Thermodynamic framework for identifying free energy inventories of enzyme catalytic cycles

Stephen D. Fried and Steven G. Boxer

Department of Chemistry, Stanford University, Stanford, CA 94305-5012

Contributed by Stephen G. Boxer, June 12, 2013 (sent for review February 24, 2013)

Pauling’s suggestion that enzymes are complementary in structure to the activated complexes of the reactions they catalyze has provided the conceptual basis to explain how enzymes obtain their creative catalytic powers, and has served as a guiding principle in drug design for over 50 y. However, this model by itself fails to predict the magnitude of enzymes’ rate accelerations. We construct a thermodynamic framework that begins with the classic concept of differential binding but invokes additional terms that are needed to account for subtle effects in the catalytic cycle’s proton inventory. Although the model presented can be applied generally, this analysis focuses on ketosteroid isomerase (KSI) as an example, where recent experiments along with a large body of kinetic and thermodynamic data have provided strong support for the noncanonical thermodynamic contribution described. The resulting analysis precisely predicts the free energy barrier of KSI’s reaction as determined from transition-state theory using only empirical thermodynamic data. This agreement is suggestive that a complete free energy inventory of the KSI catalytic cycle has been identified.

E nzymes are generally viewed as devices that use protein architecture to precisely position several key functional groups and electrostatic partners with respect to the substrate (1–4). This structure-based approach is complemented by thermodynamic models, which provide abstract but quantitative descriptions of these interactions, including those that escape visual intuition. In particular, the model shown in Fig. 1, first advanced by Kurz (5), has served as a dominant paradigm for conceptualizing enzyme catalysis for 50 y. This thermodynamic cycle formalizes Pauling’s hypothesis that enzymes are complementary in structure to the transition states of the reactions they catalyze (6, 7) and, combined with physical models to describe the enzyme–transition state interactions, it can help determine what leads to enzymes’ catalytic effects (8).

According to an elementary formulation of transition-state theory (TST), a rate constant, $k$, is determined by the position of a pseudoequilibrium (represented by $K^*$) to the activated complex (9). In this depiction, chemical reactions are treated as equilibria between ground states and transition states (Fig. 1A). An enzyme’s accelerated rate ($k_{cat} > k_{uncat}$) is represented by a relatively large value of $k_{cat}$ compared with $k_{uncat}$ (Fig. 1B). This implies that the enzyme must bind its cognate transition state very tightly, so the enzyme’s dissociation constant to the transition state ($K_D^{TS}$) is very small (2, 5). In this model, the overall rate constants are taken to be proportional to $K^*$. By solving the thermodynamic cycle of Fig. 1A, one obtains

$$k_{cat}/k_{uncat} = K_D^{S}/K_D^{TS},$$

where $k_{cat}$ and $k_{uncat}$ are the enzyme-catalyzed and uncatalyzed unmolecular rate constants, respectively, and $K_D^{S}$ and $K_D^{TS}$ are dissociation constants of the enzyme to substrate and transition state, respectively. Eq. 1 implies that the rate enhancement (the ratio of the enzymatic and uncatalyzed rate constants) is equal to the differential binding factor of the transition state. Despite its great conceptual power, Eq. 1 seems to fall short of accounting for actual enzymes’ catalytic power. Whereas it is often asserted that inhibitors serve as analogs of transition states (10–14), the differential binding factors for good inhibitors are normally around $10^{3−7}$ (13–15), much smaller than the fantastic rate enhancements ($k_{cat}/k_{uncat}$) on the order of $10^{12−17}$ (10–12) achieved by natural enzymes. Kurz’s simple model has been a useful tool for rationalizing why tight-binding inhibitors sometimes resemble the transition states of their target enzymes’ reactions (10), but it has been less useful for understanding the origin of enzyme catalysis.

In the present work, we argue for the third case. In laying out this argument, we extend the classical thermodynamic framework for the reaction catalyzed by ketosteroid isomerase (KSI) that incorporates insights from recent studies (25–27). This analysis of KSI reveals that subtle processes other than differential binding are at work in KSI’s catalytic cycle; when these processes are accounted for, one can obtain an estimate for the activation barrier that is consistent with kinetics (28) and computation (16, 29). This model effectively bridges the divide between the rate ratio and the binding ratio, without invoking dynamic factors.

Background on KSI

KSI is among a class of enzymes that elicit the ability to abstract protons (30) from an apparently nonacidic carbon acid (5-androsten-3,17-dione, $pK_a = 12.7$; ref. 31) using a weak general base (Asp40, $pK_a = 3.75$; ref. 32), as shown in Fig. 2A. These $pK_a$s and all other thermodynamic parameters that will be used in this analysis are compiled in Table S1. KSI is a highly proficient enzyme, because in its absence the isomerization of androstene proceeds about a trillion times more slowly (11). In the following, the uncatalyzed reaction will refer to the “chemically filtered” reference reaction—that is, the nonenzymatic reaction that proceeds by the same mechanism (17). In KSI’s case, this corresponds to the slow acetate-catalyzed isomerization of androstene in solution (Fig. 2B), which involves a high-energy dienolate intermediate (33). In KSI, this intermediate is strongly stabilized by an oxanion hole (OAH) in the active site (34). The OAH consists of supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1310964110/-/DCSupplemental.

