THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 7, pp. 4367–4376, February 14, 2014

Enhanced Amino Acid Selection in Fully Evolved Tryptophanyl-tRNA Synthetase, Relative to Its Urzyme, Requires Domain Motion Sensed by the D1 Switch, a Remote Dynamic Packing Motif*[S]

Received for publication, November 27, 2013, and in revised form, December 27, 2013 Published, JBC Papers in Press, January 6, 2014, DOI 10.1074/jbc.M113.538660

Violetta Weinreb[‡], Li Li[‡], Srinivas Niranj Chandrasekaran[‡], Patrice Koehl[§], Marc Delarue[¶], and Charles W. Carter, Jr.[‡1]

From the [‡]Department of Biochemistry and Biophysics, CB 7260, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7260, the [§]Department of Computer Science, University of California, Davis, California 95616, and the [¶]Dynamique Structurale des Macromolecules, Unité de Mixte de Recherche 3528 du Centre National de la Recherche Scientifique, Pasteur Institute, 75015 Paris, France

Background: Amino acid selection by tryptophanyl-tRNA synthetase (TrpRS) requires intermodular coupling.
Results: Dynamic repacking of four side chains increases amino acid specificity 500-fold in contemporary TrpRS by reducing pocket size near the transition state.
Conclusion: An ancient tertiary packing motif not only activates the catalytic Mg^{2+} ion during catalysis, but also determines cognate amino acid specificity.
Significance: Allosteric enforcement of specificity increases robustness to mutation.

We previously showed (Li, L., and Carter, C. W., Jr. (2013) *J. Biol. Chem.* 288, 34736–34745) that increased specificity for tryptophan *versus* tyrosine by contemporary *Bacillus stearothermophilus* tryptophanyl-tRNA synthetase (TrpRS) over that of TrpRS Urzyme results entirely from coupling between the anticodon-binding domain and an insertion into the Rossmann-fold known as Connecting Peptide 1. We show that this effect is closely related to a long range catalytic effect, in which side chain repacking in a region called the D1 Switch, accounts fully for the entire catalytic contribution of the catalytic Mg^{2+} ion. We report intrinsic and higher order interaction effects on the specificity ratio, $(k_{cat}/K_m)_{Trp}/(k_{cat}/K_m)_{Tyr}$, of 15 combinatorial mutants from a previous study (Weinreb, V., Li, L., and Carter, C. W., Jr. (2012) *Structure* 20, 128–138) of the catalytic role of the D1 Switch. Unexpectedly, the same four-way interaction both activates catalytic assist by Mg^{2+} ion and contributes −4.4 kcal/mol to the free energy of the specificity ratio. A minimum action path computed for the induced-fit and catalytic conformation changes shows that repacking of the four residues precedes a decrease in the volume of the tryptophan-binding pocket. We suggest that previous efforts to alter amino acid specificities of TrpRS and glutaminyl-tRNA synthetase (GlnRS) by mutagenesis without extensive, modular substitution failed because mutations were incompatible with interdomain motions required for catalysis.

A major unplumbed wonder of protein structure is the amazing extent to which different parts of a protein or enzyme respond appropriately to what happens elsewhere. The nearly simultaneous behavior of multiple parts is called cooperativity, and is a quintessential enabling characteristic that makes proteins such flexible and powerful catalysts. The drive to understand what "cooperativity" actually means, starting with the work of Perutz (1, 2) on the allosteric behavior of hemoglobin, has driven the greatest advances in enzymology and protein science.

Allostery, and linkage are closely related, but distinct concepts. *Allostery* refers to the intramolecular communication responsible for cooperativity; *linkage* refers to the detailed energetic coupling underlying such mechanisms. Non-additivity of perturbations to linked modules or residues, A and B, measures the energetic coupling of the *interaction*, A × B.

Bacillus stearothermophilus tryptophanyl-tRNA synthetase (TrpRS)[2] has recently opened important new windows on cooperative behavior. We began these studies hoping to identify specific mechanisms responsible for the very large non-additivity (~-6.0 kcal/mol) with which Mg^{2+} accelerates TrpRS-catalyzed tryptophan activation. The fact that the metal accelerates tryptophan activation 10^5-fold in the presence of TrpRS, but less than 10-fold in water (3), implies that *metal-protein coupling* is responsible for virtually the entire catalytic effect of Mg^{2+}. Yet, the catalytic metal ion interacts only with ATP and water molecules, and is not bound by any protein residues.

Combinatorial mutagenesis of active-site lysine residues physically coupled to the metal in crystal structures showed, by elimination, that this effect must arise from outside the active

* This work was supported, in whole or in part, by National Institutes of Health Grant 40906 from the NIGMS.
[S] This article contains supplemental Movie S1.
[1] To whom correspondence should be addressed. Tel.: 919-966-3263; Fax: 919-966-2852; E-mail: carter@med.unc.edu.

[2] The abbreviations used are: TrpRS, tryptophanyl-tRNA synthetase; GlnRS, glutaminyl-tRNA synthetase; PheRS, phenylalanyl-tRNA synthetase; ABD, anticodon-binding domain; CP1, connecting peptide 1; PreTS, pre-transition state; aaRS, aminoacyl-tRNA synthetase.

site (4). Instead, we identified a long range allosteric influence (5) from a widely distributed and highly conserved core-packing motif (6) remote from the active site that accounted entirely for the catalytic assist by Mg^{2+}. This motif is the most extensive of four discrete locations where Delaunay tessellation patterns change during TrpRS catalysis, suggesting the name "D1 Switch" (7). Repacking within this motif mediates shear developed by relative movement of the essentially rigid Rossmann-fold and anticodon-binding domains (ABD) during catalysis.

We recently showed that all three separate TrpRS functions: catalysis of amino acid activation, specific amino acid recognition, and tRNA acylation, depend in quite similar ways on dynamic interactions between the active site, the ABD, and an insertion into the Rossmann-fold called Connecting Peptide 1 (CP1). We restored the CP1 and ABD modules individually to the TrpRS Urzyme, a functional representation of ancestral Class I aaRS, showing that the two restored domains are linked by a quantitatively similar coupling energy (about −6 kcal/mol) in both selection of tryptophan *versus* tyrosine and tRNA aminoacylation (8).

Previously (9) we had attempted to change the specificity ratio to favor activation of tyrosine using mutagenesis of residues whose direct interaction with substrate tryptophan suggested that they were responsible for discriminating against tyrosine. That effort failed: mutations weakened catalytic proficiency, without favoring tyrosine activation.

Here we show instead that combinatorial mutagenesis of D1 Switch residues affects the specificity ratio, $(k_{cat}/K_m)_{Trp}/(k_{cat}/K_m)_{Tyr}$, sufficiently to account for the increased specificity of full-length TrpRS. Moreover, we illustrate the sequential structural changes involved in long range coupling by constructing a minimum action path (10) to connect the three crystal structures that represent stable states along the path. Repacking of the four mutated D1 Switch residues occurs twice, coincident with the transition states for induced-fit and catalysis. Each change is followed by reductions in the volume of the tryptophan binding pocket, consistent with the experimental specificity data.

The quantitative agreement of the simulated most probable path with the modular (8) and point mutant coupling energies extends our analysis of cooperativity, allostery, and linkage in TrpRS enzymology to unprecedented levels of detail. As long range interaction between side chains is responsible for amino acid specificity, no obvious pattern of point mutations can actually change specificity unless they are also compatible with the necessary domain movement, making specific amino acid selection between competing canonical amino acids more robust to mutation.

