

## Combining multi-mutant and modular thermodynamic cycles to measure energetic coupling networks in enzyme catalysis

Charles W. Carter, Jr., Srinivas Niranj Chandrasekaran, Violetta Weinreb, Li Li, and Tishan Williams

Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7260, USA

(Received 7 November 2016; accepted 21 December 2016; published online 26 January 2017)

We measured and cross-validated the energetics of networks in *Bacillus stearothermophilus* Tryptophanyl-tRNA synthetase (TrpRS) using both *multi-mutant* and *modular* thermodynamic cycles. Multi-dimensional combinatorial mutagenesis showed that four side chains from this “molecular switch” move coordinately with the active-site  $Mg^{2+}$  ion as the active site preorganizes to stabilize the transition state for amino acid activation. A *modular* thermodynamic cycle consisting of full-length TrpRS, its Urzyme, and the Urzyme plus each of the two domains deleted in the Urzyme gives similar energetics. These dynamic linkages, although unlikely to stabilize the transition-state directly, consign the active-site preorganization to domain motion, assuring coupled vectorial behavior. © 2017 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). [<http://dx.doi.org/10.1063/1.4974218>]

### INTRODUCTION

One of our oldest and most inclusive insights about proteins is that protein folding coordinates the behavior of the amino acids linked together in the polypeptide chain. Proposals implicating the resulting cooperative protein motions, or “dynamics,” in enzyme catalysis are commonplace.<sup>1–11</sup> A central notion underlying such proposals is that low frequency motions—loop closing, domain motion, etc.—transmit the kinetic energy along directions aligned with chemical changes, thereby contributing to the overall rate acceleration.

Such proposals have been challenged, however, for three main reasons.<sup>12</sup> (i) Frequencies of domain motions are many orders of magnitude slower than those of the chemical reactions. (ii) Forces that can be imposed with such motions are substantially weaker than those necessary to stabilize the transition states. (iii) In order to definitively implicate motions in catalysis, it is necessary to map the accurate free energy landscapes of chemical reactions and how they integrate with the protein-dynamical changes, and cases where these have been done do not reveal a significant catalytic contribution.

Nonetheless, many enzymes, especially NTP (nucleotide triphosphate)-dependent motors, synthetases, polymerases, and signaling GTPases, exhibit multi-state, *vectorial* behavior<sup>13–19</sup> in which biological function depends critically on coupling catalysis to conformational changes. Purposes served by such coupling include many of the most interesting aspects of biology. For this reason, the creation of, and transitions between, these multiple states are biologically relevant. It is therefore important to determine the mechanisms by which conformational transitions can be coupled to active-site chemistry.<sup>13,14,20</sup>

Our work on *Bacillus stearothermophilus* tryptophanyl-tRNA synthetase (TrpRS) occupies a specialized niche amidst these questions. We try to answer such questions as:

- (i) What remote parts of a protein exert a substantive influence on catalysis?
- (ii) What techniques are necessary to verify these effects?

- (iii) What is the magnitude of these influences?
- (iv) How cooperative are they?
- (v) What is their functional significance?

TrpRS furnishes an attractive laboratory model for exploring long-range coupling comprehensively.

Our interest in these questions was initially stimulated by communications with Arieh Warshel, who has argued forcefully that the essence of catalysis is electrostatic stabilization of the transition state<sup>21–23</sup> and cannot arise from dynamical effects. Reviewing an earlier paper of ours,<sup>24</sup> he challenged our contention that long-range forces might be involved in catalysis, and argued that the TrpRS rate acceleration was due primarily to electrostatic stabilization by the single Mg<sup>2+</sup> ion.

Testing that idea experimentally, we found that only about a third of the overall transition-state stabilization could be attributed to the Mg<sup>2+</sup> ion.<sup>25</sup> This posed two questions:

- (i) What accounts for the metal-independent two-thirds of the transition-state stabilization?
- (ii) How does metal-protein coupling work if Mg<sup>2+</sup> does not interact with the protein, has very little effect on the rate in water, but accelerates the rate 10<sup>5</sup>-fold in the presence of TrpRS?

The first question remains to be answered, although we have taken some steps<sup>26,27</sup> in that direction. We tentatively answer the second question here. We conclude that Warshel is correct that domain movements themselves cannot directly stabilize transition states. However, what he calls the “active site pre-organization”<sup>28</sup> appears to occur only during the domain movement, which has the effect of synchronizing catalysis with domain movement. The resulting conditional dependence of catalysis is what couples catalysis to domain movements,<sup>20,29,30</sup> enabling the vectorial behavior. Readers may wish to refer to Figs. 10 and 11, which are discussed in detail later on, and provide a useful visualization of the proposed reaction path.

## RESULTS

TrpRS uses Adenosine 5' triphosphate (ATP) to activate the carboxyl group of the amino acid tryptophan to form Trp-5' adenylate plus pyrophosphate (PPi). This reaction is the *sine qua non* of protein synthesis, and therefore of fundamental biological significance. Domain movements along the TrpRS structural reaction profile were elucidated by solving the crystal structures of TrpRS without ligands<sup>31</sup> and complexed to substrates—tryptophan, ATP;<sup>32</sup> non-reactive substrate analog combinations—tryptophanamide plus ATP;<sup>32</sup> tryptophan plus Adenosine-5' monophosphate (AMP) plus PPi;<sup>33</sup> a transition-state analog—adenosine-5'tetraphosphate;<sup>34</sup> and products—trp-5'AMP,<sup>35,36</sup> trp-5'sulfoamyl-AMP, trp-methenyl-AMP;<sup>37</sup> representing different states along the reaction path. These crystal structures group into three distinct, narrowly defined conformations that correspond to nodes connected by the three canonical stages of enzyme catalysis: induced-fit active-site assembly, catalysis, and product release.<sup>38</sup> In this respect, the three resulting “states”—open, pre-transition state (PreTS), and Products—correspond functionally and structurally to the three conformations observed in the successive stages of the chemical reactions of the ATP by the catalytic F1-ATPase  $\beta$ -subunit.<sup>39</sup> Specifically, an open conformation is associated with the unliganded state, a closed, twisted conformation is associated with the ATP bound state, and a closed, untwisted conformation is associated with the aminoacyl-AMP-bound state. We previously commented on the possible relationship between the three-state behavior and free energy transduction.<sup>38</sup>

The three TrpRS conformations are produced by the relative rigid-body movement of the anticodon-binding domain (ABD) relative to the catalytic domain.<sup>33</sup> The induced-fit and catalytic transitions are rate-limited by high-energy conformational transition states<sup>24,40</sup> along the conformational free energy surface. We identified the structures responsible for these high-energy conformations using searches for the most probable paths that connect the respective ground-state conformations.<sup>30,41</sup> An important barrier to domain movement associated with induced-fit is a change in aromatic side chain packing within the D1 switch, a motif located in the N-terminal crossover connection of the Rossmann fold.<sup>41,42</sup>

Seeking to understand the evolutionary sources of TrpRS catalysis, we have experimentally deconstructed the monomer (Fig. 1) into three nested, catalytically active components—Protozyme (46 residues), Urzyme (130 residues), and Catalytic Domain (204 residues)—which, together with the anticodon-binding domain (ABD) form the full-length TrpRS monomer.<sup>26</sup> The Protozyme and Urzyme retain a considerable catalytic activity and hence likely represent ancestral forms of contemporary aminoacyl-tRNA synthetases. Full-length synthetases combine the Urzyme with two additional modules added later in evolution: connecting peptide 1 (CP1), which completes the catalytic domain and provides the context for editing domains in the larger synthetases specific for aliphatic side chains; and the anticodon-binding domain, which enables recognition of the tRNA anticodon (Fig. 1(a)). The CP1 module does not actually move as a rigid body with the Urzyme, but moves relative to it. We, and others, have used computational methods to identify the correlated motions of the ABD domain and CP1 in several of the Class I synthetases.<sup>40,43</sup>

CP1 and the ABD together form a “cone” that is fitted around the Urzyme (Fig. 1(b)). The active-site opening is accessible only through this cone, which therefore appears to physically constrain the relative positions of the Urzyme’s modular ATP and amino acid binding sites. The active site is thus sensitive to the configuration of the CP1 and ABD modules in full-length TrpRS. We have measured the impact of these interactions on catalysis<sup>20,25,29</sup> and specificity.<sup>30,44</sup>

### Thermodynamic cycles measure the coupling energies

A key concept underlying this work is that of “energetic coupling.” Coupling implies the existence of sources of free energy that enforce joint behaviors; such joint behaviors enable new, higher levels of functionality. It is generally true that coupling energies across a very broad spectrum of circumstances can be assessed from the non-additivity of expected contributions. A useful example is that the intrinsic proton angular momentum, or spin, cannot be explained using only the intrinsic properties of its three constituent quarks, but instead requires an energetic coupling between them and the gluons that hold them together.<sup>45</sup> In nuclear physics, the data come from an inelastic electron and proton scattering experiments.<sup>46</sup> In structural

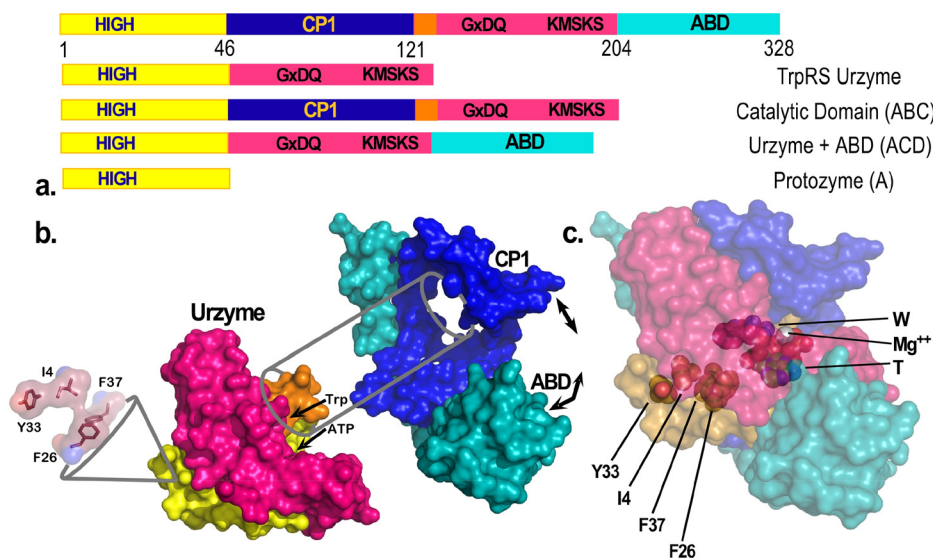


FIG. 1. TrpRS hierarchy of functional modules. (a) Schematic of sequential order of TrpRS modules: Protozyme (yellow), Urzyme, (yellow plus magenta), CP1 (dark blue), and anticodon-binding domain (teal). (b) Fitting of the D1 Switch into the Urzyme, and of the Urzyme into the framework imposed by CP1 (blue) and the ABD domain (teal). The D1 Switch forms the helix-strand corner of the first crossover connection (yellow). The second crossover connection of the Rossmann fold (red) follows the CP1 insertion and is part of the Urzyme. Correlated motions of CP1 and the ABD are indicated by arrows. (c) Relationship between the D1 Switch residues and the active-site. W is tryptophan, T is ATP.

biology, non-additivities are evaluated quantitatively for combinatorial perturbations using a device that in a biophysical context is called a thermodynamic cycle (Fig. 2).

