is ET-A. The value of β (Figure 4) is 0.88, similar to but slightly lower than that for case 3.

Further Plots May Be Constructed with All the Permutations and Combinations of the Rate and Equilibrium Constants. For example, the above results indicate that the binding energy change on going from ETA to ET-A·PP is only about 90% of that on going from ETA to ET-A since the β values are about 10% lower for cases 4–6. This may be shown directly by a plot of log (k_3/k_{-3}) against log (k_3K_{pp}/k_{-3}) (Fersht et al., 1986a). The energy changes are equivalent to

$$G_{ET\text{-}A\cdot PP} - G_{ETA} = \text{constant} + \beta(G_{ET\text{-}A} - G_{ETA}) \quad (25)$$

which is equivalent to

$$\Delta G_{T\text{-}A\cdot PP} - \Delta G_{TA} = \text{constant} + \beta(\Delta G_{T\text{-}A} - \Delta G_{TA}) \quad (26)$$

The slope of this plot is seen directly to be 0.9.

Another example is the fraction of the binding energy of ATP used in going from the ET complex to the ETA complex relative to ET going to ET-A. This is found from a plot of log $(1/K'_a)$ against log $(k_3K_{pp}/k_{-3}K'_a)$. This is again a reasonable plot of slope $\beta = 0.12$ (data not shown). The point for Cys-51 deviates from this line, the mutation Thr → Cys-51 being the closest example of a uniform binding energy change (Ho & Fersht, 1986). This deviation is seen again in the following section.

Direct Calculation of β *Values from Difference Energy Diagrams.* The apparent contributions of the free energies of binding of the various side chains to the binding energies of each state have been calculated previously from the differences in the free energy profiles for the reactions of mutant and wild-type enzymes (Wells & Fersht, 1986; Ho & Fersht, 1986). These may be used to construct individual β values for each mutant. For example, as illustrated in Figure 5, if the apparent contribution of the binding energy of a side chain in the ternary complex ETA is $\Delta\Delta G_{ETA}$, in the transition state is $\Delta\Delta G_{E[T\text{-}A]^*}$, and in the ET-A complex is $\Delta\Delta G_{ET\text{-}A}$, as defined by Wells and Fersht (1986), then in an example exactly analogous to case 4 (eq 19 and 20)

$$\beta = (\Delta\Delta G_{E[T\text{-}A]^*} - \Delta\Delta G_{ETA})/(\Delta\Delta G_{ET\text{-}A} - \Delta\Delta G_{ETA}) \quad (27)$$

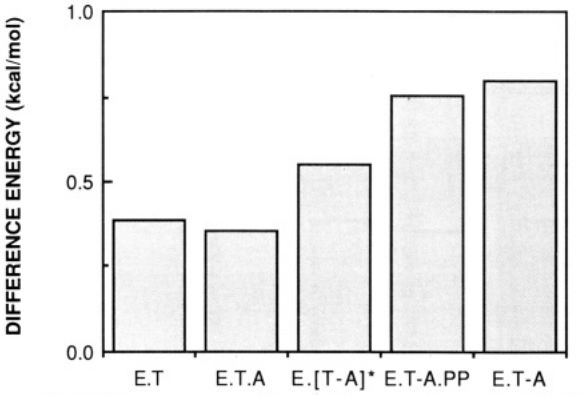

FIGURE 5: Difference energy diagram between the mutant TyrTS-(His→Asn-48) and wild-type enzyme.

The values of β derived for individual mutants in Table I are in acceptable agreement with those derived from the slopes of the linear free energy plots. For example, the plot for case 4 (eq 19 and 20) gives a slope of 0.71 whereas the direct calculation gives a range of values averaging 0.61. Similarly, the plot relating to eq 25 and 26 gives a slope of 0.9 whereas the direct calculations give an average of 0.84. Comparison of the individual values in Table I with the values from the linear free energy plots shows how the plots smooth out the individual values for each mutant and shows the overall trend. The small differences between the average values calculated from the individual mutants and the slopes of the linear free energy plots arise from the differences in the methods of statistical analysis. Linear regression tends to weight values