Author contributions: S.D.F. designed research; S.D.F. performed research; S.D.F. contributed new analytic tools; S.D.F. and S.G.B. analyzed data; and S.D.F. and S.G.B. wrote the paper.

The authors declare no conflict of interest.

1To whom correspondence should be addressed. E-mail: sboxer@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1310964110/-/DCSupplemental.

www.pnas.org/cgi/doi/10.1073/pnas.1310964110
of a tyrosine (Tyr16) and a protonated aspartic acid (Asp103) that donate hydrogen bonds to the carbonyl oxygen of androstene (35). Tyr16 is very close to two additional tyrosines (Tyr32, Tyr57), which form a hydrogen-bond network (35).

Fig. 2A introduces nomenclature conventions that will be used throughout the discussion: First, androstene, the steroid substrate, is abbreviated with only two (A and B) rings and, second, free energies of reaction (ΔG’s) are expressed over arrows in terms of logarithmic equilibrium constants (i.e., pKs). Logarithmic equilibrium constants are linear with free energy, as expressed by Eq. 2:

$$\Delta G_i = 2.303RT \mu K_i.$$  

[2]

The KSI-catalyzed and uncatalyzed isomerization of androstene share the same chemical mechanism, passing through a presumed dienolate intermediate (36, 37), so that the present analysis can focus on energetics. In Fig. 2C, an annotated reaction coordinate diagram for the first chemical step is given. ΔG’s (in kcal·mol⁻¹) for various processes associated with formation of the first transition state (TS₁) and the intermediate (I) are provided, as well as the associated binding/rate constants from which the ΔG’s are determined. As judged from interpreting chemical kinetics with TST, the free energy barrier of the first enzyme-catalyzed reaction is 10.3 kcal·mol⁻¹ (28, 38) (this barrier is evaluated from the “microscopic” rate constant for the process E+S → E+I, which is not the same as kcat from the Michaelis–Menten model, which corresponds to E+S → E+P), which is substantially less than the barrier of the same reaction (20.3 kcal·mol⁻¹; ref. 33) occurring in solution with acetate as the base. [This barrier is based on the pseudo-unimolecular rate calculated from the bimolecular rate constant and the maximal base concentration.] Using the differential binding formalism (Fig. 1) and treating the binding constant of an inhibitor (E + In = E-In) as a model for transition-state binding (E + TS₁ = E+TS₁), a prediction for the enzyme-catalyzed activation barrier can be formulated by closing a thermodynamic cycle on the reaction coordinate diagram (blue arrow, Fig. 2C). This analysis uses the empirical values for Kcat, Kcat (which stands in for Kcat), and Kcat (which corresponds to binding constants for substrate and inhibitor and the rate constant for the uncatalyzed solution reaction, respectively). The prediction from this calculation (represented by PP in Fig. 2C) is that the enzyme-catalyzed reaction would possess an activation barrier of 14.6 kcal·mol⁻¹—which is significantly higher than that observed. This disagreement is in fact very common and easy to rationalize on account of the fact that the inhibitor may only roughly recapitulate the energetics associated with the true transition state (19, 20).

In addition to stabilizing its transition state, KSI is also known to stabilize its dienolate intermediate. The “internal” equilibrium constant for KSI’s first reaction (i.e., the Kcat for E+S → E+I) has been experimentally shown to be between 0.01 and 0.3, implying the intermediate lies only 1–3 kcal·mol⁻¹ higher in free energy relative to the reactant in the active site (39, 40). In contrast, forming the intermediate is highly endergonic in solution, with a free energy change of 10.9 kcal·mol⁻¹. A Pauling-like formalism could also be applied to describe this effect. By combining the empirical values for Kcat, Kcat (binding constants for substrate and intermediate, respectively), and Kcat for S = I in solution, one should in principle be able to estimate the Kcat for E+S = E+I (and therefore its ΔG’s). This analysis furnishes a ΔG of 6.2 kcal·mol⁻¹, which is again not in agreement with the measured value. This disagreement is more difficult to rationalize, because in this case the thermodynamic cycle is rigorous and does not use one molecule to act as a surrogate for something else. We hypothesized that this inconsistency is due to underlying processes inherent in the catalytic cycle that are not represented by binding effects, and so not captured by Kurz’s model. As we show in the following, these effects come to light upon using a detailed thermodynamic cycle (Fig. 3A).