A thermodynamic cycle is a  $2^N$  factorial experiment that tests all possible combinations of specific perturbations for their impact on kinetic or thermodynamic behavior. Perturbations that have non-additive effects in contexts of other perturbations imply coupled interactions. Use of multi-mutant cycles was introduced in biochemistry by Jencks<sup>47,48</sup> and exploited by Fersht soon after the development of directed mutagenesis enhanced their feasibility.<sup>49–51</sup> High-order thermodynamic cycles have demonstrated three-dimensional coupling interactions capable of changing the direction of conformational equilibria by up to  $10^5$ -fold.<sup>52</sup>

The  $^{32}\text{PPI}$  exchange reaction used to assay the aminoacyl-tRNA synthetases is done under equilibrium conditions, so both  $K_M$  and  $k_{\text{cat}}$  are related to the thermodynamic values,  $\Delta G_{K_M}$  and  $\Delta G_{k_{\text{cat}}}$ . Free energies are additive, so linear regression methods can solve the simultaneous equations necessary to estimate and assess the statistical significance of coefficients for main and higher-order interaction effects.<sup>20</sup> The latter represent the accessible coupling energies for a given experimental design. Experimental perturbations according to a full-factorial design can thus be used to estimate both the strength and significance of energetic coupling between parts of a protein during enzyme-catalyzed reactions.

In the course of this work, we were led to investigate a series of apparently unrelated combinatorial perturbations. To our surprise, three qualitatively quite different thermodynamic cycles give the same quantitative energetic coupling during TrpRS catalysis.

### The TrpRS/ $\text{Mg}^{2+}$ cycle; $\text{Mg}^{2+}$ ion accelerates amino acid activation $\sim 10^5$ -fold in the TrpRS-catalyzed reaction, but only 5-fold in water

Prompted by the suggestion that the TrpRS catalysis should be essentially eliminated in the absence of  $\text{Mg}^{2+}$  ion, we prepared extensively depleted stocks of enzyme, assay buffer, and reagents.<sup>25</sup> The resulting assays exhibited reduced, but still substantial rate accelerations over the uncatalyzed rate, which we initially attributed to residual remaining  $\text{Mg}^{2+}$  ion (Fig. 3(b)).

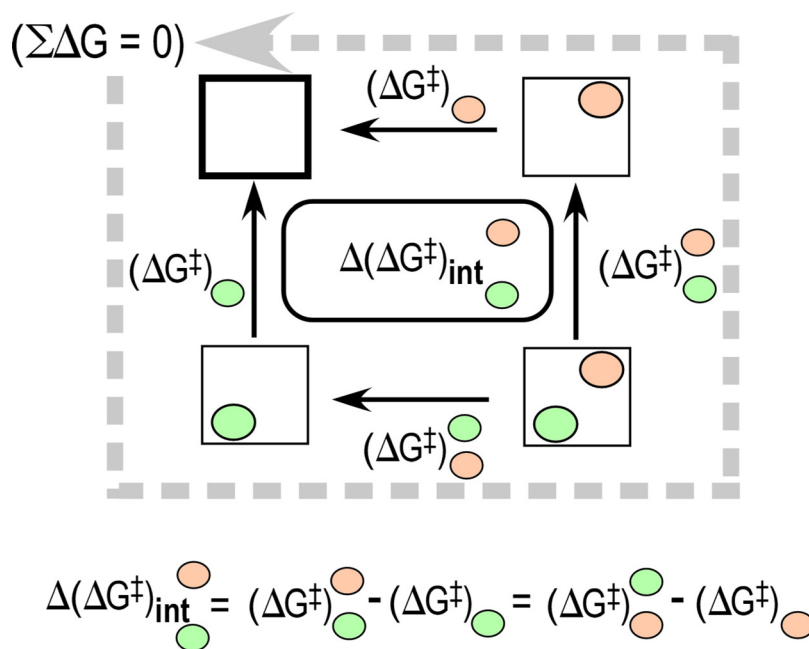


FIG. 2. Schematic of a double mutant thermodynamic cycle. The factorial design of this experiment emphasizes measurement of the impact of mutation at one site (green ellipse) in the context of mutation at a second site (salmon ellipse). Rate changes between nodes along the edges are converted to free energies,  $\Delta G^\ddagger = -RT \ln(k) = -0.592 * \ln(k)$ . As the total free energy change is zero for the circuit, non-additivities of opposite vertical and horizontal edges are necessarily equal and furnish estimates of the coupling energies between the two perturbed sites.

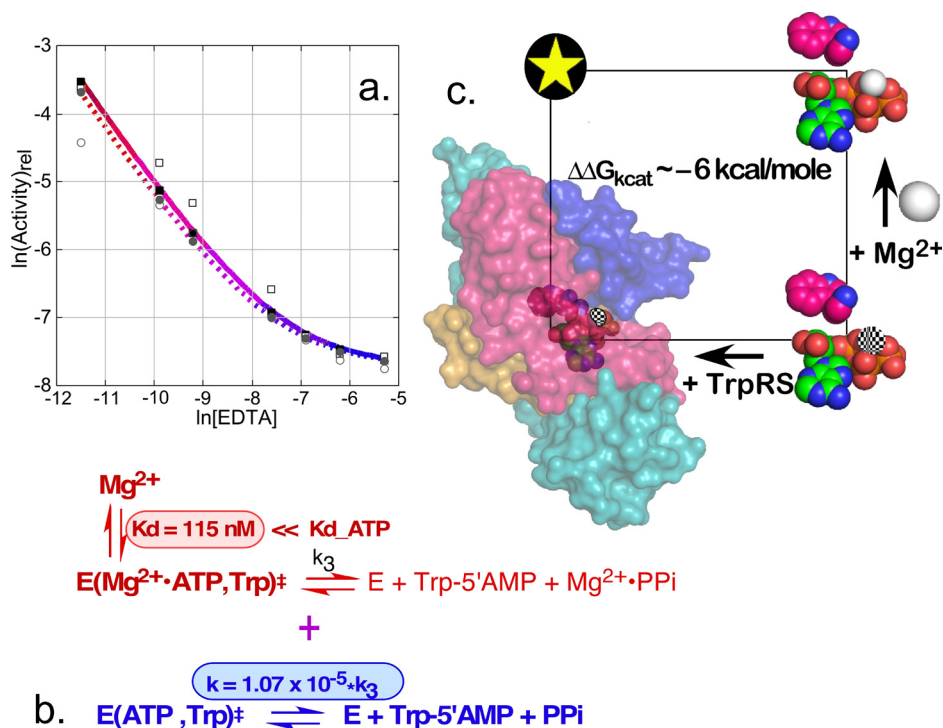


FIG. 3. The TrpRS- $\text{Mg}^{2+}$  cycle. (a) Experimental data supporting the conclusion that TrpRS induces a limiting rate acceleration in the absence of  $\text{Mg}^{2+}$ , achieved by adding increasing amounts of EDTA to remove residual metal. Red coloring of the decay curve indicates residual catalytic assist by  $\text{Mg}^{2+}$  ion; Blue coloring indicates a plateau of catalysis by metal-free TrpRS. (b) Nonlinear fitting of the [EDTA]-dependence of the logarithm of TrpRS activity yields two parameters,  $K_D^\ddagger$ , the dissociation constant for  $\text{Mg}^{2+}$  ion from the transition state (highlighted in red) and the maximum rate acceleration given by saturating  $\text{Mg}^{2+}$  concentrations (highlighted in blue) (adapted with permission from Weinreb and Carter, Jr., *J. Am. Chem. Soc.* **130**, 1488–1494 (2008). Copyright 2008 American Chemical Society). (c) Tryptophan activation is accelerated to very different extents by the addition of either TrpRS ( $10^9$ -fold) or  $\text{Mg}^{2+}$  (5-fold). The two together induce an overall acceleration of  $10^{14}$ -fold. The  $\Delta G_{\text{int}} = 6.4 \text{ kcal/mole}$ .  $R^2 = 0.999$  for the model, with  $P(F) = 0.0006$ .

Titration of the assays with increasing amounts of ethylene diamine tetraacetate (EDTA) caused an additional rate reduction, but that reached a plateau beyond which there was, surprisingly, no further decrease.<sup>25</sup>

Iterated solution of the simultaneous binding equilibria using known dissociation constants for the  $\text{Mg}^{2+}\cdot\text{ATP}$  and  $\text{Mg}^{2+}\cdot\text{EDTA}$  equilibria allowed us to fit the activity data at two TrpRS concentrations to a nonlinear model with two parameters: the proportional decrease in rate in the absence, relative to that in the presence of  $\text{Mg}^{2+}$  ion ( $1.1 \times 10^{-5}$ ) and the dissociation constant of  $\text{Mg}^{2+}$  from the transition-state complex (120 nM)<sup>25</sup> (Fig. 3(c)). The former result implies that the metal accounts for only about 33% of the transition-state stabilization free energy, and is unexpected for several reasons: (i) It implies that the TrpRS active site alone is capable of  $\sim -12 \text{ kcal/mole}$  of transition-state stabilization free energy, even in the absence of the metal; (ii)  $\text{Mg}^{2+}$  ion accelerates ATP-dependent reactions in water by only  $\sim 5$ -fold.<sup>53</sup> The two-factor thermodynamic cycle mapping the effects of TrpRS and  $\text{Mg}^{2+}$  ion thus exhibits a non-additivity of  $-6.4 \text{ kcal/mole}$ , implying that the metal and protein are coupled by that energy in the transition state for ATP-dependent tryptophan activation.<sup>25</sup>

$\text{Mg}^{2+}$  has been observed in crystal structures only of the PreTS state. In those structures (1MAU, 1M83), it interacts directly only with three phosphate oxygen atoms and two water molecules, and not with any TrpRS functional groups. The absence of protein-metal contacts necessarily implied that the energy coupling them together arises from as yet unidentified sources that, in turn, interact with the  $\text{Mg}^{2+}$  ion indirectly, via the ATP phosphate oxygens. We first sought to identify these unknown coupling interactions within the active site, via combinatorial mutagenesis of the active-site lysine residues that are physically coupled to the metal because



they compete with the metal for electrostatic interactions with phosphate oxygen atoms (Fig. 4).<sup>29</sup> Those experiments showed that the  $Mg^{2+}$ -protein interactions within the active-site actually retard the reaction,<sup>29</sup> forcing us to look further afield for the sources of catalytic assist by  $Mg^{2+}$ .

