BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

A Tale of Two Isomerases: Compact versus Extended Active Sites in Ketosteroid Isomerase and Phosphoglucose Isomerase

Srinivas Somarowthu,[†] Heather R. Brodkin,[‡] J. Alejandro D'Aquino,[‡] Dagmar Ringe,[‡] Mary Jo Ondrechen,*[,†,§] and Penny J. Beuning*[,†,§]

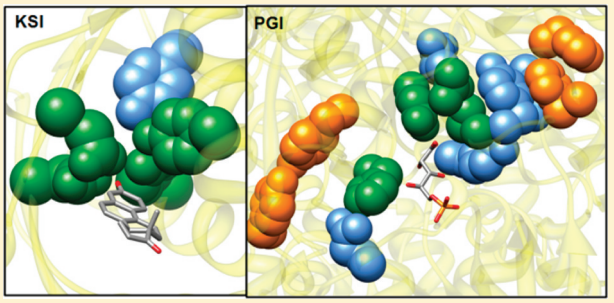
[†]Department of Chemistry and Chemical Biology, Northeastern University, Boston, Massachusetts 02115, United States

[‡]Departments of Biochemistry and Chemistry and Rosenstiel Basic Medical Sciences Center, Brandeis University, Waltham, Massachusetts 02454-9110, United States

[§]Center for Interdisciplinary Research on Complex Systems, Northeastern University, Boston, Massachusetts 02115, United States

S *Supporting Information*

ABSTRACT: Understanding the catalytic efficiency and specificity of enzymes is a fundamental question of major practical and conceptual importance in biochemistry. Although progress in biochemical and structural studies has enriched our knowledge of enzymes, the role in enzyme catalysis of residues that are not nearest neighbors of the reacting substrate molecule is largely unexplored experimentally. Here computational active site predictors, THEMATICS and POOL, were employed to identify functionally important residues that are not in direct contact with the reacting substrate molecule. These predictions then guided experiments to explore the active sites of two isomerases, *Pseudomonas putida* ketosteroid isomerase (KSI) and human phosphoglucose isomerase (PGI), as prototypes for very different types of predicted active sites. Both KSI and PGI are members of EC 5.3 and catalyze similar reactions, but they represent significantly different degrees of remote residue participation, as predicted by THEMATICS and POOL. For KSI, a compact active site of mostly first-shell residues is predicted, but for PGI, an extended active site in which residues in the first, second, and third layers around the reacting substrate are predicted. Predicted residues that have not been previously tested experimentally were investigated by site-directed mutagenesis and kinetic analysis. In human PGI, single-point mutations of the predicted second- and third-shell residues K362, H100, E495, D511, H396, and Q388 show significant decreases in catalytic activity relative to that of the wild type. The results of these experiments demonstrate that, as predicted, remote residues are very important in PGI catalysis but make only small contributions to catalysis in KSI.

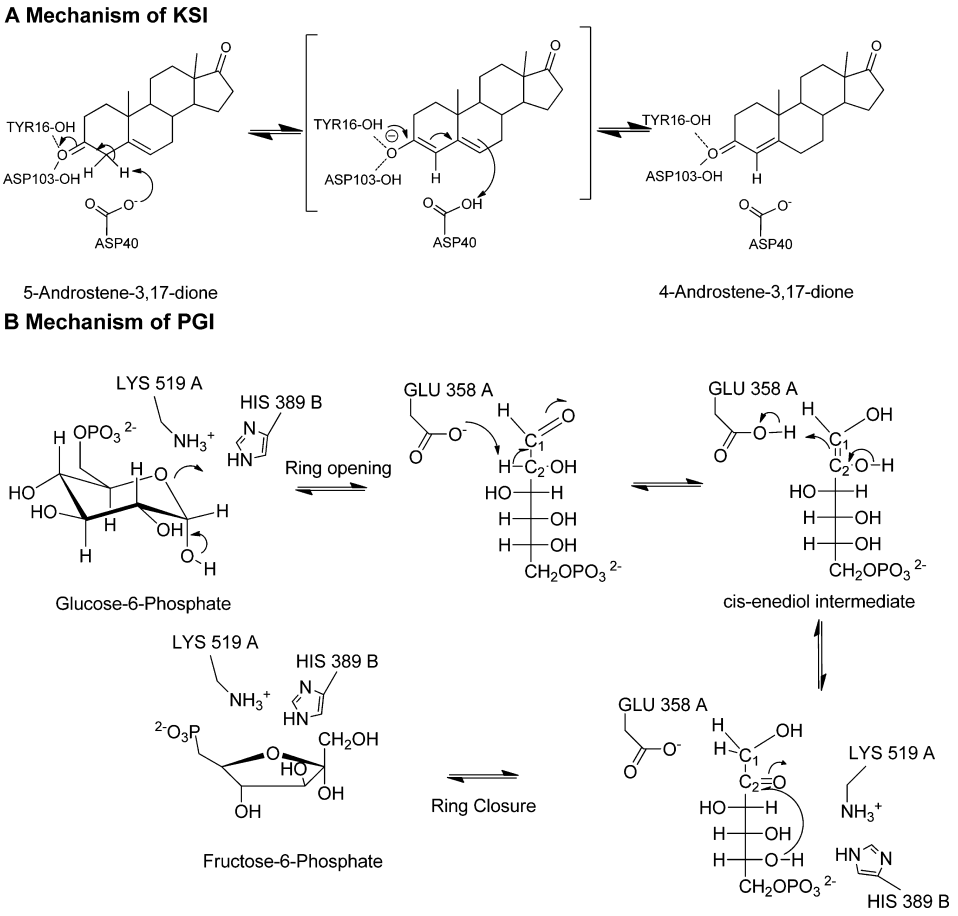


Reactions catalyzed by enzymes are typically many orders of magnitude faster than the corresponding uncatalyzed reactions,[1,2] with high substrate specificity and regio- and stereoselectivity. As enzymes have evolved to catalyze biological reactions inside living cells, they generally work well under mild conditions and in aqueous solution. These properties of enzymes have made them attractive for many medicinal and industrial applications, including green chemistry.[3,4] To be optimally useful in commercial applications, the catalytic and physical properties of enzymes, including their stability, activity, specificity, and solubility, often need to be modified. The largest barrier for the rational design of enzymes is the lack of a detailed theoretical basis for the high catalytic efficiency and exquisite specificity of enzymes. Much of the prior work addressing this issue[1,2,5,6] has focused on the amino acid residues that are involved in catalysis through direct spatial contact with the reacting substrate molecule(s). The role of remote residues, i.e., residues that are not in the immediate area of the reacting substrate molecules but nevertheless contribute to the progress of the reaction, in catalysis has not been extensively explored previously. In many enzymes, especially non-metalloenzymes, the chemical steps are conducted by the catalytic amino acid residues, which often have chemical properties (e.g., shifted pK_a) different from those of the corresponding amino acids free in solution. These unusual properties may be the result of interactions of these catalytic residues with other residues in the enzyme. Identifying and understanding the interactions between residues that contribute to catalysis could point the way toward systematic modulation of the desired properties of catalytic residues, and thereby the development of better control of enzyme activity and specificity.

We previously developed THEMATICS[7] (Theoretical Microscopic Anomalous Titration Curve Shapes) and POOL[8] (Partial Order Optimum Likelihood), computational predictors of residues involved in catalysis and binding. Both utilize the three-dimensional structure of the query protein as input.

Received: July 13, 2011
Revised: September 14, 2011
Published: October 4, 2011

Scheme 1. Mechanisms of KSI and PGI



Strikingly, in some cases, these methods predict the participation in catalysis by residues not in direct contact with the reacting substrate molecule. Briefly, the THEMATICS method solves the Poisson−Boltzmann equations and yields the electrical potential function for the protein and then computes the theoretical microscopic titration curves for all of the ionizable residues in the protein structure. When the protonation equilibria of multiple ionizable residues are coupled, the titration curves of these residues will have perturbed shapes. It is the shapes of these titration curves, and not the predicted pK_a shift, that form the basis for THEMATICS predictions. The residues with the most perturbed titration curve shapes tend to be active site residues, and thus, spatial clusters of residues with the greatest anomalies in the theoretical titration curve shapes are predicted to be functionally important. POOL is a machine learning method based on a monotonicity-constrained maximum likelihood technique. POOL utilizes THEMATICS and other input features to place all of the residues in a protein structure in rank order according to their probability of functional importance. While THEMATICS predicts only the seven ionizable amino acid types, POOL utilizes environment variables to predict all 20 amino acid types. Typically, the set of residues in the top 8−10% of the POOL ranking is taken as the active site prediction.

In a preliminary study, we conducted an analysis of nitrile hydratase from *Pseudomonas putida*,[9] in which second- and third-shell residues predicted by THEMATICS were mutated using site-directed mutagenesis and the resultant variants were analyzed for activity. Four of five single-residue variants with conservative mutations at predicted positions, three in the second shell and one in the third shell, showed activity decreased by 1−2 orders of magnitude. This shows that the active site of nitrile hydratase is built in multiple layers, where interactions between first-shell residues (residues in direct contact with substrate) and residues behind them in the second and third shells are important for catalysis.

In this study, to test whether second- and third-shell residues participate as predicted, we chose two enzymes for experimental analysis, ketosteroid isomerase (KSI) from *P. putida* and human phosphoglucose isomerase (PGI). KSI and PGI both catalyze isomerization reactions as shown in Scheme 1, but they have very different degrees of remote residue participation predicted by THEMATICS. For KSI, all of the residues that THEMATICS predicts to be participating in catalysis are, with one exception, in direct contact with the reacting substrate. On the other hand, for PGI, THEMATICS predicts participation in catalysis by multiple layers of residues in the first, second, and third shells surrounding the reacting substrate molecule. These two enzymes represent two extremes in the predicted degree of remote residue participation and thus serve as good test cases for the question at hand.

