Assessing positioning in enzymatic catalysis via ketosteroid isomerase conformational ensembles

Filip Yabukarski[1*], Justin T Biel[6], Margaux M Pinney[1], Tzanko Doukov[7], Alex S Powers[2,4,5,], James S Fraser[6], Daniel Herschlag[1,3,5*]

[1]Department of Biochemistry, [2]Department of Chemistry, [3]Department of Chemical Engineering, [4]Department of Computer Science, [5]Stanford ChEM-H, Stanford University, Stanford, California 94305, United States, [6]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, United States, [7]Stanford Synchrotron Radiation Light Source, SLAC National Accelerator Laboratory, Menlo Park, California 94025, United States.

*correspondence:
fyabukar@stanford.edu
herschla@stanford.edu

Abstract

Our physical understanding of enzyme catalysis has been limited by the scarcity of data for the positioning and motions of groups in and around the active site. To provide foundational information and test fundamental catalytic models, we created conformational ensembles from 45 PDB crystal structures and collected new 'room temperature' X-ray crystallography data for ketosteroid isomerase (KSI). Ensemble analyses indicated substantial pre-positioning and minimal conformational heterogeneity loss through the reaction cycle. The oxyanion hole and general base residues appear conformationally restricted, but not exceptionally so relative to analogous non-catalytic groups. Analysis of surrounding groups and mutant ensembles provide insight into the balance of forces responsible for local conformational preferences. Oxyanion hole catalysis appears to arise from hydrogen bond donors that are stronger than water, without additional catalysis from geometrical discrimination, more distal effects, or environmental alterations. The presence of a range of conformational sub-states presumably facilitates KSI's multiple reaction steps.

Introduction

Given the central role of enzymatic catalysis for life, there have been extensive efforts to develop a deep understanding of enzyme mechanisms (Benkovic et al., 2008; Blow, 2000; Fersht, 1985; Hammes et al., 2011; Knowles, 1991; Kraut et al., 2003; Ma et al., 2000; Warshel, 1998; Wolfenden and Snider, 2001). Studies to elucidate the origins of enzymatic catalysis have routinely identified active site groups that facilitate the substrate's chemical transformation, such as the general base and oxyanion hole hydrogen bond donors shown in **Figure 1A and B** for ketosteriod isomerase (KSI) (Bryan et al., 1986; Fersht, 1985; Henderson, 1970; Kim et al., 1997; Kraut, 1977; Pollack, 2004; Raines et al., 1986; Robertus et al., 1972; Walsh, 1979). Yet, these and other so-called "catalytic groups" provide little catalysis outside of the context of a folded protein; the positioning of catalytic groups by the folded enzyme and of substrates within the active site via binding interactions are needed to achieve optimal, or even measurable, catalysis. Accordingly, proposals to account for enzymatic rate enhancements over the past century have consistently invoked the importance of residues beyond the active site and, in particular, the importance of positioning of active site groups for catalysis (Bruice, 1970, 1976; Haldane, 1930; Henderson and Wang, 1972; Hur and Bruice, 2003; Jencks, 1975; Klinman, 2015; Menger, 2005; Milstien and Cohen, 1970; Nowak and Mildvan, 1972; Pauling, 1946; Storm and Koshland, 1970; Warshel, 1998). However, these catalytic proposals have been difficult to evaluate, perhaps due to technical limitations.

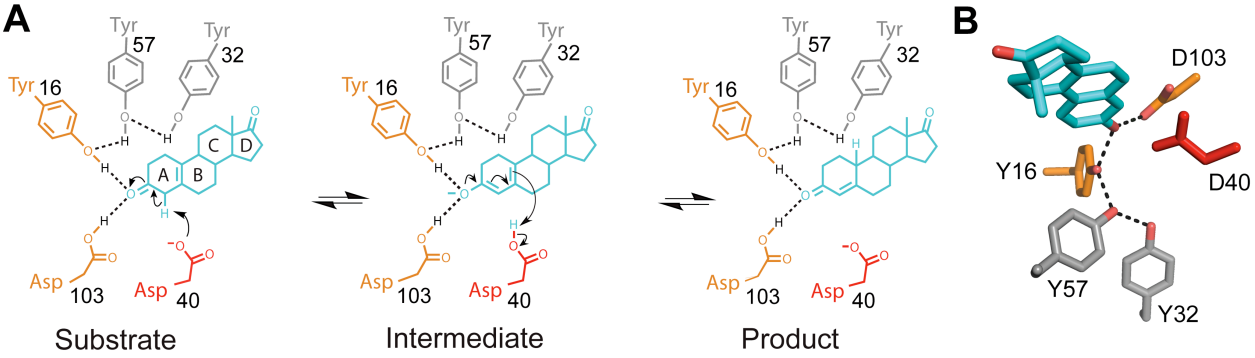

Figure 1. The Ketosteroid Isomerase (KSI) reaction. Reaction mechanism and schematic depiction of the active site (**A**) and its three-dimensional organization (**B**) (PDB 1OH0 (Kim et al., 1997)). KSI catalyzes double bond isomerization of steroid substrates utilizing a general acid/base D40 (which we refer to herein as a general base, for simplicity), and an oxyanion hole composed of the side chains of Y16 and D103 (protonated); general base and oxyanion hole residues are colored in red and orange, respectively.

Classical catalytic proposals invoked static structures of shape and charge complementarity to the reaction's transition state (Eyring, 1935; Haldane, 1930; Jencks, 1969, 1975; Lienhard, 1973; Pauling, 1946; Polanyi, 1921). More recently emerging models have emphasized the occurrence of and importance of structural dynamics in protein function and enzyme catalysis (Boehr et al., 2009; Cooper and Dryden, 1984; Fraser et al., 2009; Henzler-Wildman and Kern, 2007; Klinman, 2015; Klinman and Kohen, 2013; Peng, 2009; Smock and Gierasch, 2009; Tokuriki

and Tawfik, 2009). For example, enzymes commonly contain loops that close to interact with substrates, and emerging results provide evidence that the most reactive sub-states in enzymatic reactions are not the most populated (Callender and Dyer, 2015; Klinman, 2009, 2013; Reddish et al., 2014). It has been also suggested that the ability to adopt reactive conformations may be dynamically linked to surface residues and solvent (Agarwal, 2004, 2019; Agarwal et al., 2002, 2004; Meadows et al., 2014; Ramanathan and Agarwal, 2011) and further proposed that catalytic rates are gated by collective dynamic events of the protein (Silva et al., 2011; Suarez and Schramm, 2015).

Considering both positioning and dynamics, it has been recognized that:

> "For catalysis, flexible but not too flexible, as well as rigid but not too rigid, is essential. Specifically, the protein must be rigid enough to maintain the required structure but flexible enough to permit atomic movements as the reaction proceeds." (Hammes et al., 2011)

In other words, *too much* flexibility will hinder catalysis, with the extreme being an unfolded, inactive enzyme, and *too much* rigidification would prevent the atomic motions needed to rearrange bonds in the course of the chemical reaction. Inherently, all atoms move due to thermal excitation and enzymes are held together by weak, noncovalent forces (Dill, 1990; Pace et al., 2014); nevertheless, we do not know how much a folded protein dampens these motions or whether it could rigidify an active site so much that reaction would be repressed. Indeed, the plethora of proposals for enzymatic catalysis reflects the importance and complexity of enzyme action and perhaps also our limited means of testing positioning and dynamics—the core elements of these proposals.

Key to resolving these fundamental questions is experimental information about precision *vs.* heterogeneity. In particular, we want to know the extent of positioning of residues in the active site prior to substrate binding, once the substrate is bound, and when transition state or transition state-like interactions are made. This information will be critical in evaluating catalytic mechanisms. For example, is positioning precise enough to discriminate geometrically between optimal ground and transition state orientations? Is the positioning of active site and substrate groups sufficiently restricted to maintain local binding interactions such as hydrogen bonds, or are the motions extensive enough such that individual interactions are regularly made and broken within an enzyme-ligand complex? Most broadly, determining how much and what type of positioning is needed to achieve the high catalytic efficiency of natural enzymes and how natural enzymes position catalytic groups is information that will ultimately be needed to design enzymes that rival those present in Nature.

Here we investigate positioning and conformational heterogeneity in an enzyme without complicated loop or domain closures needed for catalysis, as these larger-scale motions are not found in most enzymes and when they are, they represent only a small fraction of the active site interactions (Gutteridge and Thornton, 2005). Further, we chose an enzyme without catalytic

domain and loop closures in order to focus on smaller-scale positioning and heterogeneity that is characteristic of all enzymes. Positioning and heterogeneity on this smaller-scale must be understood to assess and understand active site positioning and its relation to catalysis but has been difficult to evaluate because of limitations of traditional approaches.

We utilize emerging techniques that now allow us to evaluate the precision in positioning of active site residues and the extent to which enzyme conformational ensembles change through the reaction cycle. We begin by testing two general models that are difficult to assess via traditional approaches. We then determine the conformational heterogeneity of the groups directly involved in catalysis to provide foundational information about those residues, and we evaluate the interactions responsible for limiting and allowing their motions. Finally, we use our dynamic information, along with new and prior functional data, to evaluate contributions to catalysis.

The first model we test derives from Wolfenden's suggestion in 1976 that enzymes undergo changes in shape when forming the enzyme-substrate complex and undergo additional subsequent changes to give maximal complementarity to the transition state, where stabilization is greatest (Wolfenden, 1976). However, analysis of 60 enzymes with apo and ligand-bound structures revealed little structural change upon formation of enzyme-substrate complexes (Gutteridge and Thornton, 2005), providing evidence against this proposal which we refer to as the "Static Gradual Adaption Model". Here, we extend this model, progressing from average structures to dynamic ensembles to test whether alignment of enzyme groups might proceeds through the reaction coordinate to become most precise in the transition state (**Figure 2A**). This "Dynamic Gradual Adaptation Model" aligns with the widely discussed perspective dating back to Haldane, Polanyi, Pauling, and Jencks that enzymes are most complementary in charge and shape to their transition states so that a better fit in and stronger interactions to the transition state may lead to increased order (Haldane, 1930; Jencks, 1969, 1975; Lienhard, 1973; Pauling, 1946).

A second general dynamic model that we test derives from the observation by Shoichet and colleagues (Shoichet et al., 1995; Zhang et al., 1995b) that active site residues can be destabilizing to the folded protein, consistent with evolutionary selection of these residues for catalysis rather than for overall stability, and with Vallee and Williams' "entatic state" proposal that active site groups are distorted from their most stable free conformation (by folding free energy) to more closely match the conformational or electronic needs of a reaction's transition state (Vallee and Williams, 1968; Williams, 1972). In accord with these observations and model, more precise positioning of active site residues relative to residues throughout the enzyme would destabilize the folded enzyme, from the loss in conformational entropy, but could contribute to catalysis by increasing the probability of optimal transition state interactions. We refer to this as the "Dynamic Entatic State Model" (**Figure 2B**) and we test it herein by comparing the conformational heterogeneity of analogous active site and non-active site groups.

Testing these models and obtaining foundational information about active site positioning and conformational heterogeneity requires the conceptual perspective of the protein conformational energy landscape and resultant conformational ensembles, as introduced by Frauenfelder and coworkers (Austin et al., 1975; Frauenfelder et al., 1979, 1991) and as is now commonly invoked (e.g., Boehr et al., 2009; Fraser et al., 2011; Keedy, 2019; Smith et al., 1990; Wand et al., 2013). Testing these models also requires new experimental tools to obtain ensembles, which have been limiting until recently.

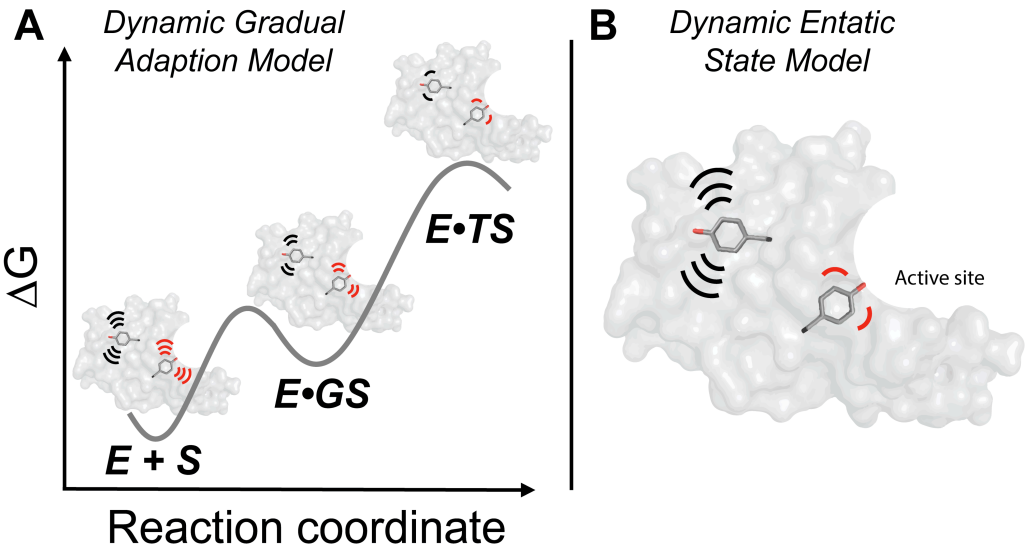

Figure 2. Dynamic models for enzyme catalysis. Both panels show an enzyme with two highlighted tyrosine residues, a 'non-catalytic' tyrosine (black) in the enzyme core (grey) representing the non-catalytic residues, and a 'catalytic' tyrosine (red) to represent catalytic residues in the active site. (**A**) The Dynamic Gradual Adaption Model. Both non-catalytic and catalytic tyrosine residues become more conformationally restricted as the reaction proceeds, but with greater conformational restriction occurring in the transition state. (**B**) The Dynamic Entatic State Model. Folding energy and local interactions provide greater restrictions and more precise positioning of the catalytic tyrosine (red, representing active site residues), relative to a non-catalytic tyrosine (black, representing non-catalytic residues). The restriction of active site residues (reduced conformational heterogeneity) is "paid for" with folding free energy and is used to enhance catalysis according to this model. In both panels, motions are schematically depicted by the motion lines.

NMR in particular has been a powerful method for identifying changes in mobility of different regions and across different states (Boehr et al., 2006; Frederick et al., 2007; Henzler-Wildman and Kern, 2007; Kumar et al., 2018; Marlow et al., 2010; Sekhar and Kay, 2019; Takeuchi et al., 2015). Nevertheless, NMR dynamic data do not typically provide information on the direction and absolute extent of motion (Kleckner and Foster, 2011; Kovermann et al., 2016). Molecular dynamics approaches can provide detailed conformational ensembles and atomic motions, but computational outputs may or may not reflect actual physical behavior and sampling of the conformational landscape may not be complete (Childers and Daggett, 2018; Clarage et al., 1995; van Gunsteren et al., 2018). Traditional X-ray crystal structures provide information about atomic positions and are experimental, but only indirectly provide ensemble properties, as a

static X-ray model represents a conformational snapshot (Brünger and Nilges, 1993; Furnham et al., 2006; Holton et al., 2014; Rejto and Freer, 1996).

Fortunately, two X-ray crystallographic approaches have recently emerged that can provide the atomic ensemble information needed to test the above models: high sequence similarity PDB structural ensembles (referred to as 'pseudo-ensembles' herein) (Best et al., 2006; Zoete et al., 2002) and multi-conformer models from X-ray data obtained at temperatures above the protein's glass transition (referred to as 'room temperature' or 'RT' X-ray diffraction in the literature and herein) (van den Bedem et al., 2009; Fraser et al., 2009, 2011; Keedy et al., 2015a). In principle, these approaches are complementary. Pseudo-ensembles provide information about residues and functional groups that move in concert (i.e., coupled motions) but require dozens of structures and can be subject to conformational distortions due to cryogenic (cryo, 100 K) freezing (Halle, 2004; Juers and Matthews, 2001; Keedy et al., 2014; Sandalova et al., 1999). RT multi-conformer models lack direct information about coupled motions but provide information at temperatures at which harmonic and anharmonic protein motions are enabled and are typically obtained from single crystals, allowing ensemble information about new complexes or mutants to be readily obtained (Keedy et al., 2015b; Lewandowski et al., 2015; Ringe and Petsko, 2003; Tilton et al., 1992). However, to date, ensembles from both approaches have not been obtained for the same system and compared. Herein we took advantage of both cryo and RT X-ray crystallography to provide complimentary, high-resolution ensemble information and insights not possible from static structures or lower resolution and less extensive data.

We built pseudo-ensembles and RT-ensembles for the enzyme KSI, choosing KSI because of our ability to obtain diffraction data of high-resolution, because of the accumulated wealth of structural and mechanistic information for this enzyme, because of the lack of catalytic loop or domain closures, and because of KSI's use of two very common catalytic features—an oxyanion hole and a general base (**Figure 1**). We used 45 literature cryo X-ray KSI structures and we obtained X-ray data at 250 K and 280 K for KSI in its Apo state and with bound ground and transition state analogs to resolutions of 1.15 – 1.55 Å. With these data we were able to test the broad models for enzymatic catalysis described above in **Figure 2**, obtain fundamental insights into the precision with which catalytic groups are positioned, and evaluate catalytic proposals.

We observe that the KSI catalytic groups exhibit considerable conformational heterogeneity, yet appear more positioned than average, but not extraordinarily so relative to non-catalytic groups. The multiple states of the catalytic groups and bound ligands likely allows KSI to catalyze multiple reaction steps for multiple substrates. The catalytic groups move anistropically, up to about the order of a bond length (~1 Å) and down to about the scale of a bond vibration (~0.2 Å). The extent of conformational sub-states in the oxyanion hole suggests that geometric discrimination between the sp^2 ground state and sp^3 transition state is unlikely (Simón and Goodman, 2010); instead, KSI appears to achieve oxyanion hole catalysis, relative to solution reactions, by utilizing hydrogen bond donors that are stronger than water, and these hydrogen bonds appear to be well-maintained. Nevertheless, the observed highly restricted motion in

particular dimensions raises the possibility that other enzymes may be able to utilize geometric discrimination despite their ensemble nature. We also evaluated the origin of the high effective molarity of the KSI general base of 10^3–10^5 M (Lamba et al., 2016), revealing a paradox between the high efficiency general base catalysis and its conformational leeway. Analysis of the residues surrounding the catalytic groups allowed us to build and test models for the interplay of forces responsible for conformational restrictions and motions, information that may be needed to meet the ultimate challenge of the routine design of new, highly-efficient enzymes. Our in-depth analysis of a model enzyme provides a basis for future studies that bring an ensemble perspective to bear on traditional structure–function studies.