### Identification of the D1 switch

To locate these distant interactions, we used a combination of two bioinformatic filters—one geometric, the other energetic—to identify side chains likely to influence conformational equilibria (Fig. 5).<sup>24</sup> First, we quantitatively identified the local packing interactions that were invariant throughout the conformational cycle using Delaunay tessellation.<sup>54–56</sup> Eliminating from consideration all invariant tetrahedra of nearest neighbors—those within rigid bodies—greatly reduced the number of such simplices that are “dynamic” and change during catalysis. Then, we used Rosetta<sup>57–59</sup> to identify point mutants that might stabilize the high-energy PreTS state, relative to the OPEN and PRODUCTS ground states. The intersection of Venn diagrams identified by both Delaunay tessellation and by the multi-state design algorithm was surprisingly small,<sup>24</sup> consisting of one large, centrally located dynamic packing motif and three smaller peripheral motifs. We called these dynamic motifs “Switches” D1–D4, from “Delaunay.”

These bioinformatic filters are discussed further in the Methods section. Their combined effects reduced the number of candidate residues to a small enough number that experimental perturbation was feasible. The majority of the residues that survived this two-stage filter were all part of the D1 switch, a packing motif at the C-terminus of the first alpha helix of the Rossmann fold. We subsequently identified this packing motif as one of the most widely distributed motifs in the proteome.<sup>42</sup> The widespread phylogenetic distribution and key location of

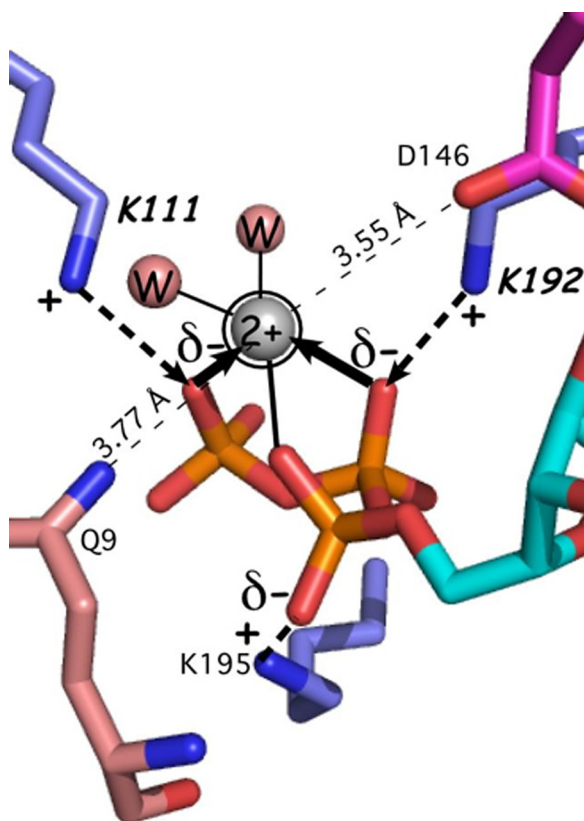


FIG. 4. Coordination of the active-site  $Mg^{2+}$  involves no protein residues. The metal ion is seen only in TrpRS PreTS crystal structures,<sup>32</sup> where it assumes an evidently strained configuration in which distances from each of three phosphate oxygen atoms are longer than those in  $Mg^{2+}$ -ATP.<sup>85</sup> The competition between  $Mg^{2+}$  and the lysine residues also weakens TrpRS affinity for ATP.<sup>24</sup>

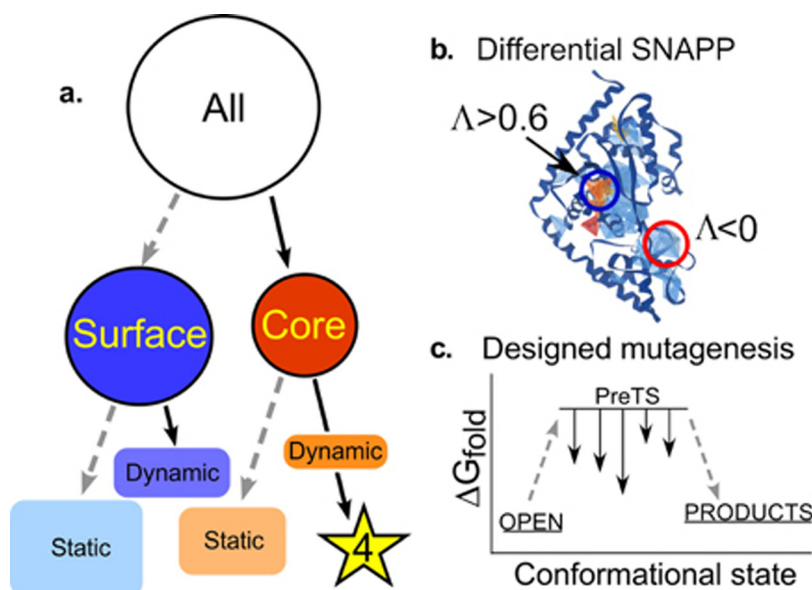


FIG. 5. Identification of dynamic contacts with a two-stage bioinformatics filter. (a) Schematic of sequential identification of dynamic and static, core and surface interactions resulting in the identification of dynamic core residues. (b) Differential Delaunay tessellation and likelihood scoring identify simplices that rearrange partners in different conformations. Of the 250 simplices in the TrpRS monomer, 70 have high SNAPP log-likelihoods ( $\Lambda > 0.6$ ; red) and 150 have negative log-likelihoods ( $\Lambda < 0.0$ ; blue). The majority of those residues identified by SNAPP analysis are embedded within rigid bodies and these do not rearrange during catalysis. Of those “dynamic” simplices that do rearrange, 4 and 28, respectively, of each type, fall within the D1 Switch and also have mutations suggested by Rosetta in (c). (c) Protein multi-state design<sup>59</sup> is used to identify the side chains whose mutation should stabilize the PreTS conformation, relative to the OPEN and PRODUCTS ground-states. Roughly 65% of the mutations suggested by Rosetta occur either in high SNAPP log-likelihood core tetrahedra (blue) or in negative log-likelihood tetrahedra, which are found at the protein surface (red).

this motif in the first crossover connection of the Rossmann fold reinforced our surmise that the residues engaged in dynamic switching in this region might be responsible for the long-range coupling to the active-site  $\text{Mg}^{2+}$  ion. Although the D1 Switch motif consists of seven residues, we decided to mutate only four, because we sought a full factorial design, and 15 variants was an accessible number to construct and assay. Moreover, because the Rosetta multi-state design algorithm had given us the specific mutations for each site—I4V, F26L, Y33F, F37I—and because these mutations all consistently reduced differences between the Rosetta scores of the PreTS and Product structures, we used these specific residues, rather than alanine mutations.

### Combinatorial mutagenesis of D1 switch residues; long-range coupling to the catalytic $\text{Mg}^{2+}$ ion

Alan Fersht used the term “multi-mutant cycle analysis” to describe the factorial directed point mutagenesis designs.<sup>51</sup> Such experiments measure the coupling energy between individual residues and/or active site ligands. Extensive combinatorial multi-mutant cycle analysis of active-site functional groups subsequently showed that none of these coupling energies between protein and metal resides in the active-site residues,<sup>29</sup> creating a strong argument for functionally relevant effects arising from parts of the protein distant from the active-site.

In order to detect coupling to the active-site metal, we first verified that the catalytic assist by  $\text{Mn}^{2+}$  ion was a suitable perturbation by showing that the metal substitution reduced activity to  $\sim 15\%$  of that with  $\text{Mg}^{2+}$  and that the concentration dependencies of catalysis by the two metals were very similar over a  $10^{-5}$ -fold range of metal concentrations, both exhibiting a maximal effect at approximately the same concentrations.<sup>29</sup> All possible combinations of the four D1 mutations were constructed, purified, and their active fractions were determined by active-site titration.<sup>60</sup>

The resulting experimental design involved 32 separate combinations of point mutant and metal ( $Mg^{2+}$  vs.  $Mn^{2+}$ ) and was thus perhaps the most ambitious of its kind yet attempted.<sup>19,61</sup> All variants were assayed by ATP-dependent Michaelis-Menten steady-state kinetic analysis. Approximately, three-fold replication of the experimental data afforded a fit in which essentially all the main and higher-order interaction free energies of transition-state stabilization could be estimated with high statistical significance.<sup>20</sup> Overall, the regression model for  $\Delta G_{\text{cat}}$  has coefficients with only modest magnitudes, of both signs, for the main effects and intermediate-order interactions. The 5-way interaction between the four permuted sites and the metal ion, however, contributes  $\sim -5$  kcal/mole. The 5-dimensional multi-mutant cycle showed that  $\sim 85\%$  of the catalytic effect of  $Mg^{2+}$  ion can be attributed to the 5-way interaction between the metal and four residues that mediate the shear during domain movement (Fig. 6).

The regression model is a variant of principle component analysis, and its coefficients furnish a detailed histogram of contributions to transition-state stabilization (Fig. 6(b)). The dominant five-way interaction term has a definitive interpretation: at the moment of transition state stabilization, *motions of all four D1 switch side chains and the active-site  $Mg^{2+}$  ion are tightly coupled, such that their simultaneous movement can reduce the free energy of the transition state.*<sup>20</sup> This interpretation implies that the change in side-chain packing that occurs during domain motion occurs essentially simultaneously with some, as yet undocumented, change in the configuration of the  $Mg^{2+}$  ion in  $Mg^{2+}$ -ATP and that this configurational change allows the metal ion to interact more tightly with the transition state. Thus, this multi-mutant cycle is in substantial quantitative agreement with the conclusions drawn from the two-way thermodynamic cycle involving the intact TrpRS enzyme and the  $Mg^{2+}$  ion described in the section on Identification of the D1 switch.

### A novel, modular thermodynamic cycle

If the interpretation of high-order coupling described in the section on Combinatorial mutagenesis is correct, then it should also be the case that the two domains that rotate relative to the active site should also be coupled together in the transition state by similar free energies.<sup>44</sup> We recently confirmed the prediction quantitatively, by constructing and carrying out a modular

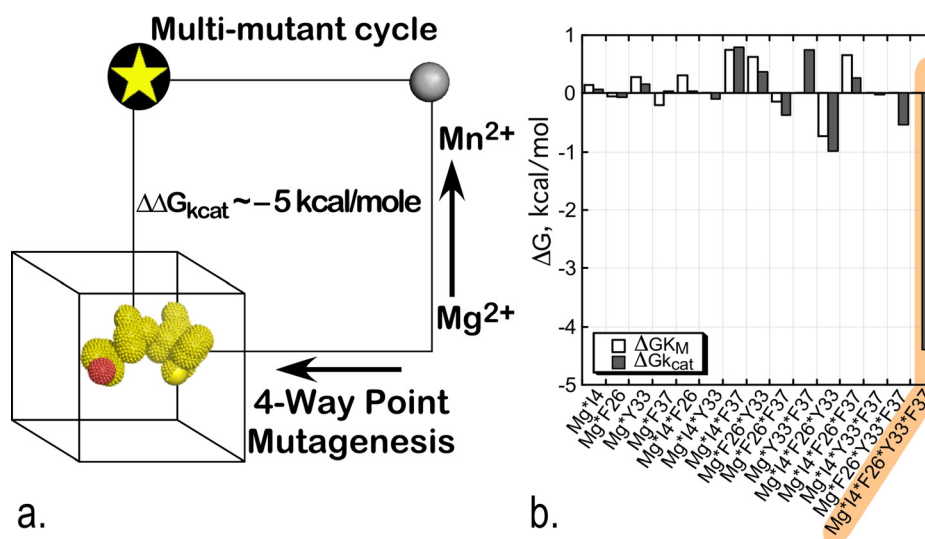


FIG. 6. Combinatorial mutagenesis of four D1 Switch residues identify a strong, 5-way interaction (shaded amber) implying that all four mutated residues move coordinately with the active-site  $Mg^{2+}$  ion in the transition state for tryptophan activation.<sup>20</sup> (a) Schematic of the 5-way thermodynamic cycle. (b) The combinatorial mutagenesis of the D1 residues, I4, F26, Y33, and F37, together with the substitution of  $Mg^{2+}$  with  $Mn^{2+}$  reveals that all five moieties are moving coordinately as the chemical transition state is stabilized.  $R^2 = 0.95$  for the model, with  $P(F) < 0.0001$ , and Student t-test probabilities range between  $< 0.001$  to  $0.05$  for coefficients  $> 0.15$  kcal/mole.



thermodynamic cycle.<sup>44</sup> Modular cycles are enabled for TrpRS by the characterization of Urzymes<sup>62–64</sup> (and potentially of protozymes<sup>26</sup>). As the CP1 and ABD domains are definitively deleted, the TrpRS Urzyme itself plays the role of a molecular knockout. Its remaining catalytic activity thus measures the impact of simultaneously losing both modules. We therefore constructed each modular perturbation necessary to complete the cycle by adding back either CP1 or the ABD to measure the effects of each modular deletion separately (refer to Fig. 1).