Δ^5-3-Ketosteroid isomerase (KSI, EC 5.3.3.1) has been studied both computationally[10−12] and experimentally[13−15] because of its high catalytic activity, with a rate approaching the diffusion-controlled limit. KSI catalyzes the isomerization of Δ^5-3-ketosteroid to Δ^4-3-ketosteroid, involving C−H bond cleavage and formation through a dienolate intermediate

dx.doi.org/10.1021/bi201089v | *Biochemistry* 2011, 50, 9283−9295

(Scheme 1A).[16] The formation of the enolate intermediate was established previously as the rate-determining step;[17] however, Hawkinson et al.[18] showed that no single chemical step can be clearly rate-limiting when 5-androstene-3,17-dione is the substrate. KSI has a metabolic role in the degradation of steroids in bacterial species that can live on steroids as their sole carbon source.[19] High-resolution crystal structures have been reported for two orthologous enzymes from *Commamonas testosteroni* (ctKSI)[20] and *P. putida* (ppKSI) Biotype B.[21] Site-directed mutagenesis experiments identified D40 (ppKSI numbering) as the catalytic base responsible for proton transfer.[22] Y16[22] and D103[23−25] are identified as residues responsible for stabilizing the transition state.

Phosphoglucose isomerase (PGI, EC 5.3.1.9) is an example of a moonlighting protein.[26] Inside the cell, PGI catalyzes the reversible isomerization of glucose 6-phosphate and fructose 6-phosphate, an essential step of glycolysis and gluconeogenesis. Extracellular PGI has many other functions, including acting as a neuroleukin,[27,28] an autocrine motility factor,[29] and a maturation factor.[30] Human PGI (hPGI) is of particular medical interest, as a deficiency in hPGI leads to non-spherocytic hemolytic anemia.[31,32] Also, PGI is used as a biomarker because of its high level of expression observed in patients with certain cancers.[33] Crystal structures with and without ligand bound are reported for both PGI from mammals (rabbit,[34−38] human,[39] and mouse[40]) and bacteria (*Bacillus stearothermophilus*[41,42]). The proposed mechanism of isomerase activity is shown in Scheme 1B.[36,43] Crystal structure analysis of rabbit,[36,38] human,[44] and mouse[40] PGI in complex with mimics of the *cis*-enediol intermediate and site-directed mutagenesis of bacterial PGI identified E358 (where residues are numbered according to the UniProt sequence of hPGI) as the catalytic base; R273 helps to stabilize the intermediate in the isomerization step, and H389 and K519 are most likely involved in the ring-opening and closing steps.

We identified second- and third-shell residues that are predicted by THEMATICS and POOL to contribute to catalytic activity for KSI and PGI. We constructed site-directed variants of these enzymes to test our prediction that relatively few residues outside the first shell of KSI contribute to activity, whereas several second- and third-shell residues contribute to PGI activity. Some residues in the second shell of KSI and PGI that were not predicted to be functionally important were also included as negative controls. In addition, one first-shell residue in hPGI, H389, was predicted to be important for catalysis but had not been studied previously in human PGI; therefore, H389 was also included in our study. Our experimental results support our predictions, as only one mutation in KSI leads to a significant decrease in activity, and this effect is relatively small. However, mutations of predicted residues in both the second and third shells lead to dramatic decreases in activity in PGI.

MATERIALS AND METHODS

Theoretical Methods. *Functional Site Prediction and Shell Analysis.* The X-ray crystal structures of wild-type KSI from *P. putida* (PDB entry 1OH0)[21] and wild-type human PGI (PDB entry 1IAT)[39] (where residues are numbered according to the UniProt sequence) were preprocessed using the molecular modeling software YASARA[45] to add any missing atoms. THEMATICS[7] calculations were performed as described previously,[46] except that in this case for the identification of remote residues, a decreased cutoff of 0.96 was used instead of 0.99.[46] For POOL[8] calculations, along with THEMATICS features, INTREPID[47] evolutionary scores are included.[48] INTREPID scores were obtained from the Berkeley phylogenomics web server (http://phylogenomics.berkeley.edu/intrepid/). First-shell residues, the residues in contact with a bound ligand, were identified using the Ligand-Protein Contacts (LPC) server.[49] Second-shell and third-shell residues, i.e., residues in contact with the first shell and second shell, respectively, were identified using the Contacts of Structural Units (CSU) server.[49]

Electrostatics Calculations. MCCE (MultiConformer Continuum Electrostatics),[50,51] which allows conformational flexibility of amino acid side chains, was implemented to predict pK_a values. Structures of wild-type hPGI and ppKSI were used as input for the MCCE calculations. Because of the size of hPGI with more than 1100 residues, the "quick" method was used for the calculations. Mean field analysis was performed at a pH equal to the MCCE-predicted pK_a for each of the catalytic residues. In addition, to compare the results from MCCE, we also employed UHBD[52] to predict pK_a values, to verify that the two methods predict the same trends. UHBD calculations were performed assuming a protein internal dielectric constant of 20; HYBRID[53] was employed to calculate the mean net charge as a function of pH for each ionizable residue in the protein structure, and from this theoretical titration curve, the pK_a values for each of these residues are computed.

Experimental Methods. *Mutagenesis.* A plasmid expressing human PGI and expression strain *Escherichia coli* DF2145 were obtained from M. Meng.[54] Plasmids expressing KSI[55] were gifts from K. Y. Choi. Mutations were constructed by using a QuikChange site-directed mutagenesis kit (Agilent) and confirmed by DNA sequencing analysis (Massachusetts General Hospital DNA Core, Cambridge, MA).

Protein Expression and Purification. KSI. Wild-type KSI and variants were purified using the published protocol[13] except that an affinity column was not employed. Briefly, expression of wild-type KSI and variants was conducted in BL21(DE3) cells in 1 L of Luria broth supplemented with 100 μg/mL ampicillin at 37 °C. Cells were disrupted by sonication, and the lysate was clarified by centrifugation at 24610*g* for 60 min. The supernatant was loaded onto a 10 mL DEAE ion-exchange column (GE Healthcare) pre-equilibrated with buffer containing 20 mM potassium phosphate (KPO_4), 1 mM EDTA, 1 mM DTT, and 50 mM NaCl (pH 7) and eluted with a linear gradient from 50 to 500 mM NaCl. Fractions containing KSI were pooled and concentrated to ∼1−2 mL with 5 kDa MWCO Vivaspin 6 (Vivaproducts). The protein sample was then loaded onto a Superdex 75 size exclusion chromatography column [26 mm × 70 cm (GE Healthcare)] and eluted with buffer containing 40 mM potassium phosphate (KPO_4), 1 mM EDTA, and 1 mM DTT (pH 7). Protein concentrations were determined by the Bradford assay (Bio-Rad).

PGI. Wild-type hPGI and variants were purified from *E. coli* strain DF2145 in 1 L of Luria broth supplemented with 100 μg/mL ampicillin according to the published protocol.[54] Clarified supernatant was then loaded onto a 10 mL MonoQ ion-exchange column (GE Healthcare) equilibrated in buffer A [10 mM triethanolamine (pH 7.5), 1 mM EDTA, phenylmethanesulfonyl fluoride (PMSF), and 0.1% β-mercaptoethanol]. The flow-through containing hPGI was collected and loaded onto a 5 mL SP ion-exchange column (GE Healthcare) pre-equilibrated in buffer A and eluted with a linear gradient from 0 to 0.75 M NaCl in buffer A. Fractions containing hPGI

were pooled and concentrated, and the buffer was exchanged with 20% glycerol in buffer A using 5 kDa MWCO Vivaspin 6 (Vivaproducts). Protein concentrations were determined by the Bradford assay (Bio-Rad).

Steady-State Enzyme Kinetics. Kinetic parameters for wild-type and variant enzymes were determined by nonlinear regression using GraphPad. Reported values for k_{cat} and K_M are from at least three independent trials.

KSI. Activity of KSI was measured spectrophotometrically[56] by following the conversion of 5-androstene-3,17-dione to 4-androstene-3,17-dione (Steraloids Inc.) at 248 nm. Reactions were conducted at 21 °C in a buffer containing 34 mM potassium phosphate (KPO_4), 2.5 mM EDTA (pH 7), various substrate concentrations, and enzyme as described previously.[24,25]

PGI. A coupled assay[54] was employed to measure the activity of hPGI with fructose 6-phosphate (Sigma) as the substrate. Reactions were conducted in a 500 μL cuvette at 21 °C in the presence of 20 mM HEPES buffer (pH 7.5) 10 units of glucose-6-phosphate dehydrogenase (Worthington), 1 mM β-NAD^+ (Sigma), enzyme, and various substrate concentrations.

CD Spectroscopy. Before the CD measurements, buffers were exchanged into 50 mM potassium phosphate (pH 7.5) using Zeba Spin (Pierce Scientific) buffer-exchange columns. Ultraviolet CD spectra (200−300 nm) were recorded using a JASCO J-715 CD spectrometer with protein concentrations of 0.1 mg/mL in a cuvette with a path length of 0.1 cm. Spectra were recorded at a scan speed of 50 nm/min and averaged over nine scans at room temperature.

Thermofluor Stability Assay. Reaction mixtures were assembled in 96-well optical plates with each sample (16 μL total volume) containing 15 μM protein, storage buffer [10 mM triethanolamine (pH 7.5), 1 mM EDTA, phenylmethanesulfonyl fluoride (PMSF), 0.1% β-mercaptoethanol, and 20% glycerol], and 25× Sypro Orange (Invitrogen), and then the plate was sealed with an optical adhesive film (Applied Biosystems).[57,58] The temperature was increased from 25 to 75 °C in an increment of 0.2 °C using an iCycler iQ5 Real Time PCR system (Bio-Rad). The change in fluorescence was detected with a built-in CCD camera with an excitation wavelength of 490 nm and an emission wavelength of 575 nm. The melting temperature (T_m) is obtained by fitting the data to the sigmoidal Boltzmann model equation using curve-fitting software XLfit 5, a Microsoft Excel add-on program. The relationship among the temperature, melting temperature, and fluorescence intensity is given by the expression intensity = $A + (B - A)/[1 + e^{(T_m - \text{temperature})/D}]$, where A and B are pre- and post-transitional fluorescence intensities, respectively, and D is a slope factor. Reported values are from at least three independent trials.

