

# Supporting Information

Kraut et al. 10.1073/pnas.0911168107

## SI Text

**Effect of the Tyr16Phe Mutation on KSI Catalysis.** Previous studies reported that the Tyr16Phe mutation results in 2000-fold and 50,000-fold reductions in  $k_{\text{cat}}$  for pKSI (1) and tKSI (2), respectively. These studies both used the substrate, 5-AND, for which chemistry is not cleanly rate limiting (3). The observed difference in the effect of this mutation on the activity of KSI from different species may reflect differences in the rate-limiting step for isomerization of 5-AND. When a substrate, 5(10)-EST, for which chemistry is rate-limiting is used (4), the  $k_{\text{cat}}$  reductions of the Tyr16Phe mutation are within two-fold for pKSI (20,000-fold) and tKSI (11,000-fold). In the present study, we use 5(10)-EST to ensure that chemistry is rate limiting for the wild-type and mutant enzymes investigated.

**Energetic Contributions of Active Site Hydrogen Bonds to KSI Catalysis.** The hydrogen bonds donated by Tyr-16 and Asp-103 to the dienolate intermediate and dienolate-like transition states are expected to be energetically favorable, relative to hydrogen bonds donated by water because Tyr and Asp have lower  $pK_a$  values than water and are thus stronger hydrogen bond donors. As shown by Hine and Jencks, stronger hydrogen bond donors preferentially interact with stronger hydrogen bond acceptors, relative to interactions of each with water (5, 6). In addition, pre-positioning of the Tyr and Asp hydrogen bond donors can further favor hydrogen bonding within this and other active sites, relative to water molecules in solution that must rearrange to optimize transition state hydrogen bonding (7, 8). Indeed, Tyr-16 and Asp-103 appear to be precisely positioned to interact with the incipient oxyanion of the reacting substrate (9).

Although repositioning of strong hydrogen bond donors can provide an energetic advantage for active site hydrogen bonds in KSI and other enzymes, other active site features can be unfavorable relative to those in solution. For example, it is possible that oxyanions such as the one formed in the KSI reaction are stabilized by three hydrogen bonds in aqueous solution (10) rather than the two hydrogen bonds in the KSI active site. Although there is no way to derive energetic values by simply counting hydrogen bonds (11, 12), each of the two enzymatic hydrogen bonds may provide greater transition state stabilization than each individual solution hydrogen bond but could, in principle, provide only limited or no net rate advantage upon summation of their individual contributions. Finally, hydrogen bond energetics are strongly affected by the surrounding environment and understanding the electrostatic properties of the active site and of protein interiors is one of the most vexing challenges of modern-day protein science (13–15).

## Materials and Methods.

**Materials.** All reagents were of the highest purity commercially available. Equilenin and 5(10)-EST were purchased from Steraloids and 2-fluoro-4-nitrophenol was purchased from Sigma. All buffers were prepared with reagent grade chemicals or better.

**KSI Mutagenesis, Expression, and Purification.** Quik-Change (Stratagene) site-directed mutagenesis was used to introduce the Tyr16Phe and Asp40Asn mutations into pKSI and tKSI genes encoded on pKK223-3 plasmids, which were confirmed by sequencing mini-prep DNA from DH5  $\alpha$  cells on an ABI3100 capillary sequencer. Proteins were expressed and purified as previously described (8).

As the other Tyr-16 mutants did not express from the pKK223-3 plasmid, wild-type pKSI was cloned into a pET21c vector using the NdeI and HindIII restriction sites, and the Tyr16Ser, Tyr16Ser/Asp40Asn, Tyr16Thr, Tyr16Ala, Tyr16Gly mutations were introduced and confirmed as above. The mutant proteins were expressed at 37 °C from LB cultures (containing 50  $\mu\text{g}/\text{ml}$  carbenicillin) of BL21(DE3) cells transformed with the appropriate pET21c plasmid. Cells were grown to an  $\text{OD}_{600}$  of approximately 0.5 and protein expression was induced with 0.5 mM IPTG followed by 3 h of further growth. Cells were harvested by centrifugation (4000 rpm for 20 min).