Results

Evaluating structural changes through the KSI catalytic cycle and testing the Static Gradual Adaption Model

To evaluate structural changes during the KSI catalytic cycle and test Wolfenden's Static Gradual Adaption Model, we took advantage of the 94 crystallographically-independent KSI molecules from the 45 cryo crystal structures available in the Protein Data Bank (PDB, **Berman et al., 2000**) (**Table S1**). Visually, there were no discernable differences between Apo, ground-state-analog bound (GSA-bound), and transition-state-analog bound (TSA-bound) structures, species that represent the major KSI catalytic states (**Figure 3A, Figure supplement 1B-E**). To quantitatively compare these states, we determined backbone root mean square deviations (RMSDs) for each cryo KSI structure relative to the highest resolution WT Apo, GSA-bound, and TSA-bound cryo structures. Because cryo cooling can induce structural changes (Halle, 2004; Keedy et al., 2014; Sandalova et al., 1999; Scheidig et al., 1999; Teeter et al., 2001), we also obtained Apo, GSA-bound and TSA-bound KSI structures from high-resolution RT X-ray diffraction data (**Table S3**), which we also compared with each of the highest resolution Apo, GSA- and TSA-bound cryo structures. In all cases, the RMSDs were below 1.0 Å, and below 0.5 Å when two short loops, encompassing residues 62–65 and 91–96, were excluded (**Figure 3B, C, Figures supplement 2–3, Tables S4-5**). The 91–96 loop interacts with the substrate but affects binding and catalysis <2-fold (Yun et al., 2005, Schwans et al., 2009); the 62–65 loop interacts with the 91–96 loop but does not interact with the substrate directly. Given these small effects, we focus on the folded enzyme core and the remaining ~10^{11} fold catalysis herein (Radzicka and Wolfenden, 1995).

The universal high structural concordance for regions other than the two loops, with RMSDs of <0.5 Å (**Figure 3B, Figures supplement 2-3**), and the absence of substantial conformational changes from cryo-cooling suggested that there are at most modest structural differences through

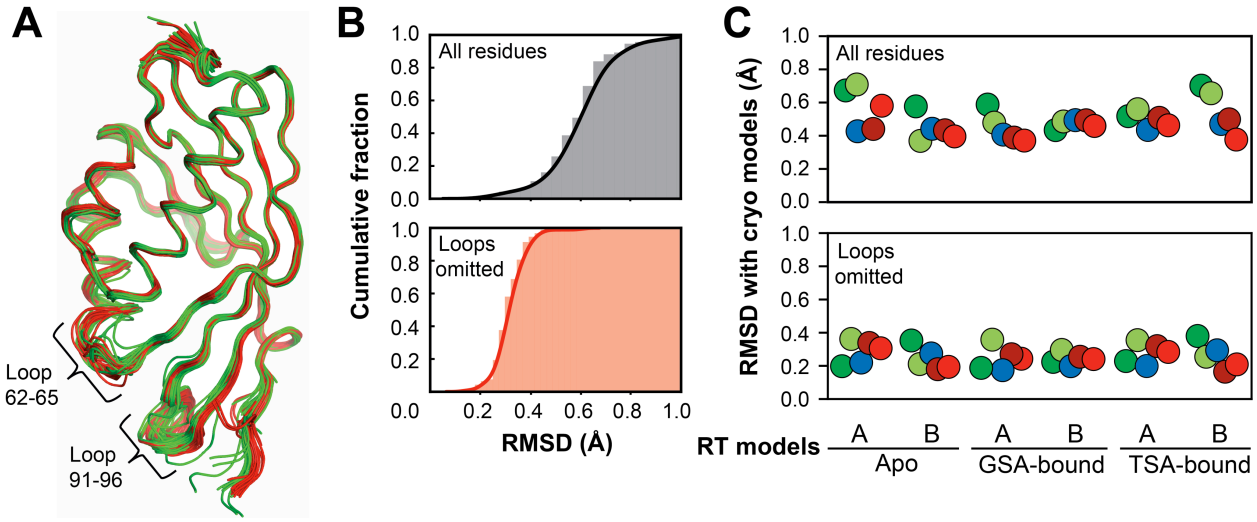

Figure 3. Comparing KSI crystal structures to test the Static Gradual Adaptation Model. (**A**) Alignment of all 94 KSI crystallographically-independent molecules from the 45 PDB crystal structures (in ribbon): Apo (green), GSA-bound (blue) and TSA-bound (red). The positions of the two flexible loops are noted. Throughout we evaluate individual monomers of the KSI dimer (see **Materials and Methods**). (**B**) Cumulative fraction of backbone RMSDs of all KSI structures *vs.* the highest resolution WT-TSA bound structure (PDB 1OH0, molecule A) shown for the entire sequence (top, grey) and excluding the 62–65 and 91–96 loops (bottom, orange). (**C**) Backbone RMSDs of Apo, GSA-bound and TSA-bound KSI single-conformation structures obtained at RT (280 K, this study; x-axis) *vs.* the highest-resolution cryo (100 K, from the PDB; y-axis). There are two crystallographically-independent molecules for each enzymes state, for both RT and cryo structures, except for the GSA-bound cryo structure, for which there was only one GSA-bound molecule. Each circle represents the RMSD between each of the two independent KSI molecules from the RT structures (A and B, x-axis) and each of the independent molecules from the cryo structures; RMSD with each of the two molecules from the cryo Apo structure (PDB 3VSY) in light and dark green; with the single molecule from the GSA-bound cryo structure (PDB 5KP4) in blue; and with the two molecules from the TSA-bound cryo structure (PDB 1OH0) in red.

the reaction cycle and that KSI is highly preorganized for catalysis. These highly similar structures provide further evidence against the Static Gradual Adaptation Model for KSI, similar to conclusions from prior analyses of crystal structures for 60 enzymes, where structural differences between Apo and ligand-bound states were shown to be modest (RMSD < 1 Å) and no larger than differences between two Apo forms of the same enzyme (Gutteridge and Thornton, 2005). Nevertheless, enzymes may change conformational heterogeneity through a reaction without changing their average structures, as proposed in the Dynamic Gradual Adaption Model (**Figure 2A**).

Conformational heterogeneity through the KSI catalytic cycle and tests of the Dynamic Gradual Adaptation Model

Obtaining KSI conformational ensembles. To build experimental conformational ensembles for KSI, we first built pseudo-ensembles composed of a large number of traditional X-ray crystal structures, in which each crystal structure (including structures with different point mutations and different crystallization conditions) is considered to correspond to a local minimum on the

native potential energy surface (Best et al., 2006; Elber and Karplus, 1987; Zhang et al., 1995a; Zoete et al., 2002) (**Figure 3A, Figure supplement 1B-D**). We were able to build an Apo and a TSA-bound pseudo-ensemble for KSI (**Table S2**). The individual structures range in resolution from 1.1 to 2.5 Å (**Figure supplement 1A**) and neither inclusion of only high-resolution structures (≤2.0 Å) nor random omission of structures substantially altered the analyzed ensemble properties (**Figures supplement 4-6**).

The degree of motion extracted from pseudo-ensembles has been shown to agree well with estimates of motion from solution NMR (Best et al., 2006), but cryo-cooling could alter rotamer distributions or the directions of motions in ways that are not necessarily captured by NMR observables (Kleckner and Foster, 2011; Kovermann et al., 2016). Additionally, the scarcity of cryo GSA-bound KSI structures (n = 1) prevented us from obtaining a pseudo-ensemble for this state, prompting us to use RT X-ray diffraction to capture GSA-bound state dynamic information, as a single RT crystallographic dataset obviates the need to obtain many independent traditional cryo structures.

Ensembles to evaluate conformational heterogeneity and test the Dynamic Gradual Adaptation Model. To provide an overall quantitative measure of the conformational heterogeneity for each residue within our pseudo-ensembles, we defined an atomic mean deviation parameter (MDev). For a given atom in a structure, the MDev describes the average displacement of equivalent atoms within the ensemble of structures, with lower and higher values representing smaller and larger positional fluctuations, respectively, corresponding to less or more conformational heterogeneity (see **Materials and Methods**).

In the Apo state, MDevs for the backbone (Cα) and side chains (Cβ) in each state were below 0.5 Å, with exceptions only in the 62–65 and 91–96 loops (**Figure 4A,** top panel, **Figure supplement 8A**, **9A, 10A** top panels). The MDevs for the catalytic residues were below average for the enzyme (dotted line) and on the lower end of observed values, and the MDevs for substrate binding residues were close to the average (**Figure 4A,** top panel, **Figures supplement 8A, 9A, 10A**:). Thus, our pseudo-ensemble analysis is consistent with a preorganized active site, but it does not indicate extreme or unusual positioning therein. We provide atomic-level analysis of this preorganization and the catalytic implications in the following sections.

Returning to the Extended Gradual Adaptation Model (**Figure 2A**), we wanted to assess changes in conformational heterogeneity through the reaction cycle, as mimicked by the progression from Apo to KSI with bound ground and transition state analogs (**Figure 1A**). There were sufficient cryo X-ray structures to build Apo and TSA-bound pseudo-ensembles and to compare Apo state MDevs to those in the TSA-bound state (**Figure 4A** and **Figure supplement 8**). The MDev values appeared to be similar across the entire structure, as seen qualitatively by comparison of the top and bottom panels in **Figure 4A** and quantitatively in **Figure 4B** by the difference in MDev between the states (note the difference in scale in **Figure 4A** *vs.* **4B**) and in **Figures 4C**

by the strong correlation of the MDev values for the Apo and TSA-bound states (R^2 = 0.80; **Figures supplement 6, 9D**). Nevertheless, the slope of this correlation was less than 1 (slope = 0.75), crudely suggesting an overall dampening of ~25% in conformational heterogeneity of the enzyme core upon binding of the TSA (**Figures 4C**). This modest dampening in the TSA complex is supported by smaller MDevs with the TSA bound, with an average reduction ($\Delta MDev_{Apo\text{-}TSA}$) of 0.05 Å per residue (**Figure 4B, D, Figures supplement 9B, C and 10D**). We also obtained similar results for a homologous KSI ($KSI_{homolog}$) from another organism for which there were fewer but nevertheless sufficient number of available cryo X-ray structures (**Figure supplement 11**).

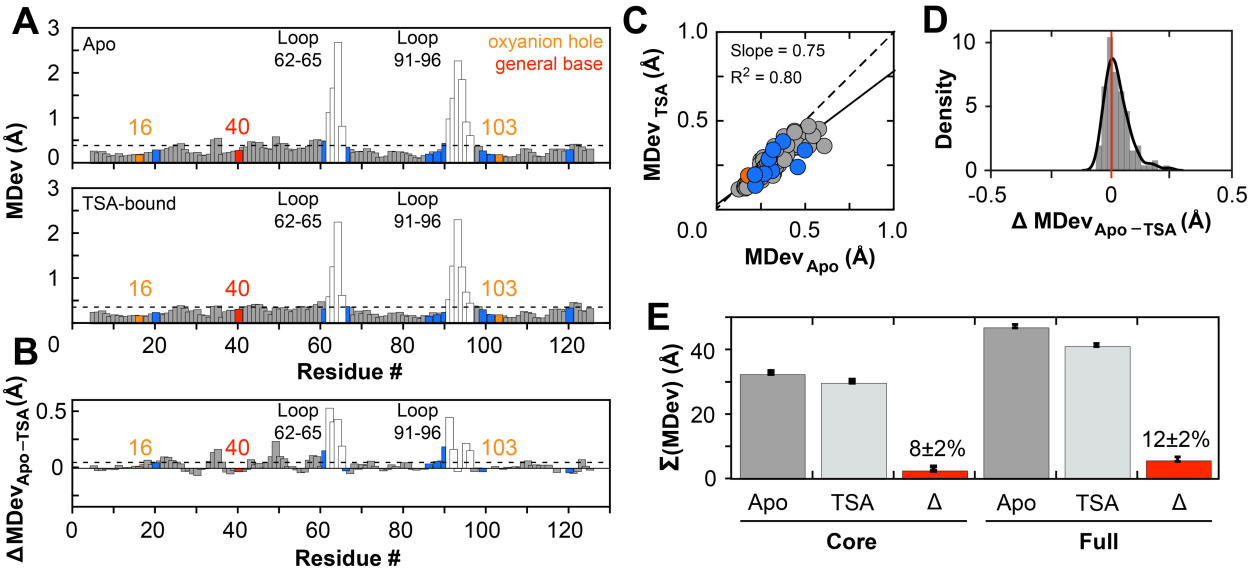

Figure 4. Assessing conformational heterogeneity through the KSI catalytic cycle via pseudo-ensembles. (**A**) Cα MDevs for KSI Apo (top) and TSA–bound (bottom) states. Dashed lines represent the average MDev. The flexible 62–65 and 91–96 loops are shown as white bars, Y16 and D103 in orange, D40 in red, and binding residues in blue. (**B**) The difference Cα MDev values between the Apo and TSA-bound states ($MDevs_{Apo\text{-}TSA}$), such that positive values indicate lower MDevs for the TSA-bound state. Note the difference in y–axis range relative to panel (A). (**C**) Correlation plot of Apo and TSA-bound Cα MDevs (excluding loops 62–65 and 91–96). The dashed line of slope 1 represents the expectation for no difference in average MDevs between the two states. (**D**) Histogram of MDev differences from panel B ($MDevs_{Apo\text{-}TSA}$) for the enzyme core (i.e., loops excluded). (**E**) Sum of Cα MDevs for Apo (dark grey bars), TSA-bound (light grey bars) and their difference (Δ, red bars). Errors were estimated using bootstrap analysis and error propagation (see **Figure supplement 7** and **Materials and Methods**). Analysis of side-chain MDevs (using Cβ) gave analogous results (**Figures supplement 9 and 10**).

We further evaluated conformational heterogeneity and tested the Extended Gradual Adaptation Model using our RT X-ray data. To estimate conformational heterogeneity from our RT X-ray data we obtained multi-conformer models for each catalytic state (**Table S3**). We then calculated crystallographic order parameters, S^2, that report on local conformational heterogeneity by capturing bond vector motions (**Figure supplement 12**) (Fenwick et al., 2014). We used S^2, rather than MDevs, because our RT multi-conformer models contain additional information

within each state that is captured by S^2 and because the limited number of conformational states in multi-conformer models limits the utility of MDev comparisons. Further, crystallographic S^2 values agree well with S^2 values obtained by solution NMR (Fenwick et al., 2014). S^2 values range from 0, for a completely unrestrained bond vector, to 1, for completely rigid bond vector, and we used the complementary ($1-S^2$) disorder parameter, so that values increase with increasing heterogeneity, analogous to the MDev values described above.

Even though ($1-S^2$) and MDev are different measures of heterogeneity, we observed similar reductions in overall heterogeneity of 10% and 8±2% in the core and 12% and 12±2% including the loops, from these two measures, respectively (**Figure 4E, Figure supplement 14**). We obtained information about heterogeneity in the KSI•GSA complex from RT X-ray data, information that cannot be extracted from the single cryo KSI•GSA structure (**Table S3**, **Figure supplement 12-14**). Comparisons of the KSI•GSA complex to the Apo and TSA-bound forms revealed that most of the reduction in heterogeneity occurs upon formation of the GSA complex (**Figure supplement 14**).

In summary, analysis of both cryo pseudo-ensembles and multi-conformer models from RT X-ray data provides evidence for dominant active site preorganization with only modest dynamic adaption to the transition state and thus evidence against extensive dampening as proposed in an Extended Gradual Adaptation Model (**Figure 2A**). The modest conformational heterogeneity changes arise primarily from the placement of a ligand in the active site, rather than from increased interactions with the negative charge of the oxyanionic TSA. Further, no regions, beyond the two loops that bind distal regions of the substrate, exhibit substantial structural reorganization or reconfiguration of internal conformational heterogeneity; in other words, the flexibility of the KSI catalytic core remains strikingly similar during catalysis. We address additional aspects of the interplay of positioning, dynamics and catalysis below and in the Discussion.

Testing the Dynamic Entatic State Model and evaluating the precision in positioning of catalytic residues

Proposals for the origin of enzymatic power invoke positioning of the groups involved in catalysis and transition state stabilization, but we do not know how positioned these groups are for any enzyme. One class of models invokes catalytic groups that are especially constrained, leaving them less flexible than residues not directly involved in catalysis. This added positioning would be destabilizing, due to decreased conformational entropy, and thus can be described in terms of a Dynamic Entatic State Model (**Figure 2B**).

To test the Dynamic Entatic State Model and to evaluate the precision in positioning of KSI catalytic groups, we assembled suitable ensembles. We obtained a reduced pseudo-ensemble

(n=54 KSI molecules) by excluding structures from our overall (full) pseudo-ensemble with mutations directly to the residues under analysis and mutations previously identified to alter the positioning of these residues (**Table S2**). From our RT X-ray data, we created an ensemble by combining the multi-conformer models for the KSI Apo, GSA-bound, and TSA-bound states, given the highly similar overall conformational heterogeneity of these states and their high residue-by-residue similarities (**Figures supplement 12, 14**). The conformational heterogeneity inferred from the reduced pseudo-ensemble and the RT X-ray ensemble correlated well with the conformational heterogeneity from the original pseudo-ensemble, suggesting that overall ensemble information is retained (**Figures supplement 16-17**).

Testing the Dynamic Entatic State Model: Are active site residues in the oxyanion hole specially positioned? We calculated MDev values for the catalytic atoms of Y16 and D103 (the oxyanion hole) and of D40, the general base (**Figure 1**), and, to provide a suitable comparison, we used chemically similar atoms from like residues throughout KSI. We carried out these calculations using both the reduced pseudo- and the RT-ensembles (**Figure 5**). Comparison of the catalytic and "other" groups revealed the following. The oxyanion hole catalytic groups sit at the lower end of observed MDev's but with values similar to the most constrained non-catalytic groups. The general base oxygen atom of D40 is also not unusually constrained and indeed exhibits more motion than the most fixed groups, a paradox we address below (see "*General base positioning and catalysis*"). Thus, our ensemble data provide evidence against the Dynamic Entatic State Model for KSI, and provide unique information about the positioning of residues throughout this enzyme.