Crystallization, Data Collection, and Crystallographic Refinement. Crystals of the ppKSI variant M105A were grown by vapor diffusion using the hanging drop method at 25 °C. Equal volumes of protein at 20 mg/mL ppKSI and a crystallization solution containing 1.4 M ammonium sulfate and 6.5% isopropyl alcohol were mixed. After a few days, crystal clusters appeared in the drops. These clusters were dissected and transferred to a solution containing 2.9 M sodium malonate and subsequently flash-frozen in liquid nitrogen prior to data collection.

Data were collected at beamline ID-23D at GM/CA-CAT (APS, Argonne, IL) at 100 K using a MARMosiac 300 CCD detector. Diffraction data were indexed in orthorhombic space group $C222_1$, integrated, and scaled using HKL2000.[59] Molecular replacement was conducted with PHENIX AutoMR,[60] using the coordinates of wild-type ppKSI (PDB entry 1OH0[21]) as a coordinate model. After a molecular replacement solution was obtained, the model was improved with PHENIX AutoBuild.[60] A test set of 5% of randomly chosen reflections was selected and used for the calculation of the R_{free} value. After a round of simulated annealing, multiple cycles of iterative manual building were performed with Coot[61] and refined against an amplitude-based maximum-likelihood target function in PHENIX[60] until the R_{free} could no longer be reduced. The atomic coordinates of ppKSI M105A, as determined in this study, have been deposited in the Protein Data Bank as entry 3SED.

RESULTS

Functional Site Prediction by THEMATICS and POOL.

Ketosteroid isomerase (KSI) and phosphoglucose isomerase (PGI) both belong to EC 5.3 (isomerases and intramolecular oxidoreductases), but they have very different degrees of remote residue participation predicted by THEMATICS (Table S1 of the Supporting Information).

KSI. THEMATICS and POOL calculations were performed on wild-type KSI from *P. putida* (PDB entry 1OH0[21]). As shown in Figure 1, THEMATICS predicts a compact active site

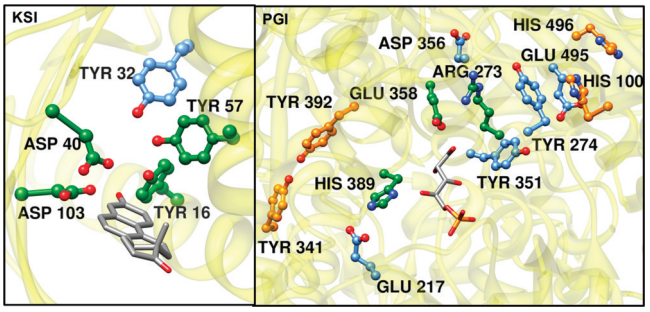


Figure 1. THEMATICS functional site predictions for ppKSI and hPGI. For KSI, THEMATICS predicts a compact active site of mostly first-shell residues, but for PGI, THEMATICS predicts a multilayered active site. Ligands are shown as gray sticks. Predicted residues are shown in ball-and-stick form. First-shell residues are colored green, and second- and third-shell residues are colored blue and orange, respectively.

for KSI with a cluster of five residues (Y16, D40, Y57, D103, and Y32), among which four residues belong to the first shell and only one residue (Y32) belongs to the second shell. POOL, which generates a ranked list of residues in order of their probability of functional importance, with rank 1 being the most likely to be functionally important, predicts very few second-shell residues in the top ranks (Table S2 of the Supporting Information). This atypical prediction, with a very low degree of participation by second-shell residues, suggests that residues outside the first shell are less likely to be important for catalysis in KSI than in other enzymes.

To improve our understanding of the nature of the active site in ppKSI, four second-shell residues (M31, E39, M84, and M105) were chosen for site-directed mutagenesis. Typically, when POOL is used to predict the functionally important residues in a given protein structure, the set of residues in the top 8% of the rank-ordered list is taken as the prediction. For KSI, the only second-shell residue in the top 8% is Y32, which

has already been tested experimentally and shown to have a small effect on catalysis.[24,62] The next highest-ranking second-shell residue, which falls just outside the top 8%, is E39, which was chosen for testing in this study. Additionally, three more residues (M31, M84, and M105) with lower POOL rankings were chosen on the basis of their proximity to the catalytic residues (Figure 2) and sequence conservation (Table 1). As shown in Figure 2, M31 is located behind the known active site residue Y16 with respect to the substrate, M84 makes contacts with the active site residue D103, M105 makes contacts with both Y16 and D103, and E39 is located adjacent to the catalytic base D40; all four are well conserved.

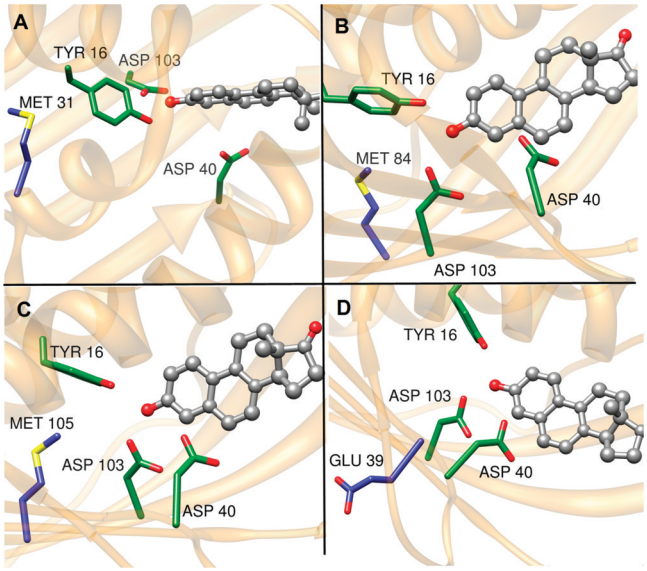


Figure 2. Cartoon diagram showing the location of four second-shell residues of ppKSI chosen for site-directed mutagenesis. (A) M31 is located behind the catalytic residue Y16. (B) M84 makes contacts with the catalytic residue D103. (C) M105 contacts both Y16 and D103. (D) E39 is located adjacent to the catalytic base D40. Second-shell residues are colored blue, and first-shell residues are colored green; the ligand is shown in ball-and-stick form.

Table 1. Shell Assignments, Distances from Each Residue to the Bound Ligand, Their Closest First-Shell Residues, and Sequence Conservation Scores for Each Remote Residue of ppKSI and hPGI Chosen for Analysis

residue	shell	distance from ligand	closest first-shell residue	ConSurf[63] normalized conservation score
		ppKSI		
M31[a]	second	10.2 Å	Y16	−1.0
E39[a]	second	7.4 Å	D40	−1.3
M84[a]	second	5.6 Å	D103	−0.63
M105[a]	second	5.4 Å	Y16 and D103	−1.2
		hPGI		
H100	third	12.6 Å	R273	−0.90
Y274	second	10.6 Å	R273	−0.88
Y341	third	14.7 Å	H389	−0.77
K362	second	10.2 Å	E358	−0.89
N386	second	12.4 Å	H389	−0.91
Q388	second	11.4 Å	H389	−0.91
H396	third	13.0 Å	K519	−0.91
E495	second	12.5 Å	R273	−0.90
D511	second	10.6 Å	E358	−0.90
N154[a]	second	11.3 Å	S160	−0.91
S185[a]	second	10.6 Å	I157	−0.91

[a]Negative controls (residues not predicted by POOL).

PGI. THEMATICS predicts four clusters for hPGI. Two clusters contain residues that are located far from the site of isomerase activity; hPGI is known to be multifunctional, and because of their location, these residues could be important for functions of hPGI other than its isomerase activity and therefore are not pursued further here. The remaining two are equivalent clusters with residues from both subunits; these clusters correspond to the two equivalent sites of isomerase activity in the dimer. THEMATICS correctly identifies three major catalytic residues (E358, H389, and R273); the only known catalytic residue that is not predicted by THEMATICS is K519. As shown in Figure 1, THEMATICS predicts a total of eight second- and third-shell residues; most of them are located around the catalytic residues, suggesting these residues may have some role in catalysis. The rank-ordered residues in the top 10% of the POOL predictions obtained using THEMATICS and INTREPID features are listed in Table S3 of the Supporting Information. Again, there are some residues in the top 10% of the POOL rankings that are located far from the isomerase active site; these may be important for some other function of hPGI. However, all four known catalytic residues occur in the top 8% of the POOL rankings. The three major catalytic residues (E358, R273, and H389) are predicted within the top 3% of the POOL rankings. Indeed, H389 is ranked #1 of 558 residues; K519 is predicted by POOL to be #39 of 558 residues, in the top 7%. Like THEMATICS, POOL predicts many second- and third-shell residues, also suggesting that these residues have a role in catalysis. To test the contributions of second- and third-shell residues to catalysis, we chose the following nine residues for site-directed mutagenesis: H100, Y274, Y341, K362, N386, Q388, H396, E495, and D511. All nine of these residues are located in the second or third shells and are in the top 10% of the POOL rankings. All but one, D511, are in the top 8% of the POOL rankings. Distances of each residue to the bound ligand are in the 10−15 Å range and are listed in Table 1, along with the shell assignments and the closest first-shell residue. Y274, K362, N386, Q388, E495, and D511 belong to the second shell and H100, Y341, and H396 to the third shell. In addition to these, two more residues, N154 and S185, were chosen as negative controls because they are residues that are not predicted by POOL and hence are expected to have no significant role in catalysis. Both N154 and S185 are second-shell residues and have high sequence conservation scores. In Table 1, ConSurf[63] conservation scores are presented, normalized such that zero is the average for all the residues in a given protein and with scores expressed in units of standard deviation; the lowest score represents the most conserved position. Thus, a residue with a score of −1 is more conserved than the average residue in that protein sequence by one standard deviation.