Thermodynamic Cycle for KSI’s First Chemical Step

Fig. 3A shows how one can determine the free energy of an enzyme-actuated process by decomposing it into thermodynamically defined steps. In this case, X stands for the transformation under consideration, E+S → E+I, which is the first chemical step of the mechanism. To determine ΔG of X, we write a thermodynamic cycle [black steps (i)–(iv)] starting with E+S, and proceed with processes that add up to form E+I. All of the steps (i)–(iv) have their ΔG’s independently defined by various empirical equilibrium constants. By working with the enzyme–substrate complex, in step (i) the substrate is removed from the enzyme’s active site. This dissociation constant is well-approximated by the Michaelis constant of KSI for a slow substrate (17 μM; ref. 41). Next, with androstene out of the enzyme, step (ii) ionizes androstene with water, which is energetically defined as the pKₐ of androstene (12.7; ref. 31). In step (iii), the general base of KSI is neutralized by water, which is also described by its pKₐ (3.75; ref. 32). Step (iv) is the favorable association of the intermediate dienolate to the enzyme active
site with the general base protonated. The thermodynamics of this step is likened to the binding of the enolic intermediate to the Asp40Asn mutant of KSI (written as KSI$_{D40N}$), whereby the Asp40Asn background simulates the neutralized (i.e., protonated) general base present during this process. It is known that the dienol binds to KSI$_{D40N}$ in (at least partially) the anionic form (36, 37). The observed binding is much tighter (K$_{\text{Enol}}^{D40N}$ = 6 nM; ref. 40) than that of the substrate, which is qualitatively consistent with Pauling’s hypothesis. In black step (iv), we also regenerate the two water molecules using the autoionization constant of water. At this stage, we see that X closes a thermodynamic cycle whose overall free energy change must be identically zero:

$$\text{pK}_{D}^{\text{Andro}} + \text{pK}_{\text{Andro}} - \text{pK}_{\text{Asp40}} - \text{pK}_{\text{N40/L-Enol}} - X = 0. \quad [3]$$

Using the known values for the equilibrium constants and the relationship between equilibrium constants and free energy, Eq. 3 allows one to calculate $\Delta G^{\circ}(X)$ to be 7.6 kcal mol$^{-1}$. This is less endergonic than the uncalyzed case by the differential binding of the intermediate over the substrate. However, the current treatment is far from recapitulating the observation that KSI renders $\Delta G^{\circ}(X)$ to be 1–3 kcal mol$^{-1}$ (40). The discrepancy is suggestive that certain processes are not being properly accounted for.

**Introducing Internal Acidity, $K_a$**

Let us reexamine step (iv) of Fig. 3A and ask a seemingly simple question: Does the experimental parameter (the binding constant of the dienol intermediate to KSI$_{D40N}$) properly represent the process being portrayed in step (iv)? Drawn in step (iv) is an ionized intermediate (dienolate) binding with a neutral enzyme to form a complex. However, in the actual binding experiment, the preponderant form of the intermediate (whose $K_a$ is 10.0) in solution is neutral (dienol). This complication will be encountered in any instance where both the enzyme (E; in this case KSI$_{D40N}$) and the ligand (L; in this case the steroid dienol) possess titratable moieties whose protonation statuses affect the binding interaction. Such a situation would arise in any enzyme catalyzing proton-transfer chemistry, and is treated in a general fashion in Fig. 3B. A priori, we consider the two binding modes on equal footing: binding of neutral ligand to ionized enzyme ($K_a^{\text{E}}$) and binding of ionized ligand to neutral enzyme ($K_a^{\text{L-}}$). Moreover, the products of these binding equilibria can interconvert via an internal proton transfer (the long hypotenusine in Fig. 3B), and the reactants of these binding equilibria can interconvert as well via exchanging protons to bulk water (the short hypotenusine). We impose only that the active site accommodate one negative charge, and so a complex between ionized ligand and ionized enzyme is not considered (35).

In any real experiment of the binding constant of L, what can be measured is an “apparent” binding constant that represents the collective affinity of ligand to enzyme, without regard to protonation states or binding modes.

$$K_{\text{app}}^{\text{E}} = ([\text{E}^-]+[\text{EH}])/([\text{L}^-]+[\text{LH}])/([\text{E}^-\cdot\text{L}^-]+[\text{E}^-\cdot\text{L}^+]). \quad [4]$$

$K_{\text{app}}^{\text{E}}$ does not correspond to any microscopic process in Fig. 3B. We can define conventional microscopic equilibrium constants for each elementary step under consideration, which are found located along the legs of the triangle in Fig. 3B. The equilibrium between E$^-$·LH and E$^-$·L$^+$ (the long hypotenusine of Fig. 3B) is referred to as the internal proton transfer. To describe this process, we will make an analogy to a different but related process along the inner hypotenuse. The proton exchange described by E$^-$·LH = E$^+$·L$^-$ is a canonical acid–base reaction, whose equilibrium position is dictated by the ratio of acid dissociation constants, $K_a^{\text{E}}/K_a^{\text{L+}}$.

The internal proton transfer between ligand and enzyme, E$^-$·LH = E$^+$·L$^-$, also depends on the relative acidity of the enzyme and the ligand; but unlike the E$^-$·LH = E$^+$·L$^+$ process, the internal exchange does not involve transfer of protons to/through water, and possibly is influenced by the enzyme’s and ligand’s properties being altered upon binding. Therefore, it is incorrect to define the enzyme’s acidity in the E$^-$·L state with...
a standard $K_a$. Instead, we refer to the enzyme’s “internal acidity” with the symbol $K_a^L$ (with a subscript $\alpha$ rather than the conventional $a$), and will use the ratio of $K_L^D/K_L^E$ to indicate the equilibrium position of the internal proton transfer.