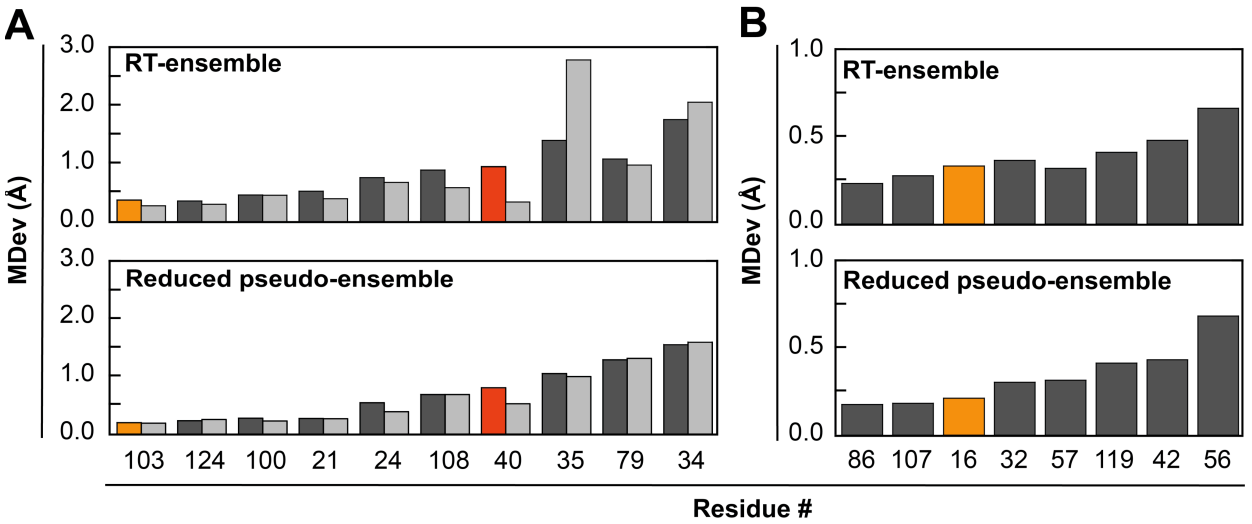

Figure 5. Testing the Dynamic Entatic State Model for KSI. Comparison of MDevs (**A**) for all Asp and Asn residues (Oδ1 atoms in light grey, O/Nδ2 atoms in dark grey) and (**B**) for all Tyr and Phe residues (OH and Cζ atoms), respectively. **Figure supplement 18** shows the positions of these residues mapped onto the KSI structure and side-chain atom nomenclature. Y16 and D103 are in orange, and D40 is in red. The upper panels show RT-ensemble results and the lower panels, the results from the reduced pseudo-ensemble. **Figure supplement 19** shows analogous results for comparisons of side chain dihedral angles.

Evaluating the precision in positioning of catalytic residues. Given our limited knowledge of the motions of and restrictions on catalytic residues, and the likely need for this information to understand catalysis and the interplay of protein forces, we looked more closely at the motions of these residues. **Figure 6** shows the ensemble distributions for the catalytic residues obtained from both ensemble methods (see also **Figures supplement 20-21**). The motions are strikingly anisotropic, indicative of differential constraints, a result that is perhaps not surprising given the non-homogeneity of protein interiors. There are motions of ~0.5–1 Å for each catalytic atom in two dimensions and considerably less motion in the third dimension (**Figures supplement 20** and **21A, B**). Thus, depending on the direction, the catalytic atoms are restricted to ~0.2 Å, slightly more than the scale of bond vibrations, or allowed up to 1 Å motion, slightly less than a typical carbon-carbon bond length. We address potential origins for the allowed and restricted motions along with catalytic implications in the following section.

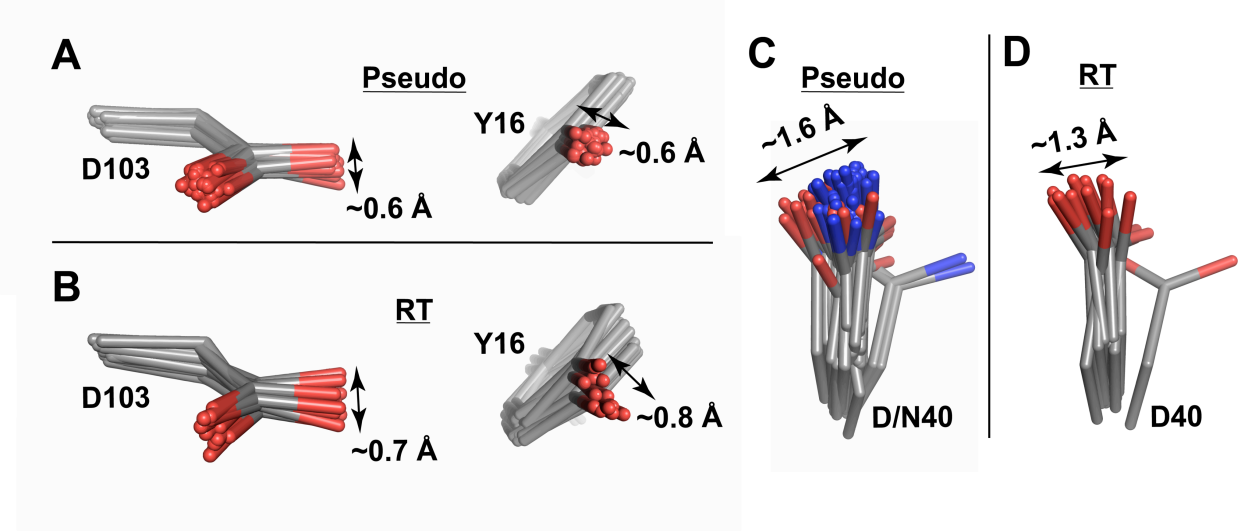

Figure 6. The precision in positioning of KSI catalytic groups. The oxyanion hole (Y16 and D103) and general base (D40) pseudo-ensembles (**A** and **C**, respectively) and corresponding RT-ensembles (**B** and **D**, respectively). Phenylalanine residues at position 16 are omitted, and chlorine atoms in chemically modified tyrosine residues have been omitted for clarity (see also **Figures Supplement 15, 20**). The D40 pseudo-ensemble contains both aspartate and asparagine residues, as asparagine mimics the protonated (intermediate) state of the general base and increases the affinity for TSAs (Kraut et al., 2006; Petrounia and Pollack, 1998). Including only aspartate or only asparagine residues does not substantially alter the range or observed motions (**Figure Supplement 21C**).

Our KSI data allow us to compare the cryo and RT-ensembles for the same protein, which is of interest because of evidence for conformational restrictions and effects at temperatures below the ~180–220 K glass transition (Fraser et al., 2009; Keedy et al., 2014, 2015b, 2018). As noted above, the overall extent of motion, evaluated at the Cα and Cβ of each residue, is highly similar (**Figure supplement 16**). In line with previous observations of restricted motions at temperatures below the glass transition, the comparisons of the side chain ensembles for the catalytic residues in **Figure 6** reveal differences. Most notably, the Y16 oxygen distribution is more isotropic in two dimensions for the pseudo-ensemble, suggesting that temperatures below the glass

transition may restrict certain KSI motions (Fenwick et al., 2014; Fraser et al., 2011; Keedy et al., 2014).

For the general base, there is a state with the general base rotated away from the main catalytic state in both ensembles (**Figure 6C, D** and **Figure supplement 21**). In the pseudo-ensemble, the rotated general base state is present in a structure bound to a single-ring TSA and may have been stabilized by interactions specific to the bound TSA. In the RT-ensemble, the rotated state is observed in the GSA-bound state. Intriguingly, the presence of this general base conformation correlates with the presence of a half-bound GSA pose (with the ligand only halfway into its binding site) that disappears upon cryo-freezing (i.e., in frozen crystals only the main catalytic general base state and the full-bound GSA pose are observed) (**Figure supplement 30)**. Perhaps this state is lost because it is energetically unfavorable at lower temperatures and the conformational rearrangement is sufficiently rapid to occur during cryo-cooling (Halle, 2004).

Catalytic implications of KSI active site positioning

Beyond testing the broad catalytic models above, our KSI ensembles provide a window into molecular behaviors that are central to its catalysis and raise paradoxes and new questions. In addition, this ensemble information allows us to begin to evaluate the interactions and forces that are responsible for positioning and thus function, information that will needed to effectively design new enzymes.

Oxyanion hole positioning and catalysis

Are hydrogen bonds to the transition state analog maintained in the enzyme ensemble and, if so, how? The wide range of observed positions of the Y16 and D103 oxygen atoms, spread over nearly 1 Å, in the simplest scenario might result in hydrogen bonds that vary considerably in length and even transiently rupture (**Figure 6A, B**). However, prior ^{1}H NMR data strongly suggest that these hydrogen bonds are stably maintained, and their chemical shifts provide evidence for short KSI hydrogen bonds with lengths in the range expected based on small molecule studies (Harris and Mildvan, 1999; Jeffrey and Yeon, 1986; Kraut et al., 2006; Mildvan et al., 1999; Pinney et al., 2018; Sigala et al., 2009; Zhao et al., 1996). Our pseudo-ensembles exhibit a narrow range of O•O hydrogen bond lengths, of about ±0.1 Å, much smaller than the displacements of oxyanion hole groups (**Figure 7A**). This narrow range suggests that the two oxyanion hole hydrogen bonds are simultaneously maintained, consistent with prior NMR and site-directed mutagenesis evidence for stable hydrogen bonds and hydrogen bond coupling (Kraut et al., 2006, 2010; Pinney et al., 2018; Schwans et al., 2016; Sigala et al., 2008, 2009). The ability to maintain these hydrogen bonds is also supported by our RT X-ray data, as the explicit sub-states within our TSA-bound multi-conformer model each allow short hydrogen

bond distances of lengths similar to lengths in cryo structures and from solution ^{1}H NMR (**Figure 7B, Figure supplement 22**).

To evaluate how the narrow distribution of hydrogen bond lengths is maintained, we aligned all KSI TSA-bound structures on the oxyanion and we color coded the hydrogen bond donors, D103 and Y16 (**Figure 7C**). In the simplest case, the oxyanion hydrogen bond donors would move in concert with the oxyanion and these elements would closely overlay when aligned on the oxyanion. Instead, **Figure 7C** shows a range of donor positions. It would also be possible for the hydrogen bond donors to move in a concert with each other, relative to the oxyanion. If this were the case then we would see corresponding color-clustering of Y16 and D103 in **Figure 7C** which

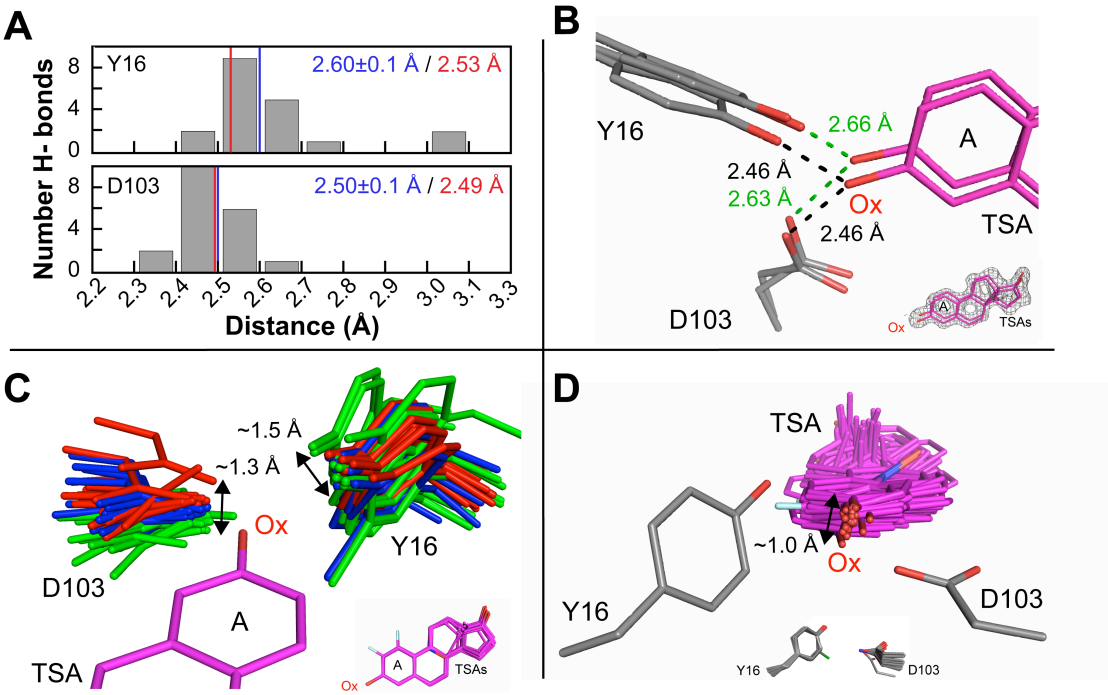

Figure 7. The KSI oxyanion—oxyanion hole ensemble. (**A**) Distribution of Y16 (top) and D103 (bottom) hydrogen bond distances from an ensemble of KSI crystal structures of variants with WT-like activity and with the same bound TSA (equilenin, n=19, **Table S2**). The mean hydrogen bond lengths and their standard deviations from the cryo crystal structure distances are shown in blue. The hydrogen bond distances obtained by solution ^{1}H NMR for D40N KSI bound to equilenin are show in red (see **Figure supplement 22** for ^{1}H NMR spectra from which distances are obtained). (**B**) The TSA-bound RT multi-conformer model shows that oxyanion hole Y16 and D103 and the bound TSA can make hydrogen bonds from different orientations. While multi-conformer models do not allow to unambiguously identify the exact hydrogen bonded sub-states, possible hydrogen bond lengths between Y16/D103 and the TSA are within the range obtained by cryo-structures and solution ^{1}H NMR. The inset shows the TSA (purple sticks) and the experimental electron density (grey mesh, contoured at 1σ). (**C**) The KSI oxyanion hole as "seen" by 36 bound TSAs (equilenin and various phenols) in cryo crystal structures (**Table S2**). Structures have been color-coded in three groups (green, blue and red) according to the D103 position in space relative to the TSA oxyanion (Ox). All TSAs have been aligned on the A ring but only one (PDB 1OH0) is shown for clarity. The inset shows the 36 aligned TSAs. (**D**) The bound TSA as "seen" by the hydrogen bonding oxygens of Y16 and D103. Y16 and D103 have been aligned such that their hydrogen bonding oxygens overlay but only one Y16/D103 set (PDB 1OH0) is shown for clarity. The inset shows all aligned Y16 and D103 hydrogen bonding groups.

is aligned relative to the oxyanion, but this correspondence is also not observed. The Y16 and D103 positions are not coordinated in a simple fashion and have an extensive and less regular ensemble of conformational states, with multiple side chain and ligand poses (**Figure 7C**, **D**) that nevertheless allow the hydrogen bonds to be maintained at near-optimal geometries.

The multiple observed side chain and ligand poses (**Figure 7B-D**) are the outcome of several competing potentials and forces, including Y16 and D103 side chain rotations, C-O(H) bond rotations, positioning of oxyanion within the oxyanion hole, and the energetics of the hydrogen bonds themselves. Apparently, the hydrogen bond potential energy wells are sufficiently deep that the hydrogen bonds are maintained and sufficiently steep, with the surroundings sufficiently mobile, to maintain near-optimal hydrogen bond geometries across a range of local conformational states.

Is there geometric discrimination in the KSI oxyanion hole? A widely adopted perspective on enzyme catalysis holds that the catalytic power of enzymes can be understood in terms of transition state complementarity, as opposed to the complementarity for a ground state (Eyring, 1935; Haldane, 1930; Pauling, 1946; Polanyi, 1921). This complementarity must encompass both electrostatics and geometry—the wrong charge in the right place will not stabilize a transition state, nor will the right charge placed too far away. Contributions from electrostatic and geometric factors have been proposed, but these contributions are very difficult to parse experimentally (Kraut et al., 2003, 2006). Preferential transition state stabilization on *geometrical grounds* has been proposed for several enzymes, including for KSI, and for oxyanion-hole containing proteases proposed to preferentially stabilize the sp^3 transition state oxyanion over the sp^2 ground state carbonyl oxygen (e.g. Blake et al., 1967; Fersht, 1985; Ford et al., 1974; Koshland, 1953; Kraut, 1977; Kraut et al., 2006; Robertus et al., 1972; Secemski et al., 1972; Simón and Goodman, 2010; Walsh, 1979). However, the mobility within the KSI's oxyanion hole renders it unlikely that the oxyanion hole provides preferential sp^3 stabilization (**Figures 6, 7B-D**), a conclusion that is underscored by the range of oxyanion hole angular orientations observed in our pseudo-ensemble (**Figure supplement 23**). Additional support for an absence of a geometric energetic effect is provided by the results described in the next section, which suggest equivalent hydrogen bond energetic behavior for a water oxygen and a carbonyl oxygen atom that have similar electron densities but different hybridization (sp^3 vs. sp^2, respectively) (**Table S59**).

Why is substrate binding not reduced by mutations in the oxyanion hole? A simple *energetic* expectation is that mutations to Y16 and D103 would weaken substrate binding—i.e., the oxyanion hole hydrogen bonds that provide substantial oxyanion stabilization and catalysis would also contribute to substrate binding, albeit less so due to the lower charge density of the carbonyl oxygen relative to an oxyanion (**Figure 1A**). Remarkably though, oxyanion hole mutations, even to bulky hydrophobic side chains (e.g., Y16F), have at most a two-fold effect on substrate binding (**Table S59**) (Kraut et al., 2010; Schwans et al., 2016). *Structurally*, one would

expect the oxyanion hole conformational ensemble to narrow upon formation of hydrogen bonds to the substrate and other ground state analogs (GSAs), but no major narrowing is observed; the Apo KSI oxyanion hole conformational ensemble is similar to that for bound states, as determined from our analysis of pseudo-ensembles and RT multi-conformer models (**Figure 4, Figures supplement 8, 9, 11, 14**).