Fraction of Total Residues Predicted, by Shell. For KSI, the residues predicted by THEMATICS constitute 22% of all first-shell residues, less than 2% of all second-shell residues, and none of the third-shell residues. For PGI, THEMATICS predicts 16% of the residues in the first shell, 7% of those in the second shell, and 3% of those in the third shell. A summary of the numbers of residues predicted by THEMATICS and

dx.doi.org/10.1021/bi201089v | *Biochemistry* 2011, 50, 9283–9295

Table 2. Steady-State Kinetic Parameters of ppKSI and Its Variants

KSI	shell	POOL percentile	k_{cat} ($\times 10^4$ s^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})	relative k_{cat}	relative k_{cat}/K_M
WT			3.6 ± 0.5	89 ± 12	4.1×10^8	1	1
M31A	second	74	1.8 ± 0.3	89 ± 18	2.0×10^8	0.49	0.49
E39Q	second	91	1.9 ± 0.4	110 ± 42	1.7×10^8	0.51	0.41
M84A	second	82	3.4 ± 0.9	81 ± 12	4.2×10^8	0.93	1
M105A	second	88	0.24 ± 0.03	44 ± 5	5.4×10^7	0.067	0.13

Table 3. Steady-State Kinetic Parameters of hPGI and Its Variants

PGI	shell	POOL percentile	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})	relative k_{cat}	relative k_{cat}/K_M
WT			500 ± 53	290 ± 48	1.7×10^6	1.00	1
H389L	first	99	<0.04	–	–	–	–
K362A	second	95	<0.04	–	–	–	–
Q388A	second	95	45 ± 6	500 ± 44	9.0×10^4	0.090	5.3×10^{-2}
E495Q	second	97	1.6 ± 0.1	230 ± 32	6.7×10^3	3.2×10^{-3}	3.9×10^{-3}
D511N	second	90	1.3 ± 0.8	180 ± 47	7.4×10^3	2.6×10^{-3}	4.4×10^{-3}
Y274F	second	96	240 ± 62	300 ± 50	0.79×10^6	0.47	0.46
N386A	second	95	470 ± 41	780 ± 300	0.61×10^6	0.94	0.36
H100L	third	96	0.6 ± 0.01	210 ± 41	2.7×10^3	1.2×10^{-3}	1.6×10^{-3}
H396L	third	99	21 ± 2	290 ± 51	7.2×10^4	0.042	4.2×10^{-2}
Y341F	third	98	340 ± 41	300 ± 97	1.1×10^6	0.67	0.65
N154Q	second	75	470 ± 52	1040 ± 130	0.45×10^6	0.93	0.26
S185A	second	75	360 ± 19	730 ± 57	0.49×10^6	0.72	0.29

POOL, sorted by shell, for KSI and PGI is given in Table S4 of the Supporting Information.

Effect of Second- and Third-Shell Residue Mutations on Catalysis. Results of kinetic assays for wild-type (WT) KSI and PGI and their variants are listed in Tables 2 and 3, respectively.

KSI. Among the four variants of KSI tested, M31A, E39Q, M84A, and M105A, the largest decrease in activity was observed for M105A, which showed a 15-fold decrease in k_{cat} and a 7.6-fold decrease in catalytic efficiency (k_{cat}/K_M). The other variants did not show significant decreases in activity, with <3-fold decreases in k_{cat} and in k_{cat}/K_M (Table 2). M105 is located behind catalytic residues Y16 and D103 with respect to the substrate. The catalytic role of both Y16 and D103 is to form hydrogen bonds with the C3 oxygen atom of the substrate and to stabilize the negative charge developed on the intermediate. Mutating M105 to A decreases the steric volume and may change the positions of the side chains of Y16 and D103, making them less available for hydrogen bonding. The crystal structure of this variant was determined to test for such structural changes in the region of the active site.

Structure of ppKSI M105A. The crystal structure of the second-shell variant ppKSI M105A was determined to identify any structural changes induced by the mutation. Details of the structure determination and refinement parameters are listed in Table 4. The electron density for ppKSI M105A was continuous throughout the entire model with the exception of the first two N-terminal residues and the last four C-terminal residues where no electron density was observed. The electron density for residues W92, N93, G94, and Q95 was initially very weak. This region was truncated, and subsequently, the backbone was rebuilt manually. As the phases improved, the side chain of N93 was built; no electron density appeared for the side chains of W92 and Q95 so they were modeled as Ala, and residue G94 was deleted from the model. The complete refined model includes one copy of the protein from residues 3–93 and residues 95–127 and 39 water molecules. The protein exhibits good geometry as determined by PROCHECK;[64] 98% of the residues are in the favored regions of the Ramachandran plot, with no outliers. The average overall temperature factor for the protein is 14.2 Å^2. After initial

Table 4. Summary of Crystal Parameters and Data Collection and Refinement Statistics for ppKSI M105A

Data Collection	
beamline	APS, GM/CA-CAT, ID-D
wavelength (Å)	0.95
space group	$C222_1$
cell constants	a = 35.1 Å
	b = 95.0 Å
	c = 72.3 Å
	$\alpha = \beta = \gamma = 90°$
total no. of reflections	153171
no. of unique reflections	26517
resolution limit (Å)	1.3 (1.3–1.35)[a]
completeness (%)	88.4 (65.2)[a]
redundancy	5.8 (3.1)[a]
$I/\sigma I$	23.1 (1.5)[a]
R_{merge} (%)	4.1 (38.1)[a]
Refinement	
resolution range (Å)	1.3–22.6
R_{free} test set size	1871
R_{cryst} (%)	20.9
R_{free} (%)	23.1
no. of atoms	
total	995
protein	956
water	39
overall B factor (Å^2)	14.2
rmsd	
bond lengths (Å)	0.010
bond angles (deg)	1.1

[a]Data for the highest-resolution bin are given in parentheses.

rounds of refinement, the side chains of residues Q7, E8, M13, E18, Q30, Q44, H48, E51, Q52, R58, Q59, R87, E89, M90, C97, E109, R112, T115, M116, and E122 had a significant amount of negative electron density, presumably as a result of radiation damage that occurred during data collection. For this reason, their occupancies were adjusted to a value that was consistent with the electron density. Radiation damage during data collection is a common occurrence in high-energy X-ray beams.[65] Extraneous electron density was observed around the side chain of C69, consistent with the oxidation product sulfinic acid (CDS). Residue C69 was modeled as a pair consisting of Cys and CDS with occupancies of 0.4 and 0.6, respectively.

The overall structure of ppKSI M105A is similar to that of wild-type ppKSI. Superimposing the coordinates of wild-type ppKSI (PDB entry 1OH0)[21] and ppKSI M105A (this work) yields an rmsd of 0.5 Å over 123 α-carbon atoms (Figure 3).

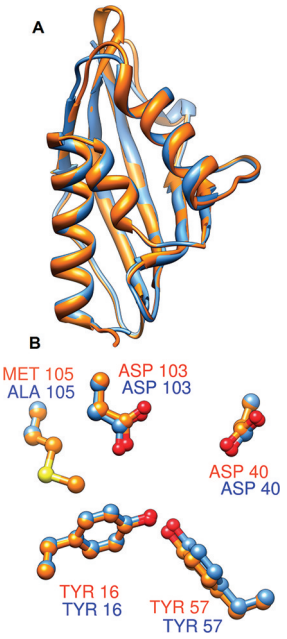


Figure 3. Crystal structure of the M105A variant of ppKSI. (A) Structural overlap of WT ppKSI (orange) and the M105A variant (blue). (B) Overlap of known catalytic residues and M105. Refinement statistics are listed in Table 4.

Despite the overall similarity between the wild type and the M105A variant ppKSI forms, two loop regions (residues G63, G64, and K65 and residues A92, N93, A95, and P96) show some differences. The loop region defined by residues A92–P96 has shifted by approximately 3 Å, while the loop region between residues G63 and K65 has shifted by approximately 2 Å in the structure of the M105A variant. Overall, the mutation at residue 105 does not alter the structure significantly, and therefore, the change in activity is most likely not caused by structural changes.

PGI. As described above, 11 second- and third-shell residues were chosen for mutagenesis and kinetics analysis to test for their participation in catalysis of hPGI. In addition, H389L, a POOL-predicted first-shell residue that had not been studied previously for hPGI, was included in the study. Kinetic parameters for wild-type PGI and variants were determined and are summarized in Table 3. Overall, four of the six predicted second-shell residues, and two of the three predicted third-shell residues, showed a significant effect on k_{cat}. The six hPGI variants that showed significant effects on k_{cat}, K362A, H100L, E495Q, D511N, H396L, and Q388A, were also analyzed using CD spectroscopy (Figure 4), and their thermal stabilities were determined (Table 5). The variants are discussed below in detail.

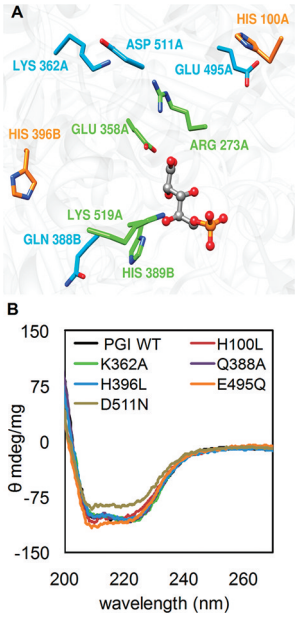


Figure 4. (A) Location of the six second-shell (blue) and third-shell (orange) residues of hPGI at which mutations resulted in significant effects on catalysis. First-shell residues are colored green, and the ligand is shown in ball-and-stick form. (B) CD spectra of the six second- and third-shell variants whose positions are shown in panel A.