Table I: Direct Calculation of β Values from Difference Energy Plots[a]

mutant	β	
	$(\Delta\Delta G_{E[T\cdot A]^*} - \Delta\Delta G_{ETA})/(\Delta\Delta G_{ET\cdot A} - \Delta\Delta G_{ETA})$	$(\Delta\Delta G_{ET\cdot A\cdot PP} - \Delta\Delta G_{ETA})/(\Delta\Delta G_{ET\cdot A} - \Delta\Delta G_{ETA})$
Cys → Gly-35	0.78	0.92
Cys → Ser-35	0.69	0.90
His → Gly-48	0.78	0.77
His → Asn-48	0.45	0.87
Thr →Gly-51	0.66	0.66
Thr → Ala-51	0.43	0.70
Thr → Cys-51	0.2	0.91
Thr → Pro-51	~0.89	~0.99

[a] The values of β in the second column are equivalent to eq 19 and 20 and those in the third column to eq 25 and 26.

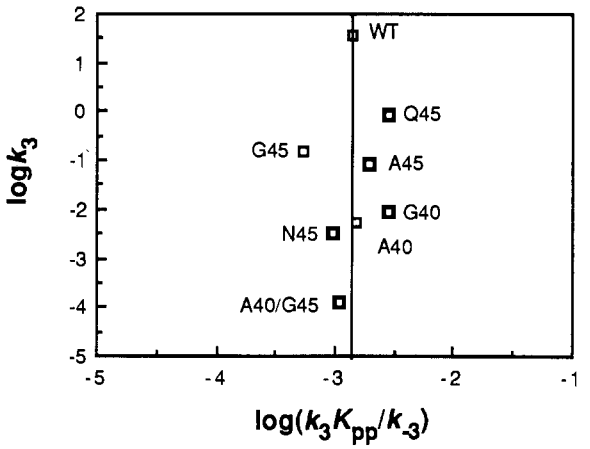

FIGURE 6: Linear free energy plot for mutations in the site for binding the γ-phosphate of ATP in the transition state.

at the extreme ends of the lines, in this case TyrTS(Cys→Gly-35) and TyrTS(Thr→Pro-51), both of which tend to higher values of β in the individual calculations. Further, calculations of individual β values that involve small values of $\Delta\Delta G$ may result in large errors.

(B) Mutations in the Transition-State Binding Site for the γ-Phosphate for ATP (Thr → Ala-40, His → Gly-45, His → Ala-45, His → Asn-45, His → Gln-45). (Case 4) Relationship between k_3 and k_3K_{pp}/k_{-3}. These mutants exhibit a different pattern of behavior from the previous ones (Figure 6). The value of k_{-3} cannot be determined for these mutants because K_{pp} is too high. But the ratio k_{-3}/K_{pp} may be accurately measured and so some of the cases can be analyzed. The slope β in Figure 6 of the logarithmic plot is infinite, that is, $\gg 1$. This is consistent with Thr-40 and His-45 being true examples of transition-state stabilizers—see below.

Discussion

Application of Linear Free Energy Relationship Analysis Requires Careful Design of Mutants and Is Aided by Structural Knowledge. Individual difference energy diagrams are the staple fare for the analysis of energy changes on mutation. The finding of linear free energy relationships correlating a series of mutational experiments is a bonus. We have described in this paper how the structure and activity of a large number of mutant enzymes fit linear free energy relationships with a remarkable consistency—the fit is as good as in many examples from physical organic chemistry. These data comprise *all* of the results of the pre-steady-state kinetics of active site mutants of the tyrosyl-tRNA synthetase that have been published over the past 5 years (Winter et al., 1982; Fersht et al., 1984, 1985; Wells & Fersht, 1986; Ho & Fersht, 1986; Leatherbarrow et al., 1986). But, it must not be considered that these constitute a set of *random* mutants. All the mutants were deliberately and carefully chosen a priori to be in the class that we now term nondisruptive deletions—we have made just small perturbations in the structure to form a highly selected set.