By using the appropriate conservation equations and mass action expressions, we can derive expressions for the microscopic binding constants in terms of the apparent binding constant and the acidity constants:

$$K_L^{DH} = K_D^{app} \left(1 + K_L^E/K_L^E^\alpha \right)/\left(1 + [H^+]^\alpha \right) \left(1 + K_L^f^\alpha/K_L^f \right)/\left(1 + [H^+]^\alpha \right)$$

$$K_L^{EL} = K_D^{app} \left(1 + K_L^E/K_L^E^\alpha \right)/\left(1 + [H^+]^\alpha \right) \left(1 + K_L^a^\alpha/K_L^a \right)/\left(1 + [H^+]^\alpha \right).$$  \([5a,5b]\)

This model is capable of reproducing the pH dependence of the apparent dissociation constant (42), and it also predicts (correctly) the ratio of the internal proton transfer is largely invariant to pH (36). Fig. 3B shows that the internal proton transfer does not exchange protons with the bulk aqueous phase, and so must be independent of pH, excluding secondary effects resulting from pH-induced structural changes to the protein. The capacity of proteins to “self-buffer” an internal proton transfer against changes in external pH has been previously observed in green fluorescent protein, where the protonation state can be readily observed from the intrinsic chromophore (43). We take these observations as a validation of Eq. 5.

### Defining Differential Acidity

Returning to the original context of Fig. 3A and B, it becomes apparent that one of the steps was not assigned a correct thermochemical description. The step that was written down as (iv) really corresponds (in light of Fig. 3B) to $K_L^{DH}$ (in blue), and not the measured binding constant of $K_L^{DH}$ we used initially. To convert from what is known ($K_D^{app}$) to what is needed ($K_L^{DH}$), one must have measures of the enzyme’s active site’s standard acidity to water ($K_L^E$) and internal acidity to the ligand ($K_L^E^\alpha$), as prescribed by Eq. 5. The essential concept that we build upon is that these two acidity constants are not equal; in fact, KSI’s active site is several orders of magnitude more acidic in the apo state. This feature (that is, $K_L^E \neq K_L^E^\alpha$) is what we describe with the term “differential acidity,” in analogy to differential binding (which implied $K_L^D \neq K_L^E$).

A recent study detailed that Tyr57 (part of the OAH) in ligand-free KSI-DaDNN has an unusually low $pK_a$ of 6.3 ± 0.1, originally evidenced by $^{13}$C NMR studies of labeled tyrosines and confirmed by UV spectroscopic titration (26). Because tyrosine normally has a $pK_a \approx 10.5$, this study highlighted how hydrogen-bond networks can severely perturb a residue’s acidity (35). With this additional knowledge in hand, the preponderant mechanism of ligand (intermediate) binding to KSI-DaDNN at pH 7.2 (the standard condition for this analysis) involves a neutral dienol and an ionized OAH given their $pK_a$s of 10.0 and 6.3, respectively. Therefore, $K_L^{DH}$ is expected to be a good estimator of $K_L^{DH}$, but very different from $K_L^{DH}$.

Classical interpretations of acidity would posit that because $pK_a$ of Asp40 is 6.3, a second acid whose $pK_a$ is 6.3 would exist at the equilibrium point with KSI, with each half-neutral/half-ionized. This picture would be true for the reaction $E^- + LH = EH + L^-$, but is incorrect for the reaction $E^+ + LH = EH^+ + L^-$, which is not strictly governed by $pK_a$s. In fact, infrared spectroscopy studies have shown that a ligand of $pK_a = 10.0$ is 30% 2% ionized at equilibrium when bound to the enzyme (27), so $pK_a = 9.7$, in agreement with the estimate of $pK_a$ from a ligand-titration study (25). The fact that KSI’s active site is significantly more acidic in the apo state ($pK_a = 6.3$) than in the liganded state ($pK_a = 9.7$) in conjunction with the way that these terms enter into $K_L^{DH}$ (Eq. 5) demonstrates the subtle but significant effect the proton inventory exercises on the thermodynamic cycle.

### Amending the Thermodynamic Model

As shown by the blue steps in Fig. 3A, either $K_L^{DH}$ or $K_L^{EL}$ can be used in $K_D^{app}$’s place to complete the cycle. On one hand, we can simply substitute $K_D^{app}$ with $K_L^{DH}$ [blue (iv)], which by Eq. 5 is 1,900-fold lower than $K_L^{DH}$ at the common pH of 7.2. The factor by which $K_L^{EL}$ differs from $K_L^{DH}$ depends dramatically on the enzyme’s two acidity parameters, as shown in Fig. S1. We recalculate the free energy change for intermediate formation [$\Delta G_f ^{(X)}$] using Eq. 3, except replacing the apparent affinity for the intermediate to KSI-DaDNN ($pK_a^{D40N/Endo-app}$, 8.2) with the affinity for the anionic intermediate to neutral KSI-DaDNN ($pK_a^{D40N}$, 11.5). This substitution leads to a $\Delta G_f ^{(X)}$ of 3.0 kcal-mol$^{-1}$, which is in good agreement with the experimental value of the internal equilibrium constant (40).