Table 5. Melting Temperatures of WT hPGI and hPGI Variants

	T_m (°C)
WT	54.8
Q388A	50.8
H396L	48.4
E495Q	54.5
D511N	48.4
H100L	49.9
K362A	48.9

The first-shell variant H389L showed a significant effect on catalysis, as expected. No activity was detected for this variant, corresponding to a reduction in k_{cat} of more than 10^4-fold compared to the wild-type value. The mutation to Ala at the corresponding position in the *B. stearothermophilus* enzyme, H311, shows a 10^3-fold reduction in k_{cat}.[66]

K362 is predicted by POOL to be important for catalysis [top 5% (Table S3 of the Supporting Information)]. K362 is a second-shell residue located close to known catalytic residues E358 and R273 (Figure 4). The K362A mutation resulted in a drastic decrease in the activity of hPGI. Even at a very high enzyme concentration (1 μM), the activity was almost undetectable; hence, individual kinetic parameters could not be determined, and only the observed limit of detection of the activity is reported. K362 is located adjacent to α-helix 17 on which the catalytic residue E358 is located, forming a salt bridge with another second-shell residue, D511 (vide infra). Mutating K362 to Ala removes the charge at this position and also

dx.doi.org/10.1021/bi201089v | *Biochemistry* 2011, 50, 9283–9295

disrupts the salt bridge between K362 and D511. One reason for the drastic decrease in activity could be the loss of this salt bridge, causing α-helix 17 to be more mobile, such that E358 is less likely to be in the proper position for catalysis relative to the other essential residues. The CD spectrum of hPGI K362A can be superimposed on that of wild-type hPGI, suggesting that the effect on catalysis is not due to major overall structural changes (Figure 4), while the stability is decreased by ~6 °C relative to that of wild-type hPGI (Table 5). Mean field analysis of the catalytic base E358 in wild-type hPGI shows that K362 is involved in strong electrostatic interaction with E358 and shifts its pK_a by more than 2 pH units. The absence of the Coulomb interaction in the K362A variant could be another reason for the drastic decrease in activity.

H100 is a third-shell residue located 12 Å from the active site (Figure 4). It is among the top 5% of POOL-ranked residues (Table S3 of the Supporting Information) and hence is predicted to be important for catalysis. Mutating H100 to L had no effect on K_M but resulted in a >800-fold decrease in k_{cat} relative to that of wild-type hPGI (Table 3). The H100L mutation decreased the stability of the enzyme by ~5 °C (Table 5). H100 interacts with several second-shell residues, mainly with E495 (Figure 4) and Y274; the E495Q mutation also has a significant effect on catalysis (vide infra). It is interesting to note that while H100 is a third-shell residue that interacts with the catalytic residues only indirectly, the H100L mutation has a larger effect on the catalytic efficiency than some of the second-shell residues that interact directly with the catalytic residues. The CD spectrum of the H100L variant is similar to that of wild-type PGI (Figure 4); hence, the variant is folded, and its effect on catalysis is probably not due to a major overall structural change. Mutating H100 to L does change the electrostatic and hydrogen bonding network around the active site, and these changes are the likely source of the decrease in catalytic efficiency by 3 orders of magnitude.

E495 is a second-shell residue predicted to be in the top 3% of the POOL ranking (Table S3 of the Supporting Information) and is therefore thought to be important for catalysis. The conservative E495Q mutation in hPGI resulted in a >300-fold decrease in k_{cat}, but the K_M is essentially the same as that of wild-type PGI (Table 3). Also, the E495Q mutation has no effect on the thermal stability of the enzyme, as measured by thermofluor (Table 5), and the CD spectrum of this variant is very similar to that of wild-type hPGI (Figure 4). E495 is located behind R273 with respect to the substrate (Figure 4); the side chain of E495 interacts with the catalytic residue R273 and is also located close to H100 (vide supra), R104, and W277. Mutating E495 to Q eliminates the ability to form a charge at this position but preserves the dipolar and hydrogen bonding interactions, as well as the steric bulk. Therefore, the observed 250-fold reduction in catalytic efficiency for the E495Q mutation suggests that E495 contributes to electrostatic properties within the active site and that these contributions are important to catalysis; this is discussed further below.

In hPGI, D511 is a second-shell residue located behind catalytic residues R273 and E358 with respect to the substrate (Figure 4). It is ranked in the top 10% by POOL (Table S3 of the Supporting Information). The D511N variant showed a decrease in k_{cat} of >300-fold (Table 3) and a decrease in the stability of the enzyme by 6 °C (Table 5). The CD spectrum of the D511N variant is slightly changed compared to that of wild-type PGI (Figure 4). D511 forms a salt bridge with K362; disruption of this salt bridge is probably a major reason for the decreased activity and decreased stability.

H396 is a third-shell residue located near the interface of the two subunits. It is near the top of the POOL rankings (Table S3 of the Supporting Information) and is predicted to be important for catalysis. The H396L variant shows a 24-fold lower k_{cat} than the wild type, while its K_M remains essentially the same as that of wild-type PGI (Table 3). This mutation resulted in a 6 °C decrease in T_m (Table 5), whereas the CD spectrum is essentially superimposable with that of wild-type hPGI (Figure 4). The side chain of H396 forms a hydrogen bond with the backbone oxygen atom of Y391, stabilizing α-helix 29. This, together with the observed loss of stability, suggests that structural and dynamic effects are a part of the role that H396 plays in catalysis.

Q388 is among the top 8% of POOL residues (Table S3 of the Supporting Information) and is predicted to be important for catalysis. Analysis of the Q388A variant of PGI resulted in a 10-fold decrease in k_{cat}, a 19-fold decrease in the catalytic efficiency (Table 3), and a 4 °C decrease in T_m (Table 5). There is a slight increase in the K_M of the Q388A variant compared to that of wild-type PGI. The CD spectrum of the Q388 variant is essentially superimposable with that of wild-type hPGI (Figure 4). Q388 is located next to catalytic residue H389. The side chain oxygen atom of Q388 is approximately 5 Å from the ring nitrogen atom (Nε) of His389, but Q388 also forms hydrogen bonds with N428, Q432, and Y341. Disrupting these hydrogen bonds by mutating Q388 to A could shift the side chain position of H389 away from the substrate and therefore lead to the decrease in activity and stability.

Effects of Second- and Third-Shell Residues on the pK_a Values of Active Site Residues. The location, environment, and electrostatic interactions of an ionizable residue can affect the pK_a of that particular residue. A large shift in the pK_a of active site residues in variants compared to the wild type suggests that there is a change in the electric field at the active site, which can then lead to altered activity. Continuum electrostatics calculations were performed to predict the pK_a values of active site residues. For each variant, a homology model was built with YASARA (http://www.yasara.org) using the corresponding wild-type enzyme as the template structure. MCCE[50] calculations were performed on the homology models, in a fashion similar to that used for the wild type. While pK_a is the best metric for gaining insight into the changes in the electrostatic interactions and electric field environment of the active site, the predicted or computed pK_a can be difficult to interpret because of the perturbed titration curves of active site residues.[7] For this reason, titration curve moments[46,67] were also computed for the active site residues of each second- or third-shell variant and wild-type enzyme using THEMATICS. Curve moments (especially μ_3 and μ_4) can provide valuable information about the electrostatic interactions of a given residue,[67] in addition to the pK_a.

MCCE-predicted pK_a values for the catalytic base (D40 in ppKSI and E358 is hPGI) are listed in Table 6, and the effects of second- and third-shell residue mutations on the pK_a of the catalytic base E358 of hPGI, as computed by both MCCE and UHBD, are listed in Table 7. We note that the computed pK_a values of active site residues have inherently large errors, but the relative pK_a values for the wild type and single-point variants are more meaningful. Our analysis therefore focuses on the shifts in the computed pK_a of the catalytic base upon mutation of a nearby residue. The mean field approximation

Table 6. Predicted pK_a Values for Catalytic Residues in ppKSI and hPGI[a]

catalytic residue	MCCE-predicted pK_a	residues influencing the pK_a
D40 (KSI)	4.1	–
E358 (PGI)	−7.8	R96, R104, D161, **K211**, E217, **R273**, **Q354**, D356, K362, E495, K497, D511, **K519**, E382, **H389**

[a]Residues contributing more than 1 pK_a unit are listed in the rightmost column. Bold residues belong to the first shell; underlined residues belong to the second shell.

performed at a pH equal to the pK_a of the residue in question allows identification of the other residues that contribute to the free energy of protonation of that residue. Those residues that cause a pK_a shift of the catalytic residue by more than 0.5 pK_a unit are listed in Table 6. As shown in Table 6, the predicted pK_a for D40 in ppKSI is not shifted from its free solution value, and there are no other residues that influence the pK_a of D40. On the other hand, the pK_a of hPGI E358 is shifted, and there are many other residues that influence the pK_a, as identified through the mean field approximation. These results strongly suggest that second- and third-shell residues are important in the control of the pK_a of E358 and that this is one mechanism by which remote residues influence catalysis in hPGI. Both methods predict a very low pK_a and a highly perturbed titration curve for E358. A major shift in the pK_a of E358 relative to that of the wild type is observed for variants E495Q and K362A. Note that even though there are differences in the pK_a values, as computed by MCCE and UHBD, the trend is similar; i.e., according to both methods, K362A and E495Q are the variants with the greatest shifts in the computed pK_a values of the catalytic residue E358 compared to that of the wild type. Also, the curve moment μ_4 is reduced ~2-fold for these two variants. Clearly, the computed titration curve of the catalytic base differs the most from that in WT in those variants that have the greatest effect on catalysis. After K362A and E495Q, the H100L variant showed the next largest shift in the pK_a of E358 according to MCCE. However, the titration curves of E358 are similar in shape for the H100L variant and wild-type hPGI; both exhibit highly perturbed theoretical titration behavior, as shown by the large μ_4 values.

DISCUSSION

Understanding the catalytic efficiency of enzymes is a fundamental question in biochemistry and a subject of intense study, in part because of the many current and potential applications in medicine and industry. Linus Pauling[68] was one of the first to propose an explanation for the catalytic power of enzymes, in terms of transition-state stabilization by enzyme active sites. Since then, several other theories that describe the origins of enzyme catalysis have been proposed with terms such as electrostatic preorganization,[69] low-barrier hydrogen bonds,[70] near-attack conformations,[71] entropic effects, dynamics,[72] and quantum tunneling.[73] Rapid progress in experimental and computational investigations has enhanced our understanding of how enzymes affect catalysis. One example of an application of this understanding is the recent report of the design of artificial enzymes using computational methods.[74,75]

Despite the major progress and breakthroughs in the field of enzymology, designed or engineered enzyme catalysts generally do not match naturally evolved enzyme catalysts in terms of efficiency. For example, the efficiency of the artificially designed Kemp eliminase[74] is very low compared to those of naturally occurring enzymes. When this designed enzyme was subjected to directed evolution, the resulting protein with multiple mutations showed improved catalytic efficiency by >200-fold, and most of the required mutations were not in direct contact with the reacting substrate molecules but located away from the catalytic residues.

