JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

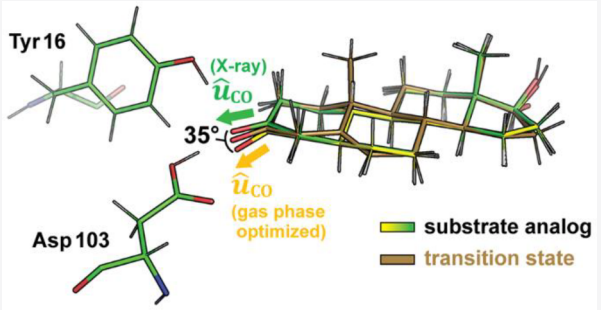
A Preorganized Electric Field Leads to Minimal Geometrical Reorientation in the Catalytic Reaction of Ketosteroid Isomerase

Yufan Wu, Stephen D. Fried, and Steven G. Boxer*

Cite This: J. Am. Chem. Soc. 2020, 142, 9993−9998
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Electrostatic interactions play a pivotal role in enzymatic catalysis and are increasingly modeled explicitly in computational enzyme design; nevertheless, they are challenging to measure experimentally. Using vibrational Stark effect (VSE) spectroscopy, we have measured electric fields inside the active site of the enzyme ketosteroid isomerase (KSI). These studies have shown that these fields can be unusually large, but it has been unclear to what extent they specifically stabilize the transition state (TS) relative to a ground state (GS). In the following, we use crystallography and computational modeling to show that KSI's intrinsic electric field is nearly perfectly oriented to stabilize the geometry of its reaction's TS. Moreover, we find that this electric field adjusts the orientation of its substrate in the ground state so that the substrate needs to only undergo minimal structural changes upon activation to its TS. This work provides evidence that the active site electric field in KSI is preorganized to facilitate catalysis and provides a template for how electrostatic preorganization can be measured in enzymatic systems.



Downloaded via STANFORD UNIV on June 11, 2020 at 22:11:47 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

Electrostatic stabilization has been widely discussed as an important feature that endows enzymes with their signature high catalytic proficiencies.[1−9] One (simple) way to frame this hypothesis is that enzymes create a particular electrostatic environment in their active sites, which preferentially stabilizes the charge distribution of the transition state (TS) more than the ground state (GS) to accelerate the reaction[4]

$$\Delta\Delta G^{\ddagger} = -((\vec{F}_{enz,TS}\cdot\vec{\mu}_{TS}) - (\vec{F}_{enz,R}\cdot\vec{\mu}_{R})) \quad (1)$$

where $\vec{\mu}_R$ is the reactant's dipole moment, $\vec{\mu}_{TS}$ is the transition state's dipole moment, $\vec{F}_{enz,R}$ is the electric field the environment exerts on the reactant dipole, and $\vec{F}_{enz,TS}$ is the electric field the environment exerts on the transition state dipole. It should be noted that eq 1 exactly treats the simple case where the substrate is a point-dipole; in more realistic scenarios, $\Delta\Delta G^{\ddagger}$ is a sum of several terms for each bond dipole that changes during the reaction coordinate in which the fields correspond to the fields projected onto the given bond dipoles at their respective positions.

Conceptually, electrostatic stabilization can be divided into two limiting cases.[6] In one case, where the dipole moment of the substrate does not reorient upon activation, preferential stabilization of the TS can be achieved because the magnitude of the dipole moment in the TS is larger than that in the GS, which here we call a *scaling effect*. In a second case, where the dipole moment of the substrate reorients but does not change its magnitude upon activation, preferential stabilization can be achieved because the orientation of the dipole moment in the TS is better aligned with the electric field of the enzyme, which we call an *orientational effect*.

Previous work in our lab has demonstrated a significant role of electrostatic stabilization in the catalytic proficiency of the model enzyme ketosteroid isomerase (KSI).[1,2] KSI catalyzes the isomerization of 5-androstenedione to 4-androstenedione via a dienolate intermediate with a rate acceleration (k_{cat}/k_{uncat}) of approximately $10^{7.5}$ (corresponding to a barrier reduction, $\Delta\Delta G^{\ddagger}$, of 10.2 kcal mol^{-1}), where k_{uncat} is the unimolecular rate constant of an "uncatalyzed" reaction that proceeds through the same nominal mechanism.[3,10] An oxyanion hole composed of tyrosine 16 (Tyr^{16}) and aspartic acid 103 (Asp^{103}) in the active site directly interacts with the carbonyl group of the substrate that undergoes a charge rearrangement in KSI's rate limiting step (Figure 1A upper). Tyr16 is further hydrogen-bonded (H-bonded) with tyrosine 57 and tyrosine 32 to form an extended H-bond network (Figure S1A).[11] Using a combination of vibrational Stark effect (VSE) spectroscopy, solvatochromic data and molecular dynamic (MD) simulations, the electric field experienced by the

Received: January 28, 2020
Published: May 7, 2020

© 2020 American Chemical Society

https://dx.doi.org/10.1021/jacs.0c00383
J. Am. Chem. Soc. 2020, 142, 9993−9998

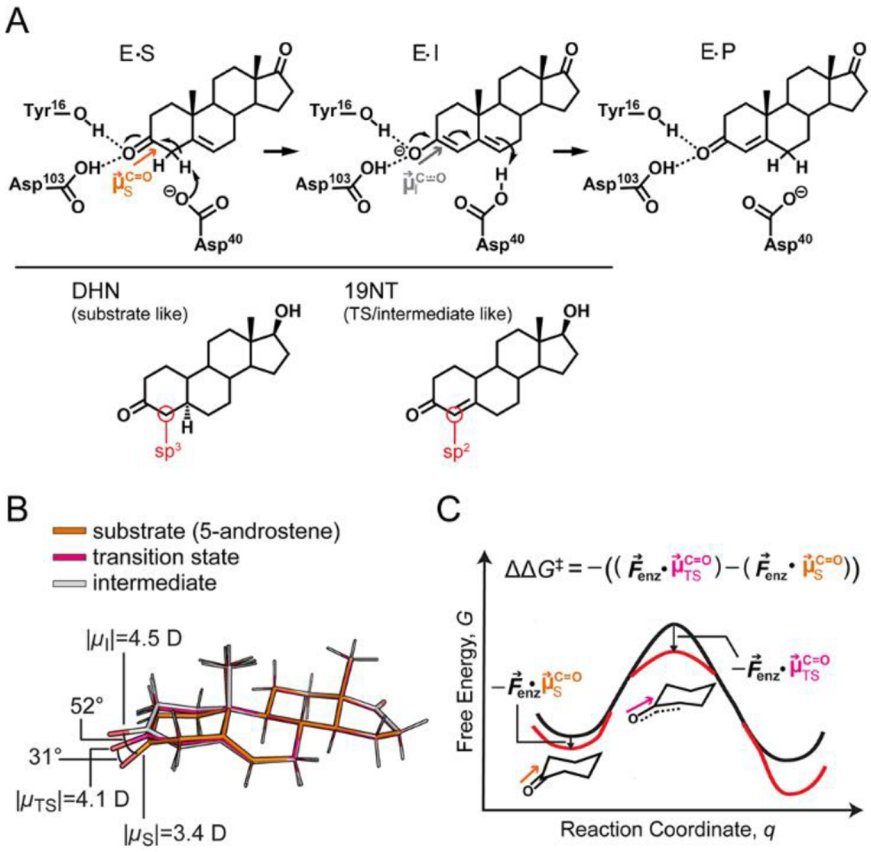


Figure 1. Electrostatic stabilization in the context of ketosteroid isomerase (KSI). (A) KSI catalyzes the isomerization of 5-androstenedione to 4-androtenedione via a dienolate intermediate. The negative charge on the oxygen in the transition state (TS) and intermediate state (I) is stabilized by the oxyanion hole. As the reaction proceeds from the GS to the TS and I, the hybridization of the C4 carbon (red circle) of the substrate changes from sp^3 to sp^2. 5α-dihydronandrolone (DHN) and 19-nortestosterone (19NT) are used to mimic the two states, respectively. (B) *Ab initio* calculations on the substrate, TS, and I of KSI's reaction in the gas phase suggests electrostatic and geometric changes occur along the reaction coordinate. (C) The electric field of KSI, $\vec{F}_{enz}$, preferentially stabilizes the TS because of its larger bond dipole. Furthermore, the angle change shown in part B could result in the C−O dipole in the TS becoming better aligned with $\vec{F}_{enz,TS}$ leading to additional stabilization. Equation 1 encompasses both the scaling effect and orientational effect of electrostatic catalysis.

carbonyl group of a TS/I analog, 19-nortestosterone (19NT), was found to be very large when bound in the KSI active site. Moreover, this active site electric field (as probed by the carbonyl group of 19NT) is linearly correlated with the free energy barrier of the catalyzed reaction, $\Delta G^{\ddagger}$, across wild-type KSI and several KSI mutants with canonical and noncanonical amino acid substitutions (Figure S1B).[1,2]

In previous work modeling the role of electrostatics in catalysis KSI, we employed the simplifying assumption that the orientation of the substrate's C═O does not change significantly during KSI's catalysis and that the (constant) C═O orientation could be modeled with 19NT. In this model, eq 1 simplifies to the following expression

$$\Delta\Delta G^{\ddagger} = -|\vec{F}_{enz}|(|\vec{\mu}_{TS}| - |\vec{\mu}_{R}|) \quad (2)$$

from which one would expect to see a linear correlation between activation barrier and active site electric field. This model assumes that reorientation of the substrate's C═O dipole upon formation of the TS is minimal and electrostatic stabilization in KSI works entirely through the scaling effect. Under this assumption, VSE measurements showed that the carbonyl bond of the substrate increases in magnitude by 1.1 D upon passage to the TS and electrostatic stabilization contributes to 70% of KSI's barrier reduction (Figure S1B).[1]

In the following, we aim to provide a more complete and accurate description of electrostatic catalysis in KSI by including orientational effects. As shown in Figure 1B, in addition to developing a bigger C═O dipole moment, *ab initio* calculations on the isolated, full substrate in the gas phase suggest that the carbonyl bond of the substrate also undergoes a significant angle shift as the reaction proceeds (see SI Methods). Gas-phase optimized structures of the steroid ligand corresponding to the GS, TS, and dienolate intermediate suggest that a C═O dipole reorientation of up to 31° could also be exploited by KSI to maximize the preferential electrostatic interaction with the TS (Figure 1B and 1C). Because geometrical reorientation of the C═O dipole of the TS results mainly from the change in hybridization of the neighboring carbon (C4) from sp^3 to sp^2, we hypothesized that binding 5α-dihydronandrolone (DHN) and 19NT to KSI could mimic the geometry of the enzyme•substrate complex in the GS and the TS/I states, respectively (Figure 1A lower). We obtained crystal structures of the two inhibitors in complex with WT KSI to assess reorientation of the substrate's C═O dipole along the reaction coordinate. Furthermore, the electric fields experienced by the C═O of 19NT and DHN when bound to KSI were calculated with MD simulations from the crystallographic coordinates and estimated experimentally with

https://dx.doi.org/10.1021/jacs.0c00383
J. Am. Chem. Soc. 2020, 142, 9993−9998

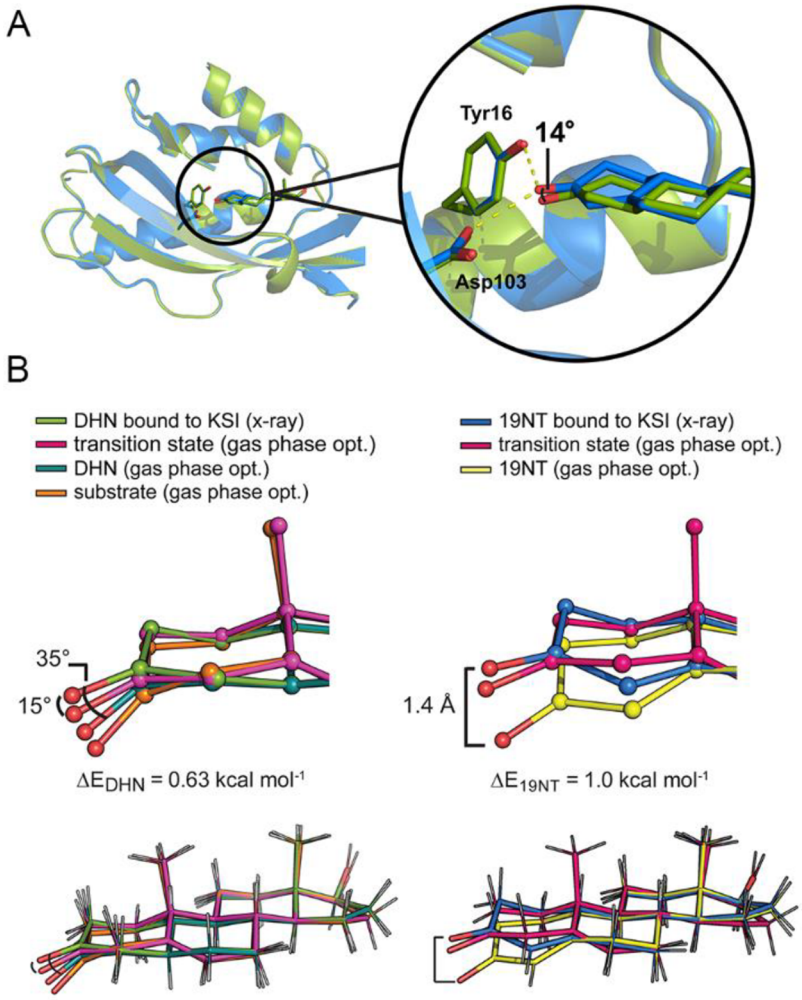


Figure 2. Crystal structures of wild type KSI bound with DHN (6UFS) and with 19NT (5KP4) show perturbed geometry of the ligands. (A) The structure of KSI bound to DHN (green, 1.5 Å) is globally aligned with that of KSI bound to 19NT (blue, 1.7 Å), giving an RMS deviation of 0.177 Å. Within this alignment frame, the angle shift between the carbonyl of DHN and the carbonyl of 19NT is about 14°, smaller than the predicted difference of 31° from the optimized gas-phase geometries (Figure 1B). (B) The C═O bonds of DHN and 19NT get distorted from their gas-phase optimal geometries to more closely resemble the TS structure when bound to KSI. A small energy penalty (ΔE) is calculated from the energy difference between the gas-phase optimal structure of the ligand and the gas-phase structure "constrained" to the crystal coordinates (Figure 3 and Figure S3).