Before embarking on the rigorous analysis, we satisfied ourselves that we were dealing with nondisruptive deletants by a number of criteria. First, the structures of the two mutants at the two extremes of the plots in Figure 4 have been solved by X-ray crystallography [TyrTS(Thr→Pro-51), Brown et al. (1986); TyrTS(Cys→Gly-35), M. Fothergill (unpublished data from this laboratory)], and there are insignificant changes in structure other than the change in side chain and the consequent alteration in solvation. Second, the complete free energy profiles of all the mutants have been determined (Wells & Fersht, 1986; Ho & Fersht, 1986; R. J. Leatherbarrow and A. R. Fersht, unpublished results), and there are no serious perturbations of the energies of the ground-state complexes (E·Tyr, E·Tyr·ATP, E·Tyr-AMP·PP, and E·Tyr-AMP) other than those expected. For example, mutations in the ATP-binding site do not alter the dissociation constant of tyrosine and vice versa, except where there are long-range electrostatic effects. A rare example of the dissociation constant of tyrosine being affected by a mutation in the ATP site is in TyrTS(His→Asn-48) (Figure 5). Here, an electrostatic interaction with the carboxylate of tyrosine is lost on mutation.

Difference energy diagrams are also invaluable for deciding which points to plot on what graphs. For example, it is obvious on examination of Figure 5 that there is a consistent increase in binding energy on going from ETA to E[T-A]* to ET-AA·PP to ET-A. Conversely, in Figure 6, the binding energy changes are seen only in transition-state binding. These clearly represent different phenomena and must be treated separately. If all the experiments were plotted on the same graph without use of chemical knowledge for selection, then the trends in the plots would be obscured. The selection of data for linear free energy plots by use of structural information and accumulated chemical knowledge is entirely precedented by physical organic chemistry. For example, in experiments involving substituted benzene rings, substituents would be segregated into one class consisting of meta and para substituents, and a second class involving ortho substituents because of the known steric effects in the latter. Further, it would be decided whether or not to allow for bond conjugation in choosing the substituent constant for para derivatives. The free energy plots for the mutant enzymes thus fall into two groups, as expected from the difference energy diagrams (Figures 4 and 6).

The first consists of residues that bind to the side chain of tyrosine and the ribose ring of ATP, where the binding energy increases throughout the reaction. Values of β were obtained on the variation of those side chains, which are consistent with the following (Figure 7). On binding ATP to the ET complex, only 12% of the final binding energy of adenosine in the ET-A complex is realized. There is an improvement in binding energy on going from the ternary complex ETA to the transition state E[T-A]* such that 84% of the final binding energy in ET-A is realized. The intermediate ET-A·PP invokes about 91% of the energy. These are, of course, average values for the side chains analyzed.

The second category consists of mutants at positions 40 and 45, which were postulated previously to be involved in binding the transition state only. Mutation of these hardly affects the equilibrium constant between ETA and ET-A and so there is an ill-defined relationship between rate and equilibrium

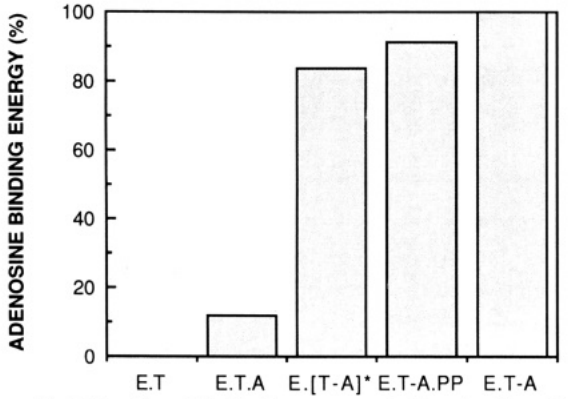

FIGURE 7: Utilization of the binding energy of adenosine throughout the reaction. The initial state is the E·T complex, and the final state is the E·T-A complex, which is taken as the 100% point for realization of the binding energy of adenosine. The relevant interactions are those with Cys-35, His-48, Thr-51, and their mutants. The heights of the bars are 100 times the β value for the relevant plots.

constant that is equivalent to a value of β tending to infinity.

When Should Binding Changes Conform to Linear Free Energy Relationships? Binding energy changes during a reaction may be classified into *uniform*, where a change in structure affects the energy level of each state by the same amount, and *differential*, where each state is altered by different amounts (Albery & Knowles, 1977). These may be further elaborated, as in Figure 8. Suppose the energy levels in differential binding change linearly with reaction coordinate between enzyme-bound substrate, transition state, and product so that the binding energy of the transition state is intermediate between the two ("linear differential"). Suppose, for example, the energy of the EI complex is lowered by a mutation and the ES complex is unaffected. If the linear change in energy is maintained on mutation, the energy level of the E[S]* complex will be lowered by a smaller amount, depending on how much it resembles S or I in its interactions with the enzyme. If it is midway between the two, as in Figure 8, the change in energy will be 50% of that of the change in the EI complex. Thus, linear differential binding will give rise to a linear free energy relationship.