EXPERIMENTAL PROCEDURES

Mutagenesis and Protein Purification—Mutations were constructed using GeneTailor (Invitrogen). The expression plasmid pET11a containing WT TrpRS was methylated by DNA methylase and used as template for mutagenic PCR. Primers were designed for both directions, with a 15–20-nucleotide overlap for efficient circularization. The PCR product was used to transform DH5α *Escherichia coli* cells, and plated on LB plates with ampicillin. Plasmids were sequenced to confirm the mutations.

Mutant proteins were expressed in *E. coli* BL21(DE3)pLysS, with both ampicillin and chloramphenicol. Simultaneous 3-liter cultures of two mutants were resuspended in 20 mM HEPES, 0.3 M NaCl, 10 mM 2-mercaptoethanol, 30 mM imidazole (pH 7.6) and rapidly frozen. The cell paste was sonicated five times for 20 s with a Fisher Scientific Sonic Dismembrator and cleared by centrifugation at 27,800 × *g* for 30 min at 4 °C. Nickel-nitrilotriacetic acid beads (HisPur Ni-NTA Resin, Thermo Scientific; 1 ml) were added to the supernatant, followed by 1 h of shaking at 4 °C. After centrifugation the beads were washed with 20 ml of lysis buffer and shaken 10 min at 4 °C. Bound enzyme was eluted three times with lysis buffer plus 0.3 M imidazole. Purified mutant proteins were cleaved with a 1:10 ratio of tobacco etch virus protease overnight at 4 °C, diluted to reduce the imidazole concentration below 30 mM, and passed over the nickel resin again to bind uncleaved protein and tobacco etch virus. The eluate was concentrated using an Amicon PM10 Ultra membrane and stored at −20 °C in 50% glycerol. All proteins were >95% homogeneous as judged by gel electrophoresis.

Active-site Titration—The reaction mixture contained 50 mM HEPES (pH 7.5), 10 mM $MgCl_2$ 100 mM KCl_2, 1 mM DTT, 10 μM [γ-^{32}P]ATP, 2.5 mM Trp, 10 units/ml of inorganic pyrophosphatase. We added 3 μM enzyme and incubated the mixture at 37 °C. At specific time points the reaction was quenched by adding 3 μl of the reaction mixture to 6 μl of 400 mM sodium acetate (pH 5.0), 0.1% SDS. 3 μl of this mixture was spotted on a thin layer chromatography plate. To separate [^{32}P]ATP from [^{32}P]PP_i we pre-soaked and eluted the TLC plates with 750 mM KH_2PO_4 (pH 3.5), 4 M urea buffer (11). The TLC plate was dried, exposed 15–30 min, depending on the variant, and scanned on a Typhoon scanner and quantified with ImageJ software. All assays included data points well into the steady state (linear with time).

PreTS crystal structures for ~8 mutant proteins including the quadruple mutant reveal very little structural variance from the native structure, Protein Data Bank code 1MAU (12). Active titers for the various mutant enzymes range from 0.3 to 0.95 with a mean value of 0.51 ± 0.17. The structural and enzymatic evidence that the mutant enzymes are properly folded is thus very strong.

Michaelis-Menten Kinetics—PP_i-exchange assays were done at 37 °C and initiated with 10 μl of enzyme to 190 μl of assay mixture: 0.1 M Tris-Cl, 0.01 M potassium fluoride, 5 mM $MgCl_2$, 2 mM ATP, 10 mM 2-mercaptoethanol (pH 8.0) plus 2 mM $^{32}PP_i$ at a specific radioactivity between 1×10^5 and 2×10^5 cpm/mol. Varying enzyme concentrations (100–400 μM) and incubation times (15 min for Trp, 20 min for Tyr) were used, depending on the activity level, which was determined by range finding for each variant. Michaelis-Menten kinetics were examined by varying the amino acid concentrations (0.01, 0.05, 0.5, 1.0, 2.0, 5.0, 10.0, 30, 60, 90, and 150 μM tryptophan and 0.005, 0.01, 0.05, 0.1, 0.15, 0.25, 0.5, 0.75, 1.0, 1.5, and 2.5 mM tyrosine). Assays were performed on 96-well plates with a Beckman-Coulter Biomek 3000 automatic liquid-handling work station, using Whatman filter plates and a Promega Vacman vacuum manifold for filtrations. [^{32}P]ATP was eluted from charcoal with pyridine as described (19). For metal substitution assays the mixtures were made without Mg^{2+} and treated with Chelex

100 for 30 min at 4 °C to remove trace metals and supplemented with 0.9 mM $MnCl_2$. All experiments were repeated two or more times with four replicates, to improve the accuracy of regression coefficients.

Multidimensional Thermodynamic Cycles—The formal similarity between thermodynamic cycles (13) and factorial experimental design (14) can help clarify how this approach works. Briefly, as in a full factorial experiment, a multidimensional thermodynamic cycle measures the impact of each perturbation from a set in the context of all other perturbations in that set. This means that at least as many protein variants are examined as there are vertices in the corresponding thermodynamic cycle. In general, many or all of the variants are examined multiple times to estimate the experimental variance. For the fifth-dimensional thermodynamic cycle in Ref. 5, each of four sites was mutated to a single other amino acid, and the resulting 15 variants were assayed multiple times with both Mg^{2+} and Mn^{2+}. We used the same mutants here.

Key to the relationship between these experimental results and the coupling energies is the *experimental design matrix*. Each experiment corresponds to a row in this matrix, giving the state of each perturbation and its higher-order interactions. For a double mutant cycle for sites A and B, the row corresponding to variant A has the form [0, 1, 0] because variant A is a mutant at site A (labeled 0), wild type at site B (labeled 1), and has 0 also in the third column, which is the product of columns A and B and corresponds to the A × B interaction.

Without experimental data, contributions of wild-type residues and their interaction energies to system function are unknown. Multiple column vectors correspond to experimental values for dependent variables. Examples are ΔGk_{cat} and ΔGK_M from Michaelis-Menten kinetics. Free energies are more useful than rate and equilibrium constants, because they are additive, and hence can be used to construct linear models from the state columns of the design matrix. These additional column vectors are pre-pended to the design matrix.

The final element of the analysis is a column vector of coefficients, which describe how much of a given predictor is necessary for optimal agreement between the observed data and corresponding results calculated from the linear model. Premultiplying the coefficient vector by the design matrix gives a set of simultaneous linear equations specifying the calculated results. If the number of experiments is greater than the number of unknowns in these simultaneous equations, then the solution and error estimates are found by least squares.

The algebraic formulation of the previous paragraph should clarify that the highest-order mutant is but one data point of the many that need to be summed together to estimate the coefficients, emphasizing the crucial difference between the experimental result for any given variant, and the coefficient accorded to the corresponding intrinsic or interaction effect by the linear equations. Just as the calculated data results all must be summed for each coefficient, coefficients must be estimated from all experimental points. In particular, it is overwhelmingly tempting to infer the coupling energy for the highest order interaction from the observed behavior of the highest order mutant (*i.e.* the quadruple mutant in a four-way experiment). The two are, in fact, almost totally unrelated.

Minimum Action Path Simulation (10)—A most probable trajectory connecting ground state crystal structures of *B. stearothermophilus* tryptophanyl-tRNA synthetase during induced-fit and subsequent catalysis was simulated from elastic network models (10) using a version of the program PATH (developed by P. Koehl) that finds the path with minimum action that connects three states. Spring constants for elastic network models of all states were derived using the MinActionPath server.[3] Total time intervals allocated to observe transitions, and energy differences between successive states were determined from the convergence surface.[4]

The PATH program carries out a rapid search for the transition state conformation between crystallographic states by assuring structural, velocity, and energy continuity to minimize the total Onsager-Machlup action of the path. Minimum action therefore assures the most probable path.