VSE spectroscopy. The difference in the electric fields experienced by the two ligands reflects the extent to which KSI's electric field is aligned with the TS's dipole.

RESULTS AND DISCUSSION

Crystal Structures of DHN and 19NT Bound to Wild Type KSI Indicate Distortion of the Ligand Carbonyl Group toward a TS Geometry. The crystal structures of wild type KSI bound to DHN and bound to 19NT were obtained with a resolution of 1.5 and 1.7 Å, respectively. An overall RMS deviation of 0.177 Å was observed when the two structures are superimposed, suggesting minimal change in the overall enzyme architecture within the error of the structure (Figure 2A and statistics in Table S1). Focusing on the interactions between the ligands and two critical active site residues, Tyr^{16} and Asp^{103}, the structures show that the positions of the active site residues are virtually unaltered between these two states, while the two ligands (most notably, the C═O bond) assume slightly different orientations relative to the protein. The carbonyl bond of DHN is shifted approximately 14° "down" from that of 19NT, smaller than the predicted difference of 31° from the optimized gas phase geometries (Figure 1B). Importantly, when the ligand geometries taken from the crystal structures are compared with the molecules' gas-phase optimal geometries, the carbonyl bonds of both ligands are found to be distorted to more closely resemble the TS geometry (Figure 2B and Figure S2), explaining the smaller angle shift in KSI compared to in the gas phase.

To ensure that the observed geometry distortion is not an artifact from refinement of the structure, we searched the Protein Data Bank for other crystal structures that contain DHN or highly similar molecules as a bound ligand and found two structures: 5α-estran-3,7-dione (ESR) bound to KSI^{D40N} from *Comamonas testosteroni* (an orthologue of the KSI we have studied herein, referred to as *t*KSI, PDB 1OHP) and dihydrotestosterone (DHT) bound to 17β-hydroxysteroid dehydrogenase type 1 (17β-HSD1, PDB 3KLM). We reasoned that if the perturbed geometry of the carbonyl groups in the KSI complexes is functionally relevant, it would be observed in

https://dx.doi.org/10.1021/jacs.0c00383
J. Am. Chem. Soc. 2020, 142, 9993−9998

the former structure but not the latter since 17β-HSD1 catalyzes a chemical conversion on the distal D-ring of the steroid and has no activity for the carbonyl group on the A-ring. As shown in Figure S4A, the carbonyl group of ESR bound to *t*KSI is indeed distorted in a manner highly similar to that of DHN bound to KSI, while the geometry of 17β-HSD1 bound DHT resembles the gas-phase optimized geometry of DHN with mild structural alterations.

These observations are reminiscent of the "Circe effect" proposed by Jencks,[12,13] who suggested that enzymes can destabilize a substrate's reactive region (in this case, the A-ring carbonyl bond) to facilitate its positioning in the enzyme's active site by utilizing favorable distal interactions (in this case, the multiring system) to offset the former's free energy cost.[12] Using density functional theory (DFT), we estimated that the energy penalty for the bond distortion is close to 1.0 kcal mol^{-1} by comparing the energy of the gas-phase optimized geometry of the ligands (hereafter denoted as 'gas phase opt.') to the gas-phase optimized energy of the ligands but in which the 4 dihedral angles of the A-ring are constrained to their values from the crystal structures (ΔE in Figure 2B). These distortion energies are very likely well below the energetic benefit that could be gained from forming an optimal H-bond or electrostatic interaction between the carbonyl and the oxyanion hole, as will be discussed in the following. Therefore, it is not surprising that KSI's large active site electric field partially aligns the carbonyl group of the substrate to maximize stabilization of the complex. The resulting TS-like geometry of the ligands thereby provides evidence that the electric field of KSI can orient the C═O to a TS-like geometry and that this geometric preference is imposed upon a substrate-like ligand.

Computational Modeling on the KSI Complex Reveals a High Degree of Alignment between the Carbonyl Dipole of the Bound Ligand and the Electric Field. To test the hypothesis that the geometric perturbations experienced by steroids in KSI's active site are utilized for electrostatic catalysis, we carried out simulations using these crystal structures to estimate the electric field KSI projects on the C═O bond dipole in each case and additionally calculated the alignment of the local electric field vector to the C═O bond vector with the equation

$$\%\text{aligned} = \vec{F}_{enz}\cdot\hat{u}_{CO}/|\vec{F}_{enz}| \quad (3)$$

(details in SI Methods) where the numerator corresponds to the projection of the enzyme field onto the C═O bond unit vector, and the denominator is the magnitude of the enzyme field vector.[14] The structures of KSI•19NT and KSI•DHN complexes were first energy minimized with the AMBER force field, and then a short simulation was run at low temperature on the energy-minimized complexes to derive the electric field magnitude at the carbonyl of 19NT or DHN and the field's projection along the C═O vector as previously described (denoted as "X-ray" in Table 1).[14] These simulations were intentionally cold and short in order to associate an electric field with the crystallographic configuration and are not intended to reflect a thermally averaged value;[15] nevertheless, we ran simulations for ten 1 fs steps to validate the structures were not unstable.[15] Next, we removed the DHN (or 19NT) coordinates from the X-ray structures and replaced them with the DFT gas-phase optimized DHN (or 19NT) coordinates by alignment (Figure 3 and also see Figure S3 for the full flowchart explaining these manipulations). These fictitious structures, denoted in Table 1 with the heading 'gas-phase opt.'

Table 1. Calculated Active Site Electric Fields and Their Geometry Relative to the Ligand's C═O Bond Vector[a]

	DHN (Substrate like)		19NT (TS like)		TS
coordinates[a]	X-ray (1)	gas phase opt. (2)	X-ray (6)	gas phase opt. (7)	gas phase opt. (5)
Field magnitude[b] $\lvert\vec{F}_{enz}\rvert$ (MV/cm)	−139.5	−105.3	−160.4	−100.9	−151.8
Feld projection[c] $\vec{F}_{enz}\cdot\hat{u}_{CO}$ (MV/cm)	−127.6	−83.5	−141.7	−76.1	−144.6
%aligned[d]	91%	79%	88%	75%	95%
Δstabilization[e] (kcal mol^{-1})	6.6	reference	10.7	reference	/
C═O...O_{16} (Å)	2.53	/	2.57	/	2.70
C═O...O_{103} (Å)	2.76	/	2.65	/	2.48

[a]See Table S2 and Figure S3 for complete simulation scheme and results. The number of each entry corresponds to the respective species in Figure 3 and Figure S3. [b]The electric field magnitude, $\lvert\vec{F}_{enz}\rvert$, is the average of the magnitude of the enzyme's electric field at the C and O atoms of the bound ligand. [c]The electric field projected on the carbonyl is calculated by $\vec{F}_{enz}\cdot\hat{u}_{CO}$, and also equals $\lvert\vec{F}_{enz}\rvert\cdot$ %aligned, where $\lvert\vec{F}_{enz}\rvert$ is the magnitude of the enzyme's field. [d]The %aligned is calculated by eq 3. [e]Δstabilization (the stabilization energy gained from C═O distortion) is the difference of the field projection multiplied by the dipole of the carbonyl (1 MV cm^{-1} D ≃ 0.048 kcal mol^{-1}).

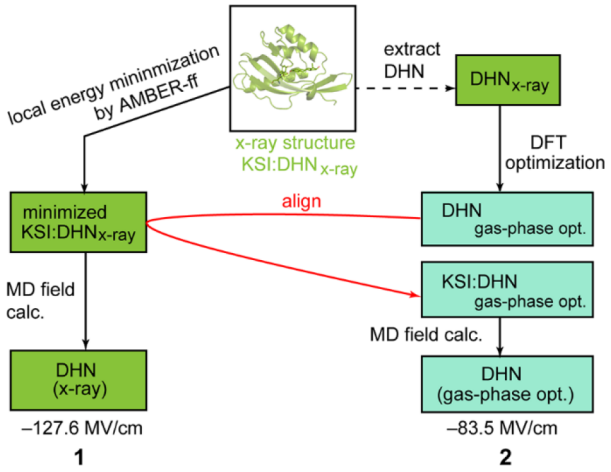


Figure 3. Modeling scheme of different KSI•DHN structures for MD calculations. Each pose is given a number whose corresponding simulation results are listed in Table 1 and Table S2. (See Figure S3 for a more detailed scheme including all calculations).

coordinates, enable us to estimate how unfavorable the gas-phase optimal geometries are in the context of the electrostatic interactions created by KSI's active site.

As shown in Table 1, by dividing the projected field with the electric field magnitude (eq 3), we found that KSI's electric field is approximately 90% aligned with the carbonyl dipole of both DHN and 19NT. The %aligned further increases to 95% when the modeled TS structure is docked into KSI by aligning it to 19NT (Figure 3 and Figure S3). Focusing on how the electric field is experienced by the carbonyl bond itself, the field projection on 19NT's carbonyl bond (−141.7 MV/cm) agrees well with that measured by VSE spectroscopy (−141.3 MV/cm, Table S3),[1] as well as with high-level quantum

https://dx.doi.org/10.1021/jacs.0c00383
J. Am. Chem. Soc. 2020, 142, 9993−9998

simulations (−152 MV/cm),[16] highlighting the principal role of electrostatics in explaining this interaction. The value is also close to the predicted field on the TS's C−O bond (−144.6 MV/cm), demonstrating 19NT to be a faithful probe for KSI's electrostatic environment in the TS because KSI perturbs 19NT's geometry to become very TS-like (Figure 2B and Figure S2). The field projection on DHN's carbonyl bond is 14 MV/cm smaller than that of 19NT (−127.6 MV/cm), which can be explained in part because the O−O distance between the carbonyl of DHN and the hydroxyl group of Asp103 is slightly longer (∼0.1 Å) than its counterpart in the 19NT complex.[16] We note that since these calculated electric fields reflect X-ray structures, they do not represent ensemble-averaged electric fields. Nevertheless, their consistency with a population from MD simulations (ref 15) and experimental rate constants (Discussion S1) suggest that they represent ground state conformations that are most likely the catalytically relevant states.[15]

As seen in Table 1, the electric fields calculated based on the ligands' actual geometries when bound to KSI are considerably larger than those based on the gas-phase optimized geometries (also see Table S2 and Discussion S1). These calculations, though based on fictitious structures, suggest a significant energetic benefit associated with bringing the C═O dipole deep in the active site and aligned with KSI's electric field. We estimate these structural rearrangements strengthen the enzyme's electrostatic interaction with DHN by ∼6 kcal mol^{-1} and with 19NT by ∼11 kcal mol^{-1} − more than enough to compensate for the small energetic cost of ∼1 kcal mol^{-1} from local ligand bond distortion. These results imply that despite what may superficially appear to be a distortion from inspection of the structure, KSI's binding mode of its substrate in fact stabilizes the C═O dipole in both the GS and TS structures, in contradistinction to the Circe effect (Figure S5).