Removal of a group involved in uniform binding will not affect the rate or equilibrium constants of the enzyme-bound states. A particular example of differential binding is when there are interactions that specifically stabilize (or destabilize) the transition state but affect the initial or final states much less, for example, the binding site for the γ-phosphate of ATP. Another example would be electrostatic interactions with charges that are generated in the transition state. As in Figure 8, alteration of these interactions need not alter the binding energies in the ES or EI complex. In the extreme case where the mutated group affects only the transition state, k_2 in eq 4 would alter, but the equilibrium constant k_2/k_{-2} would be unaffected. This generates an infinite value of β, as found in Figure 6. In less examples where the ground states were affected to some extent, high values of β of positive or negative slope would be generated. An observed value of β of greater than 1 is indicative of transition-state binding modes.

Breakdown of Linear Free Energy Relationships. The classes of binding energy changes illustrated in Figure 8 may not be apparent for various reasons even though they are occurring. One reason is that linear free energy relationships are often found to be linear portions of a more extensive nonlinear scheme. Curvature can arise from the nature of the transition state changing with changes in structure of the reagents (Jencks, 1985). The same may be true for binding

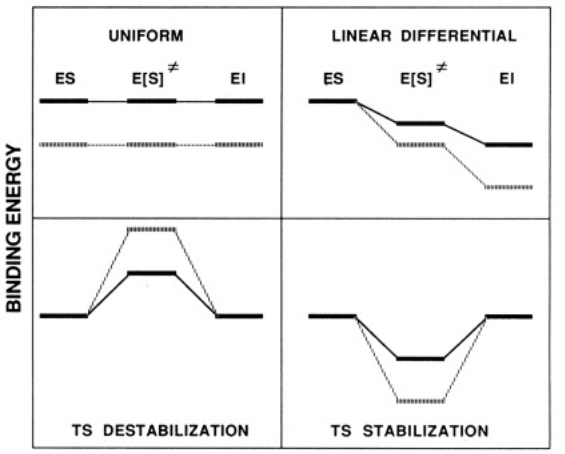

FIGURE 8: Modes of interaction of a side chain of an enzyme with a substrate (S), transition state ([S]‡), and intermediate (I) and the effects of mutation. The binding energy of the interaction is plotted as the reaction proceeds. The broken lines represent the interaction energy of the side chain in wild-type enzyme, and the solid lines represent possible energy levels after mutation of the side chain to remove the interaction. In practice, combinations of all effects may be found.

energy changes. There are, however, more likely sources of problems that obscure the existence of linear free energy relationships. The first stems from the very nature of changes in protein structure on mutation. The existence of a linear free energy relationship arises from the enzyme reacting in a uniform way to mutational change as if structural changes are part of a continuum. That is, the transition state occupies a particular position on the reaction coordinate that responds in a uniform and consistent manner to changes in structure. But a mutation involves a discrete change, and the change could be too large.

Linear free energy relationships also need not occur. There must be many residues that will not fit into a simple pattern: many mutations may cause highly specific effects, each change being unique.

Possible Artifacts in Interpretation of Fit of Mutants to Linear Free Energy Plots. Three phenomena have been encountered that give rise to results that could be misinterpreted (Lowe et al., 1987). The first is a general point about the nature of mutations. There could be effects on mutation that are propagated through the structure, as in disruptive mutations. For example, a mutation in the binding site might cause a small effect on the binding of a substrate but cause a structural change that perturbs another residue which is involved in transition-state stabilization. This would give a deviation from a linear free energy plot and possibly an anomalously high value of β that could be misinterpreted [e.g., TyrTS(Ala-38), TyrTS(Ala-78), etc.]. The second stems from the mathematical analysis. In a plot of, say, log k_3 against log (k_3/k_{-3}), the two variables are not independent. Suppose k_{-3} is measured accurately but k_3 is subject to large error. Then, the error distribution in the log–log plot consists of a line of slope 1.0 passing through the correct value of log k_3. Thus, if the value of β is close to unity, a highly incorrect value of k_3 would not be apparent—as seen for TyrTS(Ala-173) (Lowe et al., 1987). Nevertheless, the fit to the plot requires that the value of A in the Brønsted equation must also be correct. Plots of slope close to unity may well conceal artifactual data.