Combinatorial domain perturbation (Fig. 7) reveals that neither CP1 nor the ABD contribute materially to transition state stabilization by the Urzyme alone, yet the presence of both modules together in full length TrpRS boosts the catalytic effect by  $10^5$ -fold! Non-additivity in the resulting modular thermodynamic cycle measured the free energy coupling between the two mobile domains to be  $-5 \sim \text{kcal/mole}$  in aminoacylation of  $\text{tRNA}^{\text{Trp}}$ . This coupling energy closely matched that obtained in the five-way multi-mutant cycle.<sup>20,44</sup>

Multi-mutant and modular models validate one another in specificity as well as in catalysis. The 4-way interaction between I4, F26, Y33, and F37 also makes a similar contribution to the specific recognition of Tryptophan vs. Tyrosine ( $k_{\text{cat}}/K_{\text{M}})_{\text{Trp}}/(k_{\text{cat}}/K_{\text{M}})_{\text{Tyr}}$  in tryptophan activation.<sup>30</sup> Simultaneous motions of CP1 and the ABD are necessary in order to achieve a fully functional TrpRS catalysis and specific tryptophan recognition: in the absence of either domain, the overall rate acceleration falls by  $\sim 10^{-5}$ -fold, almost exactly to that of the Urzyme alone, and the free energy of discriminating between tryptophan and tyrosine falls from  $\sim 6 \text{ kcal/mole}$  to  $1 \text{ kcal/mole}$  (Fig. 8).<sup>30,44</sup>

### TrpRS multi-state behavior is coupled to catalysis and hence functionally relevant

The coupling evident in Fig. 6(b) is notable for several reasons. First, it is contrary to that expected for higher order factorial experiments<sup>65</sup> in which intrinsic effects normally dominate, and higher order interactions diminish in magnitude. That expected behavior has been verified in a comparable five-dimensional full factorial analysis of the catalytic roles of residues in the active site of alkaline phosphatase<sup>61</sup> (Fig. 9) whose catalytic side chains link only into small clusters that are themselves dominated by the intrinsic effects of individual side chains. Here (Fig. 6(b)), the dominant effect, by a substantial margin, is the highest order interaction term. Second, the highly cooperative nature of protein tertiary structures suggests the possibility that

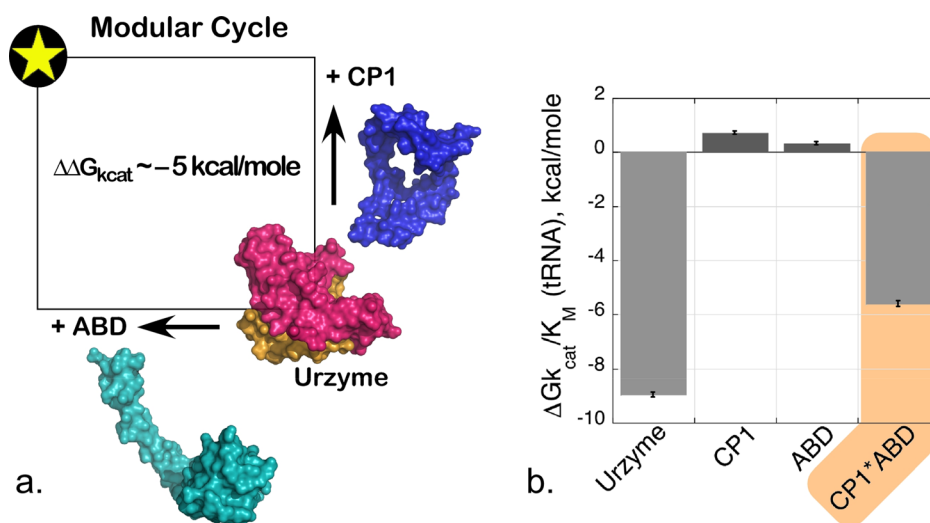


FIG. 7. Modular thermodynamic cycle implicating interdomain coupling. (a) Schematic diagram showing the two-way modular variation; adding either the CP1 or ABD domain to the TrpRS Urzyme. (b) Histogram of the free-energy contributions to  $\Delta G_{\text{cat}}/K_{\text{M}}$  for tRNA aminoacylation. Neither domain contributes favorably to catalysis in the absence of the other, and the energetic coupling free energy (shaded amber) is quantitatively the same as that observed for the D1 Switch- $\text{Mg}^{2+}$  coupling interaction.  $R^2 = 0.98$  for the model, with  $P(F) < 0.0001$ . All histograms have Student t-test probabilities  $< 0.001$ .

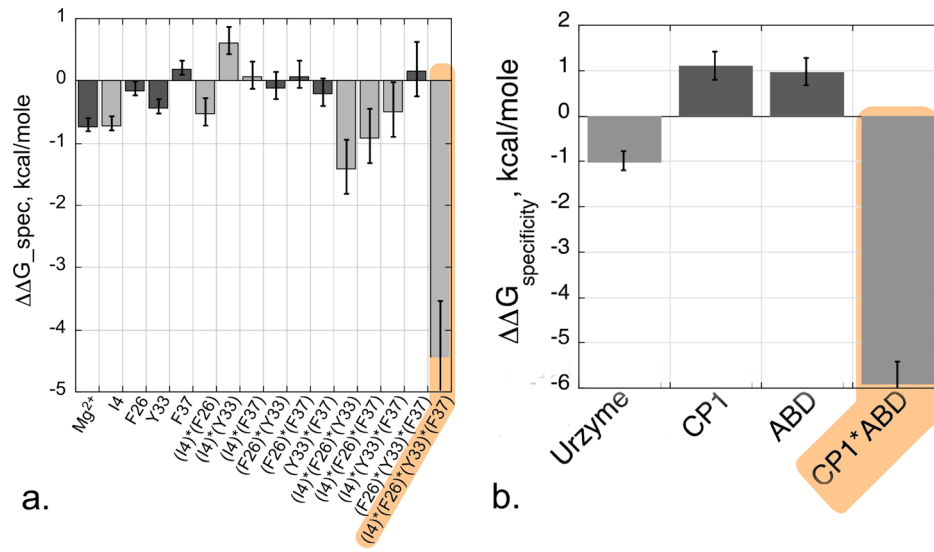


FIG. 8. Energetic coupling between D1 side chains and between domains contribute equally to specificity and catalysis. (a) Histogram comparable to that for the combinatorial mutant cycle in Fig. 6(b). (b) Histogram comparable to that for the modular cycle in Fig. 7. In both (a) and (b), the dependent variable is the specificity for tryptophan vs. tyrosine,  $\Delta G(k_{\text{cat}}/K_{\text{M-Trip}})/\Delta G(k_{\text{cat}}/K_{\text{M-Tyr}})$ .  $R^2=0.99$  for the specificity model, with P(F) and all Student t-test probabilities  $<0.0001$ .

any sufficiently high-dimensional combinatorial mutagenesis would result in high-order coupling. The pattern in alkaline phosphatase serves as an important “negative control” by showing that high-order interactions do not necessarily result from any high-dimensional mutagenesis experiment. Finally, as discussed by Sunden *et al.*, an evolutionary argument holds that high-order coupling is indeed unexpected, and hence represents the product of a complex series of selected properties.<sup>61</sup> We also have advanced similar arguments regarding the conundrum

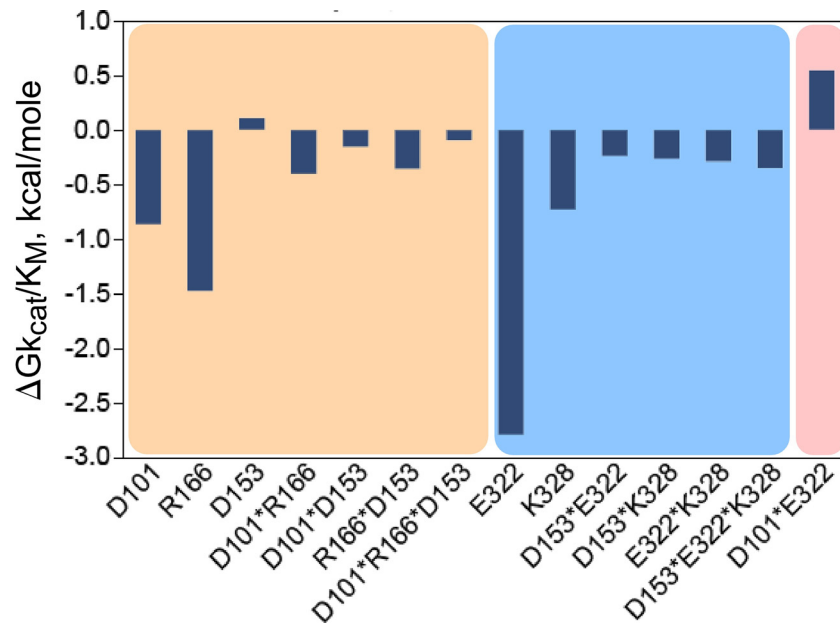


FIG. 9. Significant effects from the 5-way combinatorial mutagenesis of active-site residues in alkaline phosphatase.<sup>61</sup> In contrast to the highly coupled pattern observed in Fig. 6, high-order coupling is not observed, the intrinsic effects are dominant, and connections are limited to local networks denoted by the three shaded backgrounds.

evident in Fig. 7(b), which suggests that the full-length enzyme could not have evolved by a simple process of accumulating first one, then other of the two domains missing in the Urzyme.<sup>44</sup>

The stunning quantitative agreement between the coupling energies estimated from multi-mutant and modular combinatorial experiments with those anticipated from the characterization of metal-free catalysis<sup>25</sup> reinforces our earlier suggestions<sup>32,34,38,66</sup> that domain motions are necessary to form the TrpRS-transition state complex, and hence that differences observed between TrpRS crystal structures are relevant to the catalytic cycle. Thus, the different TrpRS crystal structures probably approximate those states connected by motions necessary for transition-state stabilization. Because the coupling energies contribute to  $\Delta Gk_{\text{cat}}$ , and hence to transition-state stabilization, they represent the transient interactions that occur only in the transition state. Hence, movement of parts of TrpRS—the D1 switch residues, the CP1 and ABD domains, and the catalytic  $\text{Mg}^{2+}$  ion—relative to one another is required to generate them, as Fersht noted when interpreting his early directed mutagenesis of the TyrRS active site<sup>50,67–69</sup> which had to be done without reference to multiple crystal structures.