The X-ray data provide a model to account for these paradoxes. Cryo X-ray structural models of Apo KSI reveal a water molecule in the oxyanion hole that is similarly positioned across the Apo KSI pseudo-ensemble (**Figure supplement 24A-C**). Because altered or exaggerated solvent localization can occur upon cryo-freezing (Halle, 2004; Sandalova et al., 1999; Scheidig et al., 1999; Teeter et al., 2001), it was important to test the presence of bound water at higher temperature. **Figure supplement 24D-E** shows parallel evidence for clearly positioned water in our RT X-ray data. Thus, a straightforward explanation for the absence of conformational narrowing upon ligand binding is that the oxyanion hole is already in a hydrogen bonded state and bound to water in the Apo enzyme. Further, if mutations to the oxyanion hole similarly weaken interactions with a bound water molecule (in the Apo state) and with a carbonyl oxygen (in GSA-bound states), then, in the simplest scenario, there would also be no change in affinity of the GSA, accounting for the counterintuitive GSA binding results (**Table S59**). In addition, the state observed at room temperature with the ground state partially in its binding pocket would likely be less sensitive to mutations in the oxyanion hole at the base of the pocket (**Figure Supplement 30A, C**). Thus, if significantly populated, this state could account at least in part for the insensitivity of ground state affinities to oxyanion hole mutations.

What restricts and permits motions of the oxyanion hole residues? Traditional X-ray crystallography models provide many insights, including identifying hydrogen bonds and hydrophobic contacts that may constrain motions. It has been noted that hydrogen bonds are more directional and thus more restricting than hydrophobic interactions in isolation (Chothia and Janin, 1975; Desiraju and Steiner, 1999; Fersht et al., 1985; Jeffrey and Saenger, 1991; Steiner and Desiraju, 1998; Watson et al., 2003). But in the crowded idiosyncratic environment of a protein interior, favored conformational states and the breadth of their distributions will be determined by multiple energetic contributors, including hydrogen bonds, van der Waals interactions, steric repulsion, and bond angle preferences, which need to be integrated over the range of states. For example, if we were to just consider static states, we would conclude that a van der Waals contact between Y16 and M116 may, or may not, help position Y16, depending on the particular crystal structure we looked at (**Figure supplement 25**).

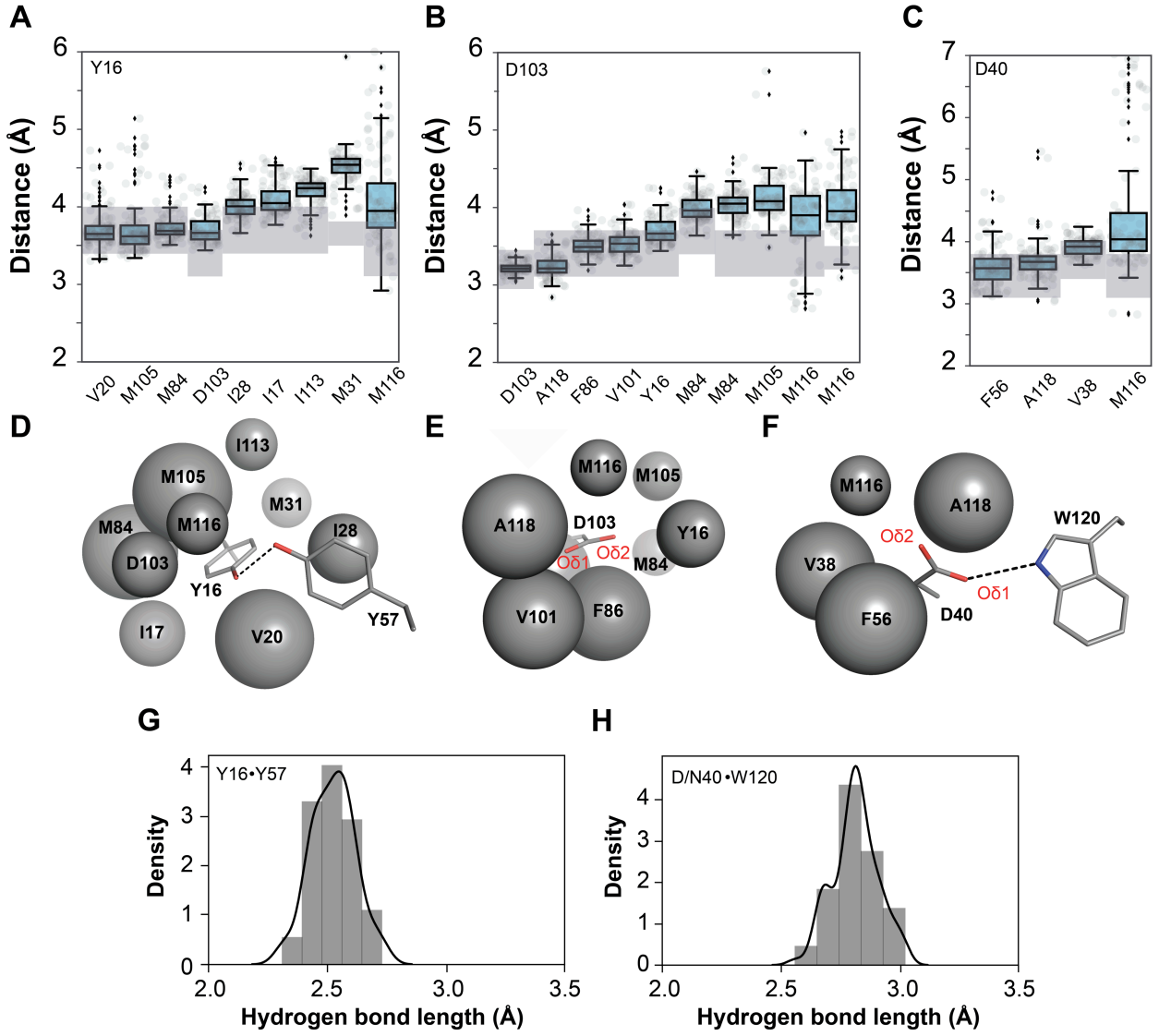

Figure 8. Ensemble analysis to evaluate the packing around the KSI catalytic groups. Ensemble distances between Y16 (**A**), D103 (**B**), and D40 (**C**) and their surrounding packing groups from all 94 independent KSI molecules from the 45 cryo crystals structures available in the PDB (**Table S1**). The boxes show the quartiles of the dataset while the whiskers extend to include the rest of the distribution. The closest atoms making van der Waals interactions have been identified and distances between specific atom pairs have been measured (see **Tables S42, 44, 46** for the list of atom pairs), hydrogen bonding lengths are presented separately (G,E) as hydrogen bonds are not simple van der Waals interactions. Two sets of distances have been reported for M84 and M116, as two distinct atoms were within similar distances. Van der Waals radii (r_{vdw}, shaded rectangles) are represented as a range because of uncertainty introduced by the absence of hydrogen coordinates in the X-ray structural models and because the oxygen r_{vdw} is orientation-dependent (see **Table S48** for a list of r_{vdw} used here). **D**, **E** and **F** schematically depict the results from panels A, B and C, respectively, with the packing atoms represented as spheres; larger spheres represent more tightly packed surroundings and Y16, D103 and D40 and the hydrogen bonding groups Y57 and W120 are represented in sticks for clarity (PDB 1OH0). (**G**, **H**) Histogram of all Y16-Y57 and D40-W120 hydrogen bonding distances from same structures as in A-C (position 40 contains both aspartate and asparagine residues).

Our ensemble data allow us to more deeply explore motions, providing initial insights into the forces that may be responsible for positioning, and, in certain instances, providing testable models for the local interactions and conformational arrangements that restrict and permit local motions. To describe packing and van der Waals interactions in an accessible form, we defined a packing distance between surrounding residues to the catalytic groups, Y16, D103, and D40, yielding, for each individual pseudo-ensemble structure, a single distance of closest approach between atoms of the catalytic side chains and surrounding groups (**Tables S42-48**). These distances, plotted together, provide ensemble-level local packing information (**Figure 8A-C**). While our ensemble contains more detailed information, the simplicity of this representation facilitates interpretation, and more detailed analyses can be carried out in the future, especially to evaluate predictions from molecular dynamics simulations.

For Y16, several groups, including M84, are situated within van der Waals contact distances, whereas M31 is consistently beyond van der Waals contact distance and M116 is highly variable in its positioning with respect to Y16 (**Figure 8A, D, Figure supplement 26**). In addition to these packing interactions, Y16 accepts a hydrogen bond from Y57, and the Y16-Y57 distance is highly conserved (**Figure 8G**). The observed longer Y16/M31 distances indicate that the collection of states with tight simultaneous packing of all residues is higher in free energy than the observed states, and we considered two simple models for the absence of close Y16/M31 packing:

> *Model I*: Packing is more favorable with residues other than M31 (e.g., M84), leading to the choice to not utilize the available van der Waals interaction energy from packing with M31.
>
> *Model II*: The Y16-Y57 hydrogen bond energy and positioning dominates to constrain Y16 away from M31 and closer to M84.

Model I predicts that breaking the Y16-Y57 hydrogen bond will leave Y16 or its replacement residue in place, whereas in *Model II* predicts a rearrangement, most simply to make closer interactions with M31. When we compared KSI sub-ensembles with the Y16-Y57 hydrogen bond to those with mutations that removed this hydrogen bond (**Figure 9A, B,** Y16-Y57 hydrogen bond intact (grey stick) and ablated (green sticks), respectively; **Table S2**), we saw that Y16 shifted its position, moving toward M31 and away from M84 without the Y16-Y57 hydrogen bond (**Figure 9C-D,** Y16-Y57 hydrogen bond intact (grey histograms) and ablated (green histograms), respectively). These results support *Model II* and an energetic and conformational trade-off between the Y16-Y57 hydrogen bond and what would otherwise be a more symmetrical packing of Y16 with both M31 and M84 (**Figure 9A-D**). Nevertheless, a range of Y16 distances is observed with respect to both M31 and M84 with the Y16-Y57 hydrogen bond removed, suggesting a rather flat local energy landscape in which interactions with M31 and M84 are similar in energy, whereas states rearranged to allow simultaneous

packing are disfavored, presumably because the cost of losing other interactions with these side chains is too high.

Removal of the Y16–Y57 hydrogen bond also appears to render the ring at position 16 more flexible (**Figure 9A, B and E**), consistent with the above interpretations. Interestingly, increased flexibility of the ring at position 16 does not result in any rearrangements or significant structural changes in the surrounding residues (**Figure supplement 27**). Further, neither ablation of the Y16 hydrogen bonding group (e.g., Y16F mutation, **Figure supplement 15**), nor increased flexibility or apparent mispositioning of Y16 upon ablation of the Y16–Y57 hydrogen bond (**Figure supplement 28**) appear to impact D103 positioning, suggesting that while Y16 and D103 hydrogen bond lengths are coupled (Pinney et al., 2018; Sigala et al., 2009), the allowed Y16 and D103 side chain orientations are not (also see **Figure supplement 15** and "*Are hydrogen bonds to the transition state analog maintained in the enzyme ensemble and, if so, how?*" above).

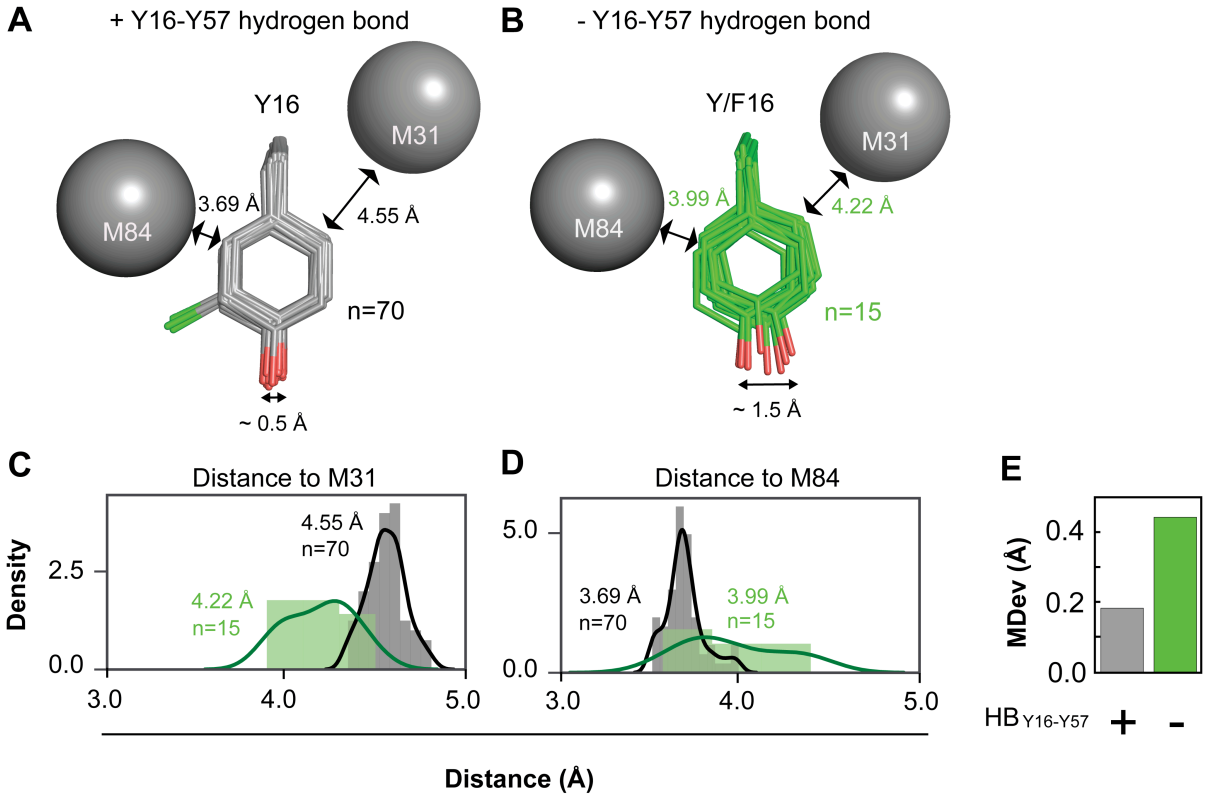

Figure 9. Testing models for the asymmetric packing of Y16. Y16 residues with intact Y16–Y57 hydrogen bond (**A**, in grey) or Y/F16 residues with the Y16–Y57 hydrogen bond ablated (**B**, in green) (**Table S2**). Residues M31 and M84 are shown as grey spheres with the mean distances between M84 and Y16 or M31 and Y16 given (n = 70 and 15 for intact and ablated Y16-Y57 hydrogen bonds, respectively). Chlorine substitutions in some of the tyrosines in A are colored in light green. (**C, D**) Histogram of the distribution of Y16–M31 and Y16–M84 distances for the ensembles from panels (**A**) and (**B**), respectively. (**E**) MDev for the Cζ atom of the phenyl ring for the ensemble from panel (A) in grey and from panel (B) in green. (See **Figure supplement 18** for definition of the Cζ atom.)

The observed conformational constraint from the Y16–Y57 hydrogen bond underscores the question of how D103 can be well-positioned in the absence of analogous hydrogen bonding. Analysis of the D103 surroundings revealed particularly close packing of the non-catalytic Oδ1 of D103 with multiple residues: F86, V101, and A118 (**Figure 8B, E**). Indeed, the RT-ensemble suggested that this oxygen atom may be more restricted than the protonated catalytic oxygen (Oδ2) that sits in the oxyanion hole (**Figure 5A**). This atypical arrangement in which hydrophobic interactions surround a carboxylic acid oxygen atom (McDonald and Thornton, 1994) is particularly intriguing as it appears to accomplish two objectives: *i.* positioning the catalytic group, and *ii.* increasing its pK_a so that the carboxylate group remains protonated and can act as a hydrogen bond donor at physiological pH. The protonated D103 side chain provides greater oxyanion stabilization than side chains with higher pK_as and lower hydrogen bonding proton charge density (**Figure 1**; (see Herschlag and Pinney, 2018; Shan and Herschlag, 1999) see "*Discussion*"). Thus, counterintuitively, interactions of a polar oxyanion atom with hydrophobic groups appear to provide important favorable interaction energy, from van der Waals interactions, presumably sterically constraining the carboxylic acid group for function and disfavoring conformational states where the carboxylic acid group can rearrange and be solvated to favor its anionic form. This model is testable, predicting that mutations that reduce packing will decrease D103's pK_a and positioning. Studies of this type will be needed to dissect the precise relationship of positioning to function.

General base positioning and catalysis

How does KSI's general base, D40, have an unusually high effective molarity without an unusually high degree of conformational restriction? When a unimolecular reaction is compared to a bimolecular reaction, the ratio of rate constants has units of molar and, in the simplest scenario, this value, referred to as the Effective Molarity (EM), represents how well aligned the unimolecular groups are for reaction (**Figure 10A**) (Kirby, 1980; Page and Jencks, 1971). The EM of ~10^3–10^5 M determined for the KSI general base, for KSI from two species, is considerably higher than typical for a positioned general base and suggested the possibility of an unusually high conformational restriction of D40 (Lamba et al., 2016).

Our ensembles allowed us to test this prediction and indicate that, while D40 is more positioned than the average aspartate or asparagine residue, its positioning is not highly constrained; indeed, it is less constrained than D103 (**Figure 5A** and **Figure 6C, D**). Further, if the general base were precisely positioned with respect to the acidic carbons of the bound ligand, then we would expect that alignment of KSI structures with bound TSAs would result in a close overlay of the general base, or possibly preferred positioning near each acidic carbon position. Instead, there is a range of general base positions and a range of distances between the general base and the carbons between which protons are shuffled in the KSI reaction with different substrates (**Figure 10B, C**). In the simplest scenario, the time spent close to each position (**Figure 10C**) would correlate

with the relative probability of the microscopic step involving that proton abstraction, a prediction that remains to be tested. Overall, these observations do not account for the observed highly effective proton abstraction by D40 so that alternative models need to be considered.

