Bond Bundle Analysis of Ketosteroid Isomerase

Timothy R. Wilson[1], Amanda Morgenstern[2], Anastassia N. Alexandrova[3], and M.E. Eberhart[4,*]

[1]*Department of Chemistry, Colorado School of Mines, 1500 Illinois St., Golden, CO 80004, USA*
[2]*Department of Chemistry & Biochemistry, UCCS, 1420 Austin Bluffs Pkwy, Colorado Springs, CO 80918*
[3]*Department of Chemistry, UCLA, 607 Charles E. Young Drive East, Los Angeles, CA 90095*
[4]*Department of Chemistry, Colorado School of Mines, 1500 Illinois St., Golden, CO 80004, USA*
[*]*Corresponding author. Email: meberhar@mines.edu*

October 20, 2022

Abstract

Bond bundle analysis is used to investigate enzymatic catalysis in the ketosteroid isomerase (KSI) active site. We identify the unique bonding regions in five KSI systems, including those exposed to applied oriented electric fields and those with amino acid mutations, and calculate the precise redistribution of electron density and other regional properties that accompanies either enhancement or inhibition of KSI catalytic activity. We find that catalytic enhancement results from promoting both inter- and intra-molecular electron density redistribution, between bond bundles and bond wedges within the KSI-docked substrate molecule, in the forward direction of the catalyzed reaction. Though the redistribution applies to both types of perturbed systems, and is thus suggestive of a general catalytic role, we observe that bond properties (*e.g.* volume vs. energy vs. electron count) can respond independently and disproportionately depending on the type of perturbation. We conclude that the resulting catalytic enhancement/inhibition proceeds via different mechanisms, where some bond properties are utilized more by one type of perturbation than the other. Additionally, we find that the correlations between bond wedge properties and catalyzed reaction barrier energies are additive to predict those of bond bundles and atomic basins, providing a rigorous grounding for connecting changes in local charge density to resulting shifts in reaction barrier energy.

Keywords: *bond bundle, bond wedge, gradient bundle analysis, gradient bundle decomposition, KSI, enzymatic catalysis, electron density analysis, QTAIM*

1 Introduction

Enzymes can accelerate chemical reactions by many orders of magnitude. Our understanding of the mechanisms responsible for this process has grown at an increasing rate over the last few decades thanks to two significant advances: i) increasingly accurate structural studies; and ii) improved computational platforms and methods that allows us to predict, among other things, the conformation and energy of folded proteins and the reaction dynamics at enzyme active sites.[1–5] This wealth of data is useful for identifying new structures and structure-property relationships that may, in turn, be incorporated into the existing chemical formalisms, known collectively as chemical intuition.

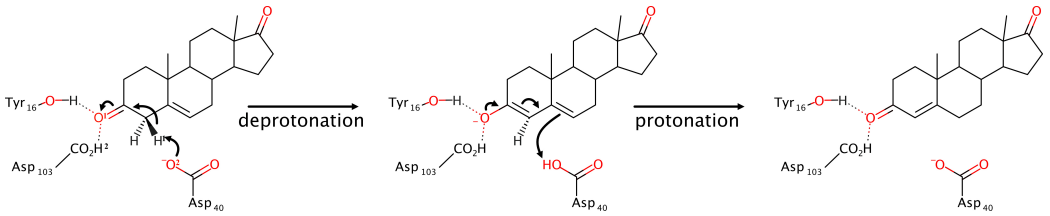
Here we investigate the model ketosteroid isomerase (KSI) enzyme, the focus of numerous experimental and computational investigations (see Reference 6 for additional references), specifically to observe the redistribution of electron charge density, $\rho(\boldsymbol{r})$ caused by amino acid mutations or applied external electric fields (EEFs). These are thought to affect KSI catalytic activity primarily through changes to its electrostatic preorganization,[7–10] though it remains unconfirmed that these two types of perturbations affect electrostatic preorganization via the same mechanism.

Electrostatic preorganization is a strong, non-uniform electric field that augments catalysis through transition state stabilization,[8] and that results from the arrangement and composition of amino acids about the active site, whose specific arrangement is maintained by the larger protein scaffolding. Using computational methods to predict electrostatic preorganization, researchers are exploring methods for its analysis, comparison, and interpretation.[6,11,12]

Our approach uses bond wedge and bond bundle analysis to reduce a system to a set of chemical bonding regions whose energy, extent, electron count, and many other properties can be accurately calculated.[13,14] Bond wedges are regions within an atom (atomic basin) where charge density accumulates as a consequence of atomic interactions. The boundaries separating bond wedges are precisely defined zero-flux surfaces whose shape and movement dictate how charge redistribution will affect bond properties via simple geometric constraints on $\rho(\boldsymbol{r})$.[15]

Our two-part investigation proceeds with a direct inspection of bond bundle property distributions in the KSI active site in the presence and absence of a uniform EEF known to enhance its catalytic activity. Here we will see that bond bundle electron counts provide an unprecedented level of precision and ease of chemical interpretation. In the second part, we explore a set of KSI systems with varying catalytic activity. The investigation reveals the reactant state bond wedge and bond bundle redistribution that facilitates the forward reaction direction, and locates the active-site regions that most strongly correspond to catalytic enhancement or inhibition, whether caused by applied EEFs or amino acid mutation.