All of the mutant proteins expressed from the pET21c vector were insoluble and thus were purified and refolded from inclusion bodies. Pellets were resuspended in 40 mM Tris•HCl, pH 8.5, 1 mM EDTA, 25% sucrose buffer and lysed by sonication or passage through a French pressure cell. Lysates were centrifuged (20,000 rpm for 20 min) and pellets were solubilized in 1% sodium deoxycholate, 20 mM Tris•HCl, pH 8.5, 200 mM NaCl, 2 mM EGTA followed by centrifugation at 8,000  $\times g$ . Protein inclusion bodies contained in the resulting pellet were washed by multiple cycles of resuspension/centrifugation in 0.25% sodium deoxycholate, 10 mM Tris•HCl, pH 8.5, followed by 20 mM Na•HEPES, pH 8.5, 500 mM NaCl, 1 mM EDTA to remove detergent. Purified inclusion bodies (which were >95% KSI as determined by SDS-PAGE) were then resolubilized in 20 mM Na•HEPES, pH 8.5, 500 mM NaCl, 1 mM EDTA, 8 M urea. Mutant proteins were refolded by dilution to 0.4 M urea in 40 mM potassium phosphate ( $\text{KP}_i$ ) at pH 7.2 followed by stirring at 4 °C for 1 h. After filtration, the refolded enzyme was purified by deoxycholate affinity chromatography as previously described (8). Final protein purity was >99% as estimated from a Coomassie-stained SDS-PAGE gel. Protein concentrations were determined using the calculated molar extinction coefficient at 280 nm (16).

KSI mutants were independently purified using three alternative strategies that gave enzymes with identical activities to enzymes purified by the method above. (i) KSI was purified from inclusion bodies as above. Instead of refolding by dilution to 0.4 M urea in 40 mM  $\text{KP}_i$ , however, KSI was refolded and purified on a HiTrap FastFlow Q column (Amersham). (ii) Cell pellets containing pKSI mutants expressed from the pET21c vector were resuspended in 40 mM  $\text{KP}_i$ , pH 7.2, 1 mM EDTA, 2 mM DTT, lysed by French press, and centrifuged. The resulting pellet was resuspended in 40 mM  $\text{KP}_i$ , pH 7.2, 1 mM EDTA, 8 M urea, diluted to 0.6 M urea in 40 mM  $\text{KP}_i$ , pH 7.2, 1 mM EDTA, incubated on ice for 20 min, and centrifuged for 20 min at 20,000 rpm. The supernatant was passed through a 0.2  $\mu\text{m}$  syringe filter and soluble, refolded pKSI was purified by deoxycholate affinity chromatography as previously described (8). (iii) Although the majority of the Tyr-16 mutant proteins expressed from the pET21c plasmid was contained within insoluble inclusion bodies, a small portion of each enzyme (1–2 mg) was contained in the soluble fraction of crude cell lysates. The Tyr16Ser and Tyr16Ala mutants were purified from the soluble fraction of clarified lysates of resuspended cell pellets as previously described (8).

**KSI Kinetics.** Reactions with 5(10)-EST were monitored continuously at 248 nm using a Uvikon 9310 spectrophotometer. A molar absorptivity of 14,800  $\text{M}^{-1} \text{cm}^{-1}$  was determined for the commercially available reaction product, 4-estrene-3,17-dione. Reactions were conducted at 25 °C in 40 mM  $\text{KP}_i$ , pH 7.2, 1 mM EDTA, and 2 mM DTT, with 2% DMSO added as a cosolvent to maintain

substrate solubility. Kinetic parameters ( $k_{\text{cat}}$  and  $K_M$ ) were determined by fitting the observed initial reaction velocity as a function of substrate concentration (2–384  $\mu\text{M}$ ) to the Michaelis–Menten equation. Reported rate constants are the average of three or more independent assays using differing enzyme concentrations varied over at least a fourfold range. Reported errors are the standard deviations in computed averages. Enzyme purified using different purification strategies (see previous section) gave identical results within error.