A

$$\text{A-B} \xrightarrow{k_1\,(\text{s}^{-1})} \text{C}$$

$$\text{A + B} \xrightarrow{k_2\,(\text{M}^{-1}\text{s}^{-1})} \text{C}$$

$$\text{Effective Molarity (EM)} = \frac{k_1}{k_2}\,(\text{M})$$

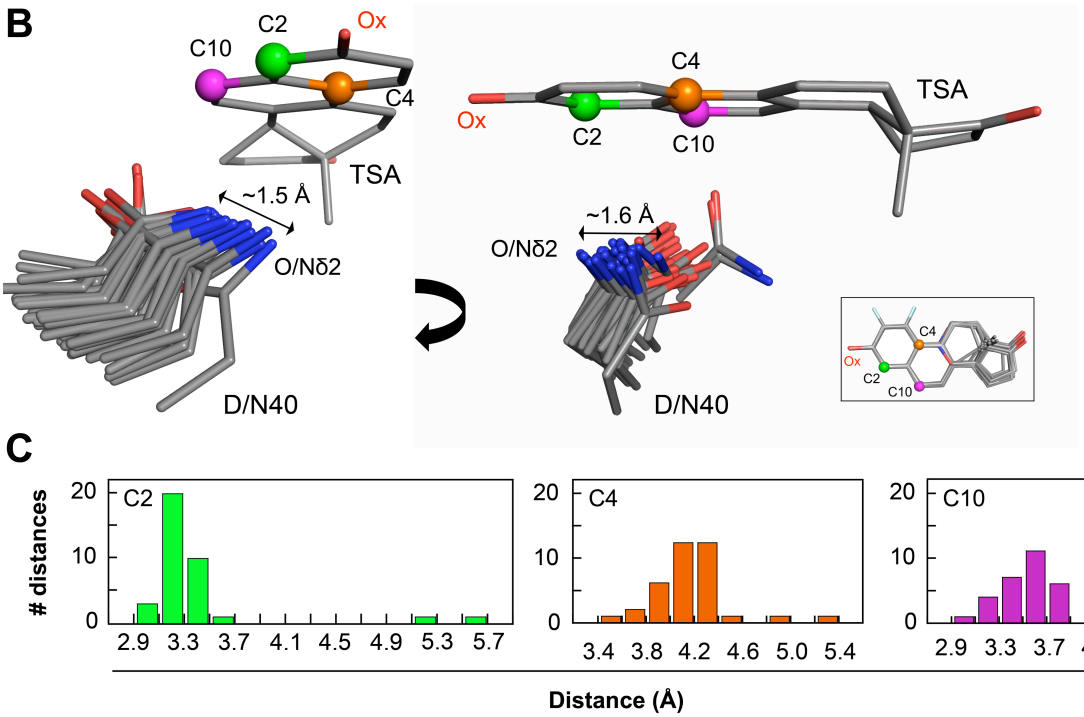

Figure 10. The high effective molarity of the general base in KSI is not due to a highly precise positioning. (A) Effective molarity (EM; in units of Molar, M) is the ratio of the rate constant for an intramolecular reaction (s^{-1}) and the rate constant for the same reaction when the reactive moieties are two separate molecules ($M^{-1}s^{-1}$) (Jencks, 1969; Kirby, 1980). (**B**) A bound TSA as "seen" by the KSI general base in KSI TSA-bound ensemble of cryo crystal structures (asparagine mimics the protonated (intermediate) state (Kraut et al., 2006; Petrounia and Pollack, 1998)) (**Table S2**). All TSAs (equilenin and various phenols) have been aligned on the A ring but only one (PDB 1OH0) is shown for clarity. The carbon positions between which protons are shuffled in KSI reactions are represented as green, magenta, and orange spheres (see **Figure 1** and **Figure supplement 29** for the reaction mechanisms). The inset shows all aligned TSAs. (**C**) Histograms of distances between the Oδ2 (Nδ2 when asparagine at position 40) and the carbon positions from (B) on the bound TSAs in each crystal structure. Bar colors correspond to carbon colors from (B).

While the most straightforward interpretation of a high EM is precise positioning, all factors that augment the intramolecular reaction contribute to this value (Jencks, 1969; Kirby, 1980; Lamba et al., 2016; Page and Jencks, 1971). In small molecule studies, destabilization of the

intramolecular ground state, as in the classical example of dimethylmaleic acid cyclization, is a frequent contributor (Kirby, 1980; Kirby and Lancaster, 1972). Similarly, co-localizing reacting groups of like charge (where one or both experiences a decrease in charge density in the transition state) can also increase reactivity and EMs (Kirby, 1980). Indeed, there is evidence for ground state destabilization for KSI, as replacement of D40 by neutral groups and mutations that increase its flexibility increase substrate affinity (Ruben et al., 2013). This destabilization presumably arises from anionic D40 abutting the steroid hydrocarbon in a largely hydrophobic environment. Nevertheless, this destabilization is modest, ~10-fold, suggesting that additional factors, yet to be identified, contribute to catalysis by D40. Below, we analyze groups responsible for positioning D40 and integrate this information with functional data.

How do the D40 surroundings position it, while still allowing a range of conformations? D40, like D103 and Y16, has multiple packing interactions (**Figure 8C, F**) and, like Y16, accepts a hydrogen bond (**Figure 8F**). Intriguingly, the hydrogen bond is accepted by the non-proton abstracting oxygen (Oδ1), and this oxyanion is conformationally more restricted than the oxygen responsible for proton abstraction (Oδ2; **Figure 11D, E**). The hydrogen bond from W120 to Oδ1 appears to act as a pivot point that allows broader conformational exploration by the more distal Oδ2 (**Figure 11F**). The considerable D40 Oδ2 flexibility arises from the angular freedom of the relatively long and weak D40 Oδ1•W120 hydrogen bond (**Figure 8H**) and from the anion-aromatic interaction with F56, which has a broader potential energy well (**Figure 11B, C**) (Jackson et al., 2007; Kapoor et al., 2016; Philip et al., 2011; Thomas et al., 1982). KSI may use the weaker anion- aromatic interactions to the general base oxygen atom, rather than a hydrogen bond, to widen the general base ensemble. The wider ensemble might more efficiently shuffle protons between different positions, while still keeping the general base close to the proton donating and receiving carbons on substrates and preventing excessive flexibility. In addition, the weaker anion-aromatic interaction with the general base oxygen atom, rather than a hydrogen bond, may enhance general base proton affinity—and thus lead to more effective proton abstraction—while still providing sufficient positioning.

While the motional freedom of the proton-abstracting oxygen atom may facilitate KSI's ability to carry out proton transfer steps at multiple substrate positions with multiple substrates (**Figure 1A, Figure supplement 29**), it does not account for its high effective molarity and high catalytic effectiveness in carrying out *each* proton transfer (Lamba et al., 2016). As noted above, there is no indication of a subset of highly favored positions corresponding to each position of proton abstraction/donation (**Figure 10B, C**), and the GSA flexibility within the KSI binding pocket renders this possibility even more unlikely (**Figure supplement 30**). In addition, $KSI_{homolog}$ replaces W120 with phenylalanine to give a second anion-aromatic interaction and no hydrogen bonding interactions to the general base (**Figure supplement 31B**). As expected, D40 exhibits greater flexibility, as determined from our $KSI_{homolog}$ pseudo-ensemble (**Figure supplement 31C-E**), but remarkably, is a more efficient catalyst, with an EM for D40 of 5×10^4 M instead of 0.1×10^4 M for KSI (Lamba et al., 2016; Schwans et al., 2016). Further, notably, mutation of

W120 to phenylalanine increases KSI catalysis with respect to WT KSI and the converse mutation in $KSI_{homolog}$ (F120W) decreases catalysis with respect to WT $KSI_{homolog}$ (**Figure supplement 31F, 32, Table S58**).

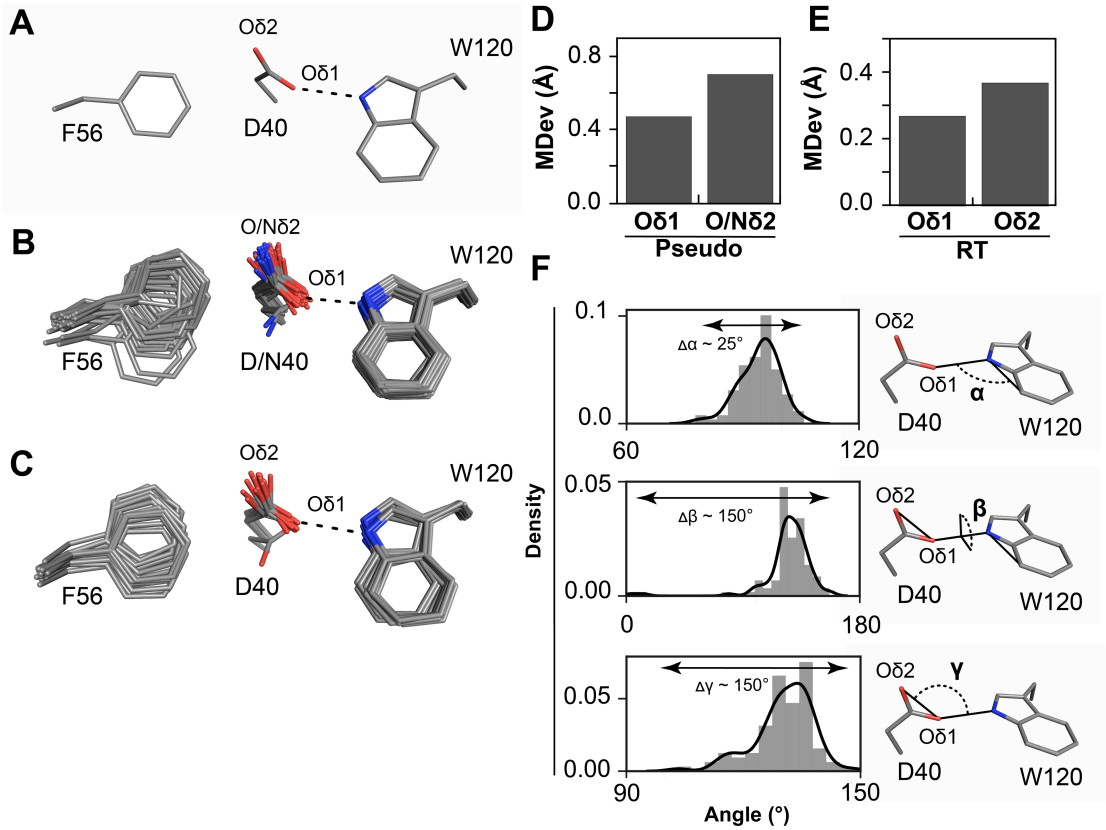

Figure 11. General base motions are modulated by a combination of anion-aromatic and hydrogen-bonding interactions. (**A**) Anion-aromatic and hydrogen bonding interactions with F56 and W120 sidechains, respectively, implicated in general base positioning (PDB 1OH0). (**B**) The full pseudo-ensemble and (**C**) RT-ensemble of the general base, F56 and W120 sidechains. (**D-E**) MDevs of the non-catalytic (Oδ1) and catalytic Oδ2 (Nδ2 when asparagine at position 40) oxygen atoms of the general base from the full-pseudo (D) and the RT-ensemble (E). (**F**) Angles characterizing the general base and W120 relative sidechain orientations obtained from the full pseudo-ensemble.

D40 catalysis is compromised by mutations to other residues that surround it and appear to aid its positioning, for both KSI and $KSI_{homolog}$ (Schwans et al., 2013, 2014). Thus, while general base flexibility appears required for function, more freedom of motion does not uniformly enhance catalysis; rather, there is presumably an optimal balance between allowing and limiting conformational motion of D40.

Discussion

General Implications

Deepening our understanding of the energetics and function of biological macromolecules demands ensemble descriptions. We built ensembles from the large number of high-resolution ketosteroid isomerase (KSI) structures in the Protein Data Bank and from high-resolution room-temperature (RT) X-ray data across a series of states mimicking the reaction coordinate. Analysis of these ensembles provided insights and tests of models not accessible from traditional X-ray structures. Further, our pseudo-ensembles from multiple cryo X-ray structures and our RT multi-conformer models were complimentary: pseudo-ensembles allowed us to observe correlated positional effects and draw new conclusions from existing data, and RT X-ray data allowed us to ensure that our conclusions were overly not affected by temperature artifacts from the glass transition and to readily obtain conformational information about a new state, in this case KSI with a bound ground state analog.

We tested two broad catalytic models, whether enzyme groups, including catalytic residues, are largely prepositioned or narrow their conformational ensemble as the reaction proceeds (the Dynamic Gradual Adaption Model, **Figure 2A**), and whether catalytic residues exhibit extraordinary positioning relative to other residues (the Dynamic Entatic State Model, **Figure 2B**). Our ensemble data provided evidence against each model for KSI. Similarly, prior NMR data suggested that most of the substrate positioning needed to form a new bond occurs in forming the enzyme•substrate complex of the UMP/CMP kinase reaction (Zeymer et al., 2016). Future ensemble determinations with additional enzymes will test the generality of these conclusions.

Our ensemble data defined the extent and nature of positioning in KSI's active site. Catalytic residues were more restricted than most other residues, but not extraordinarily so, exploring conformational space asymmetrically from about ~1 Å down to ~0.2 Å, ranging from slightly less than a typical carbon-carbon bond length to slightly more than a bond vibration.

Our ensembles allowed us to more deeply explore the interactions and forces that allow and restrict motions. Analysis of our ensembles led to initial insights and testable models for the interactions responsible for the positioning of the oxyanion hole and general base catalytic residues, insights not possible from traditional static models. Further, as ensembles more directly and extensively relate to the output of molecular dynamics and other simulations, than comparisons to one or a few rate or equilibrium measurements, experimental ensembles will be of particular value in testing and refining molecular dynamics and other modeling methods. In addition, comparisons with simulations will be of value to obtain more extensive and quantitative ensembles. Our ensembles do not extend all the way to reactive distances for proton transfer, consistent with Klinman's proposal that enzymatic reactions may routinely occur from rare

conformational poses (Klinman, 2009, 2013). Our observation of continuous local conformational distributions most simply suggests that these reactive states reside on the edges that extrapolate from our ensembles, but the complete conformational landscape cannot be derived from either X-ray approach used herein.

Combining ensemble information and information across the reaction coordinate with kinetic and thermodynamic data has allowed us to more deeply and comprehensively probe structure-function interrelationships. These approaches, extended to additional KSI variants and complexes and to other enzymes and their complexes, may usher in a new structure–function era. Insights derived from ensembles that improve our ability to connect structure to energetics and function may be needed to achieve the long-term goals of a deep understanding of enzymatic catalysis and an ability to carry out at-will design of new enzymes.

Summary of KSI Catalytic Implications

Our ensemble results, combined with decades of functional studies, provide a deep and self-consistent picture of KSI oxyanion hole energetics, account for the ability of KSI to carry out multiple proton transfer reactions, and raise new questions concerning the detailed energetics of general base catalysis. This in-depth analysis of a highly-studied model enzyme provides a prototype and powerful point of comparison for future studies that bring an ensemble perspective to bear on traditional structure–function studies.

Aligning the KSI pseudo-ensemble by the oxyanion of bound transition state analogs (TSAs) reveals many poses of the oxyanion hole and general base residues, as noted above (**Figure 7C and 10B**). The immediate implications of this range of positions are that oxyanion hole discrimination between the ground state sp^2 and transition state sp^3 geometries is highly unlikely and that the positional variability should allow proton abstraction and donation at multiple positions and with multiple substrates (**Figure supplement 21, 23, 29**). Our RT X-ray data, including data with a bound ground state analog, provide further support for these conclusions, as does a less extensive pseudo-ensemble for a related KSI ($KSI_{homolog}$, **Figures supplement 30, 31**).

The catalysis provided by KSI's oxyanion hole, relative to water, appears to arise from the presence of hydrogen bonds that are intrinsically stronger than water, rather than exceptional positioning or environmental effects. The increased strength arises because the carboxylic and phenolic protons of D103 and Y16, respectively, have greater positive charge density on their donating hydrogen atoms and deeper hydrogen bond potential wells, relative to the corresponding hydrogen bonds from water (**Figure 12**) (Dwyer, 2005; Herschlag and Pinney, 2018; Pinney et al., 2018; Shan and Herschlag, 1999; Shan et al., 1996; Taft, 2007). When D103 in KSI is substituted by asparagine (D103N) or replaced by water (within a cavity in a D103G

mutant), both mutants have their catalytic rates decreased by a similar 10 – 30 fold (Lamba et al., 2016; Schwans et al., 2016) (**Table S59**). In other words, the positioned amide is no more effective than water present in cavities created, and the water and amide protons have similar charge densities but different geometric properties and likely different positioning (**Figure 12**). Thus, their similar catalytic efficiencies suggest that oxyanion hole catalysis does not arise from preferential and exceptional positioning and that D103 (and Y16) catalysis arises from stronger hydrogen bond donation than water.