In view of the recent critique of the functional relevance of domain motions<sup>12</sup> it is worth discussing these results cautiously. The concluding sections address separate aspects of this question.

### Multi-state behavior is imposed by conformational transition states

In order for a protein to exhibit the multistate behavior, its conformation must have multiple distinct potential energy wells, separated by passes of high conformational free energy. We have used a minimum action path algorithm (PATH)<sup>70</sup> to characterize how the three TrpRS conformational states connect to one another via such conformational transition states (Fig. 10(a) (Multimedia view)).<sup>30,33,41</sup> Replica exchange discrete molecular dynamics<sup>71</sup> showed that the conformational transition state for TrpRS catalytic transition is narrowly defined structurally, and that it coincides with a saddle point in the respective conformational free energy landscape<sup>41</sup> (Fig. 10(b)). Importantly, that landscape refutes the contention<sup>72</sup> that the saddle point in the free energy surface between multiple states of a protein is relatively flat in the direction normal to the path, and hence helps assure that the most probable path does result from minimizing the Onsager-Machlup action functional expressing the energy dissipation for a stochastic, overdamped system.<sup>73</sup>

In the same study, we also compared the conformational transition states of the TrpRS Induced-Fit transition to those associated with the Pre-power stroke to rigor transition of the myosin VI converter domain<sup>74</sup> and calcium release by calmodulin. These three transition states are remarkably similar: each entails re-packing of multiple aromatic side chains, which is much slower than other types of repacking. These results, together with our observation that the volume of the tryptophan binding site assumes a minimum just prior to the conformational transition state for the catalytic transition, and hence is likely the moment at which a decision is made regarding whether or not the correct amino acid is bound<sup>30</sup> all reinforce the conformational transition state as a real, functionally relevant conformation.

### Energetic coupling renders the catalytic NTP hydrolysis contingent on the domain motion

Although TrpRS dynamic behavior appears to be closely linked to catalysis by long-range energetic coupling influences on  $k_{\text{cat}}$ , the domain movement itself probably does not actually lower the energy of the transition state. As noted,<sup>12</sup> the time frame of the motions is too slow, and the molecular forces generated are too weak to couple directly to transition-state stabilization. Several apparently unrelated observations suggest a novel, perhaps counterintuitive interpretation of the relationship between domain motion and catalysis.

Rather than contributing directly to transition state stabilization, TrpRS domain motion may contribute instead to dynamic active-site *preorganization*<sup>28</sup> (Fig. 11). By this hypothesis, motions allow the active site to non-randomly populate configurations complementary to the

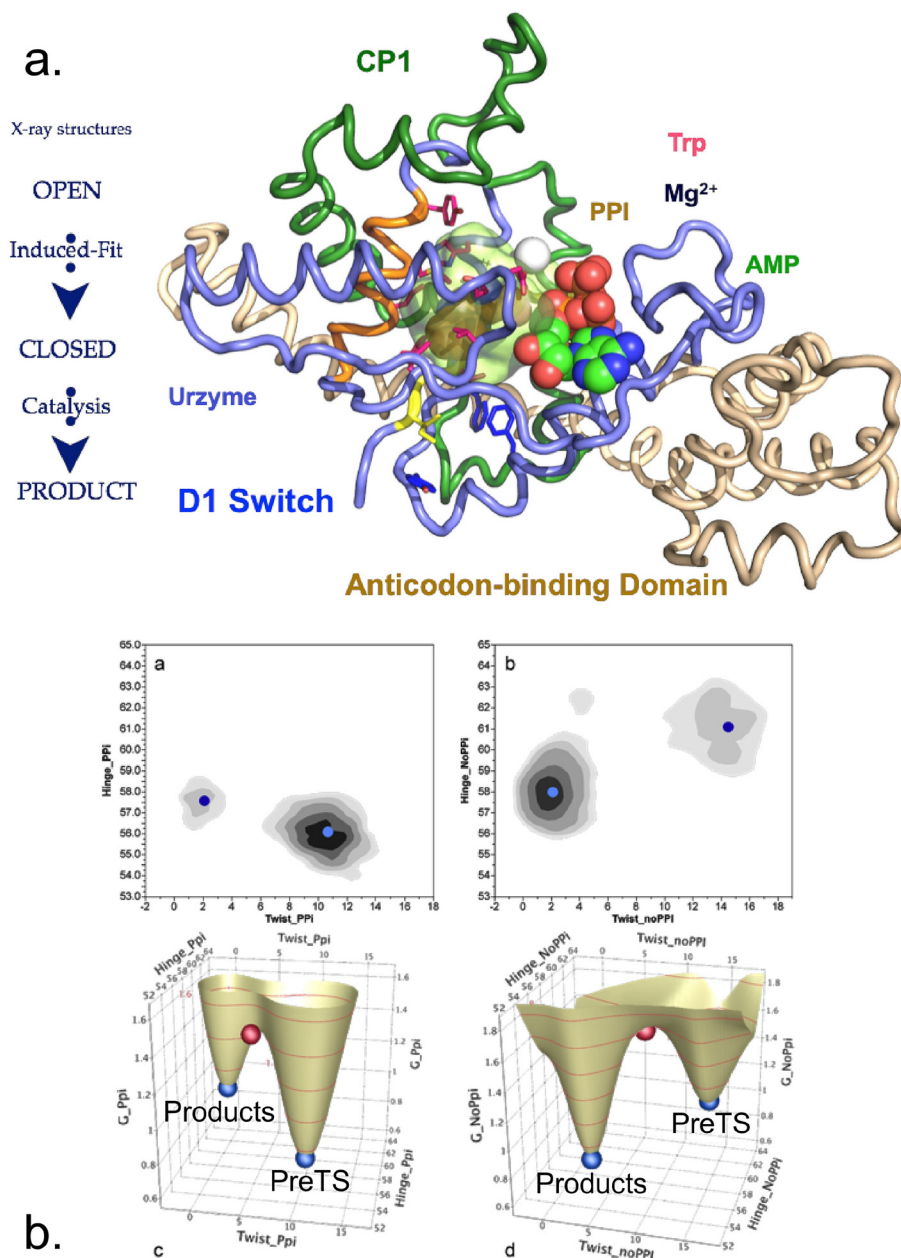


FIG. 10. Computational studies relevant to inferences about the relevance of coupling energies.<sup>41</sup> (a) PATH trajectory [adapted from Weinreb *et al.*, J. Biol. Chem. **289**, 4367–4376 (2014)] implicating the D1 Switch rearrangement in the rate-limiting step in the TrpRS induced-fit conformational change. (b) Free energy surfaces for the catalytic conformational change derived from replica exchange Directed Molecular Dynamics.<sup>71</sup> Simulation performed with a restraint to retain the bound pyrophosphate (left) show that the PreTS conformation is thermodynamically favored; that performed without restraint shows that pyrophosphate is released, and the Products conformation is favored (right). Reproduced with permission from Chandrasekaran *et al.*, Struct. Dyn. 3, 012101 (2016). Copyright 2016 AIP Publishing. (Multimedia view) [URL: <http://dx.doi.org/10.1063/1.4974218.1>].

chemical transition states. The active site is properly configured only transiently, because it coincides with the maximum rate of conformational change and hence exists only for a brief portion of the domain motion. Because preorganization occurs transiently and with a low frequency, the overall chemical reaction may effectively proceed more slowly than it would if the active-site remained pre-organized throughout the substrate residence time.

If this hypothesis is correct, then it would account for three observations that previously seemed unrelated:

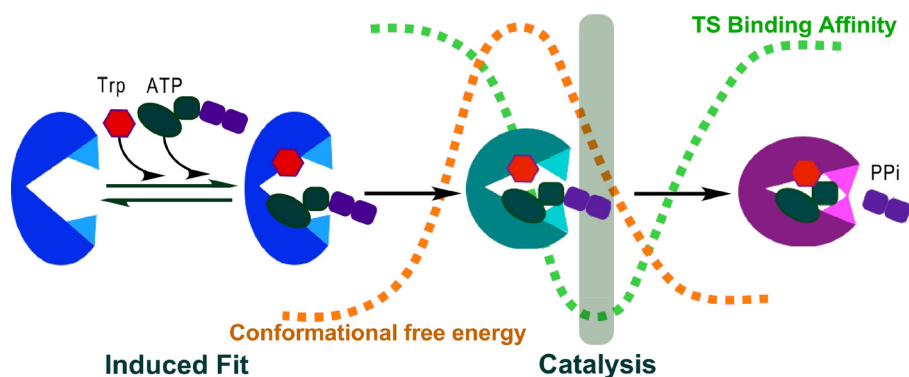


FIG. 11. Conformational coupling renders ATP utilization conditional to domain motion. Tryptophan activation by ATP is illustrated schematically, coloring the Open, PreTS, and Products conformations blue, teal, and magenta, respectively. The brown dashed curve reflects transition in conformational free energy; evidence supporting the qualitative aspects of this curve derives from both computational<sup>24,33,40,86</sup> and experimental<sup>32,34</sup> studies. The green dashed curve reflects the evolution of transition-state binding affinity, inferred from thermodynamic cycle studies summarized herein.<sup>20,30,44,87,88</sup> The offset of the two dashed curves means that high transition-state affinity occurs transiently during the domain motion (grey vertical stripe), thereby creating a conditional dependence of transition-state stabilization on domain movement, as is required for efficient free energy coupling. Dashed lines are not intended to indicate the same energy scales, and the time scales are arbitrary.

- (i) Active-site titration experiments show that the chemical step occurs in about the same time frame as the domain motions, i.e., milliseconds (Fig. 12(a)). Thus, the chemical reaction itself may actually be controlled by, i.e., retarded, by reduced time during which the active-site retains the catalytically active configuration.
- (ii) The free energy of activation for the chemical step measured in single-turnover active site titration experiments,  $\Delta G_{\text{chem}}$ , is closely correlated to the D1 Switch mutations and their higher-order interactions in almost identical fashion to that observed for the  $\Delta G_{\text{cat}}$  values (Figs. 12(b) and 12(c)). Turnover itself, estimated from the steady-state rate from such an experiment is, however, uncorrelated with the pattern of mutations. This result is unexpected in light of the evident correlation between mutant sites and  $\Delta G_{\text{cat}}$  values from steady-state experiments. It can be understood, however, by recognizing that differences in a small fraction of the total reaction time in the latter context are amplified many fold by turnover. The significance of this correlation, evidenced by the very small t-test P-values ( $P < 0.0001$ ), is that structural perturbations influence primarily the chemical step of the reaction, supporting the conclusion that acceleration of the chemical step itself is tightly coupled to the domain motion.
- (iii) Studies of the mechanism by which indolmycin inhibits bacterial TrpRSs<sup>75,76</sup> show that indolmycin affinity for TrpRS is enhanced by the presence of  $\text{Mg}^{2+} \cdot \text{ATP}$ , but not by ATP alone. We have argued elsewhere<sup>75,76</sup> that this unusual effect can be attributed to a small shift in the crystallographically observed metal position, which moves it from a high-energy (and catalytically competent) position to a low-energy (and catalytically inactive) conformation.