The third is that the behavior of an individual mutant on a plot may well deviate from the mean. Small deviations are often not apparent on log–log plots because of the compression of data when there are several orders of magnitude of spread.

For example, the difference energy diagram for TyrTS(Phe-169) indicates that there is a uniform binding energy change and not the strikingly progressive increase in binding energy as the reaction procedes. But this is not easily seen above the noise in the log–log plots. The linear free energy plots show the overall trends and smooth out small individual variations. The behavior of any one mutant is best observed from its particular difference energy diagram.

We are confident that the linear free energy plots presented in this study are not artifactual because there is so much consistency between them and the difference energy diagrams: that is, between the general trends and the behavior of individuals.

Conclusions. The demonstration that certain structural changes of the tyrosyl-tRNA synthetase cause changes in rate that may be described satisfactorily by linear free energy equations is important for the following reasons.

(i) The equations combine the data from many experiments and so systematize and bring order to them.

(ii) They collectively delineate aspects of the transition-state structure, in particular by how much the interactions in the transition state resemble those in the enzyme–substrate and the enzyme–intermediate complexes.

(iii) The equations show trends. Any mutant that significantly deviates from the trend is immediately apparent. The accompanying paper (Lowe et al., 1987) contains examples of deviations.

The last point is particularly important. The results of any one individual mutation can be challenged on the grounds that perhaps the mutation propagates a significant change in protein structure beyond the local changes immediate to the mutated side chain. This is difficult to answer experimentally even by X-ray crystallographic studies since they would not detect a series of small changes of the order of 0.1 Å. But, when a whole series of mutational experiments can be fitted to a linear free energy equation, there is overwhelming evidence that each mutant shows part of a general phenomenon of the relationship between structure and activity.

We must end with a strong warning. This type of approach requires a thorough knowledge of the consequences of mutations, in particular from structural data and from difference energy diagrams, before interpretation of results.

APPENDIX

Demonstration That Free Energy Changes in the Reactions of Mutant Enzymes Are a Consequence of Binding Energy Changes. Wolfenden (1972) showed that the dissociation constant of a transition state from an enzyme may be calculated from the ratio of the rate constants for the enzyme-catalyzed reaction and the uncatalyzed reaction in solution combined with the dissociation constant of the enzyme–substrate complex. The following extends that approach by using it with the treatment of Fersht (1974). In Figure 1, which refers to eq 4, let the free energy of the substrate in water be G_{WS}, the free energy of the enzyme–substrate complex be G_{ES}, and the free energies of the equivalent forms of the transition state be $G_{W[S]^\bullet}$ (in water) and $G_{E[S]^\bullet}$ (enzyme–transition-state complex). The energies of the different states of the intermediate, I, are denoted similarly, and the unligated enzyme has an energy G_{WE}.

The binding energy of the substrate, ΔG_S, is given by

$$\Delta G_S = G_{ES} - G_{WE} - G_{WS} \quad \text{(A1)}$$

Similarly, the binding energies of the transition state and intermediate are

$$\Delta G_{[S]^\bullet} = G_{E[S]^\bullet} - G_{WE} - G_{W[S]^\bullet} \quad \text{(A2)}$$

and

$$\Delta G_I = G_{EI} - G_{WE} - G_{WI} \quad \text{(A3)}$$

From equilibrium thermodynamics

$$RT \ln (k_2/k_{-2}) = G_{ES} - G_{EI} \quad \text{(A4)}$$

From transition-state theory

$$RT \ln (k_2/K_S) = kT/h - (G_{E[S]^\bullet} - G_{WE} - G_{WS}) \quad \text{(A5)}$$

$$RT \ln k_2 = kT/h - (G_{E[S]^\bullet} - G_{ES}) \quad \text{(A6)}$$

(where R is the gas constant, T is the temperature, ad k and h are the Boltzmann and Planck constants).