Other work has shown that mutations of second- and third-shell residues can have dramatic effects on enzyme activity. Specifically, our previous work with nitrile hydratase showed that mutations of either of two THEMATICS-predicted second-shell residues, which are involved in a hydrogen bonding interaction, resulted in a 2-order of magnitude decrease in catalytic efficiency, while in the case of one of the variants, the crystal structure was nearly identical to that of the wild type.[9] Mutation of the second-shell residue D270 in mandalate racemase led to a 10^4-fold decrease in activity that may be due to the loss of a specific rotameric state of the catalytic H297, stabilized by D270, and also to a shift in the pK_a of H297;[76] both D270 and H297 are predicted by POOL and are in the top 5% of the POOL ranking. In another example, mutation of second-shell residue Y208 and third-shell residue E129 of triosephosphate isomerase (26.5 kDa) resulted in 2400- and 30-fold reductions in catalytic efficiency, respectively, which were attributed to destabilization of a "closed," catalytically proficient conformation of the enzyme.[77] These results supply further evidence that remote residues can help in optimizing the properties of active site residues and should be taken into account in protein design. Moreover, Morley and Kazlauskas[78] in a literature survey noticed that mutations at

Table 7. Effects of Remote Residue Mutations on the pK_a of hPGI E358

	MCCE		UHBD		curve moments			
	pK_a	ΔpK_a	pK_a	ΔpK_a	μ_3	$\Delta\mu_3$	μ_4	$\Delta\mu_4$
WT	−7.8		−3.20		12.5		186.7	
H100L	−11.7	−3.91	−3.76	−0.56	16.4	3.93	198.5	11.8
Y274F	−10.8	−2.98	−3.93	−0.73	8.50	−3.99	145.3	−41.4
Y341F	−9.71	−1.91	−3.97	−0.77	10.6	−1.86	153.9	−32.8
K362A	−13.1	−5.30	−5.28	−2.08	8.03	−4.46	96.25	−90.5
N386A	−10.8	−3.00	−4.29	−1.09	8.21	−4.28	140.7	−46.0
Q388A	−11.1	−3.29	−3.92	−0.72	13.0	0.49	192.6	5.86
H396L	−8.80	−1.00	−3.63	−0.43	6.13	−6.36	122.0	−64.67
E495Q	−13.7	−5.91	−4.95	−1.75	7.32	−5.17	103.2	−83.5
D511N	−10.2	−2.37	−4.86	−1.66	10.5	−2.03	123.0	−63.7

remote residues could be sometimes as effective as mutations close to the active site in improving the catalytic activity and thermal stability of enzymes.

We note that for PGI, the second- and third-shell residues with the greatest effect on the catalytic efficiency, K362, H100, E495, and D511, are those located behind E358 or R273. These results suggest that these four residues are involved in imparting the necessary chemical and electrostatic properties to the catalytic base E358 and to R273 and in positioning them correctly. Thus, K362, H100, E495, and D511 are part of a spatially extended network that facilitates proton transfer and stabilization of the intermediate. These residues appear to be more important than those behind H389 and K519 that are involved in the more facile ring-opening and ring-closing steps.

There are many possible mechanisms by which mutations at remote residues can affect the catalytic process, including changes in the hydrogen bonding network, changes in the electric field at the active site, quantum mechanical effects, dynamic effects, or structural effects. In the case of hPGI, these results suggest that multiple such effects are in play. The MCCE- and UHBD-computed pK_a values of the catalytic base of wild-type hPGI and variants indicate that K362 and E495 have the greatest effect on the pK_a of the catalytic base. Mutations of these two second-shell residues are also observed to have the greatest decrease in k_{cat}/K_M, suggesting that control of the protonation behavior of the catalytic base is one of the functional roles of K362 and E495.

On the other hand, the minimal participation of second- and third-shell residues in KSI suggests that the electrostatic couplings within its first shell are sufficiently strong that residues outside the first shell are not required. This is plausible, given the highly hydrophobic active site of KSI; the very low effective dielectric constant within the active site of KSI would correspond to stronger Coulomb forces than in the more hydrophilic PGI active site. The very strong coupling between hydrogen bonds observed within the active site of KSI by proton NMR and deuterium isotope effects[79] constitutes further evidence of strong interactions within the first shell of KSI.

The catalytic importance of K362 and D511 observed here for hPGI also lends support to the suggestions made previously[37,38] that conformational changes, including motion of α-helix 23, are important for catalysis.[34,38,39] On the basis of structural changes observed with bound ligands in rabbit PGI, Lee et al. reported[38] that motion of "a helix containing amino acids 513−520 moves in toward the substrate to form additional hydrogen bonds with the substrate" as a key part of the catalytic mechanism, wherein substrate binding can be described by an induced fit model. Later, Davies and Muirhead made a similar inference from rabbit PGI structures with different bound ligands and suggested that multiple conformational shifts are required for catalysis; in particular, they inferred that, in a given turnover cycle, α-helix 23 moves toward the substrate binding pocket before isomerization and then away from the pocket following isomerization, so that the catalytic K519 (K519 in human and K518 in rabbit) is brought into contact with the substrate and then retreats to release the product.[37] In hPGI, one might view α-helix 23 as part of an extended helical structure starting with a 3/10-turn at residues 511−513, followed by an α-helix running from residue 514 to 529, with a slight bend at residue 520 and with the catalytic K519 located in the middle. At the N-terminal end of this extended helix is D511, which forms a salt bridge with K362. The catalytic importance of K362 and D511 observed here, together with the previous structural observations of Lee et al.[38] and Davies and Muirhead,[37] suggests that the K362−D511 couple is vital to the positioning of the catalytic K519 and also to the conformational changes that allow the entry of substrate and subsequent release of product. Mutation of K362 or D511 to residues that cannot form the salt bridge thus may disrupt the sampling of conformational states that is needed for efficient PGI catalysis. We note that THEMATICS predicts that both the charged and uncharged states exist with significant populations in equilibrium for both K362 and D511. Therefore, the K362−D511 salt bridge in wild-type hPGI is able to break to allow motion of the extended helix and then to re-form to relock it into position.

In contrast, the rapid catalysis by KSI that approaches the diffusion-controlled limit has been described in terms of electrostatic preorganization, wherein the dipoles within the active site are already substantially prearranged to stabilize the transition state[10] and the substrate binds according to a lock and key model. While the relatively small size of KSI (14.4 kDa) is one possible explanation for the low level of participation by residues outside the first shell, we note that there are other relatively small enzymes in which a remote residue has been shown to be critical for catalysis, such as T80 in HIV protease (a dimer of 11 kDa subunits),[80,81] where T80 appears to be essential for motion of a hinge region that allows access to the active site.

This report provides evidence that enzyme active sites can exist either as compact or as multilayered assemblies of residues, although more examples are needed to understand more generally the extent and prevalence of multilayered enzyme active sites and the roles of remote residues. Specifically, ppKSI has a very compact active site consisting mostly of first-shell residues, whereas hPGI has an extended multilayered active site with significant participation by certain second- and third-shell residues. These results also suggest that second- and third-shell residues affect catalysis in PGI through electrostatic coupling with the catalytic residues in the first shell and also by participation in conformational changes that are important to the catalytic mechanism. Perhaps more importantly, these results show that computational methods, specifically THEMATICS and POOL, are robust enough to identify functionally important residues that are located away from the active site.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

THEMATICS and POOL predictions for ppKSI and hPGI (Tables S1−S4). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*M.J.O.: phone, (617) 373-2856; fax, (617) 373-8795; e-mail, m.ondrechen@neu.edu. P.J.B.: phone, (617) 373-2865; fax, (617) 373-8795; e-mail, beuning@neu.edu.

Funding

We acknowledge financial support from the National Science Foundation (MCB-0843603 to M.J.O. and P.J.B. and Career Award MCB-0845033 to P.J.B.), the National Institutes of Health (GM32415 to D.R.), and a New Faculty Award from the Camille and Henry Dreyfus Foundation (P.J.B.). S.S. was

supported partially by a Dissertation Completion Fellowship from the Northeastern University Office of the Provost. P.J.B. is a Cottrell Scholar of the Research Corporation for Science Advancement.

■ ACKNOWLEDGMENTS

We thank Prof. April Gu (Civil and Environmental Engineering, Northeastern University) for use of the real-time PCR instrument and Prof. John Engen and Dr. Thomas Wales (Northeastern University) for use of the CD instrument. We thank Prof. Kwan Yong Choi (Pohang University of Science and Technology, Pohang, South Korea) and Prof. Menghsiao Meng (National Chung Hsing University, Taichung, China) for providing plasmids expressing ppKSI and hPGI, respectively. We thank Prof. Connie Jeffery (University of Illinois, Chicago, IL) and the reviewers for valuable comments.

■ ABBREVIATIONS

KSI, ketosteroid isomerase; ppKSI, *P. putida* ketosteroid isomerase; PGI, phosphoglucose isomerase; hPGI, human phosphoglucose isomerase; LPC, Ligand-Protein Contacts; CSU, Contacts of Structural Units; MCCE, MultiConformer Continuum Electrostatics; UHBD, University of Houston Brownian Dynamics; EDTA, ethylenediaminetetraacetic acid; HEPES, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid; DTT, dithiothreitol; PDB, Protein Data Bank; PMSF, phenylmethanesulfonyl fluoride; rmsd, root-mean-square deviation; SDM, site-directed mutagenesis.

■ REFERENCES

(1) Wolfenden, R., and Snider, M. J. (2001) The Depth of Chemical Time and the Power of Enzymes as Catalysts. *Acc. Chem. Res. 34*, 938−945.

(2) Zhang, X., and Houk, K. N. (2005) Why Enzymes Are Proficient Catalysts: Beyond the Pauling Paradigm. *Acc. Chem. Res. 38*, 379−385.