### Coupling catalysis to conformational changes creates vectorial behavior

Enzyme mechanisms, especially those that couple NTP hydrolysis to work and/or information, use sophisticated dynamic networks to transduce the active-site chemistry into domain motions that change the macromolecular binding affinities. The dual energetic coupling of catalysis to domain motion and D1 switch repacking with  $\text{Mg}^{2+}$  ion movement renders catalytic utilization of ATP conditional, or contingent upon the domain motion. TrpRS domain motions apparently constitute a timing mechanism.

The variety of different TrpRS X-ray structures has identified each conformation with a distinctive spectrum of ligand affinities,<sup>32–34</sup> including those ligand configurations associated with the transition state. Decelerating the chemical step from what it might be in a more static active site, synchronizing chemical and mechanical events and thereby



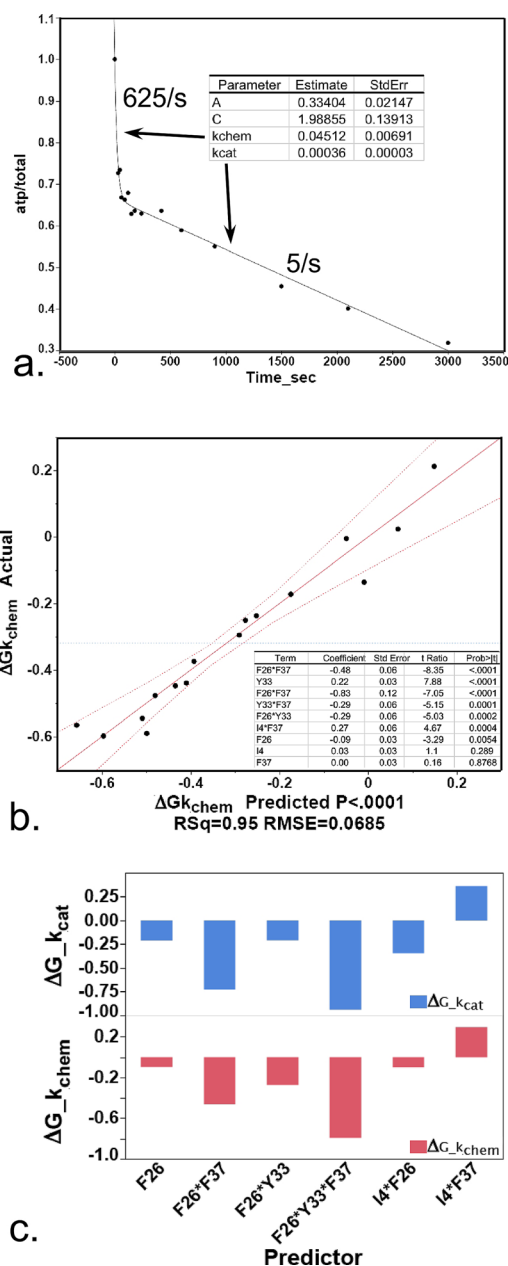


FIG. 12. (a) Measurement of the rate of the chemical step of TrpRS-catalyzed tryptophan activation by active-site titration.<sup>60,89</sup> The fitted line,  $y = A + \exp(-k_{\text{chem}}*t) - k_{\text{cat}}*t + C$ , yields estimates for the first-order rate of the chemical step,  $k_{\text{chem}}$  and the linear rate corresponding to product release and turnover. The actual turnover number, 5/s, and the ratio  $k_{\text{chem}}/k_{\text{cat}}$  furnishes an estimate of the concentration independent rate of the chemical step, which is 625 times faster than the steady-state rate, and is on the order of a few milliseconds. (b) Regression model for  $\Delta G_{k_{\text{chem}}}$  values for WT and the 15 D1 Switch mutant enzymes show that the relevant dependent parameter in pre-steady state active-site titration experiments is the intrinsic rate of the chemical step. The inset table shows estimated energetic contributions of four main effects, i.e., from each D1 site, together with five two-way interactions, in order of decreasing statistical significance.  $R^2 = 0.95$  and  $P(F) < 0.0001$ . (c) Histograms of regression coefficients for the contributions of each factor for  $\Delta G_{k_{\text{cat}}}$  and  $\Delta G_{k_{\text{chem}}}$ .

optimizing the efficiency of free energy transfer. Controlling the pre-organization of the catalytic metal ion can thus be seen as part of an orchestrated sequence that couples these differential binding changes associated with ATP binding, hydrolysis, and product release, to the domain configuration, ensuring that ATP hydrolysis occurs if and only if the conformation changes.<sup>20</sup>

Our recent demonstration that the equilibrium for this catalytic domain movement is driven only by the release of pyrophosphate<sup>41</sup> appears to be an additional mechanism ensuring that catalysis is coupled to (subsequent) domain movement, presumably regulated by tRNA binding. Using PPi release to change the overall conformational equilibrium resembles the behavior observed for myosin, for which the power stroke is initiated by orthophosphate release, whereas the release of adenosine-diphosphate happens only as the power stroke proceeds toward the rigor state.<sup>77-79</sup> As both orthophosphate and pyrophosphate are much more strongly solvated by water than is ATP,<sup>80</sup> an early release of the phosphoryl group may be more widely used to orchestrate the vectorial domain motion in transducing enzymes.

Serial conditionality furnishes qualitatively new insights into the mechanism of chemical free energy transduction and vectorial processes. Perhaps the clearest discussion of these underlying issues is that by Jencks,<sup>13,14</sup> who outlined how sequential changes in ligand-binding affinity lead to profound changes in the overall free energy of ATP hydrolysis, and recognized the necessity for understanding what he called the “coupling rules.” As the structural landscape of transducing enzymes was as yet unknown, Jencks was mute about what mechanisms actually linked conformational changes to catalysis, and this issue remains unresolved.<sup>81</sup> Identifying<sup>24,40</sup> and mutating<sup>29</sup> key residues associated with the conformational transition state(s), together with modular thermodynamic cycles<sup>44</sup> to measure energetic coupling between domains, now enhance the understanding of how idiosyncratic structural and mechanistic details implement the coupling rules. These methods should also be useful in studying a large number of other transducing NTPases.

The accompanying paper<sup>82</sup> describes an improved analysis of TrpRS conformational transition states whose computationally derived thermodynamic and kinetic parameters correlate strongly with structures and kinetic data for all sixteen D1 switch variants, unifying and further validating the results summarized here.

## METHODS

### Identifying potential couples

A full-length enzyme has almost innumerable separate interactions. Consider only  $2^N$  experiments obtained by constructing all possible combinations of mutations for one type of substitution per residue in TrpRS (328 residues). This amounts to  $2^{328} = e^{1900}$  different variants. Clearly, the number of variants must be massively reduced in order even to begin to evaluate long-range energetic coupling in an enzyme. Fortunately, several aspects of domain motion simplify the problem if sufficient structural information is available for the conformational variations along the reaction path.

Many conformational changes consist of rigid-body motions. As these rigid bodies preserve their internal structure, amino acid residues within them can be ignored to a first approximation in designing the combinatorial cycles. Delaunay tessellation discriminates the invariant tertiary interactions from those that change during the structural reaction profile. The remaining, dynamic, interactions constitute a far smaller reservoir of candidates for permutation studies. Rigid body changes therefore entail either hinge-bending or shear, both of which impact small numbers of residues, reducing the number to be considered for combinatorial point mutagenesis.

Algorithms developed for multistate protein design<sup>57-59</sup> furnish a second, complementary source of useful, residue-specific estimates of free energies associated with the domain motion when crystal structures are available for multiple states. Multi-state design identifies the side chain mutations that alter the relative conformational energies. Dynamic residues identified as potential sites of mutation to alter conformational equilibria form an overlapping, but distinct set of residues.

The Venn diagram of sidechains identified by the Delaunay SNAPP (Simplicial Network Analysis of Protein Packing)<sup>54-56</sup> and Rosetta energy function tests (Fig. 5) reduces the number of mutants to a manageable number.

Finally, the Rosetta energy function identifies the specific point mutants that have the greatest impact. This contribution cannot be ignored, because it greatly reduces the number of possible mutations at each site to a single amino acid. The process described here thus identifies specific mutations to a sufficiently small number of sites, to compose a sensible set of combinatorial experiments.

### Regression modeling simplifies the data analysis

Regression analysis of experimental results recorded in a factorial design is implemented in most statistics programs. We use JMP statistical discovery software (JMP),<sup>83</sup> which combines a wide range of state-of-the-art statistical and presentation modules from the SAS system with a very user-friendly interface. Linear models require the use of free energies, which are additive. Thus, all rates and equilibrium constants must be converted first into Gibbs energies. Then, the Gibbs energies contributing either to  $k_{cat}$ ,  $K_M$ , or  $k_{cat}/K_M$  can be expressed as the linear combination (1)<sup>84</sup>

$$\begin{aligned} \Delta G_{obs} = & C + \sum_{main} \Delta G_i^* [i] + \sum_{2-way} \Delta G_{ij}^* [i]^* [j] + \sum_{3-way} \Delta G_{ijk}^* [i]^* [j]^* [k] \\ & + \Delta G_{ijkl}^* [i]^* [j]^* [k]^* [l] + \sum_{main} \Delta G_{i-M}^* [i]^* [M] + \sum_{2-way} \Delta G_{ij-M}^* [i]^* [j]^* [M] \\ & + \sum_{3-way} \Delta G_{ijk-M}^* [i]^* [j]^* [k]^* [M] + \Delta G_{ijkl-M}^* [i]^* [j]^* [k]^* [l]^* [M], \end{aligned} \quad (1)$$

where  $i-l$  are either 1 or  $-1$  depending on whether the residue at position  $i$  is wild-type or mutant, and  $M$  is 1 or  $-1$  depending on whether the catalytic metal is  $Mg^{2+}$  or  $Mn^{2+}$ .

A decisive advantage of this approach to the analysis is that coefficients for main effects and interaction terms furnish intrinsic energetic contributions and coupling free energies from the cross terms in units of kcal/mole. A detailed example, clarifying various questions about the method as an alternative to manual calculation of all quantities was described previously<sup>51</sup> is discussed in the supplementary material for Ref. 84.

### ACKNOWLEDGMENTS

We are grateful to B. Kuhlman, J. Hermans, J. Das, R. Wolfenden, and N. Dokholyan for advice and assistance. This work was supported by NIGMS (78227 and 40906).