**KSI X-ray Crystallography.** Co-crystals of pKSI Tyr16Ser/Asp40Asn containing bound equilenin in space group  $P2_1$  were obtained using hanging drop vapor diffusion by mixing 2  $\mu\text{l}$  of 30 mg/ml KSI containing 1 molar equivalent of equilenin with 2  $\mu\text{l}$  of reservoir solution (0.9 M ammonium sulfate, 40 mM  $\text{KPi}$ , pH 7.2). Blade-shaped crystals appeared after incubation for 1 week at 20 °C. Cryoprotection was achieved by soaking crystals in a 1:1 mixture of mother liquor and 2.9 M sodium malonate followed by a direct soak in 2.9 M sodium malonate prior to flash-freezing in liquid nitrogen. Diffraction data from a single crystal maintained at 100 K were collected at beamline 8.2.2 of the Advanced Light Source (Lawrence Berkeley National Laboratory). Data were integrated and scaled using DENZO and SCALEPACK, respectively (17). Five percent of the observed data were set aside for cross-validation. Data collection statistics are summarized in Table S3.

An initial protein model was obtained by molecular replacement with Phaser (18) using the coordinates from the previously published structure of pKSI Asp40Asn (PDB code 2PZV) that had been mutated *in silico* to include the Tyr16Ser mutation. After inspection of the  $\sigma_A$ -weighted  $2F_o - F_c$  and  $F_o - F_c$  electron density maps, a complete model was constructed using Coot (19). Simulated annealing refinement was carried out using a maximum-likelihood amplitude-based target function as implemented in PHENIX (20). Further refinement and water picking were carried out with PHENIX, interspersed with manual corrections and model rebuilding in Coot. A final round of refinement in PHENIX treated each monomer as an independent TLS group. Refinement statistics are shown in Table S3. Structure factors and coordinates have been deposited in the RCSB Protein Data Bank as entry 3IPT. All structural figures were prepared using MacPyMOL (21).

The triangle-shaped electron density observed in the  $\sigma_A$ -weighted  $2F_o - F_c$  map in the cavity created by the Tyr16Ser mutation was too large to model as a single water but too weak to fit as a group of disordered waters. The density was best (though still imperfectly) described by the presence of two water molecules with occupancies split between two sites each. Because we were unable to fit the diffuse density with a specific water arrangement, the density was left empty in the final structural model. Further quantitative interpretation of the observed density would require fully unbiased electron density maps from experimentally determined phases.

To rule out other possible explanations, we considered whether the observed planar, triangular electron density could be accounted for by the presence of any other small molecule in the purification or crystallization buffers (see previous section). The only modeled species able to fit the observed density was

urea, which was used to refold KSI from inclusion bodies prior to purification. We considered it unlikely that a non-covalently bound urea molecule would remain associated with KSI throughout the extensive buffer exchanges during purification and crystallization. Furthermore, purification of soluble Tyr16Ser without refolding with urea yielded enzyme with identical activity as urea-refolded KSI (see previous section). Nevertheless, we directly tested for the presence of urea in pKSI Tyr16Ser/Asp40Asn purified under identical conditions to the crystallized protein by refolding KSI with  $^{15}\text{N}$ -labelled urea and acquiring  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectra. These spectra, shown in Fig. S2, provide no evidence for the presence of a bound urea. On the basis of our kinetic and spectroscopic results, we conclude that the observed electron density does not arise from a bound urea molecule.

**KSI NMR Spectroscopy.** One-dimensional  $^{19}\text{F}$  NMR spectra of 2-fluoro-4-nitrophenolate ( $pK_a = 6.0$ ) (22) bound to pKSI Tyr16Ser/Asp40Asn or Tyr16Phe/Asp40Asn were acquired on a 500 MHz Varian  $^{\text{UNITY}}$ INOVA NMR spectrometer (operating at 470 MHz) running VNMR v6.1C and equipped with a 5 mm pulsed field gradient switchable probe operating at ambient temperature (20 °C). Samples contained 10%  $\text{D}_2\text{O}$  as the lock solvent and chemical shifts were referenced to an external standard of trifluoroacetic acid (-76.1 ppm relative to  $\text{CFCl}_3$ ). Solution spectra of 4 mM 2-fluoro-4-nitrophenol were acquired (12 scans) in 10 mM HCl (aq), 10 mM NaOH (aq), THF, and THF containing 40 mM triethylamine. Formation of the ionized phenolate in basic solutions of water or THF was signaled by the 2.2 ppm up-field shift of the  $^{19}\text{F}$  resonance relative to the detected peak in solutions of HCl (aq) or THF, as previously reported (22). Spectra of the KSI-bound ligand were acquired for 300 scans and contained 2 mM enzyme, 1 mM 2-fluoro-4-nitrophenolate, 40 mM  $\text{KPi}$ , pH 7.2, 1 mM EDTA, and 2 mM DTT. Spectra were processed using a 10 Hz line broadening, and a baseline correction was applied over the observed peaks.