Large, Preoriented Electrostatic Environment in KSI Provides Its Catalytic Power over Simple Solvent. The calculations above using crystallographic structures point to an electrostatic environment in KSI's active site that is preorganized toward the TS geometry. Upon binding to KSI, the substrate-like ligand is driven by a large, oriented electric field to assume a TS-like geometry at the reactive site, so that minimal dipole reorientation needs to occur during the reactive event.

Using the projected fields on the carbonyl of 19NT and DHN (Table 1), we can estimate the partial contribution of the orientational effect to electrostatic catalysis by KSI, described at the outset. The total electrostatic contribution to the barrier reduction estimated using eq 1 is

$$\Delta\Delta G^{\ddagger} = -((\vec{F}_{enz,TS}\cdot\vec{\mu}_{TS}) - (\vec{F}_{enz,R}\cdot\vec{\mu}_{R}))$$

$$= -((142\ \text{MV/cm} \times 4.1\ \text{D}) - (128\ \text{MV/cm} \times 3.4\ \text{D}))$$

$$= -7.1\ \text{kcal mol}^{-1}$$

noting that 1 MV cm^{-1} D is approximately 0.048 kcal mol^{-1}. In contrast, the stabilization energy if only the scaling effect was considered estimated using eq 2 is

$$\Delta\Delta G^{\ddagger} = -|\vec{F}_{enz}|(|\vec{\mu}_{TS}| - |\vec{\mu}_{R}|)$$

$$= -142\ \text{MV/cm} \times (4.1 - 3.4)\ \text{D}$$

$$= -4.8\ \text{kcal mol}^{-1}$$

The angular preference of the active site electric field favoring the TS geometry therefore imparts an additional stabilization energy of 2.3 kcal mol^{-1}, corresponding to 30% of the total electrostatic contribution to the barrier reduction. On the basis of this simple analysis, KSI is able to selectively stabilize the TS significantly more than the GS by exploiting a fairly small shift in C═O orientation. This notion of "geometric discrimination" has been previously discussed,[20] although it is treated here quantitatively within an electrostatic framework. In this context, we might refer to the active site electric field as "preorganized" because it (i) optimally stabilizes the TS geometry and (ii) does not change to accommodate the GS geometry (as a bulk solvent would), but rather forces the GS to assume a more "TS-like" geometry. Note that, the catalytic effect ascribed to electrostatics here is quite similar to the value we estimated previously (7.3 ± 0.4 kcal mol^{-1})[6], though importantly, here, we have obtained the result without any extrapolations (also see Discussion S2).[17]

The preorganization of the electric field in enzymes highlights its fundamental difference from the solvent reaction field in aqueous solutions, which is instead optimized for the GS charge distribution and has to reorient to accommodate the TS as the reaction proceeds, imposing a reorganization energy.[4] The energetic cost of solvent reorganization limits the catalytic capacity of water despite the significant magnitude of its electric field (Figure S5).[4,18]

We sought to further validate this notion experimentally by measuring the electric field projection on the carbonyl of DHN via VSE spectroscopy, under which conditions the enzyme-ligand interaction is fully equilibrated. In order to obtain reliable IR spectra of the carbonyl group in the midst of the strong background from the protein amide band I, it is essential to obtain isotope-edited difference spectra, as was done for 19NT previously.[1] While it is straightforward to prepare the ^{18}O-substituted version of DHN, rapid back-exchange with bulk water, even when the sample was prepared in $D_2{}^{18}O$ (and the protein uniformly ^{13}C labeled to shift the amide I band), made it difficult to obtain definitive data (in contrast, back-exchange for 19NT is very slow). The results of our attempts to perform these measurements are presented in Discussion S3 and Figure S6.

■ CONCLUSIONS

In summary, this work clarifies a number of important aspects about electrostatic catalysis. Perturbed substrate geometry has long been considered a key factor in promoting catalysis;[19−21] however, the belief that such distortions implied a ground-state destabilization contribution could not be easily reconciled with computation and theory.[22] Here, we show that the catalytic effects associated with altered substrate geometry are best interpreted within an electrostatic framework, where they serve to prime a substrate for movement to a transition state, which can nevertheless be stabilizing in the ground state thanks to large active site electric fields. In KSI, this orientational effect provides a significant catalytic contribution, which is impressive in that the geometric changes during this reaction are actually quite small. The benefits of a preorganized electrostatic environment would be expected to be even more profound in cases where a substrate's geometrical reorientation along the reaction coordinate are more dramatic, as is the case in many other enzymatic transformations.

https://dx.doi.org/10.1021/jacs.0c00383
J. Am. Chem. Soc. 2020, 142, 9993−9998

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.0c00383.

Experimental procedures, crystallographic data, complete information on the computational modeling scheme and results, raw IR spectra of DHN, and further evaluation on the energy profile of the isomerization reaction in KSI and in solution (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Steven G. Boxer − *Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States;* orcid.org/0000-0001-9167-4286; Email: sboxer@stanford.edu

Authors

Yufan Wu − *Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States;* orcid.org/0000-0002-4050-9413

Stephen D. Fried − *Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States;* orcid.org/0000-0003-2494-2193

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.0c00383

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported, in part, by NIH Grant GMR35-118044 (to S.G.B.). Use of the Stanford Synchrotron Radiation Lightsource (SSRL), SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (including P41GM103393). S.D.F. was supported by a Junior Research Fellowship from King's College, Cambridge (UK) at the time of this work. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH.

■ REFERENCES

(1) Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme electric fields power catalysis in the active site of ketosteroid isomerase. *Science* **2014**, *346*, 1510−1513.

(2) Wu, Y.; Boxer, S. G. A critical test of the electrostatic contribution to catalysis with noncanonical amino acids in ketosteroid isomerase. *J. Am. Chem. Soc.* **2016**, *138*, 11890−11895.

(3) Radzicka, A.; Wolfenden, R. A proficient enzyme. *Science* **1995**, *267*, 90−93.

(4) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. Electrostatic basis for enzyme catalysis. *Chem. Rev.* **2006**, *106*, 3210−3235.

(5) Hanoian, P.; Liu, C. T.; Hammes-Schiffer, S.; Benkovic, S. Perspectives on electrostatics and conformational motions in enzyme catalysis. *Acc. Chem. Res.* **2015**, *48*, 482−489.

(6) Fried, S. D.; Boxer, S. G. Measuring electric fields and noncovalent interactions using the vibrational Stark effect. *Annu. Rev. Biochem.* **2017**, *86*, 387−415.

(7) Zoi, I.; Antoniou, D.; Schwartz, S. D. Electric fields and fast protein dynamics in enzymes. *J. Phys. Chem. Lett.* **2017**, *8*, 6165−6170.

(8) Herschlag, D.; Natarajan, A. Fundamental challenges in mechanistic enzymology: progress toward understanding the rate enhancements of enzymes. *Biochemistry* **2013**, *52*, 2050−2067.

(9) Menger, F. M.; Nome, F. Interaction vs. preorganization in enzyme catalysis. A dispute that calls for resolution. *ACS Chem. Biol.* **2019**, *14*, 1386−1392.

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

(11) Fafarman, A. T.; Sigala, P. A.; Schwans, J. P.; Fenn, T. D.; Herschlag, D.; Boxer, S. G. Quantitative, directional measurement of electric field heterogeneity in the active site of ketosteroid isomerase. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E299−308.

(12) Jencks, W. P. Binding energy, specificity, and enzymic catalysis: the circe effect. *Anvances Enzymol. Relat. Sub.* **2006**, *43*, 219−410.

(13) Page, M. I.; Jencks, W. P. Entropic contribution to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *68*, 1678−1683.

(14) Fried, S. D.; Bagchi, S.; Boxer, S. G. Measuring electrostatic fields in both hydrogen-bonding and non-hydrogen-bonding environments using carbonyl vibrational probes. *J. Am. Chem. Soc.* **2013**, *135*, 11181−11192.

(15) Welborn, V. V.; Head-gordon, T. Fluctuations of electric fields in the active site of the enzyme ketosteroid isomerase. *J. Am. Chem. Soc.* **2019**, *141*, 12487−12492.

(16) Wang, L.; Fried, S. D.; Markland, T. E. Proton network flexibility enables robustness and large electric fields in the ketosteroid isomerase active site. *J. Phys. Chem. B* **2017**, *121*, 9807−9815.

(17) Natarajan, A.; Yabukarski, F.; Lamba, V.; Schwans, J. P.; Sunden, F.; Herschlag, D. Comment on "Extreme electric fields power catalysis in the active site of ketosteroid isomerase. *Science* **2015**, *349*, 936.

(18) Fried, S. D.; Boxer, S. G. Response to comments on "Extreme electric fields power catalysis in the active site of ketosteroid isomerase". *Science* **2015**, *349*, 936−936.

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

(20) Robertus, J. D.; Kraut, J.; Alden, R. A.; Birktoft, J. J. Subtilisin: a stereochemical mechanism involving transition-state stabilization. *Biochemistry* **1972**, *11*, 4293−4303.

(21) Fastrez, J.; Fersht, A. R. Demonstration of the acyl-enzyme mechanism for the hydrolysis of peptides and anilides by chymotrypsin. *Biochemistry* **1973**, *12*, 2025−2034.

(22) Warshel, A. Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *J. Biol. Chem.* **1998**, *273*, 27035−27038.

Supporting information for:

"A Preorganized Electric field leads to Minimal Geometrical Reorientation in the Catalytic Reaction of Ketosteroid Isomerase"

Yufan Wu, Stephen D. Fried and Steven G. Boxer

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, United States

Department of Chemistry, John Hopkins University, Baltimore, Maryland 21218, United States

Department of Chemistry, Stanford University, Stanford, California 94305-5012, United States

Contents

Materials and Methods.

Materials. 5α-Dihydronandrolone (DHN) was purchased from Steraloids (Newport, RI). Anhydrous solvents and deuterium oxide (99.5% D) were purchased from Acros Organics and Cambridge Isotopes. Sulfur trioxide-pyridine complex was purchased from Sigma-Aldrich. $^{13}C_6$-D-glucose (99%), $H_2{}^{18}O$ (97%) and $D_2{}^{18}O$ (98% D, 97% O) were purchased from Cambridge Isotopes.

Preparation of uniformly ^{13}C labeled KSI. Preparation of wild type KSI was carried out using BL21(AI) cells (Invitrogen) in M9 minimal medium, augmented with 1mM $MgSO_4$, 0.1mM $CaCl_2$, 0.2% $^{13}C_6$-D-glucose and MEM Vitamin Solution. Protein expression was induced at OD~0.8 with 1 mM IPTG and the cells were grown for another 7-9 hours at 37°C before harvest. The protein was purified with Ni-NTA affinity chromatography and anion exchange chromatography (GE Healthcare) as described previously.[1]

Synthesis and purification of 5α-Dihydronandrolone sulfate. 10 mg of 5α-Dihydronandrolone (DHN) was dissolved in 160 µL of dimethylformamide and reacted with 30 mg of sulfur trioxide-pyridine complex. The reaction vial was flushed with nitrogen gas for 3-5 minutes to eliminate unwanted moisture and sealed. The reaction was allowed to run overnight under constant stirring. The progress of the reaction was monitored with thin-layer chromatography (EtOAc : MeOH : H_2O, 7 : 2 : 1, v/v). Due to the small scale of the reaction, the product DHN sulfate was isolated using reverse phase high performance liquid chromatography (RP-HPLC) on a C18 column with water : acetonitrile (15 : 85, buffered with 5mM ammonium acetate at pH = 6.5) as the mobile phase. The retention time of the DHN sulfate was 28 minutes (at 45%B). A small peak corresponding to the unreacted DHN was observed at 38 minutes (75%B). The purity of the isolated DHN sulfate was assessed via ^{1}H-NMR (Fig. S7).