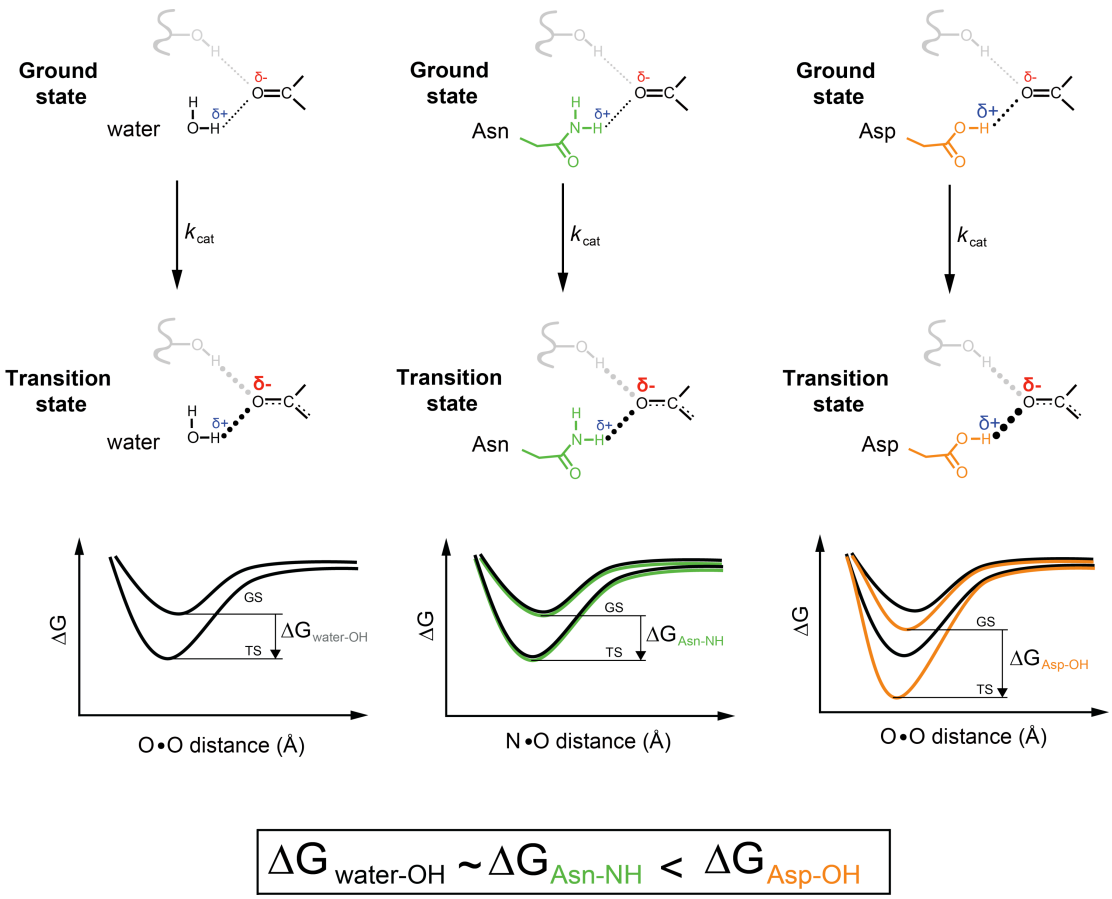

Figure 12. KSI oxyanion hole catalytic model. To provide catalysis, relative to water, the oxyanion hole of KSI uses hydrogen bonds that are stronger than water, rather than exceptionally precisely positioned hydrogen bonds. During the KSI reaction, the amount of negative charge on the substrate carbonyl increases and this negative charge accumulation is stabilized by hydrogen bonds. Analogously, hydrogen bonds become stronger as the charge density on the hydrogen bond donating hydrogen increases (Dwyer, 2005; Herschlag and Pinney, 2018; Shan and Herschlag, 1996, 1999; Shan et al., 1996; Taft, 2007). Thus, water (left, in black) and the amide group of an asparagine side-chain in the oxyanion hole of KSI (D103N, middle, in green), which have similar hydrogen charge densities, provide similar extent of transition state stabilization (Shan et al., 1996); hydrogen bond strength is depicted by the size of the dots representing the hydrogen bonds. The WT KSI oxyanion hole with D103 (protonated) stabilizes the transition state to a greater extent because its hydrogen bonding hydrogen has more charge density than water and asparagine side-chain (right, in orange). In all three scenarios, hydrogen bonds shorten and strengthen in the transition state (indicated with thicker doted lines in the TS compared to GS), but the shortening and strengthening in the TS is larger when the hydrogen bond donor is an aspartic acid residue.

There have been numerous discussions of possible alterations in the energetic properties of catalytic interactions from creation of an active site environment that differs from and is sequestered from aqueous solution (Berg et al., 2002; Cleland and Kreevoy, 1994; Dewar and Storch, 1985; Gerlt and Gassman, 1993; Graham et al., 2014; Richard et al., 2014; Shan and Herschlag, 1999; Warshel, 1998; Warshel et al., 2006). However, prior studies have shown that the dependence of catalysis and TSA binding on the charge densities of the oxyanion and the oxyanion hole hydrogen bond donors is similarly shallow to hydrogen bond dependencies observed in aqueous solution; the stronger dependencies observed in non-aqueous environments such as organic solvents and the gas phase are not observed within the KSI oxyanion hole (Kraut et al., 2006; Natarajan et al., 2014; Peter Guthrie, 1996; Shan et al., 1996; Sigala et al., 2015).

There have also been discussions of the involvement of longer-range electrostatic interactions in transition state stabilization. From the standpoint of a static structure one can readily imagine substantial energetic effects from interactions with charges and partial charges that are strategically positioned throughout an enzyme, although they would need to be asymmetrically placed and oriented to favor the transition state. In essence, these longer-range electrostatic interactions would be minimally diminished by dielectric-type mitigating effects because the fixed protein interior would not reorient in response to changes in or build-up of charge and thus would minimally lessen energetic interaction energies. An ensemble description, however, leads to a different perspective.

Given that there are multiple conformational sub-states, if there were substantial, directed long-range electrostatic interactions, the sub-states would be expected to respond to the charge changes that occur through the reaction cycle; this conformational response would diminish the observed electrostatic energies, but only partially if highly restricted. Despite observing a range of positions for residues throughout KSI (i.e., conformational sub-states), we see no substantial change in their average position or their range of motion in going from Apo to the TSA-bound state (**Figure 4, Figures supplement 8-10, 14**). Most simply, this absence of substantial rearrangements suggests that the active site charge changes are not strongly felt throughout the protein, which correspondingly suggests minimal long-range energetic contributions. It is possible that rearrangements occur but are too subtle to observe in our ensembles. However, effects from surrounding water might be expected to mitigate any potential enhanced electrostatic interaction energies; even though active sites can be solvent inaccessible, they are not far removed from solvent and electrostatic screening by water is highly effective even through intervening space. The additional observations of essentially no difference in catalysis, hydrogen bond properties, or hydrogen bond energetics for KSI substrates and ligands with one *versus* the standard four steroid rings, despite enhanced active site access to water upon removal of three steroid rings, also suggests that any potential long-range electrostatic interactions are already effectively ameliorated by the response of solvent and protein groups (Kraut et al., 2006; Schwans et al., 2009; Shan and Herschlag, 1999; Sigala et al., 2007).

In contrast to longer-range, through-space effects, there is evidence for local, conformational and energetic effects, and our ensembles and sub-ensembles allow us to visualize, better understand, and develop and test models for these effects. First, mutations lead to rearrangements in the most represented conformational sub-states, with changes largely proximal to the site of mutation (**Figure 9, Figures supplement 15, 27, and 28**). In addition, the hydrogen bond network from Y32 and Y57 to Y16 alters the energetics and physical properties of the Y16•oxyanion hydrogen bond (Kim et al., 2000; Pinney et al., 2018). These effects are modest, <10 fold on catalysis and <0.2 Å in hydrogen bond length, and diminish steeply for the more remote Y32 interaction, but they are large enough to aid catalysis and function and be subject to selective pressure (Pinney et al., 2018 and Pinney et al. submitted).

Returning to the KSI general base, D40, our ensemble data have revealed a paradox: a high effective molarity (EM), traditionally interpreted as highly precise positioning, for a group that is not particularly positionally restricted (**Figures 5, 6C-D, 10B, Figures supplement 19, 21**). D40's flexibility, enhanced by anion-aromatic interactions (**Figure supplement 31**), and the multiple active site ligand poses, including poses of a bound ground state analog, (**Figure 7B, D, Figure supplement 30**) presumably allow KSI to carry out proton transfers at multiple steroid positions for multiple substrates, but nevertheless raise the question of how KSI manages this flexibility while achieving high catalytic efficacy. We speculate that the local environment enhances the reactivity of D40, in part by the previously-identified ground state destabilization (Ruben et al., 2013) and in part by mechanisms that remain to be elucidated but that may arise from desolvation. In particular, the reaction of a desolvated general base that directly abstracts a proton from an acidic carbon may be favored over a reaction proceeding via and requiring alignment of an intervening water molecule (Jencks, 1969; Kirby, 1980; Page and Jencks, 1971).

Most broadly, combining ensemble information and information across the reaction coordinate with kinetic and thermodynamic data allows us to more deeply and comprehensively probe structure-function relationships. Overall, our observation of limited restriction of D40 positioning introduces a new structure-function perspective and the need for new models to account for high catalytic effectiveness of KSI's general base. While ensembles are a critical component of understanding macromolecular function, our ability to resolve current questions and distinguish models and resolving the general base paradox will also require complementary functional and energetic studies.

Materials and Methods

KSI expression and purification. The ketosteroid isomerase enzymes from *Pseudomonas putida* (pKSI, referred to herein as KSI, UniProt P07445) and *Comamonas testosteroni* (tKSI, referred to herein as $KSI_{homolog}$, UniProt P00947) were expressed and purified as previously described with minor modifications (Kim et al., 1994; Lamba et al., 2016). KSI W120F and $KSI_{homolog}$ F120W (KSI numbering) variants were obtained using standard mutagenesis protocols and the presence of the desired mutations was confirmed via DNA sequencing (see **Table S60**). Briefly, BL21 cells transformed with plasmid carrying the desired KSI construct were grown at 37 °C to OD 0.5−0.6 in LB media (EMD Millipore Corp, Billerica, MA, USA) containing 50 μg/mL carbenicillin (Goldbio, St Lousi, MO, USA), and protein expression was induced with 1 mM isopropyl-β-D-1-thiogalactopyranoside (Goldbio, St Lousi, MO, USA). After induction, cultures were grown for 10−12 h at 37 °C. Cells were harvested by centrifugation at 5000 g for 30 min at 4 °C and lysed using sonication. Lysed cells were centrifuged at 48000 g for 30 min at 4 °C. Enzymes were purified from the soluble fraction, first using an affinity column (deoxycholate resin) followed by a size exclusion chromatography column (SEC) Superdex 200. Prior to the purification of each enzyme, the affinity column, FPLC loops, and SEC column were washed with 40 mM potassium phosphate (JT Baker, Omaha, NE, USA), 6 M guanidine (JT Baker, Omaha, NE, USA), pH 7.2 buffer, and then equilibrated with 40 mM potassium phosphate, 1 mM sodium EDTA, 2 mM DTT (Goldbio, St Lousi, MO, USA), pH 7.2 buffer.

KSI solution kinetics. KSI Michaelis–Menten parameters were obtained by monitoring the 5(10)-estrene-3,17-dione ((5(10)-EST), Steraloids, Newport, RI, USA) reaction at 248 nm (extinction coefficient 14,800 M^{-1} cm^{-1}) in a PerkinElmer Lambda 25 spectrophotometer. Reactions were measured at 25 °C in 4 mM sodium phosphate, pH 7.2 buffer with 2% DMSO (JT Baker , Omaha, NE, USA) added for substrate solubility. Low buffer concentrations were used to minimize the background reaction rate. Values of k_{cat} and K_M were determined by fitting the initial rates as a function of substrate concentration to the Michaelis–Menten equation. Typically, seven to eight substrate concentrations, varying from 2 to 600 μM, were used for each mutant. The k_{cat} and K_M values were averaged from two independent measurements using different enzyme concentrations varied over 2–3 fold. Averaged values and errors representing the standard deviations are given in **Table S58**.

KSI ^{1}H solution Nuclear Magnetic Resonance. The ^{1}H NMR spectrum of KSI D40N bound to equilenin was acquired at the Stanford Magnetic Resonance Laboratory using an 800 MHz Varian ^{UNITY}INOVA spectrometer running VNMRJ 3.1A and equipped with a Varian 5 mm triple resonance, pulsed field gradient ^{1}H[^{13}C,^{15}N] cold probe, as previously described (Kraut et al., 2006). The sample contained 1.0 mM KSI and 2.0 mM equilenin (Steraloids, Newport, RI, USA) in 40 mM potassium phosphate (pH 7.2), 1 mM sodium·EDTA, 2 mM DTT, and 10% DMSO-d_6 (v/v) (Cambridge Isotope Laboratories, Tewksbury, MA, USA). DMSO-d_6 served as the deuterium lock solvent and prevented freezing at low temperatures. The spectrum was

obtained in a 5 mm Shigemi symmetrical microtube at –3.5 °C, following temperature calibration with a 100% methanol standard. The 1331 binomial pulse sequence was used to suppress the water signal with a spectral width of 35 ppm (carrier frequency set on the water resonance) and an excitation maximum between 14-18 ppm (Turner, 1983). The data was processed using 10 Hz line broadening and baseline correction applied over the peaks of interest. Chemical shifts were referenced internally to the water resonance.

Protein crystallization and X-ray data collection. All enzymes were crystallized as previously described (Pinney et al., 2018). Briefly, enzyme were crystallized by mixing 1 µL of enzyme at 1 mM and 1 µL of crystallization solution (17-23% PEG 3350 (Hampton Research, Aliso Viejo, CA, USA) and 0.2 M $MgCl_2$ (JT Baker, Omaha, NE, USA)) in a vapor diffusion hanging drop setup at room temperature. For crystallization of KSI bound to the transition state analog (equilenin) or the ground-state analog (4-androstenedione (Steraloids, Newport, RI, USA)), equilenin or 4-androstenedione were first dissolved in methanol (JT Baker, Omaha, NE, USA) at 20 mM and 40 mM concentration, respectively. Each ligand was then mixed with enzyme to achieve final concentrations of 1 mM enzyme and 2 mM equilenin or 4 mM 4-androstenedione (10% methanol in the final enzyme-ligand solution). As a ground-state analog, 4-androstenedione binds more weakly than the transition-state analog equilenin, and thus higher concentration was used to achieve higher occupancy. Crystals typically appeared after 24-72 h. Prior to data collection, crystals with minimum dimensions 0.2 x 0.2 x 0.2 mm were transferred from the crystallization solution to paratone N oil (Hampton Research, Aliso Viejo, CA, USA) where excess crystallization solution was stripped and crystals where then either frozen in liquid nitrogen for 100 K data collection and then mounted on the goniometer or directly mounted on the goniometer for 250 K or 280 K data collection. Data collection temperature was controlled using a N_2 cooler/heater. Single-crystal diffraction data were collected at SSRL, beamline BL9-2, using wavelengths of either 0.787 Å or 0.886 Å. See Supplementary file **Table S3** for diffraction data statistics.

Crystallographic data processing and model building. Data processing was carried out with in-house scripts: http://smb.slac.stanford.edu/facilities/software/xds/#autoxds_script. Briefly, data reduction was done using the XDS package (Kabsch, 2010), scaling and merging was done using *Aimless* (Collaborative Computational Project, Number 4, 1994; Evans and Murshudov, 2013) and structure factor amplitudes were obtained using *Truncate* (Collaborative Computational Project, Number 4, 1994; French and Wilson, 1978). Initial phases were obtained via molecular replacement using *PHASER* (McCoy et al., 2007) and the PDB entry 3VSY as a search model. Model building was carried out with the program *ARP/wARP* (Langer et al., 2008) and manually in *Coot* (Emsley and Cowtan, 2004). Traditional, single conformation models, in which major alternative side chain and backbone conformations were modeled, were refined manually after visual inspection with *Coot* and using *phenix.refine* (Afonine et al., 2012). Torsion-angle simulated annealing (as implemented in *phenix.refine*) was used during the initial

stages of refinement. Riding hydrogens were added in the late stages of refinement and their scattering contribution was accounted for in the refinement. Ligand restraints were generated using the *GRADE* server (http://grade.globalphasing.org/cgi-bin/grade/server.cgi). Model quality was assessed using *Molprobity* (Chen et al., 2010) as implemented in *phenix.refine* and via the PDB Validation server (https://validate-rcsb-2.wwpdb.org/). See Supplementary file **Table S3** for refinement statistics.

Multi-conformer models were obtained from the 250 K diffraction datasets, using previously described methods (van den Bedem et al., 2009; Fenwick et al., 2014; Keedy et al., 2015a, 2015b, 2018). As a large body of work identified the 180–220 K temperature range as an inflection point above which various protein motions are activated, providing strong evidence that at and above 250 K both harmonic and anharmonic protein motions are enabled (Doster et al., 1989; Frauenfelder et al., 1979; Keedy et al., 2015b; Lewandowski et al., 2015; Tilton et al., 1992), and as the 250 K diffraction data was of higher resolution than our 280 K data (**Table S3**), we used the higher-resolution 250 K to obtain multi-conformer models of Apo, GSA-bound, and TSA-bound KSI. Briefly, the program *qFit* was used to obtain multi-conformation models (van den Bedem et al., 2009; Keedy et al., 2015a) using as input the traditional single-conformation models obtained above after removing the riding hydrogen atoms. Subsequent to the automated multi-conformer model building, ill-defined water molecules were deleted and alternative protein side and main chain conformations and orientations were edited manually after visual inspection in *Coot* and based on the fit to the electron density (Emsley et al., 2010). Models were subsequently refined with *phenix.refine* (Afonine et al., 2012). Riding hydrogen atoms were added in the late stages of refinement and their scattering contribution was accounted for in the refinement. Final multi-conformer model quality was checked by *MolProbity* (Chen et al., 2010) and via the PDB Validation server (https://validate-rcsb-2.wwpdb.org/). See Supplementary file **Table S3** for refinement statistics.

Crystallographic order parameters calculation. Crystallographic order parameters, S^2, were obtained from the 250 K multi-conformer models as previously described (Fenwick et al., 2014). These order parameters include both harmonic and anharmonic contributions as captured by the crystallographic atomic displacement parameters (B-factors) and by the occupancies of alternative rotameric states, respectively; these values correlate well with solution NMR-derived S^2 (Fenwick et al., 2014). The analysis was applied to the bond most closely associated with the first side-chain dihedral angle (χ_1), using Cβ—H for all amino acids other than Gly and Cα—H for Gly. Because S^2 varies from 0 to 1 as a measure of order, we used 1-S^2 as a measure of disorder.