The  $^1\text{H}$  NMR spectrum of the tKSI Tyr16Phe/Asp40Asn•equilenin complex was acquired (500 scans) at the Stanford Magnetic Resonance Laboratory on an 800 MHz Varian  $^{\text{UNITY}}$ INOVA spectrometer equipped with a 5 mm, triple resonance, gradient  $^1\text{H}$ ( $^{13}\text{C}/^{15}\text{N}$ ) using the 1331 binomial pulse sequence (23) to suppress the water signal, as previously described (8, 24). The NMR sample contained 300  $\mu\text{M}$  KSI, 300  $\mu\text{M}$  equilenin, 40 mM  $\text{KPi}$ , pH 7.2, 1 mM EDTA, and 10%  $\text{DMSO-}d_6$  (as lock solvent and cryogen) and was maintained at -3 °C during spectral acquisition. The spectrum was processed using a 10 Hz line broadening and base-line correction. A far-downfield peak at 17.4 ppm was observed in independently prepared samples.

**UV-vis Absorbance Spectra.** Absorbance spectra of 2-fluoro-4-nitrophenol were acquired in a microcuvette with a Uvikon 9310 absorbance spectrophotometer. Spectra of 18  $\mu\text{M}$  phenol in 10 mM HCl (pH 2) and 10 mM NaOH (pH 12) were compared with the spectrum of the same concentration phenol in 40 mM  $\text{KPi}$ , pH 7.2, 1 mM EDTA, in the presence of 500  $\mu\text{M}$  pKSI Tyr16Phe/Asp40Asn or Tyr16Ser/Asp40Asn after subtraction of the free enzyme spectrum.

- Kim SW, Choi KY (1995) Identification of active site residues by site-directed mutagenesis of delta 5-3-ketosteroid isomerase from *Pseudomonas putida* biotype b. *J Bacteriol* 177:2602–2605.
- Kuliopulos A, Mildvan AS, Shortle D, Talalay P (1989) Kinetic and ultraviolet spectroscopic studies of active-site mutants of delta 5-3-ketosteroid isomerase. *Biochemistry* 28:149–159.
- Hawkinson DC, Eames TC, Pollack RM (1991) Energetics of 3-oxo-delta 5-steroid isomerase: Source of the catalytic power of the enzyme. *Biochemistry* 30:10849–10858.
- Pollack RM, Bantia S, Bounds PL, Koffman BM (1986) Ph dependence of the kinetic parameters for 3-oxo-delta 5-steroid isomerase. Substrate catalysis and inhibition by (3s)-spiro[5 alpha-androstane-3,2'-oxiran]-17-one. *Biochemistry* 25:1905–1911.
- Hine J (1972) Structural effects on rates and equilibria. XV. Hydrogen-bonded intermediates and stepwise mechanisms for proton-exchange reactions between oxygen atoms in hydroxylic solvents. *J Am Chem Soc* 94:5766–5771.
- Stahl N, Jencks WP (1986) Hydrogen bonding between solutes in aqueous solution. *J Am Chem Soc* 108:4196–4205.
- Warshel A (1998) Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *J Biol Chem* 273:27035–27038.
- Kraut DA, et al. (2006) Testing electrostatic complementarity in enzyme catalysis: Hydrogen bonding in the ketosteroid isomerase oxyanion hole. *PLoS Biol* 4:501–519.
- Sigala PA, et al. (2008) Testing geometrical discrimination within an enzyme active site: Constrained hydrogen bonding in the ketosteroid isomerase oxyanion hole. *J Am Chem Soc* 130:13696–13708.
- Cappa CD, Smith JD, Messer BM, Cohen RC, Saykally RJ (2007) Nature of the aqueous hydroxide ion probed by x-ray absorption spectroscopy. *J Phys Chem A* 111:4776–4785.
- Baldwin RL (2003) In search of the energetic role of peptide hydrogen bonds. *J Biol Chem* 278:17581–17588.