- <sup>1</sup>T. Kern, P. Schanda, and B. Brutscher, "Sensitivity-enhanced IPAP-SOFAST-HMQC for fast-pulsing 2D NMR with reduced radiofrequency load," *J. Magn. Reson.* **190**, 333–338 (2008).
- <sup>2</sup>M. Deniziak, C. Sauter, H. D. Becker, C. A. Paulus, R. Giegé, and D. Kern, "Deinococcus glutaminyl-tRNA synthetase is a chimera between proteins from an ancient and the modern pathways of aminoacyl-tRNA formation," *Nucleic Acids Res.* **35**, 1421–1431 (2007).
- <sup>3</sup>P. E. Wright and H. J. Dyson, "Linking folding and binding," *Curr. Op. Struct. Biol.* **19**, 31–38 (2009).
- <sup>4</sup>D. D. Boehr, D. McElheny, H. J. Dyson, and P. E. Wright, "The dynamic energy landscape of dihydrofolate reductase catalysis," *Science* **313**, 1639–1641 (2006).
- <sup>5</sup>E. Rennella, R. Huang, A. Velyvis, and L. E. Kay, "<sup>13</sup>CHD2-CEST NMR spectroscopy provides an avenue for studies of conformational exchange in high molecular weight proteins," *J. Biomol. NMR* **63**, 187–199 (2015).
- <sup>6</sup>P. Vallurupalli, G. Bouvignies, and L. E. Kay, "Studying 'invisible' excited protein states in slow exchange with a major state conformation," *J. Am. Chem. Soc.* **134**, 8148–8161 (2012).
- <sup>7</sup>K. M. Flaherty, S. M. Wilbanks, C. DeLuca-Flaherty, and D. B. McKay, "Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity," *J. Biol. Chem.* **269**, 12899–12907 (1994).
- <sup>8</sup>D. McElheny, J. R. Schnell, J. C. Lansing, H. J. Dyson, and P. E. Wright, "Defining the role of active-site loop fluctuations in dihydrofolate reductase catalysis," *Proc. Nat. Acad. Sci. U. S. A.* **102**, 5032–5037 (2005).
- <sup>9</sup>J. R. Schnell, H. J. Dyson, and P. E. Wright, "Structure, dynamics, and catalytic function of dihydrofolate reductase," *Annu Rev. Biophys. Biomol. Struct.* **33**, 119–140 (2004).
- <sup>10</sup>E. Z. Eisenmesser, O. Millet, W. Labeikovsky, D. M. Korzhnev, M. Wolf-Watz, D. A. Bosco, J. J. Skalicky, L. E. Kay, and D. Kern, "Intrinsic dynamics of an enzyme underlies catalysis," *Nature* **438**, 117–137 (2005).
- <sup>11</sup>B. Schuwirth, M. A. Borovinskaya, C. W. Hau, W. Zhang, A. Vila-Sanjurjo, J. Holton, and J. H. Doudna, "Structures of the bacterial ribosome at 3.5 Å resolution," *Science* **310**, 827–833 (2005).
- <sup>12</sup>A. Warshel and R. P. Bora, "Perspective: Defining and quantifying the role of dynamics in enzyme catalysis," *J. Chem. Phys.* **144**, 180901 (2016).
- <sup>13</sup>W. P. Jencks, "Reaction mechanisms, catalysis, and movement," *Protein Sci.* **3**, 2459–2464 (1994).
- <sup>14</sup>W. P. Jencks, "Utilization of binding energy and coupling rules for active transport and other coupled vectorial processes," *Method Enzymol.* **171**, 145–164 (1989).

- <sup>15</sup>W. P. Jencks, "The utilization of binding energy in coupled vectorial processes," *Adv. Enzymol. Relat. Areas Mol. Biol.* **51**, 75–106 (1980).
- <sup>16</sup>R. D. Astumian, "Stochastic conformational pumping: A mechanism for free-energy transduction by molecules," *Annu. Rev. Biophys.* **40**, 289–313 (2011).
- <sup>17</sup>R. D. Astumian, "Thermodynamics and kinetics of molecular motors," *Biophys. J.* **98**, 2401–2409 (2010).
- <sup>18</sup>V. Ovchinnikov, B. L. Trouta, and M. Karplus, "Mechanical coupling in myosin V: A simulation study," *J. Mol. Biol.* **395**, 815–835 (2010).
- <sup>19</sup>E. Sadvovsky and O. Yifrach, "Principles underlying energetic coupling along an allosteric communication trajectory of a voltage-activated K channel," *Proc. Nat. Acad. Sci. U. S. A.* **104**(50), 19813–19818 (2007).
- <sup>20</sup>V. Weinreb, L. Li, and C. W. Carter, Jr., *Structure* **20**, 128–138 (2012).
- <sup>21</sup>A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu, and M. H. M. Olsson, "Electrostatic basis for enzyme catalysis," *Chem. Rev.* **106**, 3210–3235 (2006).
- <sup>22</sup>J. Åqvist, K. Kolmodin, J. Florian, and A. Warshel, "Mechanistic alternatives in phosphate monoester hydrolysis: what conclusions can be drawn from available experimental data?," *Chem. Biol.* **6**, R71–R80 (1999).
- <sup>23</sup>J. Åqvist and A. Warshel, "Free energy relationships in metalloenzyme-catalyzed reactions. calculations of the effects of metal ion substitutions in staphylococcal nuclease," *J. Am. Chem. Soc.* **112**, 2860–2868 (1990).
- <sup>24</sup>M. Kapustina, V. Weinreb, L. Li, B. Kuhlman, and C. W. Carter, Jr., "A conformational transition state accompanies tryptophan activation by *B. stearotherophilus* tryptophanyl-tRNA synthetase," *Structure* **15**, 1272–1284 (2007).
- <sup>25</sup>V. Weinreb and C. W. Carter, Jr., "Mg<sup>2+</sup>-free *B. stearotherophilus* tryptophanyl-tRNA synthetase activates tryptophan with a major fraction of the overall rate enhancement," *J. Am. Chem. Soc.* **130**, 1488–1494 (2008).
- <sup>26</sup>L. Martinez, M. Jimenez-Rodriguez, K. Gonzalez-Rivera, T. Williams, L. Li, V. Weinreb, S. Niranj Chandrasekaran, M. Collier, X. Ambroggio, B. Kuhlman, O. Erdogan, and C. W. J. Carter, "Functional class I and II amino acid activating enzymes can be coded by opposite strands of the same gene," *J. Biol. Chem.* **290**(32), 19710–19725 (2015).
- <sup>27</sup>P. J. Sapienza, L. Li, T. Williams, A. L. Lee, and C. W. Carter, Jr., "An ancestral tryptophanyl-tRNA synthetase precursor achieves high catalytic rate enhancement without ordered ground-state tertiary structures," *ACS Chem. Biol.* **11**, 1661–1668 (2016).
- <sup>28</sup>B. R. Prasad and A. Warshel, "Prechemistry versus preorganization in DNA replication fidelity," *Proteins: Struct., Funct. Bioinf.* **79**, 2900–2919 (2011).
- <sup>29</sup>V. Weinreb, L. Li, L. S. Kaguni, C. L. Campbell, and C. W. Carter, Jr., "Mg<sup>2+</sup>-assisted catalysis by *B. stearotherophilus* TrpRS is promoted by allosteric effects," *Structure* **17**, 1–13 (2009).
- <sup>30</sup>V. Weinreb, L. Li, S. N. Chandrasekaran, P. Koehl, M. Delarue, and C. W. Carter, Jr., "Enhanced amino acid selection in fully-evolved tryptophanyl-tRNA Synthetase, relative to its urzyme, requires domain movement sensed by the D1 switch, a remote, dynamic packing motif," *J. Biol. Chem.* **289**, 4367–4376 (2014).
- <sup>31</sup>V. A. Ilyin, B. Temple, M. Hu, G. Li, Y. Yin, P. Vachette, and C. W. Carter, Jr., "2.9 Å crystal structure of ligand-free tryptophanyl-tRNA synthetase: Domain movements fragment the adenine nucleotide binding site," *Protein Sci.* **9**, 218–231 (2000).
- <sup>32</sup>P. Retailleau, X. Huang, Y. Yin, M. Hu, V. Weinreb, P. Vachette, C. Vornrhein, G. Bricogne, P. Roversi, V. Ilyin, and C. W. Carter, Jr., "Interconversion of ATP binding and conformational free energies by Tryptophanyl-tRNA Synthetase: Structures of ATP bound to open and closed, pre-transition conformations," *J. Mol. Biol.* **325**, 39–63 (2003).
- <sup>33</sup>P. Laowanapiban, M. Kapustina, C. Vornrhein, M. Delarue, P. Koehl, and C. W. Carter, Jr., "Independent saturation of three TrpRS subsites generates a partially-assembled state similar to those observed in molecular simulations," *Proc. Natl. Acad. Sci. U. S. A.* **106**, 1790–1795 (2009).
- <sup>34</sup>P. Retailleau, V. Weinreb, M. Hu, and C. W. Carter, Jr., "Crystal structure of tryptophanyl-tRNA synthetase complexed with adenosine-5' tetraphosphate: Evidence for distributed use of catalytic binding energy in amino acid activation by class I aminoacyl-tRNA synthetases," *J. Mol. Biol.* **369**, 108–128 (2007).
- <sup>35</sup>S. Doublé, G. Bricogne, C. J. Gilmore, and C. W. Carter, Jr., "Tryptophanyl-tRNA synthetase crystal structure reveals an unexpected homology to Tyrosyl-tRNA synthetase," *Structure* **3**(1), 17–31 (1995).
- <sup>36</sup>P. Retailleau, Y. Yin, M. Hu, J. M. Roach, G. Bricogne, C. Vornrhein, P. Roversi, E. Blanc, R. M. Sweet, and C. W. Carter, Jr., "High resolution experimental phases for Tryptophanyl-tRNA synthetase (TrpRS) complexed with Tryptophanyl-5'AMP," *Acta Crystallogr.* **D57**, 1595–1608 (2001).
- <sup>37</sup>Y. Yin, Ph.D. thesis, University of North Carolina at Chapel Hill, 1995.
- <sup>38</sup>C. W. Carter, Jr., V. Ilyin, Y. Yin, X. Huang, and P. Retailleau, paper presented at the Using Crystallography to Understand Enzyme Mechanism, St. Paul, 2001.
- <sup>39</sup>J. P. Abrahams, A. G. W. Leslie, R. Lutter, and J. E. Walker, "Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria," *Nature* **370**, 621–628 (1994).
- <sup>40</sup>M. Kapustina and C. W. Carter, Jr., "Computational studies of tryptophanyl-tRNA synthetase ligand binding and conformational stability," *J. Mol. Biol.* **362**, 1159–1180 (2006).
- <sup>41</sup>S. N. Chandrasekaran, J. Das, N. V. Dokholyan, and C. W. Carter, Jr., "A modified PATH algorithm rapidly generates transition states comparable to those found by other well established algorithms," *Struct. Dyn.* **3**, 012101 (2016).
- <sup>42</sup>S. Cammer and C. W. Carter, Jr., "Six Rossmannoid folds, including the class I aminoacyl-tRNA synthetases, share a partial core with the anticodon-binding domain of a class II aminoacyl-tRNA synthetase," *Bioinformatics* **26**(6), 709–714 (2010).
- <sup>43</sup>M. Budiman, M. H. Knaggs, J. S. Fetrow, and R. W. Alexander, "Using molecular dynamics to map interaction networks in an aminoacyl-tRNA synthetase," *Proteins: Struct., Funct. Bioinf.* **68**, 670–689 (2007).
- <sup>44</sup>L. Li and C. W. Carter, Jr., "Full implementation of the genetic code by tryptophanyl-tRNA synthetase requires intermolecular coupling," *J. Biol. Chem.* **288**, 34736–34745 (2013).
- <sup>45</sup>N. Goldenfield, in *This Idea Must Die*, edited by J. Brockman (The Edge Foundation/HarperCollins, New York, 2015), pp. 25–28.
- <sup>46</sup>R. Jaffe, "Where does the proton really get its spin?," *Phys. Today* **48**(9), 24–30 (1995).
- <sup>47</sup>W. P. Jencks, "Binding energy, specificity, and enzymatic catalysis: The circe effect," *Adv. Enzymol. Relat. Areas Mol. Biol.* **43**, 219–410 (1975).