(3) Kirk, O., Borchert, T. V., and Fuglsang, C. C. (2002) Industrial enzyme applications. *Curr. Opin. Biotechnol. 13*, 345−351.

(4) Schmid, A., Dordick, J. S., Hauer, B., Kiener, A., Wubbolts, M., and Witholt, B. (2001) Industrial biocatalysis today and tomorrow. *Nature 409*, 258−268.

(5) Ringe, D., and Petsko, G. A. (2008) How Enzymes Work. *Science 320*, 1428−1429.

(6) Benkovic, S. J., and Hammes-Schiffer, S. (2003) A Perspective on Enzyme Catalysis. *Science 301*, 1196−1202.

(7) Ondrechen, M. J., Clifton, J. G., and Ringe, D. (2001) THEMATICS: A simple computational predictor of enzyme function from structure. *Proc. Natl. Acad. Sci. U.S.A. 98*, 12473−12478.

(8) Tong, W. X., Wei, Y., Murga, L. F., Ondrechen, M. J., and Williams, R. J. (2009) Partial Order Optimum Likelihood (POOL): Maximum Likelihood Prediction of Protein Active Site Residues Using 3D Structure and Sequence Properties. *PLoS Comput. Biol. 5*, 15.

(9) Brodkin, H. R., Novak, W. R. P., Milne, A. C., D'Aquino, J. A., Karabacak, N. M., Goldberg, I. G., Agar, J. N., Payne, M. S., Petsko, G. A., Ondrechen, M. J., and Ringe, D. (2011) Evidence of the Participation of Remote Residues in the Catalytic Activity of Co-Type Nitrile Hydratase from *Pseudomonas putida*. *Biochemistry 50*, 4923−4935.

(10) Kamerlin, S. C. L., Sharma, P. K., Chu, Z. T., and Warshel, A. (2010) Ketosteroid isomerase provides further support for the idea that enzymes work by electrostatic preorganization. *Proc. Natl. Acad. Sci. U.S.A. 107*, 4075−4080.

(11) Warshel, A., Sharma, P. K., Chu, Z. T., and Åqvist, J. (2007) Electrostatic Contributions to Binding of Transition State Analogues Can Be Very Different from the Corresponding Contributions to Catalysis: Phenolates Binding to the Oxyanion Hole of Ketosteroid Isomerase. *Biochemistry 46*, 1466−1476.

(12) Chakravorty, D. K., Soudackov, A. V., and Hammes-Schiffer, S. (2009) Hybrid Quantum/Classical Molecular Dynamics Simulations of the Proton Transfer Reactions Catalyzed by Ketosteroid Isomerase: Analysis of Hydrogen Bonding, Conformational Motions, and Electrostatics. *Biochemistry 48*, 10608−10619.

(13) Kraut, D. A., Sigala, P. A., Pybus, B., Liu, C. W., Ringe, D., Petsko, G. A., and Herschlag, D. (2006) Testing electrostatic complementarity in enzyme catalysis: Hydrogen bonding in the ketosteroid isomerase oxyanion hole. *PLoS Biol. 4*, 501−519.

(14) Kraut, D. A., Churchill, M. J., Dawson, P. E., and Herschlag, D. (2009) Evaluating the Potential for Halogen Bonding in the Oxyanion Hole of Ketosteroid Isomerase Using Unnatural Amino Acid Mutagenesis. *ACS Chem. Biol. 4*, 269−273.

(15) Kraut, D. A., Sigala, P. A., Fenn, T. D., and Herschlag, D. (2010) Dissecting the paradoxical effects of hydrogen bond mutations in the ketosteroid isomerase oxyanion hole. *Proc. Natl. Acad. Sci. U.S.A. 107*, 1960−1965.

(16) Ha, N.-C., Choi, G., Choi, K. Y., and Oh, B.-H. (2001) Structure and enzymology of Δ^5−3-ketosteroid isomerase. *Curr. Opin. Struct. Biol. 11*, 674−678.

(17) Kuliopulos, A., Talalay, P., and Mildvan, A. S. (1990) Combined effects of two mutations of catalytic residues on the ketosteroid isomerase reaction. *Biochemistry 29*, 10271−10280.

(18) Hawkinson, D. C., Eames, T. C. M., and Pollack, R. M. (1991) Energetics of 3-oxo-Δ^5-steroid isomerase: Source of the catalytic power of the enzyme. *Biochemistry 30*, 10849−10858.

(19) Owen, R. W., Mason, A. N., and Bilton, R. F. (1983) The degradation of cholesterol by *Pseudomonas sp.* NCIB 10590 under aerobic conditions. *J. Lipid Res. 24*, 1500−1511.

(20) Cho, H.-S., Choi, G., Choi, K. Y., and Oh, B.-H. (1998) Crystal Structure and Enzyme Mechanism of Δ^5−3-Ketosteroid Isomerase from *Pseudomonas testosteroni*. *Biochemistry 37*, 8325−8330.

(21) Kim, S. W., Cha, S.-S., Cho, H.-S., Kim, J.-S., Ha, N.-C., Cho, M.-J., Joo, S., Kim, K. K., Choi, K. Y., and Oh, B.-H. (1997) High-Resolution Crystal Structures of Δ^5−3-Ketosteroid Isomerase with and without a Reaction Intermediate Analogue. *Biochemistry 36*, 14030−14036.

(22) Kim, S., and Choi, K. (1995) Identification of active site residues by site-directed mutagenesis of Δ^5−3-ketosteroid isomerase from *Pseudomonas putida* biotype B. *J. Bacteriol. 177*, 2602−2605.

(23) Kim, S., Joo, S., Choi, G., Cho, H., Oh, B., and Choi, K. (1997) Mutational analysis of the three cysteines and active-site aspartic acid 103 of ketosteroid isomerase from *Pseudomonas putida* biotype B. *J. Bacteriol. 179*, 7742−7747.

(24) Kim, D.-H., Jang, D. S., Nam, G. H., Choi, G., Kim, J.-S., Ha, N.-C., Kim, M.-S., Oh, B.-H., and Choi, K. Y. (2000) Contribution of the Hydrogen-Bond Network Involving a Tyrosine Triad in the Active Site to the Structure and Function of a Highly Proficient Ketosteroid Isomerase from *Pseudomonas putida* Biotype B. *Biochemistry 39*, 4581−4589.

(25) Choi, G., Ha, N. C., Kim, S. W., Kim, D. H., Park, S., Oh, B. H., and Choi, K. Y. (2000) Asp-99 donates a hydrogen bond not to Tyr-14 but to the steroid directly in the catalytic mechanism of Δ^5-3-ketosteroid isomerase from *Pseudomonas putida* biotype B. *Biochemistry 39*, 903−909.

(26) Jeffery, C. J. (1999) Moonlighting proteins. *Trends Biochem. Sci. 24*, 8−11.

(27) Faik, P., Walker, J. I. H., Redmill, A. A. M., and Morgan, M. J. (1988) Mouse glucose-6-phosphate isomerase and neuroleukin have identical 3′ sequences. *Nature 332*, 455−456.

(28) Chaput, M., Claes, V., Portetelle, D., Cludts, I., Cravador, A., Burny, A., Gras, H., and Tartar, A. (1988) The neurotrophic factor neuroleukin is 90% homologous with phosphohexose isomerase. *Nature 332*, 454−455.

(29) Watanabe, H., Takehana, K., Date, M., Shinozaki, T., and Raz, A. (1996) Tumor Cell Autocrine Motility Factor Is the Neuroleukin/Phosphohexose Isomerase Polypeptide. *Cancer Res. 56*, 2960−2963.

(30) Xu, W., Seiter, K., Feldman, E., Ahmed, T., and Chiao, J. (1996) The differentiation and maturation mediator for human myeloid