**Table S1. Effects of mutations in oxyanion holes**

Enzyme	Mutation	$k_{cat}$ effect	$\Delta G$ (kcal/mol)	Reference
Subtlisin	N155L	300	3.7	(25)
Subtlisin	N155T	4000	4.9	(26)
Subtlisin	N155G	150	3.0	(27)
Subtlisin	T220V	30	2	(28)
Papain	N19A	20	1.8	(29)
Papain	N19S	180	3.0	(29)
Cutinase	S42A	450	3.6	(30)
Oligopeptidase B	Y452F	100-1000* <sup>†</sup>	2.7-4.1	(31)
Signal Peptidase	S88A	2400	4.6	(32)
Aminopeptidase A	Y471F	700	3.9	(33)
Human dipeptidyl peptidase III	Y318F	100	2.7	(34)
Cocaine Esterase	Y44F	>1500	>4.3	(35)
Prolyl oligopeptidase	Y473F	30-150*	2-3	(36)
Transpeptidase SrtA	R197A	600	3.8	(37)
enterobactin hydrolase IroE	R130K	300	3.4	(38)
Ketosteroid Isomerase	Y16F	50000	6.4	(2)
Ketosteroid Isomerase	D103A	5000	5.0	(39)

\*Varied with different substrates.

<sup>†</sup> $k_{cat}$  not directly determined; either  $k_{cat}/K_M$  or activity effect.

**Table S2. Effects of Tyr-16 mutations on pKSI-catalyzed isomerization of 5(10)-EST**

Enzyme	$k_{cat}$ (s <sup>-1</sup> )	$K_M$ ( $\mu$ M)	$k_{cat}/K_M$ (M <sup>-1</sup> s <sup>-1</sup> )	$k_{cat}$ ratio (WT/mutant)	$k_{cat}/K_M$ ratio (WT/mutant)
Wild type	9.9 (0.9)	19 (4)	$5.4 \times 10^2$ (0.7)	[1]*	[1]*
Tyr16Ala	$4.8 \times 10^{-2}$ (0.4)	30(10)	$1.5 \times 10^3$ (0.3)	200	360
Tyr16Gly	$4.0 \times 10^{-2}$ (0.4)	23 (2)	$1.7 \times 10^3$ (0.1)	250	310
Tyr16Ser	$3.8 \times 10^{-2}$ (0.5)	18 (6)	$2.2 \times 10^3$ (0.4)	260	240
Tyr16Thr	$5.2 \times 10^{-2}$ (0.3)	19 (3)	$2.7 \times 10^3$ (0.2)	190	200
Tyr16Phe	$5.4 \times 10^{-4}$ (0.1)	41 (6)	8 (1)	19,000	70,000

Standard deviations from at least three experiments at different enzyme concentrations are given in parentheses.

\*Defined as unity for comparison.

**Table S3. Crystallographic data and refinement statistics for equilenin bound to pKSI Tyr16Ser/Asp40Asn**

Dataset	
Resolution Range (Å)	43.41-1.63
Space Group	P2 <sub>1</sub>
a, Å	50.7
b, Å	72.4
c, Å	81.3
$\alpha$ , °	90.0
$\beta$ , °	91.2
$\gamma$ , °	90.0
Number unique reflections	60,156
Completeness, %	82.3
Multiplicity	4.2 (1.7)
$R_{merge}$ , %*	16.6%
$I/\sigma_{overall}$ ( $I/\sigma_{high\ res}$ )	9.7 (1.4)
Refinement statistics	
Number residues	124
Number waters	223
$R_{work}$ , % <sup>†</sup>	20.7
$R_{free}$ , % <sup>‡</sup>	24.5
rmsd bond, Å	0.031
rmsd angle, °	2.7

\* $R_{merge} = \sum_{hkl} \sum_i |I(hkl)_i - I(hkl)| / \sum_{hkl} \sum_i I(hkl)_i$

<sup>†</sup> $R_{work} = \sum_{hkl} |F(hkl)_o - [F(hkl)_c]| / \sum_{hkl} F(hkl)_o$

<sup>‡</sup> $R_{free}$  was calculated exactly as  $R_{work}$  where  $F(hkl)_o$  were taken from 5% of the data not included in refinement.