The KSI crystals obtained in this study contained two molecules in the asymmetric unit and the average of the two molecules was used for analysis. Because the total ground-state analog (4-Androstenedione) occupancy in the final refined model of the two crystallographically-

independent KSI molecules was 1.4 instead of 2.0 (corresponding to the maximum occupancy of 1.0 for each of the KSI molecules), the (1-S^2) values for each residue were corrected using the equation:

$$(1-S^2)\,GSA_{corrected} = \left((1-S^2)GSA_{observed} - 0.3x(1-S^2)Apo\right)\Big/_{0.7}$$

Ensemble building. To obtain KSI pseudo-ensembles, all KSI cryo crystal structures were downloaded from the PDB (Berman et al., 2000) and parsed into individual KSI monomers (**Table S1**). While KSI is a dimer, we focused the analysis on the individual KSI molecule as: i) all individual KSI molecules were highly similar, with RMSDs below 0.5 Å after excluding two flexible loops (see Results); ii) KSI is known to also crystallize with one molecule in the crystallographic asymmetric unit (see **Table S1**), indicating that the two molecules of the dimer are identical; and iii) each monomer has the full catalytic machinery required for catalysis. All KSI molecules were aligned using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.) on the protein backbone (N, Cα, C, O atoms) of residues 5-125 on either the highest resolution crystal structure of KSI bound to the transitions state analog equilenin (PDB 1OH0, (Kim et al., 1997)) or as otherwise indicated. Residues 1-4 (N-terminal) and 126-131(C-terminal) were excluded from the analyses because these residues appeared highly flexible, as is common for N/C-terminal residues, and were also not modeled in some of the KSI structures. No alignment gaps were allowed during the alignment, and allowing for gaps (default PyMOL alignment procedure) did not appreciably change the results (**Figure supplement 8**). The same approach was used to obtain $KSI_{homolog}$ pseudo-ensembles, with the backbone of residues 3-122 aligned on the highest resolution crystal structure of $KSI_{homolog}$ bound to the ground state analog 4-androstene-3,17-dione (PDB 3NHX). Different sets of crystal structures were used to obtain different types of pseudo-ensembles (sub-ensembles) and the structures included in the different types of pseudo-ensembles are listed in **Tables S1, 2, 20-22** and explicitly indicated in the legends of figures from the main text and supplemental information. The KSI RT-ensemble was obtained from the Apo, GSA-bound, and TSA-bound RT multi-conformer models using the same alignment procedure and the Apo state as alignment template.

Calculating mean deviations. From the ensemble of aligned crystal structures or multi-conformer models, a list of *xyz* coordinates was defined for each atom using PyMOL and standard code. For *n* copies of each particular atom in the ensemble, a coordinate list of size 3*n* was obtained for this atom. The mean deviation (MDev) was then calculated by computing the spread of these points about the center point *j*:

$$mean\ deviation = \sqrt{\min_j\left(\frac{1}{n}\sum_i \left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 + \left(z_i - z_j\right)^2\right)}$$

Here i is an index over each of the n atoms in the ensemble and the average distance of all points to a center point from atom j was calculated. The center point j was chosen to give the minimum average distance to all other points in the given ensemble of atoms and was determined by calculating the average distance between each atom in a given ensemble and all other atoms in the same ensemble of atoms.

Bootstrap analysis. A bootstrap analysis was used to estimate the errors associated with the sum of Cα MDev values of the Apo and the transition-state bound KSI pseudo-ensembles (ΣMDev Apo and TSA-bound, **Figure 4E** from the main text). Briefly, from the number of distances, n (e.g. for the full pseudo-ensemble of 94 structures (q = 94), $n = q - 1$), used to calculate the Cα MDev for a given KSI residue x (where x denotes residues 5 to 125), a random number of distances, m_1, was randomly selected and excluded from n. A second small number of distances, m_2, equal to m_1, was randomly selected from n to replace the excluded m_1 and then the sum of the MDevs for residues 5-125 was calculated (ΣMDev). The procedure was repeated 300 times and the standard deviations associated with an increasing number of bootstrapped ΣMDev (n = 2, 5, 10, 20, 30, 40, 50, 100, 200 and 300 cycles, **Figure supplement 7, Table S27**) was used to estimate the standard deviation of the ΣMDev in **Figure 4E** from the main text.

Acknowledgments. This work was funded by a National Science Foundation (NSF) Grant (MCB-1714723) to DH. FY was supported in part by a long-term Human Frontiers Science Program postdoctoral fellowship and in part by the NSF grant MCB-1714723. MMP was supported in part by an NSF Graduate Research Fellowship and in part by a Lieberman Fellowship from Stanford University. ASP was supported in part by the National Science Foundation Graduate Research Fellowship and in part by the Stanford ChEM-H Chemistry/Biology Interface Predoctoral Training Program and the National Institute of General Medical Sciences of the NIH under Award Number T32GM120007. We thank Dr. Steve Bonilla for help with the bootstrap analysis, Stanford Synchrotron Radiation Lighthouse (SSRL) and Lisa Dunn for beam time allocation and access, Dr. Corey W. Liu, Stanford Magnetic Resonance Laboratory, and Dr. Mark Kelly, UCSF Nuclear Magnetic Resonance Laboratory, for assistance with NMR spectroscopy, and members of the Herschlag lab for helpful suggestions and feedback on the manuscript. The SMRL [MMP1] 800 MHz NMR was supported in part by NIH Shared Instrumentation Grant 1 S10 RR025612-01A1.

References

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D Biol. Crystallogr. *68*, 352–367.

Agarwal, P.K. (2004). Cis/trans isomerization in HIV-1 capsid protein catalyzed by cyclophilin A: Insights from computational and theoretical studies. Proteins Struct. Funct. Bioinforma. *56*, 449–463.

Agarwal, P.K. (2019). A Biophysical Perspective on Enzyme Catalysis. Biochemistry *58*, 438–449.

Agarwal, P.K., Billeter, S.R., Rajagopalan, P.T.R., Benkovic, S.J., and Hammes-Schiffer, S. (2002). Network of coupled promoting motions in enzyme catalysis. Proc. Natl. Acad. Sci. *99*, 2794–2799.

Agarwal, P.K., Geist, A., and Gorin, A. (2004). Protein Dynamics and Enzymatic Catalysis: Investigating the Peptidyl−Prolyl Cis−Trans Isomerization Activity of Cyclophilin A [†]. Biochemistry *43*, 10605–10618.

Austin, R.H., Beeson, K.W., Eisenstein, L., Frauenfelder, H., and Gunsalus, I.C. (1975). Dynamics of ligand binding to myoglobin. Biochemistry *14*, 5355–5373.

van den Bedem, H., Dhanik, A., Latombe, J.-C., and Deacon, A.M. (2009). Modeling discrete heterogeneity in X-ray diffraction data by fitting multi-conformers. Acta Crystallogr. D Biol. Crystallogr. *65*, 1107–1117.

Benkovic, S.J., Hammes, G.G., and Hammes-Schiffer, S. (2008). Free-Energy Landscape of Enzyme Catalysis [†]. Biochemistry *47*, 3317–3321.

Berg, J.M., Tymoczko, J.L., Stryer, L., Berg, J.M., Tymoczko, J.L., and Stryer, L. (2002). Biochemistry (W H Freeman).

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

Best, R.B., Lindorff-Larsen, K., DePristo, M.A., and Vendruscolo, M. (2006). Relation between native ensembles and experimental structures of proteins. Proc. Natl. Acad. Sci. *103*, 10901–10906.

Blake, C.C., Johnson, L.N., Mair, G.A., North, A.C., Phillips, D.C., and Sarma, V.R. (1967). Crystallographic studies of the activity of hen egg-white lysozyme. Proc. R. Soc. Lond. B Biol. Sci. *167*, 378–388.

Blow, D. (2000). So do we understand how enzymes work? Structure *8*, R77–R81.

Boehr, D.D., Dyson, H.J., and Wright, P.E. (2006). An NMR perspective on enzyme dynamics. Chem. Rev. *106*, 3055–3079.

Boehr, D.D., Nussinov, R., and Wright, P.E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. Nat. Chem. Biol. *5*, 789–796.

Bruice, T.C. (1970). 4 Proximity Effects and Enzyme Catalysis. In The Enzymes, P.D. Boyer, ed. (Academic Press), pp. 217–279.

Bruice, T.C. (1976). Some Pertinent Aspects of Mechanism as Determined with Small Molecules. Annu. Rev. Biochem. *45*, 331–374.

Brünger, A.T., and Nilges, M. (1993). Computational challenges for macromolecular structure determination by X-ray crystallography and solution NMRspectroscopy. Q. Rev. Biophys. *26*, 49–125.

Bryan, P., Pantoliano, M.W., Quill, S.G., Hsiao, H.Y., and Poulos, T. (1986). Site-directed mutagenesis and the role of the oxyanion hole in subtilisin. Proc. Natl. Acad. Sci. *83*, 3743–3745.

Callender, R., and Dyer, R.B. (2015). The Dynamical Nature of Enzymatic Catalysis. Acc. Chem. Res. *48*, 407–413.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Childers, M.C., and Daggett, V. (2018). Validating Molecular Dynamics Simulations against Experimental Observables in Light of Underlying Conformational Ensembles. J. Phys. Chem. B *122*, 6673–6689.

Chothia, C., and Janin, J. (1975). Principles of protein–protein recognition. Nature *256*, 705–708.

Clarage, J.B., Romo, T., Andrews, B.K., Pettitt, B.M., and Phillips, G.N. (1995). A sampling problem in molecular dynamics simulations of macromolecules. Proc. Natl. Acad. Sci. U. S. A. *92*, 3288–3292.

Cleland, W.W., and Kreevoy, M.M. (1994). Low-barrier hydrogen bonds and enzymic catalysis. Science *264*, 1887–1890.

Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Cooper, A., and Dryden, D.T.F. (1984). Allostery without conformational change: A plausible model. Eur. Biophys. J. *11*, 103–109.

Desiraju, G., and Steiner, T. (1999). The Weak Hydrogen Bond: In Structural Chemistry and Biology (Oxford, New York: Oxford University Press).

Dewar, M.J., and Storch, D.M. (1985). Alternative view of enzyme reactions. Proc. Natl. Acad. Sci. U. S. A. *82*, 2225–2229.

Dill, K.A. (1990). Dominant forces in protein folding. Biochemistry *29*, 7133–7155.

Doster, W., Cusack, S., and Petry, W. (1989). Dynamical transition of myoglobin revealed by inelastic neutron scattering. Nature *337*, 754.

Dwyer, D.S. (2005). Electronic properties of amino acid side chains: quantum mechanics calculation of substituent effects. BMC Chem. Biol. *5*, 2.

Elber, R., and Karplus, M. (1987). Multiple conformational states of proteins: a molecular dynamics analysis of myoglobin. Science *235*, 318–321.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. *69*, 1204–1214.

Eyring, Henry. (1935). The Activated Complex and the Absolute Rate of Chemical Reactions. Chem. Rev. *17*, 65–77.

Fenwick, R.B., van den Bedem, H., Fraser, J.S., and Wright, P.E. (2014). Integrated description of protein dynamics from room-temperature X-ray crystallography and NMR. Proc. Natl. Acad. Sci. *111*, E445–E454.

Fersht, A. (1985). Enzyme Structure and Mechanism (W.H. Freeman).

Fersht, A.R., Shi, J.-P., Knill-Jones, J., Lowe, D.M., Wilkinson, A.J., Blow, D.M., Brick, P., Carter, P., Waye, M.M.Y., and Winter, G. (1985). Hydrogen bonding and biological specificity analysed by protein engineering. Nature *314*, 235–238.

Fischer, E. (1894). Einfluss der Configuration auf die Wirkung der Enzyme. II. Berichte Dtsch. Chem. Ges. *27*, 3479–3483.

Ford, L.O., Johnson, L.N., Machin, P.A., Phillips, D.C., and Tjian, R. (1974). Crystal structure of a lysozyme-tetrasaccharide lactone complex. J. Mol. Biol. *88*, 349–371.

Fraser, J.S., Clarkson, M.W., Degnan, S.C., Erion, R., Kern, D., and Alber, T. (2009). Hidden alternative structures of proline isomerase essential for catalysis. Nature *462*, 669–673.

Fraser, J.S., van den Bedem, H., Samelson, A.J., Lang, P.T., Holton, J.M., Echols, N., and Alber, T. (2011). Accessing protein conformational ensembles using room-temperature X-ray crystallography. Proc. Natl. Acad. Sci. *108*, 16247–16252.

Frauenfelder, H., Petsko, G.A., and Tsernoglou, D. (1979). Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. Nature *280*, 558–563.

Frauenfelder, H., Sligar, S., and Wolynes, P. (1991). The energy landscapes and motions of proteins. Science *254*, 1598–1603.

Frederick, K.K., Marlow, M.S., Valentine, K.G., and Wand, A.J. (2007). Conformational entropy in molecular recognition by proteins. Nature *448*, 325–329.

French, S., and Wilson, K. (1978). On the treatment of negative intensity observations.

Furnham, N., Blundell, T.L., DePristo, M.A., and Terwilliger, T.C. (2006). Is one solution good enough? Nat. Struct. Mol. Biol. *13*, 184–185.

Gerlt, J.A., and Gassman, P.G. (1993). Understanding the rates of certain enzyme-catalyzed reactions: proton abstraction from carbon acids, acyl-transfer reactions, and displacement reactions of phosphodiesters. Biochemistry *32*, 11943–11952.

Graham, J.D., Buytendyk, A.M., Wang, D., Bowen, K.H., and Collins, K.D. (2014). Strong, Low-Barrier Hydrogen Bonds May Be Available to Enzymes. Biochemistry *53*, 344–349.

Gutteridge, A., and Thornton, J. (2005). Conformational Changes Observed in Enzyme Crystal Structures upon Substrate Binding. J. Mol. Biol. *346*, 21–28.

Haldane, J.B.S. (1930). Enzymes (London; New York: Longmans, Green).

Halle, B. (2004). Biomolecular cryocrystallography: Structural changes during flash-cooling. Proc. Natl. Acad. Sci. *101*, 4793–4798.

Hammes, G.G., Benkovic, S.J., and Hammes-Schiffer, S. (2011). Flexibility, Diversity, and Cooperativity: Pillars of Enzyme Catalysis. Biochemistry *50*, 10422–10430.

Harris, T.K., and Mildvan, A.S. (1999). High-Precision Measurement of Hydrogen Bond Lengths in Proteins by Nuclear Magnetic Resonance Methods. Proteins Struct. Funct. Genet. *35*, 275–282.

Henderson, R. (1970). Structure of crystalline α-chymotrypsin. J. Mol. Biol. *54*, 341–354.

Henderson, R., and Wang, J.H. (1972). Catalytic configurations. Annu. Rev. Biophys. Bioeng. *1*, 1–26.

Henzler-Wildman, K., and Kern, D. (2007). Dynamic personalities of proteins. Nature *450*, 964–972.

Herschlag, D., and Pinney, M.M. (2018). Hydrogen Bonds: Simple after All? Biochemistry *57*, 3338–3352.

Holton, J.M., Classen, S., Frankel, K.A., and Tainer, J.A. (2014). The R-factor gap in macromolecular crystallography: an untapped potential for insights on accurate structures. FEBS J. *281*, 4046–4060.

Hur, S., and Bruice, T.C. (2003). The near attack conformation approach to the study of the chorismate to prephenate reaction. Proc. Natl. Acad. Sci. *100*, 12015–12020.

Jackson, M.R., Beahm, R., Duvvuru, S., Narasimhan, C., Wu, J., Wang, H.-N., Philip, V.M., Hinde, R.J., and Howell, E.E. (2007). A Preference for Edgewise Interactions between Aromatic Rings and Carboxylate Anions: The Biological Relevance of Anion−Quadrupole Interactions. J. Phys. Chem. B *111*, 8242–8249.

Jeffrey, G.A., and Saenger, W. (1991). Hydrogen Bonding in Biological Structures (Berlin Heidelberg: Springer-Verlag).

Jeffrey, G.A., and Yeon, Y. (1986). The correlation between hydrogen-bond lengths and proton chemical shifts in crystals. Acta Crystallogr. B *42*, 410–413.

Jencks, W.P. (1969). Catalysis in Chemistry and Enzymology (Courier Corporation).

Jencks, W.P. (1975). Binding energy, specificity, and enzymic catalysis: the circe effect. Adv. Enzymol. Relat. Areas Mol. Biol. *43*, 219–410.

Juers, D.H., and Matthews, B.W. (2001). Reversible lattice repacking illustrates the temperature dependence of macromolecular interactions. J. Mol. Biol. *311*, 851–862.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kapoor, K., Duff, M.R., Upadhyay, A., Bucci, J.C., Saxton, A.M., Hinde, R.J., Howell, E.E., and Baudry, J. (2016). Highly Dynamic Anion–Quadrupole Networks in Proteins. Biochemistry *55*, 6056–6069.

Keedy, D.A. (2019). Journey to the center of the protein: allostery from multitemperature multiconformer X-ray crystallography. Acta Crystallogr. Sect. Struct. Biol. *75*, 123–137.

Keedy, D.A., van den Bedem, H., Sivak, D.A., Petsko, G.A., Ringe, D., Wilson, M.A., and Fraser, J.S. (2014). Crystal Cryocooling Distorts Conformational Heterogeneity in a Model Michaelis Complex of DHFR. Structure *22*, 899–910.

Keedy, D.A., Fraser, J.S., and van den Bedem, H. (2015a). Exposing Hidden Alternative Backbone Conformations in X-ray Crystallography Using qFit. PLOS Comput. Biol. *11*, e1004507.

Keedy, D.A., Kenner, L.R., Warkentin, M., Woldeyes, R.A., Hopkins, J.B., Thompson, M.C., Brewster, A.S., Van Benschoten, A.H., Baxter, E.L., Uervirojnangkoorn, M., et al. (2015b). Mapping the conformational landscape of a dynamic enzyme by multitemperature and XFEL crystallography. ELife *4*, e07574.

Keedy, D.A., Hill, Z.B., Biel, J.T., Kang, E., Rettenmaier, T.J., Brandão-Neto, J., Pearce, N.M., von Delft, F., Wells, J.A., and Fraser, J.S. (2018). An expanded allosteric network in PTP1B by multitemperature crystallography, fragment screening, and covalent tethering. ELife *7*, e36307.

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.

Kim, S.W., Kim, C.Y., Benisek, W.F., and Choi, K.Y. (1994). Cloning, nucleotide sequence, and overexpression of the gene coding for delta 5-3-ketosteroid isomerase from Pseudomonas putida biotype B. J. Bacteriol. *176*, 6672–6676.

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 Δ5-3-Ketosteroid Isomerase with and without a Reaction Intermediate Analogue. Biochemistry *36*, 14030–14036.