Isotopic labeling of 5α-dihydronandrolone sulfate. To make [^{18}O]DHN sulfate, about 1.0 mg of DHN sulfate was dissolved in 50 µL of $D_2{}^{18}O$. The reaction vial (Eppendorf microcentrifuge tube) was sealed and the reaction was allowed to proceed at room temperature overnight. The labeling reaction of the carbonyl was found to proceed readily without the need of acetic acid as a catalyst. However, the reverse reaction was also found to gradually occur during the course of IR sample preparation and measurements when trace amount of $H_2{}^{16}O$ were present, preventing unambiguous detection and assignment of the DHN carbonyl peak signals, in contrast to earlier work with [^{18}O]19NT.[2]

X-ray crystallography. Crystals of wild type KSI bound with 5α-dihydronandrolone and with 19-nortestosterone were obtained using hanging drop vapor diffusion. 1 μL of 1 mM KSI (in 40mM KPi, pH 7.2), preincubated with 2 mM ligand, was mixed with 1 μL of reservoir solution (1.0-1.2 M ammonium sulfate, 40 mM KPi (pH 7.2), 1-3% isopropanol). The hanging drops were equilibrated for 1-2 days before microseeding to initiate the crystallization. Microseeding was performed by crushing low quality crystals of apo KSI-ClY^{16} variant[1] by vortexing with a glass bead, serially diluting the resulting nanocrystals (5000-10000 times), and using a cat whisker to transfer the microseeds to the equilibrated drops.[3,4] Cryoprotection was achieved by directly soaking crystals in 2.0 M sucrose prior to flash freezing in liquid nitrogen. Single crystal diffraction data were collected at the Stanford Synchrotron Radiation Laboratory BL12-2. Data were integrated and scaled using xia2 for the DHN structure and XDS for the 19NT structure, respectively.[5–7] Data collection and refinement statistics are summarized in Table S1. An initial protein model for each structure was obtained by molecular replacement with Phaser[8] using the coordinates from the previously published structure of apo KSI^{D40N}-ClY^{57} (PDB 5D81). Apo KSI was chosen as the initial model to eliminate potential model bias in the ligand binding geometry. Simulated annealing refinement was carried out using a maximum-likelihood amplitude-based target function as implemented in Phenix[9]. Further refinement was carried out with Phenix, interspersed with manual model building in Coot[10]. All structural figures were prepared using PyMOL[11].

FTIR spectroscopy. FTIR spectra were obtained on a Bruker Vertex 70 spectrometer with a liquid nitrogen-cooled mercury cadmium telluride (MCT) detector using methods very similar to those described previously.[12] The concentration of the DHN sulfate stock was calibrated against known concentrations of trimethylsilylpropanoic acid (TSP-d_4) standard and 19NT sulfate solution (ε_{248} = 14.6 mM^{-1} cm^{-1}) with ^{1}H-NMR. An appropriate amount of each ligand stock was added to the protein aliquot in the same solvent. The final liquid solution contains around 4 mM protein and 3 mM ligand. Spectra of the sample and the reference were acquired following a 10 minutes nitrogen purge and averaging over 512 scans. In our attempt to measure the IR signals using [^{18}O]DHN sulfate bound to KSI as the spectral reference, two equal amount of KSI samples were prepared first in 10 mM KPi buffer (in $H_2{}^{16}O$) and lyophilized. One sample was then dissolved in D_2O while the other in $D_2{}^{18}O$. Ligand stocks of [^{16}O]DHN sulfate or [^{18}O]DHN sulfate were directly prepared at around 50 mM in D_2O or $D_2{}^{18}O$.

Solvatochromism of DHN and DHN sulfate. The solvatochromism model for DHN was constructed using a previous method.[2,13] In brief, the vibrational frequencies of the carbonyl bond of DHN dissolved in various organic solvents and of DHN sulfate in water were measured. Next, the electric field exerted onto the carbonyl bond by each solvation environment was calculated

from MD simulations in GROMACS with the generalized AMBER force field (GAFF)[13] (details in computational methods). By plotting the calculated electric field in the simple solvents (solvent field) against the IR peak positions of DHN, a linear correlation between the absolute electric field and the carbonyl's vibrational frequency was established to be $\bar{\nu}_{C=O}$ = 0.484 $|\vec{F}_{solv}|$+ 1726. Both DHN and DHN sulfate were dissolved in DMSO and their carbonyl vibrational frequencies were found to be identical showing that remote sulfylation of the ligand did not affect the C=O's vibrational frequency. The correspondence from the solvatochromism data was then used to map the measured IR peak(s) of the C=O bond of DHN sulfate in wild type KSI to the electric field the enzyme projects on the C=O bond axis.

Computational methods

1. DFT Calculations. Optimal geometries were obtained with gas-phase DFT calculations at the B3LYP/6-311++G(d,p) level (Table S5) of the isolated steroid molecules (substrate (5AND), DHN (substrate analog), intermediate, 19NT (intermediate/TS analog)). All DFT calculations were carried out in Gaussian 03. The geometry of the TS was determined by Berny optimization starting from 5AND as described previously,[2] and additionally included an acetate molecule to serve as the proton acceptor of 5AND's 4β proton. The constrained geometries of DHN and 19NT were obtained in the same way except 4 dihedral angles of the A-ring of the ligand were fixed to those from their respective crystal structure. The four dihedral angles constrained were the two proper dihedrals (C_1–C_2–C_3–O_3 and C_5–C_4–C_3–O_3) and the two improper dihedrals (C_1–C_2–O_3–C_3 and C_5–C_4–O_3–C_3).

To constrain the dihedral angles of substrate (5AND), the dihedral angles of DHN from the KSI•DHN structure were used; to constrain the dihedral angles of the intermediate, the dihedral angles of 19NT from the KSI•19NT structure was used. The energy costs for the carbonyl angle shift observed in the X-ray structures of 19NT and DHN (Fig. 2) were estimated as the energy difference between the optimal geometry and the constrained geometry. Bond dipole moments were calculated as described previously (Table S5).[2]

2. Parameterization. Initial models of the ligands (5AND (substrate), DHN (substrate analog), 19NT (intermediate/TS analog), and intermediate) were modeled in Gaussview and geometry optimized by DFT at the B3LYP/6-311++G(d,p) level in Gaussian 03. The geometry of the TS was determined by Berny optimization starting from 5AND as described previously,[2] and additionally included an acetate molecule to serve as the proton acceptor of 5AND's 4β proton. The resulting coordinates were put into the Antechamber program of AmberTools12, which parameterizes the ligands using the GAFF force field and assigns atomic charges using the AM1-BCC method. For the organic solvent molecules, we used the GAFF parameters of Caleman et al. [14], available at www.virtualchemistry.org. Water was modeled using the TIP3P model.

3. MD simulations in solvents. A cubic box of edge length 5.2 nm was specific around the DHN solute, and filled with solvent molecules using the GENBOX utility in GROMACS and starting coordinates from Caleman et al.[14] The solvent boxes were energy minimized with 1000 steps of steepest descent, then equilibrated for 100 ps (2 fs time step) in an NPT ensemble with a reference temperature of 300 K and a reference pressure of 1 bar. In all cases, periodic boundary conditions were applied to the solvent box, and long-range electrostatics were approximated with the particle mesh Ewald (PME) method, using 1.0 nm as a cut-off. Lennard-Jones interactions were also cut off at distances exceeding 1.0 nm. All bond vibrations were constrained using the LINCS algorithm. During equilibration, the Bussi thermostat and Berendsen barostat were active. Production dynamics were carried out for 2 ns in the NPT ensemble, continuing from the final coordinates and velocities of the equilibration run. Temperature-coupling was regulated using a stochastic dynamic integrator and the Parrinello-Rahman barostat was applied. Snapshots consisting of full-precision coordinates and forces were outputted every 200 fs, and used as the basis for electric field calculations. A charge-neutralized topology file was generated for the solvation system in which the partial charges for all of the solvent atoms are set to zero (but all of the solute's atoms retain the same charges as used during dynamics) and all noncharged parameters are kept identical.

4. Electric field calculations. First, the total force on the C-atom and on the O-atom of the carbonyl/C=O of the ligand was extracted for each snapshot in the production run. Then the trajectory from the production dynamics was post-processed with the charge-neutralized topology using GROMACS' rerun utility. In the resulting trajectory, different forces are present on each atom, due to the absence of any intermolecular electrostatic interaction. Likewise, the total force on the C and O atoms was extracted from each snapshot. With this information, the total electric field experienced by the C=O group due to the environment (solvation or protein) was calculated using equations 1-3.

$$\vec{f}^{\,i}_{\text{electro}} = \vec{f}^{\,i}_{\text{tot}} - \vec{f}^{\,i}_{\text{nonelectro}} \quad (1)$$

$$\vec{F}^{\,i}_{\text{enz}} = \vec{f}^{\,i}_{\text{electro}}/q_i \quad (2)$$

$$\text{field projection} = \vec{F}_{\text{enz}} \cdot \hat{u}_{CO} = \frac{1}{2}\left(\vec{F}^{C} \cdot \hat{u}_{CO} + \vec{F}^{O} \cdot \hat{u}_{CO}\right) \quad (3)$$

In equations 1-3, i is indexed over the C-atom and the O-atom of the carbonyl/C=O group, $\vec{f}$ denotes force, and $\vec{F}$ denotes electric field. The subtraction of all non-electrostatic forces (as determined by re-running the trajectory with the charge-neutralized topology) from the total force results in a force exerted on a particular atom due only to electrostatic interactions (Eq. 1). The electrostatic force can be converted into an electrostatic field (Eq. 2) simply by dividing by the partial charge of the atom in question (either C or O of the ligand). Finally, the electric field projected onto the C=O dipole of the ligand, $\vec{F}_{\text{enz}} \cdot \hat{u}_{CO}$, is calculated by projecting the field at either

C or O onto the unit vector defining the vibration's bond axis (i.e. along the C=O bond), and then averaging the two field projections between the two atoms (Eq. 3). For solvatochromism characterization, the electric field experienced by the vibration is averaged over all snapshots to obtain the ensemble-averaged electric field, $\langle \vec{F}_{\text{solv}} \cdot \hat{u}_{CO} \rangle$. Other statistical measures of the field distribution, such as standard deviations are calculated as well. To calculate the magnitude of an electric field at a given location (rather than its projection), we used equation 4:

$$\text{magnitude} = \left|\vec{F}_{\text{enz}}\right| = \frac{1}{2}\left(\left|\vec{F}^{C}\right| + \left|\vec{F}^{O}\right|\right) \quad (4)$$

The %aligned between KSI's intrinsic electric field and the dipole of the reactive C=O group is then determined by dividing $\vec{F}_{\text{enz}} \cdot \hat{u}_{CO}$ by $\left|\vec{F}_{\text{enz}}\right|$.

5. Cold MD simulations in KSI (see Fig. S3 for visualization of workflow). All simulations employed the AMBER99SB-ILDN force field to describe the protein, the TIP3P water model to describe crystallographic water molecules, and DHN, substrate, 19NT, and intermediate were described using parameters generated from Antechamber in AmberTools12 (described above). The X-ray coordinates of the KSI•19NT (calculation **6**) and KSI•DHN (calculation **1**) were used as the starting structures. The PDB2GMX utility with its default options was employed to protonate the structures, assign disulfide linkages, and assign protonation states to ionizable moieties. Importantly, Asp40 was manually rendered deprotonated (in contrast with PDB2GMX's defaults). In the gas phase, the enzyme•ligand complex was energy minimized with 1000 steps of steepest descent. The system was not equilibrated.

Next, short molecular dynamics simulations were performed with 10 steps (1 fs time step) in the gas phase, with no periodic boundary conditions, in an NVE ensemble propagated using the velocity Verlet integrator. These simulations were carried out at a temperature of 1 K. All Coulombic and all van der Waals interactions were explicitly calculated (i.e., no cut-offs were used, and the following settings were used: pbc=no; nstlist=0; ns-type=simple). As described in sections 2 and 3, full-precision snapshots with coordinates and forces at the C-atom and O-atom of the ligand's C=O were recorded at each step, and were used to calculate electric field projections and electric field magnitudes. To create the charge-neutralized topology file for these systems, the charges on all the water molecules and enzyme residues were set to zero (but all of the ligand's atoms retained the same charges).

Calculations **2** and **7** were carried out identically to **1** and **6**, except that the ligand (either DHN or 19NT) was first extracted from the x-ray structure, geometry optimized by DFT at the B3LYP/6-311++G(d,p) level, and aligned back to DHN or 19NT in Pymol. These were then used for electric field calculations as described.

In calculations **3** and **8**, we first optimized the structure of substrate (5AND) and intermediate by DFT at the B3LYP/6-311++G(d,p) level, aligned them into the coordinates of DHN or 19NT

in their respective x-ray structures, and replaced the coordinates of DHN (or 19NT) with those of 5AND (or intermediate). These were then used for electric field calculations as described.

Calculations **4** and **9** were conducted identically to calculations **3** and **8**, except that instead of using the optimized geometries of substrate and intermediate, we used optimized geometries subject to the constraint that they mimic four A-ring dihedral angles of DHN and 19NT from their respective x-ray structures. These calculations represent an estimate of what conformations the substrate and the intermediate adopt when bound to the enzyme. The four dihedral angles constrained were the two proper dihedrals (C_1–C_2–C_3–O_3 and C_5–C_4–C_3–O_3) and the two improper dihedrals (C_1–C_2–O_3–C_3 and C_5–C_4–O_3–C_3).