The energy differences of the enzyme-bound states can be described in terms of binding energies and the energies of the reagents in solution by manipulation of eq A1, A2, and A3 (or by inspection of Figure 1). For example

$$G_{ES} - G_{EI} = \Delta G_S - \Delta G_I + G_{WS} - G_{WI} \quad \text{(A7)}$$

$$G_{E[S]^\bullet} - G_{ES} = \Delta G_{[S]^\bullet} - \Delta G_S + G_{W[S]^\bullet} - G_{WS} \quad \text{(A8)}$$

Thus, substitution of eq A7 into A4 gives

$$RT \ln (k_2/k_{-2}) = \Delta G_S - \Delta G_I + G_{WS} - G_{WI} \quad \text{(A9)}$$

and similarly

$$RT \ln k_2 = kT/h - (\Delta G_{[S]^\bullet} - \Delta G_S + G_{W[S]^\bullet} - G_{WS}) \quad \text{(A10)}$$

$$RT \ln (k_2/K_S) = kT/h - (\Delta G_{[S]^\bullet} + G_{W[S]^\bullet} - G_{WS}) \quad \text{(A11)}$$

(Note that $G_{W[S]^\bullet} - G_{WS}$ is the activation energy of the reaction in water that goes through the same transition-state structure of the substrate as on the enzyme.) These equations may be used to construct free energy relationships and calculate binding energies from kinetic data either for the reaction of a series of substrates with a common enzyme or for one substrate with a series of mutant enzymes, the equations being essentially symmetrical. For example, for a series of mutant enzymes, G_{WS}, $G_{W[S]^\bullet}$, and G_{WI} are independent of enzyme, and only the terms involving the enzyme vary. Since kT/h is also constant

$$RT \ln (k_2/k_{-2}) = \Delta G_S - \Delta G_I + \text{constant} \quad \text{(A12)}$$

$$RT \ln k_2 = \text{constant} - \Delta G_{[S]^\bullet} + \Delta G_S \quad \text{(A13)}$$

$$RT \ln (k_2/K_S) = \text{constant} - \Delta G_{[S]^\bullet} \quad \text{(A14)}$$

Equation A12 shows that the equilibrium constant k_2/k_{-2} varies solely as the differences in binding energies of S and I with the enzymes. Equation A13 shows that the rate constant k_2 varies solely according to the differences between the binding energies of the transition state [S] and S with the enzymes. Equation A14 shows that the second-order rate constant k_2/K_S varies solely as the binding energy of the transition state with each enzyme.

Caveats. The above analysis holds for all mutational changes, but artifacts can arise from two causes: (i) the more likely cause is that free energies of binding, ΔG_S, ΔG_I, and $\Delta G_{[S]^\bullet}$, contain the energies of any conformational change in the enzyme induced on binding the substrate, and mutation could cause spurious conformational changes in the enzyme; (ii) the less likely cause is that the assumption that G_{WS}, $G_{W[S]^\bullet}$, and G_{WI} are the same for all mutants could break down if the changes on mutation are so severe that the reaction proceeds by different transition states for different mutants.

Registry No. A, 56-65-5; T, 60-18-4; TyrTS, 9023-45-4; L-Cys, 52-90-4; L-His, 71-00-1; L-Thr, 72-19-5; L-Gly, 56-40-6; L-Ser, 56-45-1;

L-Asn, 70-47-3; L-Ala, 56-41-7; L-Pro, 147-85-3.

References

Albery, W. J., & Knowles, J. R. (1977) *Angew. Chem., Int. Ed. Engl. 16*, 285–293.

Brown, K. A., Vrielink, A., & Blow, D. M. B. (1986) *Biochem. Soc. Trans. 14*, 1228–1229.

Estell, D. A., Graycar, T. P., Miller, J. U., Powers, D. B., Burniv, J. P., Ng, P. G., & Wells, J. A. (1985) *Science (Washington, D.C.) 233*, 659–663.

Fersht, A. R., Shi, J. P., Wilkinson, A. J., Blow, D. M., Carter, P., Waye, M. M. Y., & Winter, G. P. (1984) *Angew. Chem., Int. Ed. Engl. 23*, 467–473.

Fersht, A. R., Shi, J. P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., & Winter, G. (1985) *Nature (London) 314*, 235–238.

Fersht, A. R., Leatherbarrow, R. J., & Wells, T. N. C. (1986a) *Nature (London) 322*, 284–286.

Fersht, A. R., Leatherbarrow, R. J., & Wells, T. N. C. (1986b) *Trends Biochem. Sci. (Pers. Ed.) 11*, 321–325.

Foster, R. J. (1961) *J. Biol. Chem. 236*, 2461–2466.