Kirby, A.J. (1980). Effective Molarities for Intramolecular Reactions. In Advances in Physical Organic Chemistry, V. Gold, and D. Bethell, eds. (Academic Press), pp. 183–278.

Kirby, A.J., and Lancaster, P.W. (1972). Structure and efficiency in intramolecular and enzymic catalysis. Catalysis of amide hydrolysis by the carboxy-group of substituted maleamic acids. J. Chem. Soc. Perkin Trans. 2 1206–1214.

Kleckner, I.R., and Foster, M.P. (2011). An introduction to NMR-based approaches for measuring protein dynamics. Biochim. Biophys. Acta BBA - Proteins Proteomics *1814*, 942–968.

Klinman, J.P. (2009). An integrated model for enzyme catalysis emerges from studies of hydrogen tunneling. Chem. Phys. Lett. *471*, 179–193.

Klinman, J.P. (2013). Importance of Protein Dynamics during Enzymatic C–H Bond Cleavage Catalysis. Biochemistry *52*, 2068–2077.

Klinman, J.P. (2015). Dynamically Achieved Active Site Precision in Enzyme Catalysis. Acc. Chem. Res. *48*, 449–456.

Klinman, J.P., and Kohen, A. (2013). Hydrogen Tunneling Links Protein Dynamics to Enzyme Catalysis. Annu. Rev. Biochem. *82*, 471–496.

Knowles, J.R. (1991). Enzyme catalysis: not different, just better. Nature *350*, 121.

Koshland, D.E. (1953). Stereochemistry and the Mechanism of Enzymatic Reactions. Biol. Rev. *28*, 416–436.

Kovermann, M., Rogne, P., and Wolf-Watz, M. (2016). Protein dynamics and function from solution state NMR spectroscopy. Q. Rev. Biophys. *49*.

Kraut, J. (1977). Serine Proteases: Structure and Mechanism of Catalysis. Annu. Rev. Biochem. *46*, 331–358.

Kraut, D.A., Carroll, K.S., and Herschlag, D. (2003). Challenges in enzyme mechanism and energetics. Annu. Rev. Biochem. *72*, 517–571.

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*, e99.

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.

Kumar, G.S., Clarkson, M.W., Kunze, M.B.A., Granata, D., Wand, A.J., Lindorff-Larsen, K., Page, R., and Peti, W. (2018). Dynamic activation and regulation of the mitogen-activated protein kinase p38. Proc. Natl. Acad. Sci. *115*, 4655–4660.

Lamba, V., Yabukarski, F., Pinney, M., and Herschlag, D. (2016). Evaluation of the Catalytic Contribution from a Positioned General Base in Ketosteroid Isomerase. J. Am. Chem. Soc. *138*, 9902–9909.

Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat. Protoc. *3*, 1171–1179.

Lewandowski, J.R., Halse, M.E., Blackledge, M., and Emsley, L. (2015). Protein dynamics. Direct observation of hierarchical protein dynamics. Science *348*, 578–581.

Lienhard, G.E. (1973). Enzymatic Catalysis and Transition-State Theory. Science *180*, 149–154.

Ma, B., Kumar, S., Tsai, C.J., Hu, Z., and Nussinov, R. (2000). Transition-state ensemble in enzyme catalysis: possibility, reality, or necessity? J. Theor. Biol. *203*, 383–397.

Marlow, M.S., Dogan, J., Frederick, K.K., Valentine, K.G., and Wand, A.J. (2010). The role of conformational entropy in molecular recognition by calmodulin. Nat. Chem. Biol. *6*, 352–358.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

McDonald, I.K., and Thornton, J.M. (1994). Satisfying Hydrogen Bonding Potential in Proteins. J. Mol. Biol. *238*, 777–793.

Meadows, C.W., Tsang, J.E., and Klinman, J.P. (2014). Picosecond-Resolved Fluorescence Studies of Substrate and Cofactor-Binding Domain Mutants in a Thermophilic Alcohol Dehydrogenase Uncover an Extended Network of Communication. J. Am. Chem. Soc. *136*, 14821–14833.

Menger, F.M. (2005). An alternative view of enzyme catalysis. Pure Appl. Chem. *77*, 1873–1886.

Mildvan, A.S., Harris, T.K., and Abeygunawardana, C. (1999). Nuclear magnetic resonance methods for the detection and study of low-barrier hydrogen bonds on enzymes. Methods Enzymol. *308*, 219–245.

Milstien, S., and Cohen, L.A. (1970). Rate Acceleration by Stereopopulation Control: Models for Enzyme Action. Proc. Natl. Acad. Sci. *67*, 1143–1147.

Natarajan, A., Schwans, J.P., and Herschlag, D. (2014). Using unnatural amino acids to probe the energetics of oxyanion hole hydrogen bonds in the ketosteroid isomerase active site. J. Am. Chem. Soc. *136*, 7643–7654.

Nowak, T., and Mildvan, A.S. (1972). Nuclearmagnetic resonance studies of selectively hindered internal motion of substrate analogs at the active site of pyruvate kinase. Biochemistry *11*, 2813–2818.

Pace, N., Scholtz, J.M., and Grimsley, G.R. (2014). Forces stabilizing proteins. FEBS Lett. *588*, 2177–2184.

Page, M.I., and Jencks, W.P. (1971). Entropic Contributions to Rate Accelerations in Enzymic and Intramolecular Reactions and the Chelate Effect. Proc. Natl. Acad. Sci. *68*, 1678–1683.

Pauling, L. (1946). Molecular Architecture and Biological Reactions. Chem. Eng. News *24*, 1375–1377.

Peng, J.W. (2009). Communication Breakdown: Protein Dynamics and Drug Design. Structure *17*, 319–320.

Peter Guthrie, J. (1996). Short strong hydrogen bonds: can they explain enzymic catalysis? Chem. Biol. *3*, 163–170.

Petrounia, I.P., and Pollack, R.M. (1998). Substituent Effects on the Binding of Phenols to the D38N Mutant of 3-Oxo-Δ^5-steroid Isomerase. A Probe for the Nature of Hydrogen Bonding to the Intermediate †. Biochemistry *37*, 700–705.

Philip, V., Harris, J., Adams, R., Nguyen, D., Spiers, J., Baudry, J., Howell, E.E., and Hinde, R.J. (2011). A Survey of Aspartate−Phenylalanine and Glutamate−Phenylalanine Interactions in the Protein Data Bank: Searching for Anion−π Pairs. Biochemistry *50*, 2939–2950.

Pinney, M.M., Natarajan, A., Yabukarski, F., Sanchez, D.M., Liu, F., Liang, R., Doukov, T., Schwans, J.P., Martinez, T.J., and Herschlag, D. (2018). Structural Coupling Throughout the Active Site Hydrogen Bond Networks of Ketosteroid Isomerase and Photoactive Yellow Protein. J. Am. Chem. Soc. *140*, 9827–9843.

Polanyi, M. (1921). On adsorption catalysis. Z. Elektrochem. *27*, 142–150.

Pollack, R.M. (2004). Enzymatic mechanisms for catalysis of enolization: ketosteroid isomerase. Bioorganic Chem. *32*, 341–353.

Radzicka, A., and Wolfenden, R. (1995). A proficient enzyme. Science *267*, 90–93.

Raines, R.T., Sutton, E.L., Straus, D.R., Gilbert, W., and Knowles, J.R. (1986). Reaction energetics of a mutant triose phosphate isomerase in which the active-site glutamate has been changed to aspartate. Biochemistry *25*, 7142–7154.

Ramanathan, A., and Agarwal, P.K. (2011). Evolutionarily Conserved Linkage between Enzyme Fold, Flexibility, and Catalysis. PLoS Biol. *9*, e1001193.

Reddish, M.J., Peng, H.-L., Deng, H., Panwar, K.S., Callender, R., and Dyer, R.B. (2014). Direct Evidence of Catalytic Heterogeneity in Lactate Dehydrogenase by Temperature Jump Infrared Spectroscopy. J. Phys. Chem. B *118*, 10854–10862.

Rejto, P.A., and Freer, S.T. (1996). Protein conformational substates from X-ray crystallography. Prog. Biophys. Mol. Biol. *66*, 167–196.

Richard, J.P., Amyes, T.L., Goryanova, B., and Zhai, X. (2014). Enzyme architecture: on the importance of being in a protein cage. Curr. Opin. Chem. Biol. *21*, 1–10.

Ringe, D., and Petsko, G.A. (2003). The “glass transition” in protein dynamics: what it is, why it occurs, and how to exploit it. Biophys. Chem. *105*, 667–680.

Robertus, J.D., Kraut, J., Alden, R.A., and Birktoft, J.J. (1972). Subtilisin. Stereochemical mechanism involving transition-state stabilization. Biochemistry *11*, 4293–4303.

Ruben, E.A., Schwans, J.P., Sonnett, M., Natarajan, A., Gonzalez, A., Tsai, Y., and Herschlag, D. (2013). Ground state destabilization from a positioned general base in the ketosteroid isomerase active site. Biochemistry *52*, 1074–1081.

Sandalova, T., Schneider, G., Käck, H., and Lindqvist, Y. (1999). Structure of dethiobiotin synthetase at 0.97 Å resolution. Acta Crystallogr. D Biol. Crystallogr. *55*, 610–624.

Scheidig, A.J., Burmester, C., and Goody, R.S. (1999). The pre-hydrolysis state of p21ras in complex with GTP: new insights into the role of water molecules in the GTP hydrolysis reaction of ras-like proteins. Structure *7*, 1311-S2.

Schwans, J.P., Kraut, D.A., and Herschlag, D. (2009). Determining the catalytic role of remote substrate binding interactions in ketosteroid isomerase. Proc. Natl. Acad. Sci. U. S. A. *106*, 14271–14275.

Schwans, J.P., Sunden, F., Lassila, J.K., Gonzalez, A., Tsai, Y., and Herschlag, D. (2013). Use of anion-aromatic interactions to position the general base in the ketosteroid isomerase active site. Proc. Natl. Acad. Sci. U. S. A. *110*, 11308–11313.

Schwans, J.P., Hanoian, P., Lengerich, B.J., Sunden, F., Gonzalez, A., Tsai, Y., Hammes-Schiffer, S., and Herschlag, D. (2014). Experimental and computational mutagenesis to investigate the positioning of a general base within an enzyme active site. Biochemistry *53*, 2541–2555.

Schwans, J.P., Sunden, F., Gonzalez, A., Tsai, Y., and Herschlag, D. (2016). Correction to "Evaluating the Catalytic Contribution from the Oxyanion Hole in Ketosteroid Isomerase." J. Am. Chem. Soc. *138*, 7801–7802.

Secemski, I.I., Lehrer, S.S., and Lienhard, G.E. (1972). A Transition State Analog for Lysozyme. J. Biol. Chem. *247*, 4740–4748.

Sekhar, A., and Kay, L.E. (2019). An NMR View of Protein Dynamics in Health and Disease. Annu. Rev. Biophys. *48*, 297–319.

Shan, S. -o., and Herschlag, D. (1996). The change in hydrogen bond strength accompanying charge rearrangement: Implications for enzymatic catalysis. Proc. Natl. Acad. Sci. *93*, 14474–14479.

Shan, S.O., and Herschlag, D. (1999). Hydrogen bonding in enzymatic catalysis: analysis of energetic contributions. Methods Enzymol. *308*, 246–276.

Shan, S., Loh, S., and Herschlag, D. (1996). The Energetics of Hydrogen Bonds in Model Systems: Implications for Enzymatic Catalysis. Science *272*, 97–101.

Shoichet, B.K., Baase, W.A., Kuroki, R., and Matthews, B.W. (1995). A relationship between protein stability and protein function. Proc. Natl. Acad. Sci. *92*, 452–456.

Sigala, P.A., Fafarman, A.T., Bogard, P.E., Boxer, S.G., and Herschlag, D. (2007). Do ligand binding and solvent exclusion alter the electrostatic character within the oxyanion hole of an enzymatic active site? J. Am. Chem. Soc. *129*, 12104–12105.

Sigala, P.A., Kraut, D.A., Caaveiro, J.M.M., Pybus, B., Ruben, E.A., Ringe, D., Petsko, G.A., and Herschlag, D. (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.

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.

Sigala, P.A., Ruben, E.A., Liu, C.W., Piccoli, P.M.B., Hohenstein, E.G., Martínez, T.J., Schultz, A.J., and Herschlag, D. (2015). Determination of Hydrogen Bond Structure in Water versus Aprotic Environments To Test the Relationship Between Length and Stability. J. Am. Chem. Soc. *137*, 5730–5740.

Silva, R.G., Murkin, A.S., and Schramm, V.L. (2011). Femtosecond dynamics coupled to chemical barrier crossing in a Born-Oppenheimer enzyme. Proc. Natl. Acad. Sci. *108*, 18661–18665.

Simón, L., and Goodman, J.M. (2010). Enzyme Catalysis by Hydrogen Bonds: The Balance between Transition State Binding and Substrate Binding in Oxyanion Holes. J. Org. Chem. *75*, 1831–1840.

Smith, J., Kuczera, K., and Karplus, M. (1990). Dynamics of myoglobin: comparison of simulation results with neutron scattering spectra. Proc. Natl. Acad. Sci. U. S. A. *87*, 1601–1605.

Smock, R.G., and Gierasch, L.M. (2009). Sending Signals Dynamically. Science *324*, 198–203.

Steiner, T., and Desiraju, G.R. (1998). Distinction between the weak hydrogen bond and the van der Waals interaction. Chem. Commun. 891–892.

Storm, D.R., and Koshland, D.E. (1970). A Source for the Special Catalytic Power of Enzymes: Orbital Steering. Proc. Natl. Acad. Sci. *66*, 445–452.

Suarez, J., and Schramm, V.L. (2015). Isotope-specific and amino acid-specific heavy atom substitutions alter barrier crossing in human purine nucleoside phosphorylase. Proc. Natl. Acad. Sci. *112*, 11247–11251.

Taft, R.W. (2007). Progress in physical organic chemistry. (New York: Wiley).

Takeuchi, K., Tokunaga, Y., Imai, M., Takahashi, H., and Shimada, I. (2015). Dynamic multidrug recognition by multidrug transcriptional repressor LmrR. Sci. Rep. *4*.

Teeter, M.M., Yamano, A., Stec, B., and Mohanty, U. (2001). On the nature of a glassy state of matter in a hydrated protein: Relation to protein function. Proc. Natl. Acad. Sci. *98*, 11242–11247.

Thomas, K.A., Smith, G.M., Thomas, T.B., and Feldmann, R.J. (1982). Electronic distributions within protein phenylalanine aromatic rings are reflected by the three-dimensional oxygen atom environments. Proc. Natl. Acad. Sci. *79*, 4843–4847.

Tilton, R.F., Dewan, J.C., and Petsko, G.A. (1992). Effects of temperature on protein structure and dynamics: X-ray crystallographic studies of the protein ribonuclease-A at nine different temperatures from 98 to 320 K. Biochemistry *31*, 2469–2481.

Tokuriki, N., and Tawfik, D.S. (2009). Protein Dynamism and Evolvability. Science *324*, 203–207.

Turner, D.L. (1983). Binomial solvent suppression. J. Magn. Reson. 1969 *54*, 146–148.

Vallee, B.L., and Williams, R.J. (1968). Metalloenzymes: the entatic nature of their active sites. Proc. Natl. Acad. Sci. U. S. A. *59*, 498–505.

van Gunsteren, W.F., Daura, X., Hansen, N., Mark, A.E., Oostenbrink, C., Riniker, S., and Smith, L.J. (2018). Validation of Molecular Simulation: An Overview of Issues. Angew. Chem. Int. Ed. *57*, 884–902.

Walsh, C. (1979). Enzymatic Reaction Mechanisms (W. H. Freeman).

Wand, A.J., Moorman, V.R., and Harpole, K.W. (2013). A surprising role for conformational entropy in protein function. Top. Curr. Chem. *337*, 69–94.

Warshel, A. (1998). Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites. J. Biol. Chem. *273*, 27035–27038.

Warshel, A., Sharma, P.K., Kato, M., Xiang, Y., Liu, H., and Olsson, M.H.M. (2006). Electrostatic Basis for Enzyme Catalysis. Chem. Rev. *106*, 3210–3235.

Watson, J., Baker, T., Bell, S., Gann, A., Levine, M., and Losick, R. (2003). Molecular Biology of the Gene (Cold Spring Harbor Laboratory Press).

Williams, R.J. (1972). The entatic state. Cold Spring Harb. Symp. Quant. Biol. *36*, 53–62.

Wolfenden, R. (1976). Transition State Analog Inhibitors and Enzyme Catalysis. Annu. Rev. Biophys. Bioeng. *5*, 271–306.

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.

Yun, Y.S., Nam, G.H., Kim, Y.-G., Oh, B.-H., and Choi, K.Y. (2005). Small exterior hydrophobic cluster contributes to conformational stability and steroid binding in ketosteroid isomerase from Pseudomonas putida biotype B. FEBS J. *272*, 1999–2011.

Zeymer, C., Werbeck, N.D., Zimmermann, S., Reinstein, J., and Hansen, D.F. (2016). Characterizing Active Site Conformational Heterogeneity along the Trajectory of an Enzymatic Phosphoryl Transfer Reaction. Angew. Chem. Int. Ed. *55*, 11533–11537.

Zhang, X., Wozniak, J.A., and Matthews, B.W. (1995a). Protein Flexibility and Adaptability Seen in 25 Crystal Forms of T4 Lysozyme. J. Mol. Biol. *250*, 527–552.

Zhang, X.J., Baase, W.A., Shoichet, B.K., Wilson, K.P., and Matthews, B.W. (1995b). Enhancement of protein stability by the combination of point mutations in T4 lysozyme is additive. Protein Eng. *8*, 1017–1022.

Zhao, Q., Abeygunawardana, C., Talalay, P., and Mildvan, A.S. (1996). NMR Evidence for the Participation of a Low-Barrier Hydrogen Bond in the Mechanism of Δ 5-3-Ketosteroid isomerase. Proc. Natl. Acad. Sci. U. S. A. *93*, 8220–8224.

Zoete, V., Michielin, O., and Karplus, M. (2002). Relation between sequence and structure of HIV-1 protease inhibitor complexes: a model system for the analysis of protein flexibility11Edited by R. Huber. J. Mol. Biol. *315*, 21–52.