The sequence of structures was computed from the normal modes of the initial and final states of the two separate transitions, which are each given 100 steps, for a total of 200 steps. Thus the midpoint determines the transition from the induced-fit domain movement to that which takes place during the catalytic step. The two domain motions can be identified by the direction in which the ABD moves: during the induced-fit transition, it moves counterclockwise in the plane of the animation, during catalysis it rotates outward toward the viewer.

Colors used in supplemental trajectory Movie S1 are as follows: the Urzyme is slate blue, CP1 is forest green, the ABD is wheat. The β-strand forming the bottom of the tryptophan-binding pocket is yellow and the specificity helix forming the top is orange, as in Fig. 2*A*. Side chains forming the tryptophan binding pocket are hot pink; D1 Switch residue Ile-4 is yellow, and residues Phe-26, Tyr-33, and Phe-37 are deep blue.

RESULTS

Modular TrpRS Construction and the Conformational Cycle—Crystallographic structures suggest that catalysis of amino acid activation entails three distinct TrpRS structural states: open, pre-transition state (PreTS), and products (Prod). Transition between the first two structures corresponds to induced-fit, between the second and third to catalysis, and the restorative step to product release stimulated by acyl-transfer to cognate $tRNA^{Trp}$. Superposition of the three representative crystal structures (Protein Data Bank codes 1MB2, 1MAU, and 1I6L) reveals that the conformational cycle (7, 12, 16, 17) is well approximated by rigid-body movements of the ABD relative to the active site (Fig. 1).

Long range catalytic activation of the Mg^{2+} ion by the D1 Switch (5) suggested that it mediates coupling between catalytic and ABD domains. The TrpRS Urzyme (18, 19) enabled us to examine this intrinsic modularity directly (8). The Urzyme contains a functional active site without either CP1 or the ABD. It therefore likely represents an ancient Class I aminoacyl-tRNA synthetase (aaRS). The ABD and CP1 are presumably more modern modules (Fig. 2) that afford a mechanism to couple

[3] M. Delarue, MinActionPath, unpublished data.
[4] S. N. Chandrasekaran, P. Koehl, M. Delarue, S. Doniach, and C. W. Carter, Jr., manuscript in preparation.

domain movement (Fig. 1) to structural changes within the active site (Fig. 2*A*). Energetic coupling (about −6 kcal/mol) between these two modules enhances all three TrpRS functions, Mg^{2+}-dependent rate acceleration, specificity, and $tRNA^{Trp}$ aminoacylation to approximately the same degree (8).

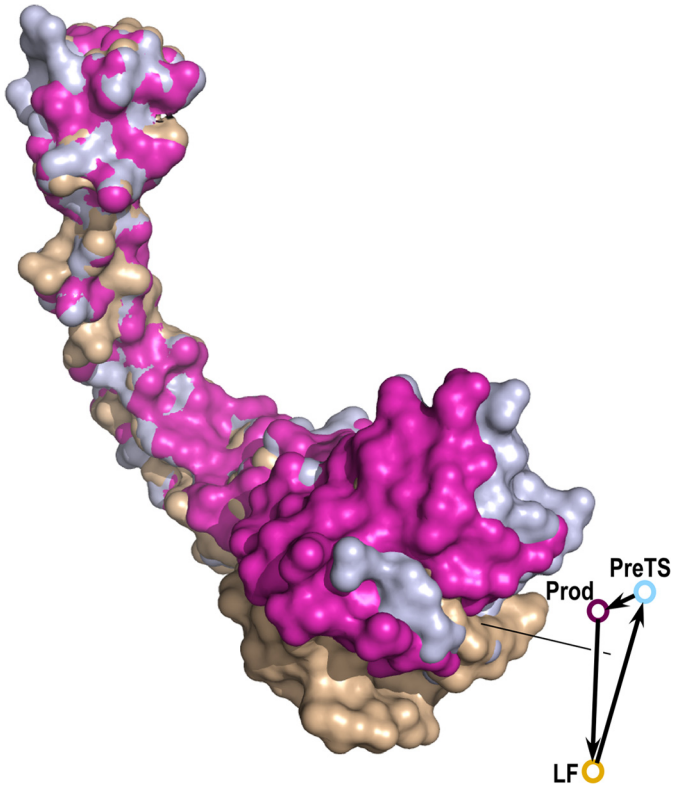

FIGURE 1. **Relative domain motion during the TrpRS catalytic cycle.** The Rossmann-fold and ABD move as rigid bodies relative to one another (12, 14, 17). Hinge bending and torsional motion of the anticodon-binding domain shown here reveal that the ABD assumes three distinct orientations relative to the Rossmann-fold domains that have been superimposed. These orientations are denoted LF (ligand-free, *wheat*), PreTS (pre-transition state; *sky blue*), and Products (*magenta*). Catalytic residues move with the ABD (11, 17).

The tryptophan binding site lies between two disjointed active site segments of the Urzyme; the N-terminal crossover and the specificity helix are separated by the CP1 module. Moreover, the CP1 module itself forms an exoskeleton that wraps around these two fragments, as suggested in Fig. 2*A*. Notably, the D1 Switch itself is a protoallosteric site contained entirely within the Urzyme (Fig. 2*B*). In this work, we extend the combinatorial mutagenesis of the D1 Switch and multidimensional thermodynamic cycles to the analysis of specificity during amino acid activation.

Multidimensional Thermodynamic Analysis of TrpRS Specificity—Michaelis-Menten parameters for all experiments are summarized in Tables 1 and 2, and illustrated graphically in Fig. 3. The three plots show the correlation between observed free energies associated with the three parameters, K_m, k_{cat}, and k_{cat}/K_m, and those calculated using appropriate linear combinations of mutant sites, which are the independent variables responsible for systematic variation. *F*-ratios for the three models are highly significant: ΔGK_M, $F = 6.4$ ($p < 0.0001$); ΔGk_{cat}, $F = 188$ ($p \ll 0.0001$); $\Delta Gk_{cat}/K_m$, $F = 33$ ($p \ll 0.0001$). Distributions of ΔGK_M for the two amino acids (Fig. 3*A*) are coextensive. However, plots for ΔGk_{cat} and $\Delta Gk_{cat}/K_m$ both clearly separate measurements for the two amino acids. They illustrate that specificity is driven largely by differences in k_{cat} for the two amino acids (Fig. 3, *B* and *C*). Tryptophan, on average, is associated with a k_{cat} value 220 times that observed with tyrosine ($\Delta G_W = -3.2$ kcal/mol).

The specificity ratios (Fig. 4) are computed from two different experimental determinations, amplifying their errors. For this reason, they are subject to higher experimental uncertainties and require extensive replication to achieve suitable estimates of variance. In view of this inherent experimental noise, we present in Tables 1 and 2 statistical summaries of the Michaelis-Menten parameters from which the specificity ratios in Table 3 were computed. The important question can be stated succinctly: have we replicated our measurements sufficiently to estimate confidently the most important coupling energies?