leukemia cells shares homology with neuroleukin or phosphoglucose isomerase. *Blood 87*, 4502−4506.
(31) Baughan, M. A., Valentine, W. N., Paglia, D. E., Ways, P. O., Simons, E. R., and Demarsh, Q. B. (1968) Hereditary Hemolytic Anemia Associated with Glucosephosphate Isomerase (GPI) Deficiency a New Enzyme Defect of Human Erythrocytes. *Blood 32*, 236−249.
(32) Beutler, E., West, C., Britton, H. A., Harris, J., and Forman, L. (1997) Glucosephosphate Isomerase (GPI) Deficiency Mutations Associated with Hereditary Nonspherocytic Hemolytic Anemia (HNSHA). *Blood Cells, Mol., Dis. 23*, 402−409.
(33) Yanagawa, T., Funasaka, T., Tsutsumi, S., Watanabe, H., and Raz, A. (2004) Novel roles of the autocrine motility factor/phosphoglucose isomerase in tumor malignancy. *Endocr.-Relat. Cancer 11*, 749−759.
(34) Arsenieva, D., and Jeffery, C. J. (2002) Conformational Changes in Phosphoglucose Isomerase Induced by Ligand Binding. *J. Mol. Biol. 323*, 77−84.
(35) Jeffery, C. J., Bahnson, B. J., Chien, W., Ringe, D., and Petsko, G. A. (2000) Crystal Structure of Rabbit Phosphoglucose Isomerase, a Glycolytic Enzyme That Moonlights as Neuroleukin, Autocrine Motility Factor, and Differentiation Mediator. *Biochemistry 39*, 955−964.
(36) Jeffery, C. J., Hardre, R., and Salmon, L. (2001) Crystal structure of rabbit phosphoglucose isomerase complexed with 5-phospho-D-arabinonate identifies the role of Glu357 in catalysis. *Biochemistry 40*, 1560−1566.
(37) Davies, C., and Muirhead, H. (2003) Structure of native phosphoglucose isomerase from rabbit: Conformational changes associated with catalytic function. *Acta Crystallogr. D59*, 453−465.
(38) Lee, J. H., Chang, K. Z., Patel, V., and Jeffery, C. J. (2001) Crystal Structure of Rabbit Phosphoglucose Isomerase Complexed with Its Substrate D-Fructose 6-Phosphate. *Biochemistry 40*, 7799−7805.
(39) Read, J., Pearce, J., Li, X., Muirhead, H., Chirgwin, J., and Davies, C. (2001) The crystal structure of human phosphoglucose isomerase at 1.6 Å resolution: Implications for catalytic mechanism, cytokine activity and haemolytic anaemia. *J. Mol. Biol. 309*, 447−463.
(40) Solomons, J. T. G., Zimmerly, E. M., Burns, S., Krishnamurthy, N., Swan, M. K., Krings, S., Muirhead, H., Chirgwin, J., and Davies, C. (2004) The crystal structure of mouse phosphoglucose isomerase at 1.6 Å resolution and its complex with glucose 6-phosphate reveals the catalytic mechanism of sugar ring opening. *J. Mol. Biol. 342*, 847−860.
(41) Chou, C.-C., Sun, Y.-J., Meng, M., and Hsiao, C.-D. (2000) The Crystal Structure of Phosphoglucose Isomerase/Autocrine Motility Factor/Neuroleukin Complexed with Its Carbohydrate Phosphate Inhibitors Suggests Its Substrate/Receptor Recognition. *J. Biol. Chem. 275*, 23154−23160.
(42) Sun, Y.-J., Chou, C.-C., Chen, W.-S., Wu, R.-T., Meng, M., and Hsiao, C.-D. (1999) The crystal structure of a multifunctional protein: Phosphoglucose isomerase/autocrine motility factor/neuroleukin. *Proc. Natl. Acad. Sci. U.S.A. 96*, 5412−5417.
(43) Rose, I. A. (2006) Mechanism of the Aldose-Ketose Isomerase Reactions. In *Advances in Enzymology and Related Areas of Molecular Biology*, pp 491−517, John Wiley & Sons, Inc., New York.
(44) Davies, C., Muirhead, H., and Chirgwin, J. (2003) The structure of human phosphoglucose isomerase complexed with a transition-state analogue. *Acta Crystallogr. D59*, 1111−1113.
(45) Krieger, E., Koraimann, G., and Vriend, G. (2002) Increasing the precision of comparative models with YASARA NOVA: A self-parameterizing force field. *Proteins: Struct., Funct., Genet. 47*, 393−402.
(46) Wei, Y., Ko, J., Murga, L. F., and Ondrechen, M. J. (2007) Selective prediction of interaction sites in protein structures with THEMATICS. *BMC Bioinf. 8*, 15.
(47) Sankararaman, S., Kolaczkowski, B., and Sjolander, K. (2009) INTREPID: A web server for prediction of functionally important residues by evolutionary analysis. *Nucleic Acids Res. 37*, W390−W395.
(48) Somarowthu, S., Yang, H., Hildebrand, D. G. C., and Ondrechen, M. J. (2011) High-performance prediction of functional residues in proteins with machine learning and computed input features. *Biopolymers 95*, 390−400.
(49) Sobolev, V., Sorokine, A., Prilusky, J., Abola, E., and Edelman, M. (1999) Automated analysis of interatomic contacts in proteins. *Bioinformatics 15*, 327−332.
(50) Mao, J. J., Song, Y. F., and Gunner, M. (2004) Multi-Conformation Continuum Electrostatics (MCCE): An efficient, accurate program for calculation of pK_as and Ems in proteins. *Biophys. J. 86*, 633A.
(51) Song, Y., Mao, J., and Gunner, M. R. (2009) MCCE2: Improving protein p*K*a calculations with extensive side chain rotamer sampling. *J. Comput. Chem. 30*, 2231−2247.
(52) Madura, J. D., Briggs, J. M., Wade, R. C., Davis, M. E., Luty, B. A., Ilin, A., Antosiewicz, J., Gilson, M. K., Bagheri, B., Scott, L. R., and McCammon, J. A. (1995) Electrostatics and diffusion of molecules in solution: Simulations with the University of Houston Brownian Dynamics program. *Comput. Phys. Commun. 91*, 57−95.
(53) Gilson, M. K. (1993) Multiple-site titration and molecular modeling: Two rapid methods for computing energies and forces for ionizable groups in proteins. *Proteins: Struct., Funct., Genet. 15*, 266−282.
(54) Lin, H. Y., Kao, Y. H., Chen, S. T., and Meng, M. (2009) Effects of inherited mutations on catalytic activity and structural stability of human glucose-6-phosphate isomerase expressed in *Escherichia coli*. *Biochim. Biophys. Acta 1794*, 315−323.
(55) Kim, S. W., Kim, C. Y., Benisek, W. F., and Choi, K. Y. (1994) Cloning, nucleotide-sequence, and overexpression of the gene coding for Δ^5-3-ketosteroid isomerase from *Pseudomonas putida* biotype-b. *J. Bacteriol. 176*, 6672−6676.
(56) Smith, S. B., Richards, J. W., and Benisek, W. F. (1980) Kinetics of isomerization of 5-androstene-3,17-dione catalyzed by Δ^5−3 keto steroid isomerase EC-5.3.3.1 from *Pseudomonas putida*. *J. Biol. Chem. 255*, 2685−2689.
(57) Ollivierre, J. N., Sikora, J. L., and Beuning, P. J. (2011) The Dimeric SOS Mutagenesis Protein UmuD Is Active as a Monomer. *J. Biol. Chem. 286*, 3607−3617.
(58) Ericsson, U. B., Hallberg, B. M., DeTitta, G. T., Dekker, N., and Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem. 357*, 289−298.
(59) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307−326.
(60) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D66*, 213−221.
(61) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126−2132.
(62) Nam, G. H., Jang, D. S., Cha, S. S., Lee, T. H., Kim, D. H., Hong, B. H., Yun, Y. S., Oh, B. H., and Choi, K. Y. (2001) Maintenance of α-helical structures by phenyl rings in the active-site tyrosine triad contributes to catalysis and stability of ketosteroid isomerase from *Pseudomonas putida* biotype B. *Biochemistry 40*, 13529−13537.
(63) Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., Pupko, T., and Ben-Tal, N. (2005) ConSurf: The projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res. 33*, W299−W302.
(64) Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) Procheck: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr. 26*, 283−291.
(65) Nave, C. (1995) Radiation Damage in Protein Crystallography. *Radiat. Phys. Chem. 45*, 483−490.
(66) Meng, M., Chane, T.-L., Sun, Y.-J., and Hsiao, C.-D. (1999) Probing the location and function of the conserved histidine residue of

phosphoglucose isomerase by using an active site directed inhibitor N-bromoacetylethanolamine phosphate. *Protein Sci. 8*, 2438−2443.
(67) Ko, J., Murga, L. F., André, P., Yang, H., Ondrechen, M. J., Williams, R. J., Agunwamba, A., and Budil, D. E. (2005) Statistical criteria for the identification of protein active sites using theoretical microscopic titration curves. *Proteins: Struct., Funct., Bioinf. 59*, 183−195.
(68) Pauling, L. (1948) Nature of Forces between Large Molecules of Biological Interest. *Nature 161*, 707−709.
(69) Warshel, A. (1998) Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites. *J. Biol. Chem. 273*, 27035−27038.
(70) Cleland, W. W., Frey, P. A., and Gerlt, J. A. (1998) The Low Barrier Hydrogen Bond in Enzymatic Catalysis. *J. Biol. Chem. 273*, 25529−25532.
(71) Bruice, T. C. (2002) A View at the Millennium: The Efficiency of Enzymatic Catalysis. *Acc. Chem. Res. 35*, 139−148.
(72) Careri, G., Fasella, P., and Gratton, E. (1979) Enzyme Dynamics: The Statistical Physics Approach. *Annu. Rev. Biophys. Bioeng. 8*, 69−97.
(73) Cha, Y., Murray, C., and Klinman, J. (1989) Hydrogen tunneling in enzyme reactions. *Science 243*, 1325−1330.
(74) Rothlisberger, D., Khersonsky, O., Wollacott, A. M., Jiang, L., DeChancie, J., Betker, J., Gallaher, J. L., Althoff, E. A., Zanghellini, A., Dym, O., Albeck, S., Houk, K. N., Tawfik, D. S., and Baker, D. (2008) Kemp elimination catalysts by computational enzyme design. *Nature 453*, 190−195.
(75) Siegel, J. B., Zanghellini, A., Lovick, H. M., Kiss, G., Lambert, A. R., St. Clair, J. L., Gallaher, J. L., Hilvert, D., Gelb, M. H., Stoddard, B. L., Houk, K. N., Michael, F. E., and Baker, D. (2010) Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science 329*, 309−313.
(76) Schafer, S. L., Barrett, W. C., Kallarakal, A. T., Mitra, B., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., and Kenyon, G. L. (1996) Mechanism of the Reaction Catalyzed by Mandelate Racemase: Structure and Mechanistic Properties of the D270N Mutant. *Biochemistry 35*, 5662−5669.
(77) Sampson, N. S., and Knowles, J. R. (1992) Segmental movement: Definition of the structural requirements for loop closure in catalysis by triosephosphate isomerase. *Biochemistry 31*, 8482−8487.
(78) Morley, K. L., and Kazlauskas, R. J. (2005) Improving enzyme properties: When are closer mutations better? *Trends Biotechnol. 23*, 231−237.
(79) Sigala, P. A., Caaveiro, J. M. M., Ringe, D., Petsko, G. A., and Herschlag, D. (2009) Hydrogen Bond Coupling in the Ketosteroid Isomerase Active Site. *Biochemistry 48*, 6932−6939.
(80) Foulkes, J. E., Prabu-Jeyabalan, M., Cooper, D., Henderson, G. J., Harris, J., Swanstrom, R., and Schiffer, C. A. (2006) Role of Invariant Thr80 in Human Immunodeficiency Virus Type 1 Protease Structure, Function, and Viral Infectivity. *J. Virol. 80*, 6906−6916.
(81) Makowski, L., Gore, D., Mandava, S., Minh, D., Park, S., Rodi, D. J., and Fischetti, R. F. (2011) X-ray solution scattering studies of the structural diversity intrinsic to protein ensembles. *Biopolymers 95*, 531−542.