Ho, C. K., & Fersht, A. R. (1986) *Biochemistry 25*, 1891–1897.

Jencks, W. P. (1985) *Chem. Rev. 85*, 511–517.

Landis, B. H., & Berliner, L. J. (1980) *J. Am. Chem. Soc. 102*, 5354–5358.

Leatherbarrow, R. J., Fersht, A. R., & Winter, G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7840–7844.

Lowe, D. M., Fersht, A. R., Wilkinson, A. J., Carter, P., & Winter, G. (1985) *Biochemistry 24*, 1506–1509.

Lowe, D. M., Winter, G., & Fersht, A. R. (1987) *Biochemistry* (following paper in this issue).

Wells, T. N. C., & Fersht, A. R. (1986) *Biochemistry 25*, 1881–1886.

Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M., & Smith, M. (1982) *Nature (London) 299*, 756–758.

Wolfenden, R. (1972) *Acc. Chem. Res. 5*, 10–18.

Structure–Activity Relationships in Engineered Proteins: Characterization of Disruptive Deletions in the α-Ammonium Group Binding Site of Tyrosyl-tRNA Synthetase†

Denise M. Lowe,‡,§ Greg Winter,‖ and Alan R. Fersht*,‡

Department of Chemistry, Imperial College of Science and Technology, London SW7 2AY, U.K., and Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, U.K.

Received January 21, 1987; Revised Manuscript Received April 27, 1987

ABSTRACT: Residues Asp-78 and Gln-173 of the tyrosyl-tRNA synthetase of *Bacillus stearothermophilus* form part of the binding site for tyrosine by making hydrogen bonds with the α-ammonium group. Asp-38 is close enough to the group to make an important electrostatic contribution. Unlike other residues in the active site that have been studied by site-directed mutagenesis, Asp-38, Asp-78, and Gln-173 are part of hydrogen-bonded networks. Each of these residues has been mutated to an alanine, and the resultant mutants have been studied by kinetics to construct the difference energy diagrams for the formation of tyrosyl adenylate. In each example, the binding of tyrosine is weakened by about 2.5 kcal mol^{-1}. But, unlike previous mutants, the dissociation of the second substrate, in this case ATP, is also seriously affected, being weakened by some 2 kcal mol^{-1} for TyrTS(Ala-78) and TyrTS(Ala-173). The energy of the transition state for the formation of tyrosyl adenylate is raised by 7.8 kcal mol^{-1} for the former and 4.5 kcal mol^{-1} for the latter mutant. Addition of these mutants to linear free energy plots constructed for the nondisruptive mutants in the accompanying study [Fersht, A. R., Leatherbarrow, R. J., & Wells, T. N. C. (1987) *Biochemistry* (preceding paper in this issue)] reveals large deviations of the data for TyrTS(Ala-38) and TyrTS(Ala-78) from the regression line. These thus belong to a different class of mutations from previous nondisruptive examples. This observation combined with the structural evidence and difference energy diagrams strongly suggests that the mutations Asp → Ala-38 and Asp → Ala-78 are disruptive in nature.

Side chains at the active site of the tyrosyl-tRNA synthetase that may possibly form hydrogen bonds with the substrates are being systematically mutated in order to analyze their contributions to catalysis (Winter et al., 1982). The mutations so far have all been in the class that may be described as nondisruptive deletions (Fersht et al., 1987). Certain residues in the binding pocket for the α-amino group of tyrosine, however, form more extensive interactions in the protein. Examination of the crystal structure of the enzyme-bound tyrosine complex (Blow & Brick, 1985; Brick & Blow, 1987) reveals that residues Tyr-169, Gln-173, and Asp-78 make hydrogen bonds with the α-ammonium group of tyrosine (Figure 1). In addition, the aliphatic portion of the side chain of Gln-173 forms one side of the binding site for the aromatic ring of the substrate tyrosine, making van der Waals' interactions. The carboxylate of Asp-38 is close enough to make an electrostatic interaction with the α-ammonium group but does not appear to make a direct hydrogen bond with it. It

† This work was funded by the Medical Research Council of the U.K.

‡ Imperial College of Science and Technology.

§ Present address: Department of Biochemical Virology, The Wellcome Research Laboratories, Beckenham, Kent BR3 3BS, U.K.

‖ Medical Research Council Laboratory of Molecular Biology.

0006-2960/87/0426-6038$01.50/0