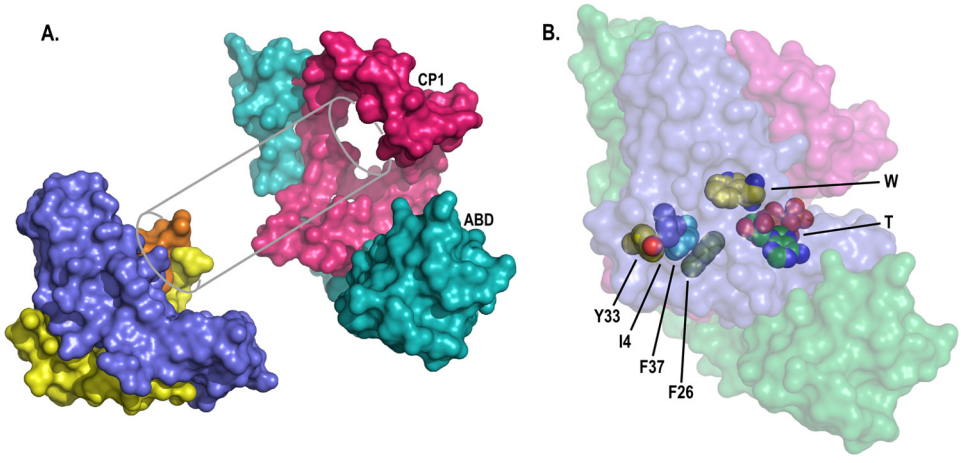

FIGURE 2. **Structural basis of allosteric behavior during amino acid activation by TrpRS.** *A,* modularity of the TrpRS monomer. The Urzyme contains three modules: the N-(*yellow*) and C (*blue*)-terminal crossover connections and the specificity helix (*orange*). It interacts snugly with both the CP1 (*magenta*) and anticodon-binding (*teal*) modules. Notably, the CP1 module circumscribes the Urzyme, constraining the relative movement of the specificity helix, relative to the N-terminal crossover connection of the Rossmann-fold, which determines the volume of the amino acid binding pocket. *B,* the main energetic communication within the Urzyme involves the D1 Switch. Four residues (Ile-4, Phe-26, Tyr-33, and Phe-37) were examined previously and are the subject of this work. The intact monomer is represented by the same colors as in *A*, except that the Urzyme is entirely *blue* and has a *transparent surface*, through which substrates tryptophan (*W*) and ATP (*T*), and D1 Switch residues are visible. The D1 Switch centroid is ~18 Å from the ATP α-phosphate and ~14 Å from the tryptophan α-carbon.

TABLE 1
Statistical properties of Michaelis-Menten parameters for TrpRS D1 Switch variants

	K_m	k_{cat}	k_{cat}/K_m	$\Delta Gk_{cat}/K_m$
	M	*s*		
Trp, Mg^{2+}				
Mean	9.43E−06	5.45	1.09E+06	−8.23
S.D.	8.45E−06	3.46	1.00E+06	−8.18
$Error_{Rel}$	0.90	0.64	0.92	0.05
Min	4.24E−07	0.86	1.60E+05	−7.09
Max	3.65E−05	18.49	3.40E+06	−8.90
Ratio/difference	86.08	21.51	21	−1.81
Trp, Mn^{2+}				
Mean	5.27E−05	4.88	1.64E+05	−7.11
S.D.	4.01E−05	2.08	1.55E+05	−7.08
$Error_{Rel}$	0.76	0.43	0.95	0.03
Min	3.50E−06	1.03	2.09E+04	−5.89
Max	1.63E−04	8.16	7.41E+05	−8.00
Ratio/difference	46.57	7.90	35	−2.11
Tyr, Mg^{2+}				
Mean	4.35E−05	0.02	3.88E+03	−4.89
S.D.	7.21E−05	0.020	8.53E+03	−5.36
$Error_{Rel}$	1.66	0.94	2.20	−0.47
Min	1.30E−06	0.004	1.44E+02	−2.94
Max	3.57E−04	0.08	4.76E+04	−6.38
Ratio/difference	274.85	22.57	331	−3.43
Tyr, Mn^{2+}				
Mean	6.05E−05	0.042	3.06E+03	−4.75
S.D.	8.30E−05	0.040	5.20E+03	−5.07
$Error_{Rel}$	1.37	0.95	1.70	−0.31
Min	1.50E−06	0.009	1.01E+01	−1.37
Max	3.00E−04	0.238	2.62E+04	−6.02
Ratio/difference	200.00	26.72	2595	−4.65

TABLE 2
Statistical properties of specificity constants k_{cat}/K_m(W)/k_{cat}/K_m(Y)

	Specificity ratio	S.D.	$Error_{rel}$	ΔG_{spec}	ΔG_{error}
Mean	926	830	0.8	3.5	−0.1
S.D.	1064	974	0.4	0.9	−0.6
Min	11	8	0.2	1.7	−1.0
Max	4192	3255	1.7	5.1	0.3
Ratio	388	430	9.6	3.0	1.3

The ultimate answer to this question lies in the statistics for the regression models. However, a simpler analysis of signal to noise from Tables 1 and 2 affords useful confirmation. Multiple determinations of the free energies for the apparent second-order rate constants in Table 1 are, on average, ~30 times the maximum relative error in multiple determinations of the same quantities. Similarly, the ratio of average specificity constant from Table 2 to the maximum relative error is 257; converted to free energies, this gives 8.6 for the range of free energies divided by the largest relative error for multiple estimates. These values imply that the range of specificities observed for the mutant proteins with Mg^{2+} and Mn^{2+} is substantially larger than the relative errors of the estimates, and hence that a significant signal can be extracted from the data, despite the increased relative error.

That conclusion is verified by regression modeling to estimate the strongest predictors of specificity from among the main and higher-order interaction effects (Table 4). Fifteen of 26 predictors have significant Student's *t* tests with *p* values <0.05. It should be noted that the *t*-statistic is the ratio of the coefficient to its standard error, so that *p* values are related to the impact of the experimental noise. In this way, the regression model apportions the experimental noise appropriately, thus highlighting the most important coefficients. Further evidence for the significance of these conclusions arises from the consistency of the implications summarized in the two following paragraphs with previously published results.

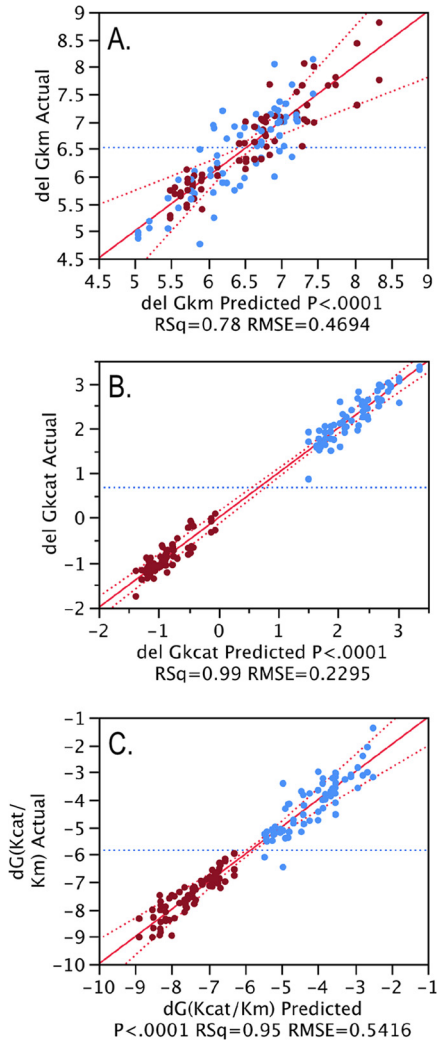

FIGURE 3. **Linear models for Gibbs-free energies associated with Michaelis-Menten parameters summarized in Tables 1 and 2.** Colors separate determinations for tryptophan (*red*) and tyrosine (*blue*). *A*, ΔGK_M, distributions for the two amino acids overlap extensively. *B*, ΔGk_{cat} distributions for the two amino acids are entirely disjoint, because of the overwhelming effect of the amino acid. *C*, $\Delta Gk_{cat}/K_m$ the distributions for ΔGk_{cat} are each broadened by the effect of ΔGK_M, but remain dominated by the effect of the amino acid. Specificity is driven largely by k_{cat}.