Alternatively, we could use a lengthier thermochemical cycle, labeled with primes (\'). After step (iii), the protonation states of the enzyme and the ligand can be swapped [blue steps (iv') and (v')], to allow for the ligand to bind under $K_L^{DH}$ [blue step (v')]. $K_L^{DH}$ is nearly identical to the apparent dissociation constant (1.3-fold larger), which makes sense because the preponderant means by which the complex is formed involves ionized enzyme and neutral ligand. Following the full cycle of Fig. 3A, we generate the following expression:

$$\Delta G_f ^{(X)} = 2.303 RT \left(pK_a^{D40N/Andro} + pK_a^{Andro} - pK_a^{Asp40} - pK_a^{DH} \right)$$

$$+ \left(pK_a^{DH} - \frac{pK_a^{DH}}{C_1} \right),$$  \([6]\)

which gives the exact same result; however, it sets the stage for a conceptually useful approximation (see SI Discussion on the approximation’s range of validity). As noted previously, $K_L^{DH}$ is approximately the same as $K_L^{DH}$. Applying this substitution and rearranging terms, Eq. 6 can be reexpressed as

$$\Delta G_f ^{(X)} = 2.303 RT \left(pK_a^{Andro} - pK_a^{Asp40} \right)$$

$$+ \left(pK_a^{DH} - \frac{pK_a^{DH}}{C_1} \right),$$  \([7]\)

Eq. 7 decomposes the free energy cost for catalyzed intermediate formation into three separate contributions. $\Delta G_f ^{(uncat)}$ is the uncatalyzed free energy cost for intermediate formation, given by the $pK_a$ mismatch between acid and base. $\Delta G_f ^{(cat)}$ is the contribution from differential binding, is responsible for 4.7 kcal-mol$^{-1}$ of intermediate stabilization. A new contribution ($\Delta G_f ^{(cat)}$) from differential acidity extends that stabilization by an additional 4.6 kcal-mol$^{-1}$, making a modification that is comparable to the energy of differential binding. The formulation in Eq. 7 shows that differential acidity directly affects the energetics of catalysis. It can further be shown using the right-hand half-cycle in Fig. 3A that the differential acidity ($pK_a^{DH} - pK_a^{E}$) equals the ratio of the two microscopic dissociation constants ($K_L^{DH}/K_L^{E}$). This ratio is independent of $K_L^{DH}$ according to Eq. 5; therefore, the differential acidity effect is completely orthogonal to the “actual” affinity of the inhibitor, and so is properly viewed as a separate contribution from differential binding. Importantly, this analysis demonstrates that a large alteration in the catalytic cycle’s energetics stems from the unusual $pK_a$ of the oxygen hole in KSI (26). Perturbed $pK_a$s in other enzyme active sites may
have similar energetic consequences in their respective cases with the framework described.

**Predicting KSI’s Rate Enhancement**

The analysis up to this point has been applied to the intermediate dienolate of KSI. However, the actual catalytic effect of KSI depends on the reduction of the activation barrier. In a more approximate fashion, a parallel thermodynamic treatment can be applied to the transition state (Fig. 4) by returning to the notion of using an inhibitor (equilenin) to serve as an analog for KSI’s transition state. The thermodynamic properties of equilenin are very similar to those of KSI’s chemical intermediate, and so a distinction between the two is not essential for this analysis. Fig. 4 treats the transition state essentially like the intermediate—a large assertion, albeit a common one (36, 37, 39–41)—which is at least supported by the fact that linear free energy analysis of KSI has demonstrated the transition state to be late along the reaction coordinate (45). The thermodynamic cycle shows a delineation of three contributions to process $X^{±}$: the uncatalyzed component, a differential binding factor, and $\Delta G_3$. Similar to Eq. 7, the thermodynamic cycle provides an approximation for $\Delta G^\ddagger$ as follows:

$$\Delta G^\ddagger(X) = \Delta G^\ddagger_{\text{uncat}} + 2.303 RT \left[ \left( \frac{pK_{a}^{D40/Andro}}{4.77} - \frac{pK_{a}^{N40/Equ-app}}{9.0} \right) + \left( \frac{pK_{a}^{E}}{6.3} - \frac{pK_{a}^{E}}{9.7} \right) \right] = 9.9 \text{ kcal} \cdot \text{mol}^{-1}. \quad [8]$$

Eq. 8 provides an estimate for the free energy barrier that is in excellent agreement with transition-state theory, which furnishes the value of 10.3 kcal mol$^{-1}$ (28). This estimate is in much better agreement than the estimate from the original analysis (14.6 kcal mol$^{-1}$), resolving the discrepancy inspired by Fig. 2C. More broadly, by accounting for differential acidity, the thermodynamic model presented adequately bridges the gap between the binding ratio and the rate ratio of Eq. 1. This agreement is suggestive that a complete free energy inventory of the KSI catalytic cycle has been identified, providing an economic solution to a problem that has led others to more complex proposals (46).