Calculation **5** was conducted as following. Using 5AND (along with an acetate molecule) as a starting point, we used Berny optimization to identify a saddle-point structure. This structure was aligned against the coordinates of 19NT and Asp40 in the KSI•19NT structure. These coordinates were used to replace those of 19NT. The resulting structure was then used for electric field calculations as described.

Supplementary discussions

1. Further evaluation of the modelling scheme of KSI complexes in MD simulations

Apart from DHN, 19NT and the TS, we also performed MD simulations on the KSI•substrate complex and the KSI•intermediate complex. As we cannot obtain direct structural data of KSI•substrate and KSI•intermediate (they are turned over on the sub-ms timescale), we have modeled the structure of KSI•substrate by optimizing the substrate structure with a gas-phase DFT calculation, but subject to the *constraint* that the A ring adopts the perturbed geometry observed in the KSI•DHN x-ray structure, and then replacing the DHN coordinates with those of the substrate by aligning them to DHN coordinates (see Fig. S3). This is referred to as the 'constrained' geometry in Table S2, and corresponds to our best estimate of what the substrate geometry looks like when bound to KSI. Likewise, for the intermediate, we carried out gas-phase DFT calculations subject to the constraint that the intermediate adopts the perturbed geometry observed in the KSI•19NT x-ray structure, and then replaced the 19NT coordinates with those of the intermediate by aligning them to 19NT coordinates (Fig. S3). As shown in Table S2, when the substrate adopts the same perturbed geometry as in DHN ('constrained'), the carbonyl bond is better aligned with the enzyme's electric field vector, resulting in a much larger projection. This result is very similar to that of DHN itself, for which actual x-ray coordinates (rather than coordinates from a constrained optimization) were used. In contrast, when the intermediate is constrained to adopt the same perturbed geometry as in 19NT, the electric field it experiences is virtually the same as if its gas-phase optimized structure were used. These results further point to the notion that the electric field in KSI is pre-oriented towards the TS/intermediate geometry. Upon substrate binding, this large pre-oriented electric field distorts the carbonyl bond of the substrate to position it towards a TS-like geometry. However, the resulting rearrangement is expected to still result in overall stabilization because of the improved alignment with the electric field.

To further validate our model, we have compared our calculated stabilization energy of the substrate to KSI to its empirical binding constant (Fig. S5). Fig. S5B shows several critical free energy changes based on several experiments: binding of substrate to enzyme (E + S ⇌ E•S, based on experimental K_M (ref. 12)), barrier crossing of substrate in enzyme (E•S → E•TS, based on experimental k_{cat} (ref. 12)), and barrier crossing of substrate in solution (S → TS, based on k_{uncat} with a 1 M reference concentration (ref. 25)). In Fig. S5A, we estimate the solvation energy of the substrate (–12.2 kcal mol^{-1}) by employing the substrate's calculated dipole (3.4 D, Table S5) and the solvent reaction field of water on DHN (–73.6 MV/cm, Tables S3 & S4). We estimate the electrostatic binding energy of substrate to KSI (relative to the gas phase) in several ways (see Fig. S5 legend) but the most realistic is using the substrate's calculated dipole (3.4 D) and the electric field calculated from the modeled structure of KSI•substrate (–112.3 MV/cm, Table S2); this

furnishes a binding energy of –18.3 kcal mol^{-1} (Table S2). The difference between these values, corresponding to a binding free energy of substrate from water to enzyme is –6.1 kcal mol^{-1} (red dashed line), which is very close to the experimental binding energy (of –5.6 kcal mol^{-1}) and even closer if one subtracts the putative substrate distortion energy to accommodate to the enzyme electric field (–6.1 + 0.63 = 5.5 kcal mol^{-1}). These calculations can be taken to support the view that the primary interactions governing substrate binding to KSI is the strong electrostatic interaction at C=O and distortion.

We further tested this scheme by estimating the transition state stabilization in solution. To do this, we invoke Marcus theory which assumes that at the transition state, the solvent reaction field is still equilibrated with the *reactant's* dipole orientation, which is 31˚ out of alignment with the TS (Fig. 1B). The stabilization energy of transition state in KSI is estimated using the TS's calculated dipole (4.1 D, Table S5) and the electric field calculated from the modeled structure of KSI•TS; this furnishes a transition state stabilization energy of –28.5 kcal mol^{-1} (Table S2). Comparing these numbers provides an estimate for the binding free energy of transition state to KSI ($\Delta G^{\circ}_{\text{bind,TS}}$ of –15.9 kcal mol^{-1}, red dashed line). This is surprisingly close to the experimentally-determined value for $\Delta G^{\circ}_{\text{bind,TS}}$, found by closing the thermodynamic cycle created by $\Delta G^{\circ}_{\text{bind,S}}$, $\Delta G^{\ddagger}_{\text{cat}}$, and $\Delta G^{\ddagger}_{\text{uncat}}$ (Fig. S5A, black dashed line), of 15.8 kcal mol^{-1}. In summary, the electric field and stabilization energies determined from our simple calculations are remarkably consistent with experiments.

2. Further discussions on the estimation of orientational effect

In order to quantify the total preferential stabilization of the TS over the GS from KSI's electric field, we employed equation (1):

$$\Delta\Delta G^{\ddagger} = -\left(\left(\vec{F}_{\text{enz,TS}} \cdot \vec{\mu}_{\text{TS}}\right) - \left(\vec{F}_{\text{enz,R}} \cdot \vec{\mu}_{\text{R}}\right)\right) \quad (1)$$

We reason that the binding modes of 19NT and DHN to KSI mimic that of the substrate and the TS based on their C_4 hybridizations, therefore, equation (1) is converted to:

$$= -(\left(\left(\vec{F}_{\text{enz}}^{\text{19NT}} \cdot \hat{\mu}_{\text{C=O}}^{\text{19NT}}\right)\left|\vec{\mu}_{\text{C=O}}^{\text{TS}}\right|\right) - \left(\left(\vec{F}_{\text{enz}}^{\text{DHN}} \cdot \hat{\mu}_{\text{C=O}}^{\text{DHN}}\right)\left|\vec{\mu}_{\text{C=O}}^{\text{5AND}}\right|\right))$$
$$= -(\left((142\ \text{MV/cm}) \times 4.1\ \text{D}\right) - \left((128\ \text{MV/cm}) \times 3.4\ \text{D}\right))$$
$$= -7.1\ \text{kcal mol}^{-1}$$

where $\left(\vec{F}_{\text{enz}}^{\text{19NT}} \cdot \hat{\mu}_{\text{C=O}}^{\text{19NT}}\right)$ is the projection of KSI's electric field along the direction of 19NT's carbonyl bond and is determined by calculating the electric field of the enzyme from MD simulations that started with the KSI•19NT crystal structure coordinates. Likewise, $\left(\vec{F}_{\text{enz}}^{\text{DHN}} \cdot \hat{\mu}_{\text{C=O}}^{\text{DHN}}\right)$ is the projection of KSI's electric field along the direction of DHN's carbonyl bond and is determined by calculating the electric field of the enzyme from MD simulations that started with

the KSI•DHN crystal structure coordinates (Table S2). Alternatively, KSI's projected field could also be derived from the vibrational frequency of the carbonyl bond via VSE spectroscopy, where the field projection on the carbonyl vector is intrinsic in the experimental measurement (Table S3, Fig. S5). $\left|\vec{\mu}_{C=O}^{TS}\right|$ and $\left|\vec{\mu}_{C=O}^{5AND}\right|$ are the magnitudes of the C=O dipole of the two ligands and are calculated from the gas-phase optimized structures of the TS and the substrate from *ab initio* calculations (Fig. 1B).

To quantify the preferential stabilization of the TS over the GS from a pure scaling effect using our previous model[2,12], the projection of KSI's electric field on the reactive carbonyl bond of the ligand is held constant while the magnitude of the bond dipole changes:

$$\begin{aligned}\Delta\Delta G^{\ddagger} &= -\left(\left(\vec{F}_{\mathrm{enz,TS}}\cdot\vec{\mu}_{\mathrm{TS}}\right)-\left(\vec{F}_{\mathrm{enz,R}}\cdot\vec{\mu}_{\mathrm{R}}\right)\right)\\ &= -(\vec{F}_{\mathrm{enz}}\cdot\hat{\mu}_{\mathrm{C=O}})(|\vec{\mu}_{\mathrm{TS}}|-|\vec{\mu}_{\mathrm{R}}|)\\ &= -|F_{\mathrm{enz}}|(|\vec{\mu}_{\mathrm{TS}}|-|\vec{\mu}_{\mathrm{R}}|)\\ &= -142\ \mathrm{MV/cm}\times(4.1-3.4)\ \mathrm{D}\\ &= -4.8\ \mathrm{kcal\ mol^{-1}}\end{aligned} \qquad (2)$$

Therefore, the inclusion of the orientational effect in our description of electrostatic catalysis in KSI adds an additional 2.3 kcal mol^{-1} of stabilization in the total barrier reduction from electrostatics.

We note that in our previous model, we derived a value of $(|\vec{\mu}_{\mathrm{TS}}|-|\vec{\mu}_{\mathrm{R}}|)$ for the C=O dipole of 1.1 D by inverting equation (2) using experimental values for $\Delta G^{\ddagger}$ and $\left|F_{\mathrm{enz}}\right|$ (Fig. S1B).[2,12] This value is 0.4 D larger than the predicted value from *ab initio* calculations. The discrepancy can be explained in that by fitting rate data to equation (2), when equation (1) is the more correct model, the orientational effect became packaged into $(|\vec{\mu}_{\mathrm{TS}}|-|\vec{\mu}_{\mathrm{R}}|)$, resulting in a larger value. On the other hand, the *ab initio* values for $|\vec{\mu}_{\mathrm{TS}}|$ and $|\vec{\mu}_{\mathrm{R}}|$ used in this study can be seen as lower limits, as these dipoles would be expected to be larger in the enzyme active site due to polarization.

3. Vibrational Stark Effect Spectroscopy studies on DHN and KSI•DHN

We first calibrated the sensitivity of DHN's C=O vibrational frequency to electric field. Analogous to previous studies on 19NT, the Stark tuning rate of the carbonyl group of DHN is calibrated using solvatochromism.[2,12] Due to the limited solubility of DHN in water, its sulfate version (DHN sulfate) was synthesized and used to determine the carbonyl's vibrational frequency in water. As shown in Fig. S6A, a systematic shift to the red is observed for DHN's carbonyl frequency in solvents with increasing polarity. The average value of the solvent field experienced by the carbonyl is calculated from MD simulations following a previous protocol (Table S4).[13] When the average value of the solvent field is plotted against DHN's carbonyl frequency, a linear

correlation is observed giving a calibration curve of $\bar{\nu}_{C=O}$ (cm^{-1}) = 1726.5 + 0.484 × $\vec{F}_{solv}$ (MV/cm), where 0.484 cm^{-1}/ (MV/cm) is the Stark tuning rate ($|\Delta\vec{\mu}_{DHN}^{C=O}|$). The sensitivity of DHN's carbonyl frequency to the electric field is smaller than that of 19NT, agreeing with previous findings that conjugated carbonyls tend to exhibit larger Stark tuning rates.[15]

To determine the electric field experienced by the carbonyl of DHN when bound to KSI, we prepared the ^{13}C form of the protein in order to shift the protein's amide I region away from the absorption region of DHN. As shown in Fig. S6C, three peaks were observed at 1660 cm^{-1}, 1681 cm^{-1}, and 1710 cm^{-1} when DHN was bound to wild-type [^{13}C]KSI. The electric fields corresponding to these peaks are -136.8 MV/cm, -93.2 MV/cm, and -34.1 MV/cm respectively. Provisionally, we have assigned the -136.8 MV/cm peak to a KSI:DHN complex that adopts a productive conformation for catalysis that resembles what we see in x-ray crystallography, because it agrees reasonably well with the calculated field based on the crystal structure (-127.6 MV/cm). The -93.2 MV/cm could correspond to a state in which DHN is not fully H-bonded to the active site, as has been seen in extensive MD simulations carried out by Welborn et al.[16] We have provisionally assigned the '-34.1 MV/cm' peak to carboxylate side-chains on KSI.[17] We attempted to confirm these assignments with isotope-edited FTIR; unfortunately, as discussed in the SI methods and shown in Fig. S6C, the oxygen atom of the carbonyl in ^{18}O-labeled DHN rapidly exchanges with solvent. Without this validation, it remains possible that these provisional assignments are not correct, but since all these derived electric fields are smaller than -141 MV/cm, we can at least conclude from these IR spectra that DHN experiences a smaller electric field than 19NT when bound to KSI.