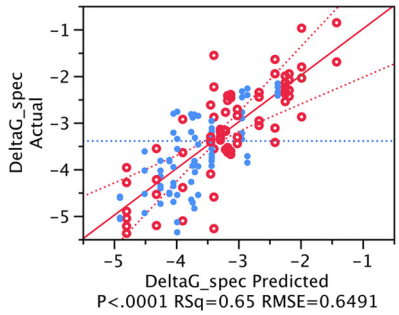

FIGURE 4. **Regression model for ΔG_{spec}, the free energy for selecting tryptophan *versus* tyrosine.** *Red open circles* are for variants with the I4V mutation, *blue closed circles* are for variants with wild-type Ile-4.

Two aspects of these coefficients stand out. First, by far the most important effect is the highest-order interaction between

TABLE 3
Specificity ratios of TrpRS variants with Mg^{2+} and Mn^{2+}

Variant	k_{cat}/K_m(W)/k_{cat}/K_m(Y); Mg^{2+}	Variant	k_{cat}/K_m(W)/k_{cat}/K_m(Y); Mn^{2+}
WT TrpRS	2192	WT TrpRS	1908
WT TrpRS	2289	WT TrpRS	587
WT TrpRS	4686	WT TrpRS	1996
WT TrpRS	4893	WT TrpRS	614
I4V	210	I4V	48
I4V	106	I4V	27
I4V	403	I4V	60
I4V	203	I4V	33
F26L	120	F26L	51
F26L	737	F26L	38
F26L	337	F26L	58
F26L	2065	F26L	43
Y33F	115	Y33F	62
Y33F	632	Y33F	50
Y33F	607	Y33F	278
Y33F	3346	Y33F	224
		Y33F	131
		Y33F	106
F37I	1629	F37I	211
F37I	859	F37I	188
F37I	605	F37I	296
F37I	2831	F37I	263
F37I	1492	F37I	543
F37I	1051	F37I	483
I4V-F26L	4790	I4V-F26L	260
I4V-F26L	3714	I4V-F26L	195
I4V-F26L	772	I4V-F26L	288
I4V-F26L	8198	I4V-F26L	216
I4V-F26L	6356		
I4V-F26L	1322		
I4V-Y33F	2225	I4V-Y33F	72
I4V-Y33F	14	I4V-Y33F	48
I4V-Y33F	126	I4V-Y33F	54
I4V-Y33F	6916	I4V-Y33F	37
I4V-Y33F	43	I4V-Y33F	43
I4V-Y33F	393	I4V-Y33F	29
I4V-F37I	2025	I4V-F37I	481
I4V-F37I	389	I4V-F37I	63
I4V-F37I	6198	I4V-F37I	426
I4V-F37I	1192	I4V-F37I	56
F26L-Y33F	722	F26L-Y33F	145
F26L-Y33F	116	F26L-Y33F	268
F26L-Y33F	2714	F26L-Y33F	328
F26L-Y33F	128	F26L-Y33F	608
F26L-F37I	1290	F26L-Y33F	393
F26L-F37I	113	F26L-Y33F	728
F26L-F37I	2444	F26L-F37I	909
F26L-F37I	215	F26L-F37I	264
		F26L-F37I	935
		F26L-F37I	271
Y33F-F37I	1338	Y33F-F37I	2067
Y33F-F37I	111	Y33F-F37I	104
Y33F-F37I	757	Y33F-F37I	5487
Y33F-F37I	2647	Y33F-F37I	275
Y33F-F37I	220	Y33F-F37I	7883
Y33F-F37I	1499	Y33F-F37I	395
I4V-F26L-Y33F	52	I4V-F26L-Y33F	31
I4V-F26L-Y33F	61	I4V-F26L-Y33F	5
I4V-F26L-Y33F	144	I4V-F26L-Y33F	125
I4V-F26L-Y33F	170	I4V-F26L-Y33F	21
I4V-F26L-F37I	81	I4V-F26L-F37I	259
I4V-F26L-F37I	256	I4V-F26L-F37I	95
I4V-F26L-F37I	309	I4V-F26L-F37I	326
I4V-F26L-F37I	972	I4V-F26L-F37I	119
I4V-Y33F-F37I	136	I4V-Y33F-F37I	4
I4V-Y33F-F37I	450	I4V-Y33F-F37I	17
I4V-Y33F-F37I	1582	I4V-Y33F-F37I	4
I4V-Y33F-F37I	5231	I4V-Y33F-F37I	17
F26L-Y33F-F37I	57	F26L-Y33F-F37I	431
F26L-Y33F-F37I	45	F26L-Y33F-F37I	1988
F26L-Y33F-F37I	647	F26L-Y33F-F37I	858
F26L-Y33F-F37I	512	F26L-Y33F-F37I	3954
I4V-Y33F-F26L-F37I	69	I4V-F26L-Y33F-F37I	26
I4V-Y33F-F26L-F37I	58	I4V-F26L-Y33F-F37I	312
I4V-Y33F-F26L-F37I	447	I4V-F26L-Y33F-F37I	16
I4V-Y33F-F26L-F37I	375	I4V-F26L-Y33F-F37I	186

TABLE 4

Regression model for contributions to the free energy of the specificity ratio

R^2 for this model is 0.65. The *F*-ratio is 8.6 with a *p* value of <0.0001. Coefficients are given in kcal/mol. The model is based on the 146 observations in Table 3, which afford 114 degrees of freedom for pure error. Most of the unexplained variance therefore arises because of the large experimental noise, as discussed under "Results."

Predictor[a]	Estimate	S.E.	*t* Ratio	Prob>|*t*|
Mg^{2+}	−0.73	0.11	−6.72	<.0001
Ile-4	−0.71	0.11	−6.53	<.0001
(Ile-4) × (Phe-26) × (Tyr-33) × (Phe-37)	−4.44	0.87	−5.10	<.0001
Tyr-33	−0.44	0.11	−4.01	0.0001
(Mg^{2+}) × (Ile-4)	0.86	0.22	3.97	0.0001
(Ile-4) × (Phe-26) × (Tyr-33)	−1.41	0.44	−3.23	0.0016
(Mg^{2+}) × (Phe-26) × (Phe-37)	1.25	0.44	2.87	0.0049
(Mg^{2+}) × (Phe-26)	−0.62	0.22	−2.87	0.0049
(Ile-4) × (Tyr-33)	0.61	0.22	2.80	0.0059
(Mg^{2+}) × (Ile-4) × (Phe-37)	−1.13	0.43	−2.59	0.0108
(Mg^{2+}) × (Ile-4) × (Tyr-33)	−1.08	0.43	−2.49	0.014
(Ile-4) × (Phe-26)	−0.52	0.22	−2.40	0.0179
(Mg^{2+}) × (Phe-26) × (Tyr-33)	0.95	0.44	2.18	0.0316
(Mg^{2+}) × (Ile-4) × (Tyr-33) × (Phe-37)	1.89	0.87	2.18	0.0316
(Ile-4) × (Phe-26) × (Phe-37)	−0.91	0.44	−2.09	0.0388
(Mg^{2+}) × (Phe-37)	−0.40	0.22	−1.83	0.0704
Phe-37	0.18	0.11	1.67	0.0982
(Mg^{2+}) × (Tyr-33)	−0.31	0.22	−1.44	0.1531
Phe-26	−0.15	0.11	−1.40	0.1631
(Ile-4) × (Tyr-33) × (Phe-37)	−0.49	0.44	−1.11	0.2675
(Tyr-33) × (Phe-37)	−0.21	0.22	−0.95	0.3462
(Phe-26) × (Tyr-33)	−0.11	0.22	−0.52	0.6046
(Mg^{2+}) × (Tyr-33) × (Phe-37)	−0.17	0.43	−0.39	0.6958
(Phe-26) × (Tyr-33) × (Phe-37)	0.15	0.44	0.35	0.7288
(Phe-26) × (Phe-37)	0.07	0.22	0.33	0.7445
(Ile-4) × Phe-37)	0.06	0.22	0.30	0.7675

[a] Multidimensional thermodynamic analysis of TrpRS specificity.

the four mutated side chains (Fig. 5*A*). The magnitude of this interaction, −4.4 kcal/mol, is remarkably similar to the five-way coupling of about −6 kcal/mol associated with the catalytic activation of Mg^{2+}, and to the interdomain coupling energies (~6 kcal/mol) estimated from the modular combinatorial analysis of contributions from CP1 and the ABD on amino acid activation, specificity, and $tRNA^{Trp}$ aminoacylation. Indeed, we propose that all effects represent the integrated, allosteric impact arising from domain motion of the ABD relative to the active site during catalysis of amino acid activation and acyl transfer.