**Discussion and Conclusion**

Differential acidity is typically not included in discussions about enzyme energetics, and we conjecture that this contribution could be significant in many other enzymes. The use of a careful thermodynamic decomposition of an enzyme-actuated process (such as Fig. 3A) over heuristic models (such as Fig. 1) may help reveal complexities in the proton/electron/atom inventory over the catalytic cycle, resulting in thermodynamic contributions that supplement the differential binding effect. KSI serves as an example to illustrate the usefulness of this framework, but the model’s general nature suggests that it could help identify free energy inventories for other enzymes as well. For the broad category of enzymes that catalyze proton transfer, a differential acidity effect will appear any time the active site and a ligand can exchange protons; then $\Delta G_3$ manifests as a difference in acidities between the apo and liganded states of the enzyme (as in Fig. 3B). For electron-transfer chemistry, one can imagine redrawing Fig. 3B where $K_a$s are replaced with redox potentials; then $\Delta G_3$ would appear as a difference in electron affinities between the apo and liganded states of the enzyme.

Most widely discussed hypotheses concerning the origins of enzyme catalysis—electrostatic stabilization, entropy trapping, approximation, desolvation, and strong hydrogen bonds—are believed to be encompassed by differential binding and be captured by $\Delta G_3$ (47). However, the present thermodynamic analysis suggests that these effects can only account for ~55% of the total barrier reduction by KSI. Furthermore, it is important to point out that computer models of KSI’s catalytic cycle have been able to accurately reproduce its barrier reduction and intermediate stabilization (16, 29), meaning that $\Delta G_3$ effects were implicitly accounted for in the different thermodynamic cycles used in those calculations. Identifying what physical interactions are responsible for differential acidity, and determining how its effect is borne out in a simulated reaction coordinate, is a new question that remains to future research.

At this stage, we can only speculate what physically underlies the differential acidity effect. One possibility is that the effect is primarily borne by the ligand; namely, the environment of the active site polarizes the electronic structure of the ligand in such a way as to modify its proton affinity. We do not favor this idea, because the equilibrium properties in the complexed state ($K_a^{±}$) appear to be roughly consistent with one’s expectations: With tyrosine’s standard $pK_a$ close to that of the ligand, finding the ligand present at 30% ionized (i.e., close to 50%) does not seem too extraordinary, and is consistent with the idea that the active site imposes a large perturbation onto the ligand’s electronic structure. Rather, what is more extraordinary is the $pK_a$ of the oxianion hole in the apo enzyme, an effect that has been attributed partially to the involvement of a bound water molecule in the hydrogen-bonding network (26). It is conceivable to imagine hydrophobically driven binding as a way of upsetting this delicate interaction, and so reverting the anomalously acidic oxianion hole to a more normative state. Tuning of active-site properties with ordered water molecules has been discussed in KSI (48) and in general (49). We provisionally attribute the differential acidity effect to changes in the active site (i.e., water displacement) accompanying ligand binding, although it is difficult to trace this effect back to physical causes based on empirical thermodynamic parameters alone. This highlights the importance of framing the origin of catalysis question in terms of a well-defined thermodynamic cycle (17), and points to the need to turn to computational methods, which have been able to identify large acidity shifts in other systems (50, 51).

This work has further implications for enzyme engineering. For many natural enzyme systems, the barrier reduction ($\Delta G^\ddagger = \Delta G_2 + \Delta G_3$) as calculated from kinetics is more dramatic than the measured differential binding to a transition-state.
analog (ΔG2). Additionally, in the pursuit of catalytic antibodies and designed enzymes, a high affinity for transition-state analogs (ΔG3) has not necessarily translated to catalytic competence, ΔG2 (26, 52). We posit that some of these discrepancies could be due to ASG, which involve proton-transfer chemistry (21, 53). These theories are necessary for advancing our understanding of enzyme catalysis, because in many cases, activation barriers calculated from thermodynamic premises fail to match those measured by kinetics. However, simple thermodynamic models such as Fig. 1 continue to see use because they excel at providing insight, and they are often easier to translate into design principles. We found that by expanding on a classic thermodynamic analysis to include one additional effect, we were able to calculate the catalytic power of KSI in agreement with kinetics, suggesting that it may not be necessary to invoke dynamic arguments to explain one highly proficient enzyme’s catalysis (54). A detailed thermodynamic analysis has revealed that a static picture suffices to explain KSI’s catalytic power, and its energetics is fully consistent with the thermodynamics of chemical equilibrium.