The observation of the multiple populations hints at the possibility of flexibility of the KSI complex in the ground state, where it is able to sample the catalytically active conformation for a fraction of the time.[16] It is also possible that alternative conformations were present in our original characterization of the KSI:19NT complexes,[2] but their spectral features were masked by the strong amide signals from the protein.

The possibility of several conformations of the enzyme-substrate complex is consistent with the notion that enzyme-substrate complexes must explore their energy landscapes to locate a catalytically active conformation.[18] In the present example, this search likely corresponds to subtle motions of the substrate sampling various poses until it finds the TS-like pose (i.e., the -136.8 MV/cm feature, which is probably very similar to the pose we captured by crystallography). Whereas this conformation would be sparingly sampled in solution, in the active site, it is stabilized by the preorganized field. In other enzymes, the search for the catalytically active state may take more time because it involves more dramatic conformational changes of the substrate and enzyme (as documented extensively by Richard and co-workers).[19,20]

Supplementary Tables

Table S1 | X-ray diffraction data collection and refinement statistics

	WT-KSI bound with dihydronandrolone (PDB# 6UFS)	WT-KSI bound with 19-nortestosterone (PDB# 5KP4)
Resolution range (Å)	36.74-1.47	36.29-1.71
Space group	$P2_12_12_1$	$P2_12_12_1$
a, Å	35.66	33.85
b, Å	73.48	72.59
c, Å	95.03	94.82
α,°	90.00	90.00
β,°	90.00	90.00
γ,°	90.00	90.00
No. unique reflections	43215	26203
Completeness, %	99.6	99.8
Multiplicity	6.3	6.6
R_{merge}, %	9.5	9.6
$I/\sigma_{overall}$	8.9(2.3)	9.6(2.1)
Refinement statistics		
No. residues	256	254
No. waters	376	138
Rwork, %	18.8	20.3
Rfree, %	22.1	23.9
rmsd bond, Å	0.006	0.006
rmsd angle, °	0.86	0.85

Table S2 | KSI active site electric fields from structural analysis and MD calculations (see Figure S3)

	coordinates	field projection[a] (MV/cm)	% aligned	stabilization[b] (kcal mol^{-1})	Δstabilization (kcal mol^{-1})	C=O…O_{16} (Å)	C=O…O_{103} (Å)	C=O angle shift
DHN	x-ray **(1)**	-127.6	91%	-20.8	7.2	2.53	2.76	reference
	gas-phase opt. **(2)**	-83.5	79%	-13.6	reference	3.18	2.79	~35°
substrate	constrained **(4)**	-112.3	88%	**-18.3**	8.6	3.10	2.60	reference
	gas-phase opt. **(3)**	-59.3	70%	-9.7	reference	3.75	2.82	~31°
19NT	x-ray **(6)**	-141.7	88%	-27.9	12.9	2.57	2.65	reference
	gas-phase opt.[c] **(7)**	-76.1	75%	-15.0	reference	3.48	2.62	~55°
inter-mediate	constrained **(9)**	-144.2	96%	-31.2	0.1	2.57	2.51	reference
	gas-phase opt. **(8)**	-144.0	92%	-31.1	reference	2.49	2.53	~8°
TS	gas-phase opt. **(5)**	-144.6	95%	**-28.5**	/	2.70	2.48	/

[a] The electric field projection is determined by using the coordinates of the ligand in the KSI active site from x-ray crystallography (or by replacing them with geometry-optimized structures, as explained in text) in short MD simulations.

[b] Stabilization energies are calculated by combining electric field projections and the C=O bond dipole magnitude (see Table S5). The bond dipole magnitudes of the substrate and the transition state are used for DHN and 19NT, respectively.

[c] We note that upon superimposing the gas-phase optimized 19NT structure onto the crystal structure, the carbonyl group was translated by ~1.4 Å (Fig. 2B). The big translation was partly due to the artifact of attempting to align the rest of the ligand's multi-ring system. It is very likely that in KSI the positioning of the A ring relative to the oxyanion hole is the main determinant of the ligand's binding mode. Therefore, we only use the alignment against the entire 19NT structure here to qualitatively demonstrate how the observed geometrical perturbation of the carbonyl leads to better alignment and in turn a larger projected field.

Table S3 | Electric fields of ligands bound to KSI as determined by Vibrational Stark Effects

	DHN (C=^{16}O) model for substrate			19NT (C=^{16}O) model for transition state		
Solvation environment	KSI		D_2O	KSI	D_2O	D_2O (in DHN's solvent sphere [d])
Peak Position $\bar{\nu}$ (cm^{-1})	1660.3 [a]	1681.4 [a]	1690.5	1590.0	1634.0	/
FWHM (cm^{-1})	9.0	10.0	25.1	3.8	40.6	/
field projection (MV/cm) [b]	-136.8	- 93.2	-74.6	-141.3	-80.4	-74.6 cos(31°)
Stabilization [c] (kcal mol^{-1})	**-22.3**	-15.2	**-12.2**	**-27.8**	-15.8	**-12.6**

[a] Both IR peaks from VSE spectroscopy are provisionally assigned to DHN – see Fig. S6.

[b] The electric field projection as determined from VSE spectroscopy (and using a solvent-based calibration curve)

[c] Stabilization energies are calculated by combining electric field projections and the C=O bond dipole magnitude (see Table S5)[d]. In rapid chemical reactions that proceed in solvent, the TS would form initially in the solvent sphere that was equilibrated with the substrate. Since the carbonyl reorients by 31° in TS, the solvent field would be out of alignment with the C=O dipole, and the projected field would be lower by a factor of cos(31°).

Table S4 | Solvatochromism data of DHN's carbonyl

Solvent	Peak Position $\bar{\nu}$ (cm^{-1})	FWHM (cm^{-1})	Electric Field Mean $\langle \vec{F}_{solv} \cdot \hat{u}_{CO} \rangle$ (MV/cm)	Electric Field Standard Deviation (MV/cm)
Hexane	1726.2	12.9	-0.01 ± 0.03	0.78
Dibutylether	1722.4	13.6	-8.85 ± 0.77	5.4
Tetrahydrofuran	1717.5	13.2	-19.6 ± 0.35	7.8
Acetonitrile	1712.4	14.5	-29.9 ± 0.86	12.9
Dichloromethane	1711.6	12.5	-31.2 ± 0.64	17.0
Chloroform	1709.4	18.2	-33.7 ± 1.2	15.1
Methyl sulfoxide	1709.3	15.9	-35.3 ± 0.45	10.9
Water	1690.5	25.1	-74.6 ± 0.58	30.0

Table S5 | Optimal geometries of ligand structures from DFT

	O charge (a.u)	C charge (a.u)	C–O (Å)	C-O dipole (D)
Substrate	-0.550	0.627	1.21	3.4
Transition State	-0.717	0.663	1.24	4.1
Intermediate	-0.802	0.685	1.26	4.5

Supplementary Figures

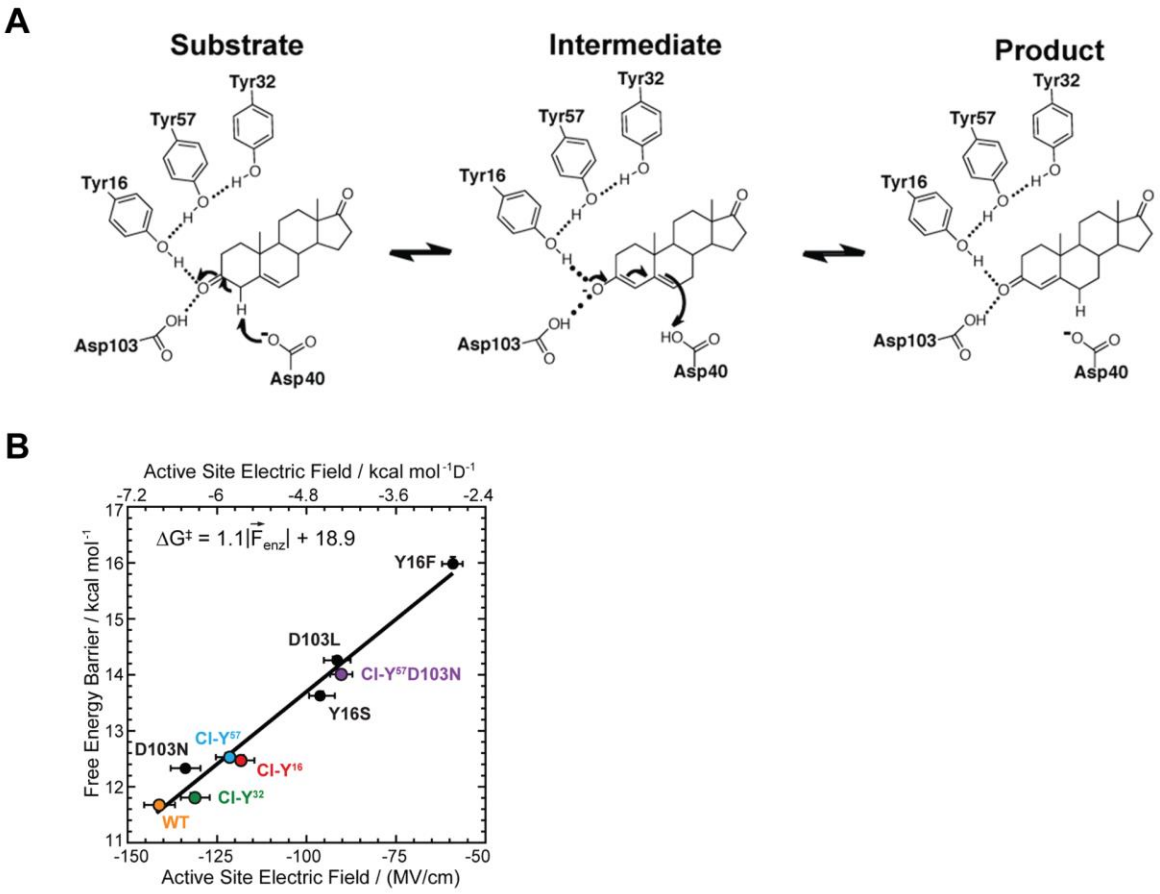


Figure S1 | Linear correlation between KSI's electric field and rate acceleration

(A) An extended H-bond network composed of the tyrosine triad (Tyr[16], Tyr[57] and Tyr[32]) and Asp[103] interacts with the carbonyl group of the substrate, stabilizing the accumulated negative charge in the TS/Intermediate state.[21] (B) The H-bond interactions enable KSI to exert a large electric field on the carbonyl. The magnitude of the electric field is linearly correlated with the free energy barrier of the catalyzed reaction in wild-type and mutants with canonical and noncanonical amino acid substitutions ($\Delta G^{\ddagger}$ was obtained from k_{cat} using transition state theory). From the slope of the linear correlation and fitting to Eq. 2, $(|\vec{\mu}_{TS}| - |\vec{\mu}_{R}|)$ was estimated to be 1.1 D. By extrapolating to zero field, the electrostatic contribution to $\Delta\Delta G^{\ddagger}$ was estimated to be 7.3 kcal mol^{-1}, contributing 70% of KSI's total barrier reduction relative to an uncatalyzed reference reaction in solution at a reference concentration of 1 M.[2,12]

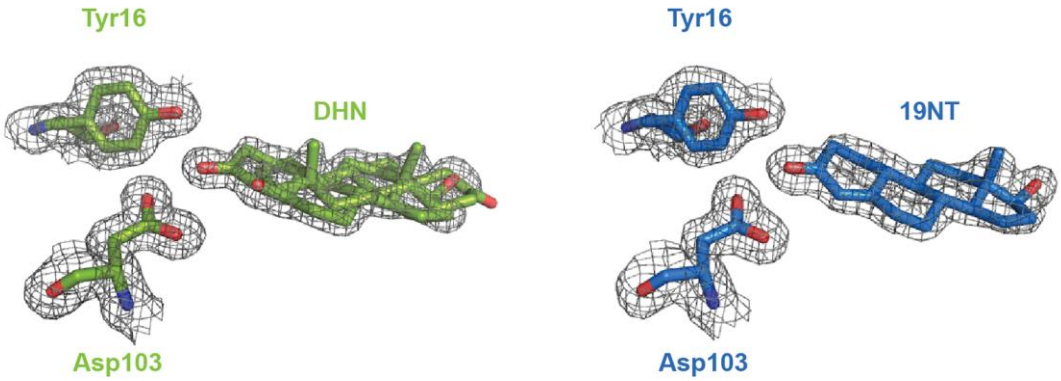


Figure S2 | Electron density maps (contoured at 0.8 σ) of the oxyanion hole residues and the bound ligand showing their relative positioning

The active site of wild type KSI bound to DHN (green) and wild-type KSI bound to 19NT (blue). The electron density map for the KSI:DHN structure shows density for the DHN ligand bound both in a forward conformation with the steroid A-ring in the oxyanion hole (65% occupancy) and a backward conformation with the D-ring in the oxyanion hole (35% occupancy). The electron density map for the KSI:19NT structure shows the forward conformation with the A-ring in the catalytic pocket.