Second, six of eight variants with wild-type Ile-4 contribute to higher specificity for tryptophan (*lightly shaded bars* in Fig. 5). Unlike the balanced contribution of the four D1 Switch residues to catalytic assist by Mg^{2+}, the higher order interactions responsible for the allosteric enhancement of specificity suggest disproportionate participation by the wild-type Ile-4 residue. We explore implications of this observation under "Discussion."

The Role of the Metal—Previously, we noted (4) that replacing Mg^{2+} with Mn^{2+} reduced the specificity for tryptophan, increasing the relative activation of tyrosine by ~3.5-fold. This value is almost precisely mirrored by the main effect of Mg^{2+} ion: −0.73 kcal/mol. Interactions involving Mg^{2+} actually lead to a net reduction of specificity: +2.9 kcal/mol. It appears, then, that unfavorable interactions between D1 residues and the metal modulate the overall net effect of D1 residues on the specificity ratio, about −8.3 kcal/mol, which implies a much higher tyrosine rejection than is observed.

DISCUSSION

Multidimensional thermodynamic cycles (13) are the "gold standard" for quantitative estimates of linkage-defining coupling energies in allostery and cooperativity. However, only rarely has the approach been extended to cycles of dimension >2, either in protein science or enzymology. Our results (5) and those of Sadovsky and Yifrach (20) illustrate two examples of functionally significant, higher order coupling interactions between individual residues. Our new paradigm, designed combinatorial mutagenesis of key dynamic residues connecting rigid bodies identified from structural studies, affords a path to new insight in other systems.

Consistency between Modular, Point Mutant Thermodynamic Cycles—We showed previously that catalytic assist by Mg^{2+} during tryptophan activation was the exclusive result of long range energetic coupling between the metal and the four D1 Switch residues mutated in this study. Implicating specific amino acids contributing to enhanced specificity in contemporary TrpRS, relative to that of the Urzyme uniquely complements that published recently (8) about the effects of adding the CP1 insertion or the ABD individually to the Urzyme. The surprising conclusion was that all three of the essential TrpRS functions, tryptophan activation, selection of tryptophan *versus* tyrosine, and $tRNA^{Trp}$ aminoacylation, benefit from approximately the same coupling energy, −6 kcal/mol between the two modules. This work presents remarkably consistent evidence that the functional intermodular communication is mediated by the D1 Switch.

A Possible Role for Ile-4—Differences between the higher order interactions illustrated in Fig. 5 and those illustrated in Fig. 4*A* of Ref. 5 merit comment. The highest order interaction is dominant by a significant margin in both distributions of coupling energies. In the allosteric coupling to catalytic assist by Mg^{2+} ion, none of the lower order interactions consistently involve the obvious, unidirectional affect of the same residue. In the present distribution, however, the Ile-4 residue is a prominent feature of the seven next largest effects (Fig. 5*A*), suggesting that this residue might serve an identifiable role in coordinating the effects of the D1 Switch in the selection of the correct amino acid. Ile-4 lies between the D1 Switch and the N-terminal β-strand that forms a "bottom" platform of the tryptophan-binding pocket, for which the specificity helix, residues 125–136, serves as the top. Thus, it occupies a unique position, relative to the tryptophan binding pocket. Fig. 2*A* illustrates that the top and bottom can be mechanically "pinched" together by the CP1 domain. Enhanced specificity arises primarily from differences in k_{cat} (Fig. 3, *B* and *C*), suggesting that this effect, and the ensuing decision of whether or not to activate the bound amino acid, occur close to the transition state.

The Observed Effects Are Associated with Changes in the Volume of the Tryptophan Binding Pocket—TrpRS catalysis proceeds by a sequence of structural states represented by three distinct crystal structures (Protein Data Bank codes 1MAW, 1MAU, and 1I6L (7, 17)). We used the recently described algorithm POVME (21) to assess the tryptophan binding pocket volumes (287 $Å^3$, 102 $Å^3$, and 92 $Å^3$) of these three states. These substantial changes suggest a possible mechanism to account for the association of specific substrate selection with domain movement, as well as the unique role of Ile-4: control over the tryptophan binding pocket volume by domain movement

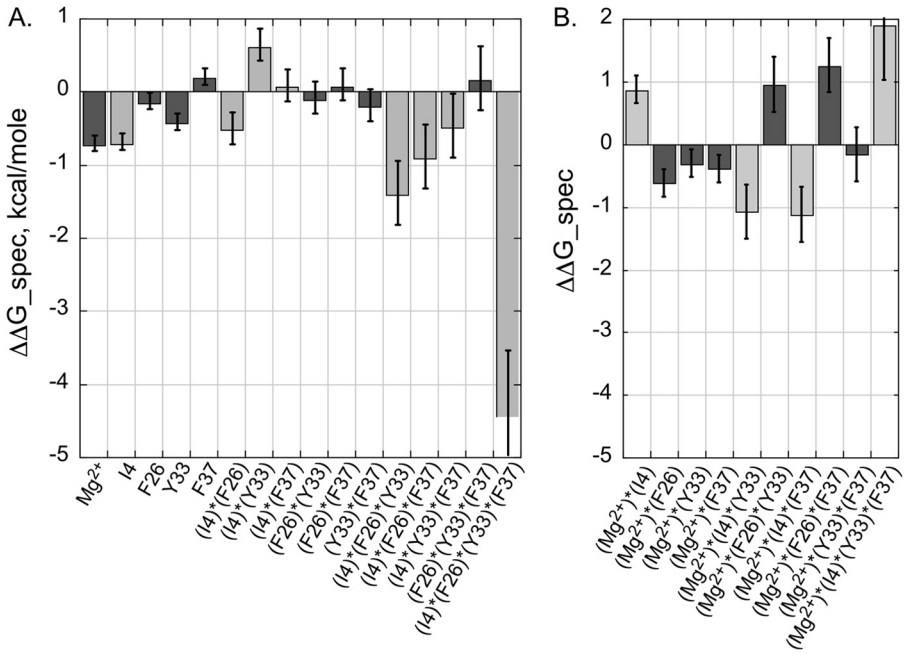

FIGURE 5. **Intrinsic and interaction effects responsible for specific selection of tryptophan *versus* tyrosine in *B. stearothermophilus* TrpRS.** *A,* intrinsic and interaction effects not involving the metal. *B,* values for terms involving coupling to Mg^{2+}. *Light bars* denote terms that include wild-type residue Ile-4.

might localize a stringent test for interactions specific for tryptophan to a point close to the chemical transition state.

Chemical transition states are so transient that their only experimental traces appear in kinetic measurements, such as those illustrated in Figs. 3 and 5. Computational methods, however, can potentially access the associated structural data. To that end, we have developed methods to construct the most probable pathways with minimum action (10, 17).[4] These structural sequences illustrate semi-quantitatively (*e.g.* Fig. 6 and supplemental Movie S1) the successive states along the path (22), together with their energies.