- <sup>48</sup>W. P. Jencks, "On the attribution and additivity of binding energies," *Proc. Nat. Acad. Sci. U. S. A.* **78**, 4046–4050 (1981).
- <sup>49</sup>E. A. First and A. R. Fersht, "Mutational and kinetic analysis of a mobile loop in tyrosyl-tRNA synthetase," *Biochemistry* **32**(49), 13658–13663 (1993).
- <sup>50</sup>E. A. First and A. R. Fersht, "Analysis of the role of the KMSKS loop in the catalytic mechanism of the tyrosyl-tRNA synthetase using multimutant cycles," *Biochemistry* **34**(15), 5030–5043. (1995).
- <sup>51</sup>A. Horovitz and A. R. Fersht, "Strategy for analyzing the co-operativity of intramolecular interactions in peptides and proteins," *J. Mol. Biol.* **214**, 613–617 (1990).
- <sup>52</sup>C. W. Carter, Jr., "High-dimensional mutant and modular thermodynamic cycles, molecular switching, and free energy transduction," *Annu. Rev. Biophys.* (in press).
- <sup>53</sup>R. B. Stockbridge and R. Wolfenden, "The intrinsic reactivity of ATP and the catalytic proficiencies of kinases acting on glucose, N-acetylgalactosamine, and homoserine: A thermodynamic analysis," *J. Biol. Chem.* **284**(34), 22747–22757 (2009).
- <sup>54</sup>C. W. Carter, Jr., B. C. LeFebvre, S. A. Cammer, A. Tropsha, and M. H. Edgell, "Four-body potentials reveal protein-specific correlations to stability changes caused by hydrophobic core mutations," *J. Mol. Biol.* **311**(4), 625–638. (2001).
- <sup>55</sup>S. Cammer, C. W. Carter, Jr., and A. Tropsha, in *Lecture Notes in Computational Science and Engineering*, edited by T. Schlick and H. H. Gan (Springer Verlag, New York, 2002), Vol. 24, pp. 477–494.
- <sup>56</sup>A. Tropsha, C. W. J. Carter, S. Cammer, and I. I. Vaisman, "Simplicial neighborhood analysis of protein packing (SNAPP): A computational geometry approach to studying proteins," *Methods Enzymol.* **374**, 509–544 (2003).
- <sup>57</sup>Y. Liu and B. Kuhlman, "RosettaDesign server for protein design," *Nucleic Acids Res.* **34**, 235–238 (2006).
- <sup>58</sup>B. Kuhlman and D. Baker, "Exploring folding free energy landscapes using computational protein design," *Curr. Opin. Struct. Biol.* **14**, 89–95 (2004).
- <sup>59</sup>A. Leaver-Fay, R. Jacak, P. B. Stranges, and B. Kuhlman, "A generic program for multistate protein design," *PLoS One* **6**(7), e20937 (2011).
- <sup>60</sup>A. R. Fersht, J. S. Ashford, C. J. Bruton, R. Jakes, G. L. E. Koch, and B. S. Hartley, "Active site titration and aminoacyl adenylate binding stoichiometry of aminoacyl-tRNA synthetases," *Biochemistry* **14**(1), 1–4 (1975).
- <sup>61</sup>F. Sundén, A. Peck, J. Salzman, S. Ressler, and D. Herschlag, "Extensive site-directed mutagenesis reveals interconnected functional units in the alkaline phosphatase active site," *eLife* **4**, e06181 (2015).
- <sup>62</sup>C. W. Carter, Jr., "Urzymology: Experimental access to a key transition in the appearance of enzymes," *J. Biol. Chem.* **289**(44), 30213–30220 (2014).
- <sup>63</sup>L. Li, V. Weinreb, C. Francklyn, and C. W. Carter, Jr., "Histidyl-tRNA synthetase urzymes: Class I and II Aminoacyl-tRNA synthetase urzymes have comparable catalytic activities for cognate amino acid activation," *J. Biol. Chem.* **286**, 10387–10395 (2011).
- <sup>64</sup>Y. Pham, B. Kuhlman, G. L. Butterfoss, H. Hu, V. Weinreb, and C. W. Carter, Jr., "Tryptophanyl-tRNA synthetase urzyme: A model to recapitulate molecular evolution and investigate intramolecular complementation," *J. Biol. Chem.* **285**, 38590–38601 (2010).
- <sup>65</sup>G. E. P. Box, W. G. Hunter, and J. S. Hunter, *Statistics for Experimenters* (Wiley Interscience, New York, 1978).
- <sup>66</sup>Y. Yin and C. W. Carter, Jr., "Quantitative analysis in the characterization and optimization of protein crystal growth," *Acta Crystallogr. D* **50**, 572–590 (1994).
- <sup>67</sup>E. A. First and A. R. Fersht, "Involvement of threonine 234 in catalysis of tyrosyl adenylate formation by tyrosyl-tRNA synthetase," *Biochemistry* **32**(49), 13644–13650 (1993).
- <sup>68</sup>A. R. Fersht, J. W. Knill Jones, H. Bedouelle, and G. Winter, "Reconstruction by site-directed mutagenesis of the transition state for the activation of tyrosine by the tyrosyl-tRNA synthetase: a mobile loop envelopes the transition state in an induced-fit mechanism," *Biochemistry* **27**(5), 1581–1587 (1988).
- <sup>69</sup>A. R. Fersht, "Dissection of the structure and activity of the tyrosyl-tRNA synthetase by site-directed mutagenesis," *Biochemistry* **26**(25), 8031–8037 (1987).
- <sup>70</sup>J. Franklin, P. Koehl, S. Doniach, and M. Delarue, "MinActionPath: Maximum likelihood trajectory for large-scale structural transitions in a coarse-grained locally harmonic energy landscape," *Nucl. Acids Res.* **35**, W477–W482 (2007).
- <sup>71</sup>D. Shirvanyants, F. Ding, D. Tsao, S. Ramachandran, and N. V. Dokholyan, "Discrete molecular dynamics: An efficient and versatile simulation method for fine protein characterization," *J. Phys. Chem. B* **116**, 8375–8382 (2012).
- <sup>72</sup>P. J. Malsom and F. J. Pinski, "Role of Ito's lemma in sampling pinned diffusion paths in the continuous-time limit," *Phys. Rev. E* **94**, 042131 (2016).
- <sup>73</sup>L. Onsager and S. Machlup, "Fluctuations and irreversible processes: Systems with kinetic energy," *Phys. Rev.* **91**(6), 1512–1515 (1953).
- <sup>74</sup>V. Ovchinnikov, M. Karplus, and E. Vanden-Eijnden, "Free energy of conformational transition paths in biomolecules: The string method and its application to myosin VI," *J. Chem. Phys.* **134**, 085103 (2011).
- <sup>75</sup>T. L. Williams and C. W. J. Carter, "Both ATP and Mg<sup>2+</sup> are required for high-affinity binding of indolmycin to human mitochondrial tryptophanyl-tRNA synthetase," *J. Biol. Chem.* (submitted).
- <sup>76</sup>T. Williams, W. Y. Yin, and C. W. Carter, Jr., "Selective inhibition of bacterial tryptophanyl-tRNA synthetases by indolmycin is mechanism based," *J. Biol. Chem.* **291**(1), 255–265 (2015).
- <sup>77</sup>R. W. Lymn, "Kinetic analysis of myosin and actomyosin ATPase," *Ann. Rev. Biophys. Bioeng.* **8**, 145–163 (1979).
- <sup>78</sup>R. W. Lymn and E. W. Taylor, "Transient state phosphate production in the hydrolysis of nucleoside triphosphates by myosin," *Biochemistry* **9**, 2975–2983 (1970).
- <sup>79</sup>R. W. Lymn and E. W. Taylor, "Mechanism of adenosine triphosphate hydrolysis by actomyosin," *Biochemistry* **10**, 4617–4624 (1971).
- <sup>80</sup>R. Wolfenden and R. Williams, "Affinities of phosphoric acids, esters and amides for solvent water," *J. Am. Chem. Soc.* **105**, 1028–1031 (1983).
- <sup>81</sup>C. Khosla, "Quo vadis, enzymology?," *Nat. Chem. Biol.* **11**, 438–441 (2015).
- <sup>82</sup>S. N. Chandrasekaran and C. W. J. Carter, "Adding torsional interaction terms to the Anisotropic Network Model improves the PATH performance, enabling detailed comparison with experimental rate data," *Struct. Dyn.* (in press).



<sup>83</sup>SAS Institute, "SAS," Cary, NC, 2007.

<sup>84</sup>V. Weinreb, L. Li, and C. W. Carter, Jr., *Structure* **20**, 128–138 (2012), see Supplementary Information.

<sup>85</sup>J.-C. Liao, S. Sun, D. Chandler, and G. Oster, "The conformational states of Mg•ATP in water," *Eur. Biophys. J.* **33**, 29–37 (2004).

<sup>86</sup>M. Kapustina, J. Hermans, and C. W. Carter, Jr., "Potential of mean force estimation of the relative magnitude of the effect of errors in molecular mechanics approximations," *J. Mol. Biol.* **362**, 1177–1180 (2006).

<sup>87</sup>V. Weinreb, B. Kuhlman, and C. W. Carter, Jr., in *American Crystallographic Association Annual Meeting*, edited by R. Angel (American Crystallographic Association, Chicago, 2010).

<sup>88</sup>V. Weinreb, L. Li, L. S. Kaguni, C. L. Campbell, and C. W. Carter, Jr., "Mg<sup>2+</sup>-assisted catalysis by *B. stearothermophilus* TrpRS is promoted by allosteric effects," *Structure* **17**, 952–964 (2009).

<sup>89</sup>C. S. Francklyn, E. A. First, J. J. Perona, and Y.-M. Hou, "Methods for kinetic and thermodynamic analysis of aminoacyl-tRNA synthetases," *Methods* **44**, 100–118 (2008).