Previous structural studies of *p*KSI and *t*KSI bound to product and transition state analog similarly observed the copresence of the two conformations, and determined that the forward orientation is the productive conformation in solution.[22,23] Therefore, we performed our analysis on the forward conformation. In both structures, the electron density around the carbonyl bond is sufficient to support a TS-like geometry modeled from the gas phase calculations.

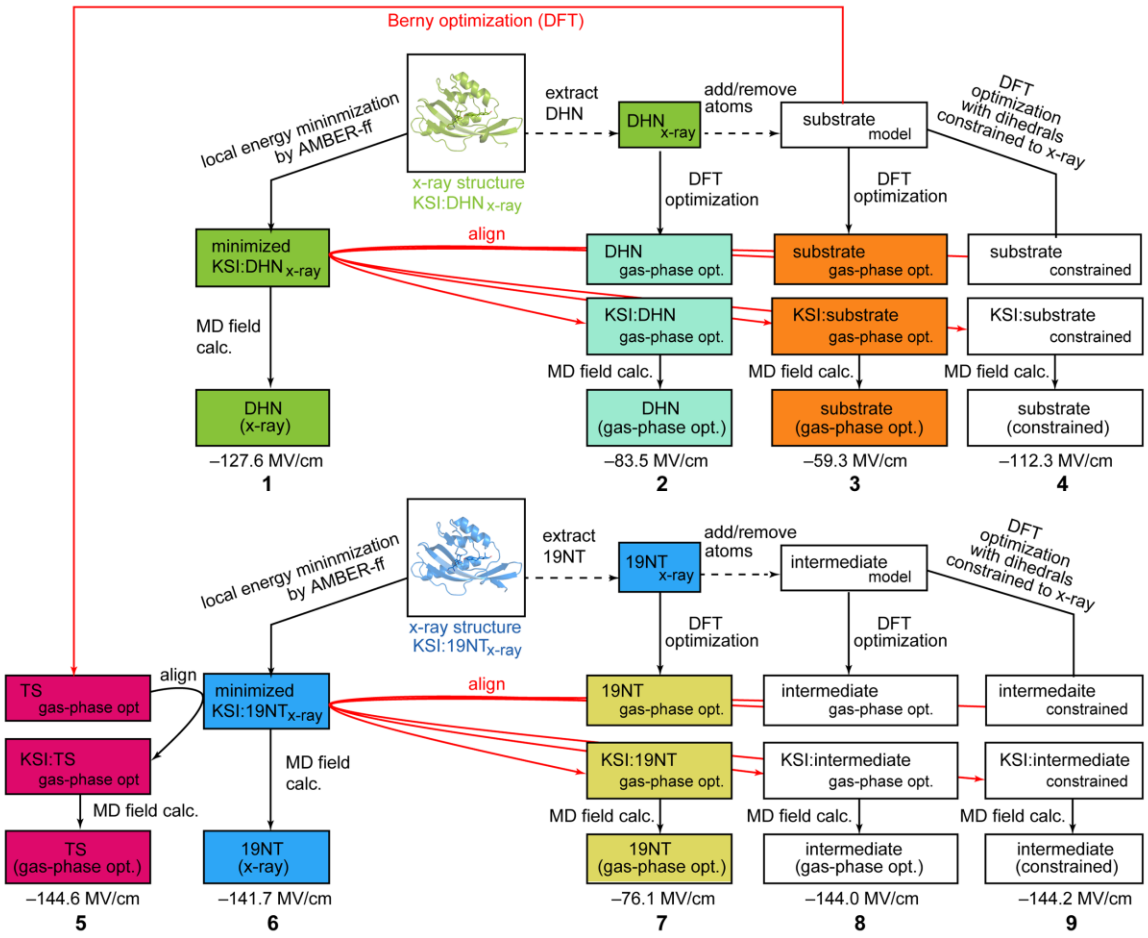


Figure S3 | Modelling scheme of different KSI•ligand complexes for MD calculations.

Each construct is given a number whose corresponding simulation results are listed in Table S2. The modelling strategy is also explained and examined in SI Methods and in supplementary discussion S1.

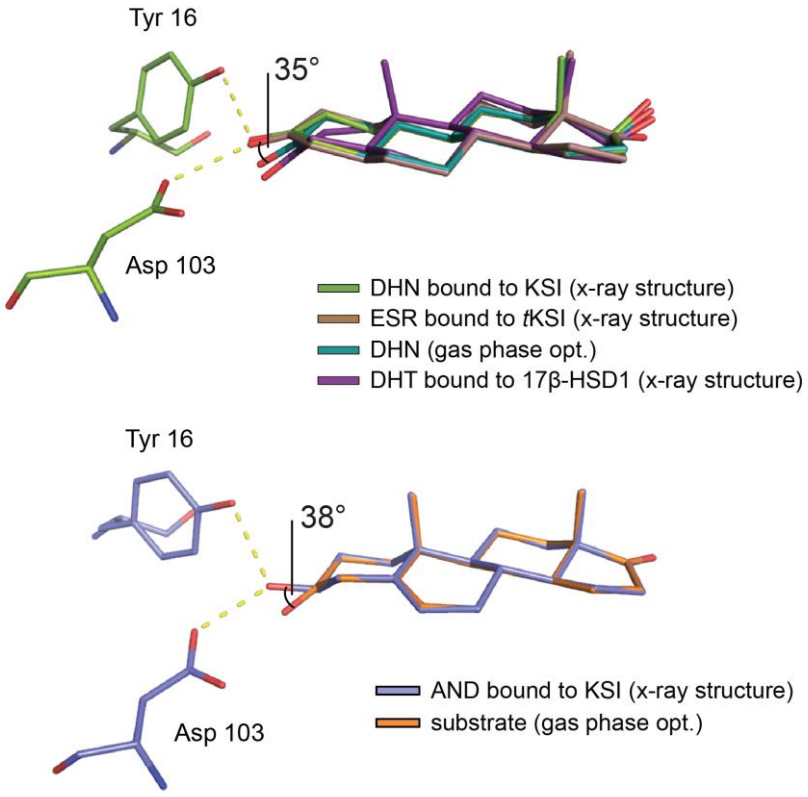


Figure S4 | Geometrical change between steroids bound to enzymes and in the gas phase.

(A) The perturbed 'TS-like' geometry for DHN bound to KSI is recapitulated by a different steroid, 5α-estran-3,7-dione (ESR), bound to a different KSI orthologue, *t*KSI, but absent in dihydrotestosterone (DHT) bound to 17β-hydroxysteroid dehydrogenase, which more closely resembles the gas-phase optimized geometry of DHN. (B) The gas-phase optimized substrate, 5-androstenedione, is compared with substrate analog, androstane-3beta-ol-17-one (AND), bound to the catalytically inactive *p*KSI^{D40N} (PDB 1E3R). The structural difference at C_3 between the substrate and AND (i.e. sp^2 hybridization to sp^3 hybridization) mimics the angle change at the carbonyl upon the formation of the TS/Intermediate (Fig. 1): AND's C-O bond is shifted from the substrate by 38°, inherently aligned with KSI's electric field (~31°). The carbonyl of DHN is distorted by the field to achieve the same optimal interaction.

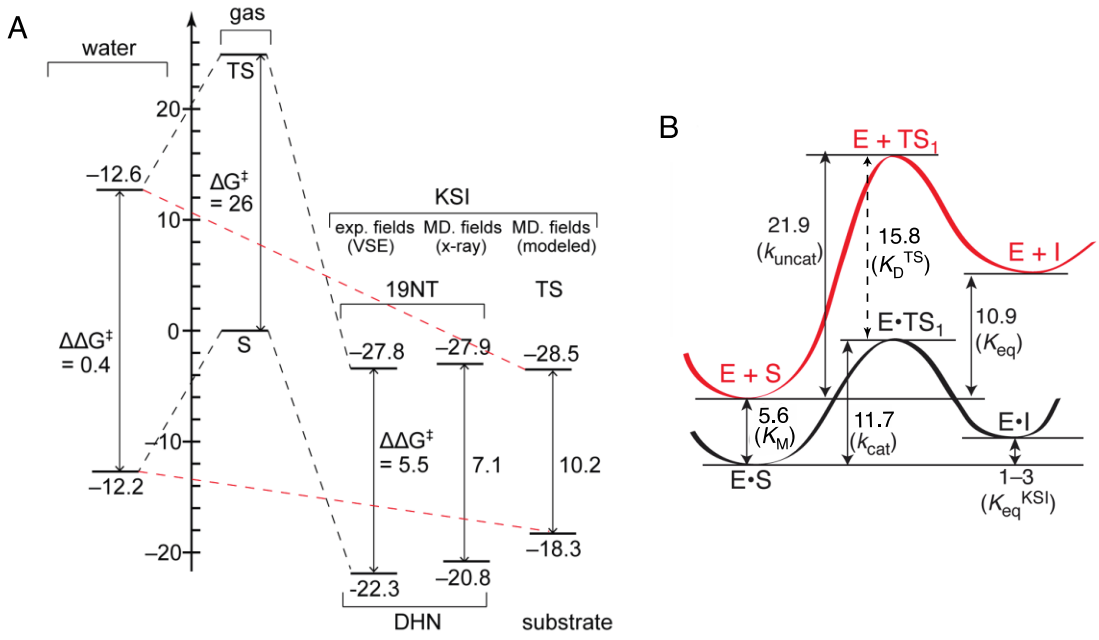


Figure S5 | Stabilization energies in different environments and the free energy profiles of the isomerization reaction in KSI and in aqueous solution

(A) The relative free energy of the substrate and the TS are estimated in gas phase, water and KSI. The stabilization energy of the TS in water is estimated under the assumption that the solvent reaction field is equilibrated with the *reactant's* dipole orientation (which is 31˚ out of alignment with the TS, Fig. 1B), whereas the stabilization energy of the substrate in water assumes 100% alignment between the solvent reaction field and the reactant dipole (Table S3). The stabilization energy in KSI is estimated in three ways: (1) from the measured electric field projections on the carbonyl of 19NT and DHN (Table S3) by VSE, which is likely an underestimation of KSI's preferential stabilization on the TS over the substrate because the actual substrate, 5-androstenedione, would not be as easily distorted to TS-structure as DHN is (Figure 2B); (2) from the MD calculated electric fields based on the actual crystal structures of the KSI•19NT and KSI•DHN complexes (Table S2); (3) from the MD calculated electric fields based on the modeled structures of the KSI•substrate and KSI•TS complexes (Table S2 in bold), which is likely an overestimation of the preferential stabilization energy because the alignment of the substrate into the DHN crystal structure coordinate is not ideal. Red dashed lines indicate the additional stabilization energies provided by the enzyme environment over solutions upon ligand binding (6.1 kcal mol^{-1} for substrate; 15.9 kcal mol^{-1} for transition state). (B) Reaction coordinate diagram for the enolization of 5-androstenedione (S) to the dienolate intermediate (I), either in KSI (black) or in solution (red). Numbers correspond to free energy differences in kcal mol^{-1}. Black arrows correspond to measured quantities and are associated with the binding/rate constants in parenthesis.[12,24–29] The figure is adapted from Fried, S.D. Ph.D. Dissertation, Stanford University, 2014.