Amino acid binding pocket volumes calculated for the trajectory using POVME (21) suggest two states along the path where these volumes assume minimum values (Fig. 6*A*), representing opportunities to validate the correct choice. Each follows closely after the corresponding conformational transition state encountered during induced-fit and catalysis (7).

Correlated movements of CP1 and the ABD impose subtle changes in the positions of both the N-terminal β-strand and the specificity helix, as suggested in Fig. 2*A*. These movements compress the volume by ~12% during the catalytic step (Fig. 6). The four D1 Switch residues rearrange in the conformational transition states that precede compression of pocket volumes (Fig. 6*A* and supplemental Movie S1), and so could couple specificity to these movements.

Allosteric Enhancement Is Widely Observed in aaRS—Both kinetic measurements (23) and indirect experiments (24) led to the early suggestion that amino acid specificity by aaRS involved the dominant role of k_{cat}. Steady-state Michaelis-Menten kinetics of GlnRS mutants showed that its specific aminoacylation of $tRNA^{Gln}$ was primarily reflected in lower mutant k_{cat} values (25, 26). Pre-steady state observations were made more recently for selectivity of GlnRS for glutamine *versus* glutamic acid (27–29). These results, and our own earlier failure to improve specificity of TrpRS for tyrosine (9) can now be interpreted more clearly in terms of allosteric enhancement by domain motion. Similarly, the amino acid binding pocket size was implicated in specific recognition of phenylalanine by PheRS (30). The generality of these observations implies that allosteric enhancement of amino acid specificity by mechanisms analogous to those described here are a general property of all or most aaRS.

Selective Advantage of Using Allostery to Enforce Specificity—The evident difficulty of changing aminoacyl-tRNA synthetase specificities by multiple obvious point mutations within the active site (9, 27–29) contrasts with the ease of creating orthogonal aaRS/tRNA pairs (31, 32). The apparent conflict merits two comments. First, the aaRSs derived for use in synthetic biology generally require sophisticated selection techniques involving suppression of nonsense mutations in essential genes by multisite mutants prepared from a "naive" amino acid binding pocket created by first mutating all relevant side chains in the amino acid binding pocket to alanine (33). Such a procedure may bypass the barrier to simple-minded mutagenesis, by selecting combinations that accommodate non-canonical amino acids without interfering overly much with the induced-fit and catalytic conformational changes. Second, whereas the engineered aaRS clearly accomplish their intended function of acylating orthogonal suppressor tRNAs with non-canonical amino acids, little has been done to define their actual specificity constants.

We are left with an important question: why off-load specific recognition to an allosteric mechanism? Several authors have suggested possible answers to this question. Cooperative transitions in proteins appear to be robust to mutation (15, 34). Thus, as the aminoacyl-tRNA synthetases depend strongly on retaining their high substrate specificities, it makes sense that allosteric mechanisms would be selected over time by the removal of point mutants that altered such specificity, whereas mutations leading to allos-

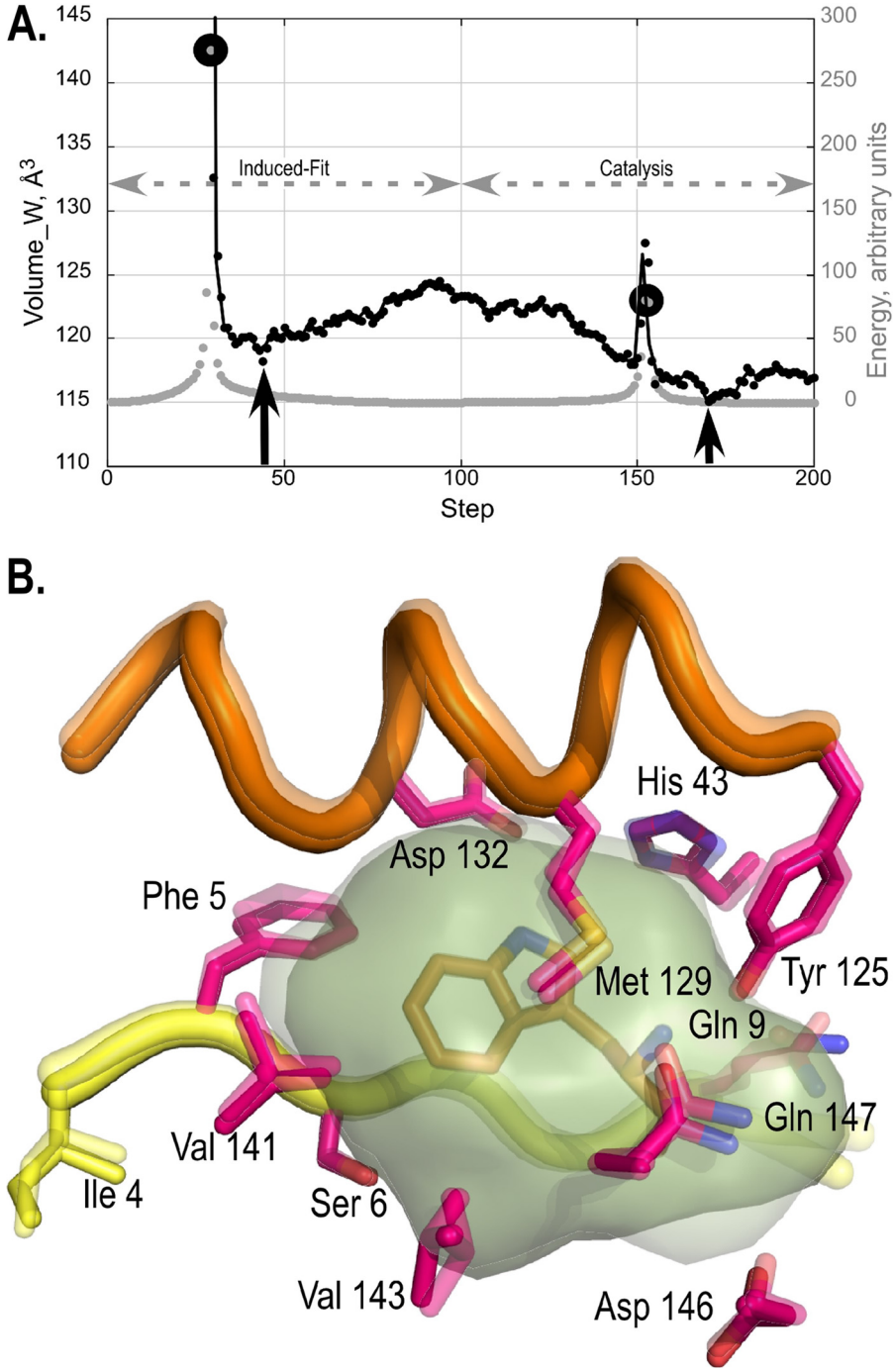

FIGURE 6. **Trajectory parameters for modeling TrpRS interactions with substrate amino acids.** *A,* a 200-step Minimum Action trajectory was computed to pass through three states provided by crystal structures (Protein Data Bank codes 1MAW, 1MAU, and 1I6L) that represent Open, Pre-transition, and Products states that demarcate the induced-fit and catalytic transitions, as indicated. Energies of states computed along the path are indicated by *gray dots*; tryptophan binding pocket volumes are given by *black dots* connected by the *black line*. Structures representing the first state with the new D1 side chain packing are indicated by *large*, *filled circles*, and coincide with the conformational transition states identified by PATH. Pocket volumes increase during both conformational transition states, and assume minimum values just thereafter, as indicated by the *bold arrows*. *B,* superposition of the tryptophan binding pocket residues from the minimum volumes following the conformational transition states associated with induced-fit (*transparent*) and catalysis as *solid*. The pocket (*smudge green*) is 12% smaller after the second transition state. The first β-strand and specificity helix are colored *yellow* and *orange* as described in the legend to Fig. 2.

teric mechanisms accumulated and began to function. Indeed, combinatorial mutagenesis of the D1 Switch residues in the TrpRS Urzyme will be a sensible extension of this work.