ACKNOWLEDGMENTS. S.D.F. thanks the National Science Foundation predoctoral fellowship program and the Stanford Bio-X interdisciplinary graduate scholarship program for support. This work is supported in part by a grant from the National Institutes of Health (GM27738).
Supporting Information

Fried and Boxer 10.1073/pnas.1310964110

SI Discussion: The Validity of Approximating $K_{D}^{LH}$ as $K_{D}^{app}$

In the main text, when entering differential acidity into the catalytic cycle, the approximation was made to replace $K_{D}^{LH}$ (the dissociation constant of neutral ligand to ionized enzyme, Eq. 6) with $K_{D}^{app}$ (the ligand’s apparent dissociation constant without regard to protonation states, Eq. 7). This reformulation allowed one to clearly distinguish differential acidity as a separate contribution from differential binding. In the specific case of ketosteroid isomerase (KSI), the approximation is well-founded, as $K_{D}^{LH}$ is indeed only 1.3-fold greater than $K_{D}^{app}$. In this section, we examine the range of validity of this approximation. Clearly, for some choice of thermodynamic parameters, $K_{D}^{LH}$ will not be close to $K_{D}^{app}$. In the general case, the differential acidity effect is represented as the factor by which $K_{D}^{LH}$ (the dissociation constant of ionized ligand to neutral enzyme) is lower than the apparent dissociation constant, $K_{D}^{app}$. The dependence of the log ratio of $K_{D}^{LH}$ to $K_{D}^{app}$ on the two acidity parameters is given in Fig. SI4. As $pK_{a}^{E}$ (the internal acidity of the enzyme active site) increases, $K_{D}^{LH}$ decreases, representing stronger effective binding and therefore greater stabilization. However, this effect plateaus (as shown by arrow 1) once $pK_{a}^{E}$ is equal to the ligand’s $pK_{a}$, or when the enzyme has equal affinity for protons as the bound ligand (1). The contribution from differential acidity also grows more pronounced as $pK_{a}^{E}$ (the standard acidity of the enzyme active site to water) decreases (as shown by arrow 2). This effect has no bound, and as the active-site oxyanion hole (OAH) becomes more and more acidic in the apo state, $K_{D}^{LH}$ continues to exponentially decrease, making the enzyme’s effective stabilization of the intermediate greater and greater.

For a large portion of the acidity parameter space, $K_{D}^{LH}/K_{D}^{app}$ is close to 1 (the yellow region of Fig. SI3B), where the approximation is valid. As $pK_{a}^{E}$ increases up to $pK_{a}^{L}$ (following arrow 1), this remains the case. However, once $pK_{a}^{E}$ exceeds $pK_{a}^{L}$, $K_{D}^{LH}$ begins to exponentially increase, which implies a deleterious effect of the differential acidity on intermediate stabilization and marks where the approximation breaks down. Combined with the result from Fig. SI4, it is clear that for catalytic purposes the optimum value of $pK_{a}^{E}$ is matched to the ligand’s $pK_{a}^{L}$, as increasing it beyond that value no longer affects $K_{D}^{LH}$.

To summarize, differential acidity depends on the enzyme’s internal acidity, $pK_{a}^{E}$, and the enzyme’s solution acidity, $pK_{a}^{L}$. For catalysis, differential acidity is optimized when $pK_{a}^{E} = pK_{a}^{L}$ and when $pK_{a}^{E}$ is as low as possible (these two principles could be considered as design criteria for a hypothetical artificial enzyme). So long as $pK_{a}^{E} < pK_{a}^{L}$, the approximation holds that $K_{D}^{LH} = K_{D}^{app}$. In this case, the contribution from differential acidity, $\Delta G_{c}$, is proportional to the simple expression $pK_{a}^{E} - pK_{a}^{L}$, in analogy to the expression for $\Delta G_{c} = pK_{a}^{E2} - pK_{a}^{E3}$.

A Note on tKSI

Interestingly, a homolog of KSI from a different organism (tKSI) does not have a remarkably low $pK_{a}^{E}$. However, unlike the variant of KSI_D40N, considered wherein bound inhibitors exist as a mixture of neutral and anionic forms (1, 2), tKSI_D40N appears to bind inhibitors (1, 3, 4) and the intermediate (3) preponderantly in the anionic form, indicating that its $pK_{a}^{E}$ is larger than 9.7. Therefore, tKSI with an unexceptional $pK_{a}^{E}$ may still exhibit a differential acidity effect if $pK_{a}^{E}$ is correspondingly larger. Considering Fig. SI4, a beneficial effect from differential acidity (that is, $K_{D}^{LH}/K_{D}^{app} < 1$) is still expected with an unexceptional $pK_{a}^{E} \sim 9$, so long as $pK_{a}^{E}$ is higher yet. However, unless $pK_{a}^{E}$ and $pK_{a}^{L}$ fortuitously take on values that render them measurable (as was the case for pKSI), the differential acidity effect will not be experimentally observable. Consideration of tKSI calls attention to the fact that differential acidity effects will not be measurable for all systems, which highlights another reason why computational methods are essential in thermodynamic analyses of enzyme catalytic cycles (5–7).