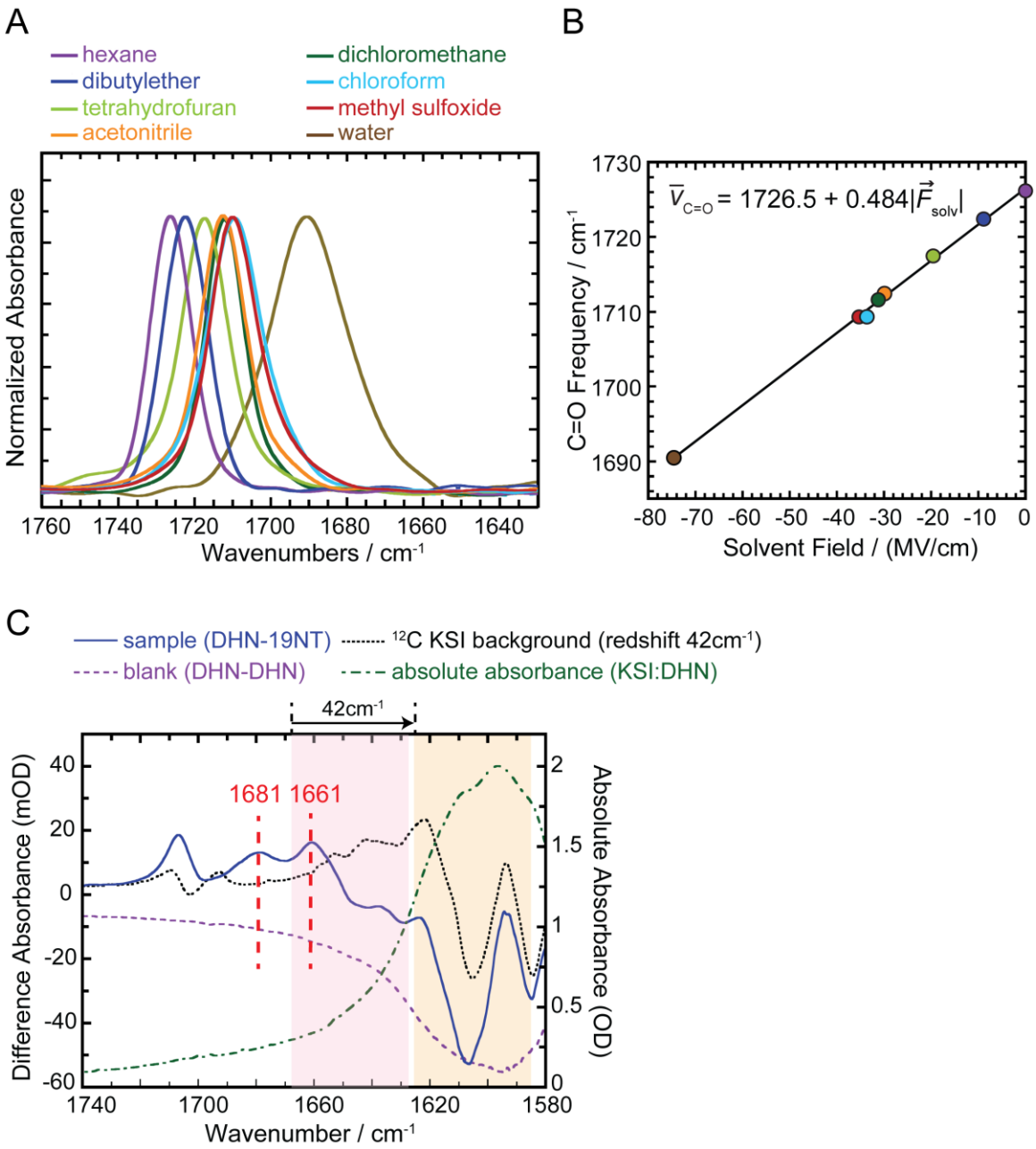


Figure S6 | Infrared spectra of DHN in different solvents and bound to wild type [^{13}C]KSI

(A) Vibrational solvatochromism for DHN's carbonyl in solvents with increasing polarity. (B) The average value of the solvent fields exerted onto the C=O of DHN display a linear correlation with the C=O stretching frequency (R^2=0.99). (C) The frequency range of the carbonyl peak(s) of DHN is determined to be between 1691 cm^{-1} and 1658 cm^{-1}, because the DHN bound to WT KSI is expected to experience an electric field larger than that in water and smaller than or similar to that experienced by TS-like ligand 19NT. Because the amide I band of ^{12}C KSI dominates the signal in this frequency range (light pink shade), we employed ^{13}C-labeling to redshift the amide I band by 42 cm^{-1} from ~1642 cm^{-1} to ~1600 cm^{-1} (light yellow shade),

leaving a relatively flat baseline in the region of interest ('KSI background'). Two peaks at 1660.3cm^{-1} (FWHW ~9 cm^{-1}, intensity ~6 mOD) and 1681.4 cm^{-1} (FWHW ~10 cm^{-1}, intensity ~4 mOD) were consistently observed above the background and were provisionally assigned as the C=^{16}O stretch of DHN sulfate bound in the active site. The peak at ~1710 cm^{-1} likely comes from the carboxyl group of amino acid side chains.[17] Comparing the absorption spectrum of the KSI complex (green, right y axis) with the *difference* absorption spectrum (blue, left y axis), we see that the total absorbance from the enzyme complex is about 25 times more intense than the absorbance from the ligand's carbonyl.

As discussed in the experimental section, we also attempted to confirm the identities of the two peaks observed with isotope labeled substrate. [^{16}O]DHN sulfate was used as the sample and [^{18}O]DHN sulfate was used as the reference. The isotope labeling on the carbonyl would ensure an otherwise identical IR background except a positive feature from C=^{16}O vibration and a negative feature from C=^{18}O vibration that are separated by approximately 30 cm^{-1}. In reality, however, the measurements were rendered unreliable by two factors: 1) The steep slope from 1600 cm^{-1} to 1640 cm^{-1} in the difference absorbance spectrum due to the ^{13}C-labeled amide signal obscured the C=^{18}O stretch of DHN sulfate; and 2) the isotope labeled carbonyl of DHN hydrolyzes back to C=^{16}O quickly with the presence of trace [^{16}O] over the course of IR sample preparation and measurements, further decreasing the IR signals.

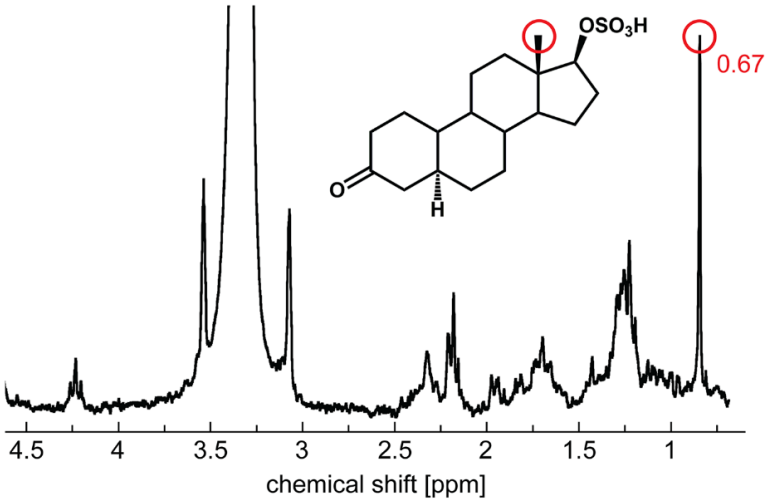


Figure S7 | ^{1}H-NMR spectrum of purified dihydronandrolone sulfate

The signal at 0.67 ppm corresponds to the three methyl group protons circled. This was used to calibrate the concentration of DHN and DHN sulfate against 19NT and trimethylsilylpropanoic acid (TSP-d_4) standard. The big peak around 3.3ppm comes from the trace H_2O in deuterated solvent dimethyl sulfoxide-d_6.[30]

References

(1) Wu, Y.; Fried, S. D.; Boxer, S. G. Dissecting proton delocalization in an enzyme's hydrogen bond network with unnatural amino acids. *Biochemistry* **2015**, *54*, 7110–7119.

(2) Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme electric fields power catalysis in the active site of ketosteroid isomerase. *Science* **2014**, *346*, 1510–1513.

(3) Oltrogge, L. M.; Boxer, S. G. Short hydrogen bonds and proton delocalization in green fluorescent protein (GFP). *ACS Cent. Sci.* **2015**, *1*, 148–156.

(4) Wilson, I. A.; Stura, E. A. Applications of the streak seeding technique in protein crystallization. *J. Cryst. Growth* **1991**, *110*, 270–282.

(5) Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 133–144.

(6) Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307–326.

(7) Gonzalez, A.; Tsai, Y. http://smb.slac.stanford.edu/facilities/software/xds/#autoxds_script 2010.

(8) McCoy, A. J.; Grosse-Kunstleve, R. W.; Storoni, L. C.; Read, R. J. Likelihood-enhanced fast translation functions. *Acta Cryst.* **2005**, *D61*, 458–464.

(9) Adams, P. D.; Grosse-Kunsleve, R. W.; Hung, L.-W.; Loerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX : building new software for automated crystallographic structure determination. *Acta Cryst.* **2002**, *D58*, 1949–1954.

(10) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Cryst.* **2004**, *D60*, 2126–2132.

(11) DeLano, W. L. http://www.pymol.org.

(12) Wu, Y.; Boxer, S. G. A critical test of the electrostatic contribution to catalysis with noncanonical amino acids in ketosteroid isomerase. *J. Am. Chem. Soc.* **2016**, *138*, 11890–11895.

(13) Fried, S. D.; Bagchi, S.; Boxer, S. G. Measuring electrostatic fields in both hydrogen-bonding and non-hydrogen-bonding environments using carbonyl vibrational probes. *J. Am. Chem. Soc.* **2013**, *135*, 11181–11192.

(14) Caleman, C.; Van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; Van Der Spoel, D. Force field benchmark of organic liquids: density, enthalpy of vaporization, heat capacities, surface tension, isothermal compressibility, volumetric expansion coefficient, and dielectric constant. *J. Chem. Theory Comput.* **2012**, *8*, 61–74.

(15) Schneider, S. H.; Boxer, S. G. Vibrational Stark effects of carbonyl probes applied to reinterpret IR and Raman data for enzyme inhibitors in terms of electric fields at the active site. *J. Phys. Chem. B* **2016**, *120*, 9672–9684.

(16) Welborn, V. V.; Head-gordon, T. Fluctuations of electric fields in the active site of the enzyme ketosteroid isomerase. *J. Am. Chem. Soc.* **2019**, *141*, 12487–12492.

(17) Barth, A. The infrared absorption of amino acid side chains. *Prog. Biophys. Mol. Biol.* **2000**, *74*, 141–173.

(18) Hammes-Schiffer, S. Catalytic efficiency of enzymes: a theoretical analysis. *Biochemistry* **2013**, *52*, 2012–2020.

(19) Richard, J. P. Protein flexibility and stiffness enable efficient enzymatic catalysis. *J. Am. Chem. Soc.* **2019**, *141*, 3320–3331.

(20) Amyes, T. L.; Richard, J. P. Specificity in transition state binding: The Pauling model revisited. *Biochemistry* **2013**, *52*, 2021–2035.

(21) Fafarman, A. T.; Sigala, P. A.; Schwans, J. P.; Fenn, T. D.; Herschlag, D.; Boxer, S. G. Quantitative, directional measurement of electric field heterogeneity in the active site of ketosteroid isomerase. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E299-308.

(22) Ha, N. C.; Kim, M. S.; Lee, W.; Choi, K. Y.; Oh, B. H. Detection of large pKa perturbations of an inhibitor and a catalytic group at an enzyme active site, a mechanistic basis for catalytic power of

many enzymes. *J. Biol. Chem.* **2000**, *275*, 41100–41106.
(23) Ruben, E. A.; Schwans, J. P.; Sonnett, M.; Natarajan, A.; Gonzalez, A.; Tsai, Y.; Herschlag, D. Ground state destabilization from a positioned general base in the ketosteroid isomerase active site. *Biochemistry* **2013**, *52*, 1074–1081.
(24) Schwans, J. P.; Kraut, D. A.; Herschlag, D. Determining the catalytic role of remote substrate binding interactions in ketosteroid isomerase. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 14271–14275.
(25) Zeng, B.; Pollack, R. M. Microscopic rate constants for the acetate ion catalyzed isomerization of 5-androstene-3,17-dione to 4-androstene-3,17-dione: a model for steroid isomerase. *J. Am. Chem. Soc.* **1991**, *113*, 3838–3842.
(26) Hawkinson, D. C.; Eames, T. C. M.; Pollack, R. M.; Energetics of 3-Oxo-Δ5-steroid Isomerase: Source of the Catalytic Power of the Enzyme. *Biochemistry* **1991**, *30*, 10849–10858.
(27) Hawkinson, D. C.; Pollack, R. M.; Ambulos, N. P. Evaluation of the internal equilibrium constant for 3-Oxo-Δ5-steroid isomerase using the D38E and D38N mutants: the energetic basis for catalysis. *Biochemistry* **1994**, *33*, 12172–12183.
(28) Pollack, R. M. Enzymatic mechanisms for catalysis of enolization: ketosteroid isomerase. *Bioorg. Chem.* **2004**, *32*, 341–353.
(29) Fried, S. D.; Boxer, S. G. Thermodynamic framework for identifying free energy inventories of enzyme catalytic cycles. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 12271–12276.
(30) Fulmer, G. R.; Miller, A. J. M.; Sherden, N. H.; Gottlieb, H. E.; Nudelman, A.; Stoltz, B. M.; Bercaw, J. E.; Goldberg, K. I. NMR chemical shifts of trace impurities: Common laboratory solvents, organics, and gases in deuterated solvents relevant to the organometallic chemist. *Organometallics* **2010**, *29*, 2176–2179.