Acknowledgment—We thank H. Fried for helpful feedback and critical comments on the manuscript.

REFERENCES

1. Perutz, M. F., Wilkinson, A. J., Paoli, M., and Dodson, G. G. (1998) The stereochemical mechanism of the cooperative effects in hemoglobin revisited. *Annu. Rev. Biophys. Biomol. Struct.* **27,** 1–34
2. Perutz, M. F. (1970) Stereochemistry of cooperative effects of hemoglobin. *Nature* **228,** 726–739
3. Weinreb, V., and Carter, C. W., Jr. (2008) Mg^{2+}-free *B. stearothermophi-*

lus tryptophanyl-tRNA synthetase activates tryptophan with a major fraction of the overall rate enhancement. *J. Am. Chem. Soc.* **130,** 1488–1494
4. Weinreb, V., Li, L., Campbell, C. L., Kaguni, L. S., and Carter, C. W., Jr. (2009) Mg^{2+}-assisted catalysis by *B. stearothermophilus* TrpRS is promoted by allosteric effects. *Structure* **17,** 952–964
5. Weinreb, V., Li, L., and Carter, C. W., Jr. (2012) A master switch couples Mg^{2+}-assisted catalysis to domain motion in *B. stearothermophilus* tryptophanyl-tRNA synthetase. *Structure* **20,** 128–138
6. Cammer, S., and Carter, C. W., Jr. (2010) 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,** 709–714
7. Kapustina, M., Weinreb, V., Li, L., Kuhlman, B., and Carter, C. W., Jr. (2007) A conformational transition state accompanies tryptophan activation by *B. stearothermphilus* tryptophanyl-tRNA synthetase. *Structure* **15,** 1272–1284
8. Li, L., and Carter, C. W., Jr. (2013) Full implementation of the genetic code by tryptophanyl-tRNA synthetase requires cooperativity between recently acquired structural modules. *J. Biol. Chem.* **288,** 34736–34745
9. Praetorius-Ibba, M., Stange-Thomann, N., Kitabatake, M., Ali, K., Söll, I., Carter, C. W., Jr., Ibba, M., and Söll, D. (2000) Ancient adaptation of the active site of tryptophanyl-tRNA synthetase for tryptophan binding. *Biochemistry* **39,** 13136–13143
10. Franklin, J., Koehl, P., Doniach, S., and Delarue, M. (2007) MinActionPath. Maximum likelihood trajectory for large-scale structural transitions in a coarse-grained locally harmonic energy landscape. *Nucleic Acids Res.* **35,** W477-W482
11. Francklyn, C. S., First, E. A., Perona, J. J., and Hou, Y.-M. (2008) Methods for kinetic and thermodynamic analysis of aminoacyl-tRNA synthetases. *Methods* **44,** 100–118
12. Retailleau, P., Hu, M., Vachette, P., and Carter, C. W., Jr. (2003) 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
13. Horovitz, A., and Fersht, A. R. (1990) Strategy for analyzing the co-operativity of intramolecular interactions in peptides and proteins. *J. Mol. Biol.* **214,** 613–617
14. Carter, C. W., Jr. (1990) Efficient factorial designs and the analysis of macromolecular crystal growth conditions. *Methods* **1,** 12–24
15. Weinkam, P., Pons, J., and Sali, A. (2012) Structure-based model of allostery predicts coupling between distant sites. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 4875–4880
16. Deleted in proof
17. Laowanapiban, P., Kapustina, M., Vonrhein, C., Delarue, M., Koehl, P., and Carter, C. W., Jr. (2009) 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
18. Pham, Y., Kuhlman, B., Butterfoss, G. L., Hu, H., Weinreb, V., and Carter, C. W., Jr. (2010) Tryptophanyl-tRNA synthetase Urzyme. A model to recapitulate molecular evolution and investigate intramolecular complementation. *J. Biol. Chem.* **285,** 38590–38601
19. Pham, Y., Li, L., Kim, A., Erdogan, O., Weinreb, V., Butterfoss, G. L., Kuhlman, B., and Carter, C. W., Jr. (2007) A minimal TrpRS catalytic domain supports sense/antisense ancestry of class I and II aminoacyl-tRNA synthetases. *Mol. Cell* **25,** 851–862
20. Sadovsky, E., and Yifrach, O. (2007) Principles underlying energetic coupling along an allosteric communication trajectory of a voltage-activated K channel. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 19813–19818
21. Durrant, J. D., de Oliveira, C. A., and McCammon, J. A. (2011) POVME. An algorithm for measuring binding-pocket volumes. *J. Mol. Graph. Model* **29,** 773–776
22. Prevost, M. S., Sauguet, L., Nury, H., Van Renterghem, C., Huon, C., Poitevin, F., Baaden, M., Delarue, M., and Corringer, P.-J. (2012) A locally closed conformation of a bacterial pentameric proton-gated ion channel. *Nat. Struct. Mol. Biol.* **19,** 642–649
23. Lam, S. S., and Schimmel, P. R. (1975) Equilibrium measurements of cognate and noncognate interactions between aminoacyl transfer RNA synthetases and transfer RNA. *Biochemistry* **14,** 2775–2780
24. Ebel, J. P., Giegé, R., Bonnet, J., Kern, D., Befort, N., Bollack, C., Fasiolo, F., Gangloff, J., and Dirheimer, G. (1973) Factors determining the specificity of the tRNA aminoacylation reaction non-absolute specificity of tRNA-aminoacyl-tRNA synthetase recognition and particular importance of the maximal velocity. *Biochimie* **55,** 547–557
25. Rogers, M. J., Adachi, T., Inokuchi, H., and Söll, D. (1992) Switching $tRNA^{Gln}$ identity from glutamine to tryptophan. *Proc. Natl. Acad. Sci. U.S.A.* **89,** 3463–3467
26. Rogers, M. J., Weygand-Durasevi, I., Schwob, E., Sherman, J. M., Rogers, K. C., Adachi, T., Inokuchi, H., and Söll, D. (1993) Selectivity and specificity in the recognition of tRNA by *E. coli* by glutaminyl-tRNA synthetase. *Biochimie* **75,** 1083–1090
27. Bullock, T. L., Uter, N., Nissan, T. A., and Perona, J. J. (2003) Amino acid discrimination by a class I aminoacyl-tRNA synthetase specified by negative determinants. *J. Mol. Biol.* **328,** 395–408
28. Bullock, T. L., Rodríguez-Hernández, A., Corigliano, E. M., and Perona, J. J. (2008) A rationally engineered misacylating aminoacyl-tRNA synthetase. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 7428–7433
29. Corigliano, E. M., and Perona, J. J. (2009) Architectural underpinnings of the genetic code for glutamine. *Biochemistry* **28,** 676–687
30. Ibba, M., Kast, P., and Hennecke, H. (1994) Substrate specificity is determined by amino acid binding pocket size in *Escherichia coli* phenylalanyl-tRNA synthetase. *Biochemistry* **33,** 7107–7112
31. Chin, J. W., Cropp, T. A., Anderson, J. C., Mukherji, M., Zhang, Z., and Schultz, P. G. (2003) An expanded eukaryotic genetic code. *Science* **301,** 964–967
32. Liu, C. C., and Schultz, P. G. (2010) Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79,** 413–444
33. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of *Escherichia coli*. *Science* **292** 498–500
34. del Sol, A., Fujihashi, H., Amoros, D., and Nussinov, R. (2006) Residues crucial for maintaining short paths in network communication mediate signaling in proteins. *Mol. Syst. Biol.* **2,** 2006.0019