Dependence of microscopic binding constants on differential acidity. Plots of Eq. 5 from the main text, showing the effect of the two acidity parameters ($K_{D}^{\text{Andro}}$ and $K_{D}^{\text{N40}}$) on the microscopic dissociation constants ($K_{D}^{\text{Andro}}$ and $K_{D}^{\text{N40}}$). The vertical axis is on a log scale and is normalized to $K_{D}^{\text{Andro}}$; pH 7.2 and $K_{D}^{\text{Andro}} = 10.0$. In A, arrow 1 shows that a higher effective affinity is achieved for increasing $K_{D}^{\text{Andro}}$ up to the $K_{D}^{\text{Andro}}$ of the ligand; arrow 2 shows that higher effective affinity accompanies a more acidic active site. In B, arrow 1 shows that $K_{D}^{\text{N40}}$ is nearly equal to $K_{D}^{\text{N40}}$ and insensitive to $K_{D}^{\text{Andro}}$ until $K_{D}^{\text{Andro}}$ surpasses the $K_{D}^{\text{Andro}}$ of the ligand.

**Table S1. Thermodynamic and kinetic parameters**

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Description</th>
<th>Value</th>
<th>Determination</th>
<th>Ref(s).</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_{D}^{\text{Andro}}$</td>
<td>Binding of substrate to KSI</td>
<td>17 μM</td>
<td>UV-vis kinetics, $K_{a}$ of a “slow” substrate</td>
<td>(1)</td>
</tr>
<tr>
<td>$pK_{a}^{\text{Andro}}$</td>
<td>$pK_{a}$ of substrate</td>
<td>12.7</td>
<td>UV-vis kinetics, detailed balance</td>
<td>(2)</td>
</tr>
<tr>
<td>$pK_{a}^{\text{Enol}}$</td>
<td>$pK_{a}$ of protonated intermediate</td>
<td>10.0</td>
<td>UV-vis kinetics, detailed balance</td>
<td>(3)</td>
</tr>
<tr>
<td>$pK_{a}^{\text{N40}}$</td>
<td>$pK_{a}$ of general base in apo enzyme</td>
<td>3.75</td>
<td>pH-titration kinetics</td>
<td>(4)</td>
</tr>
<tr>
<td>$K_{D}^{\text{Andro}}(K_{D}^{\text{Enol-app}})^{-1}$</td>
<td>Binding of intermediate to KSI$^{\text{D40N}}$</td>
<td>6 nM</td>
<td>UV-vis kinetics</td>
<td>(5)</td>
</tr>
<tr>
<td>$K_{\text{int}}$</td>
<td>The internal equilibrium constant for the process $E+\text{S} = E\text{S}$</td>
<td>0.01–0.3</td>
<td>Thermodynamic analysis</td>
<td>(5)</td>
</tr>
<tr>
<td>$pK_{a}^{\text{KSI}}$</td>
<td>$pK_{a}$ of OAH in KSI$^{\text{D40N}}$</td>
<td>6.3</td>
<td>$^{13}$C NMR, UV-vis titration</td>
<td>(6)</td>
</tr>
<tr>
<td>$pK_{a}^{\text{KSI}}$</td>
<td>$pK_{a}$ of OAH in KSI$^{\text{D40N}}$</td>
<td>6.3</td>
<td>$^{13}$C NMR, UV-vis titration</td>
<td>(6)</td>
</tr>
<tr>
<td>$K_{D}^{\text{N40}}(K_{D}^{\text{Equ-app}})^{-1}$</td>
<td>Internal acidity of OAH in KSI$^{\text{D40N}}$</td>
<td>9.7</td>
<td>UV-vis kinetics</td>
<td>(7)</td>
</tr>
<tr>
<td>$k_{\text{cat}}$</td>
<td>Rate constant for $E_{\text{S}}+\text{S} \rightarrow E\text{S}$</td>
<td>1.7 × 10^{-5} s^{-1}</td>
<td>UV-vis kinetics, intermediate partitioning</td>
<td>(9)</td>
</tr>
<tr>
<td>$k_{\text{uncat}}$</td>
<td>Rate constant for $\text{AcO}^{-}+\text{S} \rightarrow \text{AcOH}+\text{I}$</td>
<td>9.0 × 10^{-3} s^{-1}</td>
<td>UV-vis kinetics</td>
<td>(3)</td>
</tr>
</tbody>
</table>

All parameters were measured with a temperature between 20 and 25 °C and with a pH between 6.9 and 7.2, except for acid dissociation constants ($pK_{a}$), which are measured over a pH range.

*These parameters were determined for the homolog of KSI from Comamonas testosterone (tKSI), a related but distinct form of KSI from the one on which the current analysis focuses (pKSI). The two are known to use the same mechanism, and possess very similar kinetic and thermodynamic specifications when subject to the same experimental conditions. Two of the parameters for which data are not available on pKSI ($K_{\alpha}^{\text{cat}}$ and $k_{\text{cat}}$) do not affect the analysis directly because they are not used in the calculations, but are compared to calculated values. Note that the Michaelis–Menten $k_{\text{cat}}$ for tKSI and pKSI are very similar (within a factor of 1.5; ref. 4). $K_{D}^{\text{N40-Enol-app}}$ is a key parameter for which data are not available on pKSI, but its value for tKSI is very close to the value of $K_{D}^{\text{N40-Enol-app}}$ for both tKSI and pKSI, justifying its use.