We find that regional volume and energy respond in different proportions to applied EEFs than to chlorination of particular amino acids. The disproportionate response of regional properties is found to underlie differences in reaction barrier height. This suggests that regions whose properties have strong correlation with reaction barrier height and that respond proportionately to both types of perturbations play a more central role in catalytic augmentation. Our analysis method allows us to precisely quantify these changes, and because the data are real-space, regional property integrations,



Scheme 1: Steroid isomerization reaction catalyzed by KSI.

it is convenient to visualize and frame our findings within the language and concepts of chemical bonding.[16] This framing provides rigor to what Roald Hoffmann describes as the "fuzzy" application of chemical bonding.[17]

2 Background

2.1 Ketosteroid isomerase

The well-studied steroid isomerization reaction that KSI catalyzes involves the repositioning of a double C=C bond in the steroid substrate.[6,11,18–21] As shown in Scheme 1, this occurs by the removal of a proton from the secondary β-carbon, which is redeposited at the adjacent secondary carbon. Focusing on the first step, deprotonation is typically pictured as the result of a shift of charge though the substrate π system from the β-carbon to the carbonyl oxygen. In KSI this concerted atomic and electronic rearrangement is facilitated by the ideal positioning of Asp_{40}, providing a general base to receive the proton, and by the oxyanion hole that activates the carbonyl and stabilizes the charged enolate intermediate state.

Fuller *et al.* investigated the effects of applied EEFs to this process using the small-scale KSI active site shown in Figure 1, and found that a field applied parallel to the substrate carbonyl bond, pointing from O to C, augmented the electrostatic preorganization, lowering the reaction barrier, while a field in the opposite direction raised the barrier.[11] This agrees with classical intuition, that an EEF should push charge opposite the field direction, in this case combining constructively with the KSI oxyanion hole to more readily shift charge to the carbonyl oxygen—further stabilizing the enolate intermediate state—and away from the β carbon, increasing its acidity and thus facilitating deprotonation.

Hennefarth and Alexandrova were then able to show similar reaction barrier effects in KSI vari-

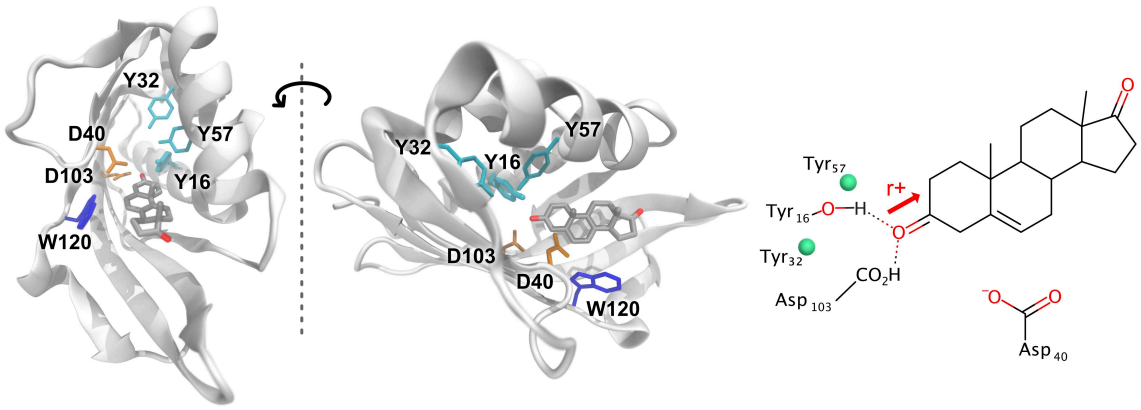


Figure 1: The full KSI protein (PDB code 1O10)[22] with docked steroid substrate shown from two angles. The tyrosine (Y or Tyr; cyan), aspartic acid (D or Asp; orange) and tryptophan (W or Trp; blue) residues included in the small scale calculation are shown relative to the substrate (colored by element). The Lewis diagram of the system is shown with the "locations" of the 3-chlorotyrosine for the KSI variant systems (Trp_{120} not shown).

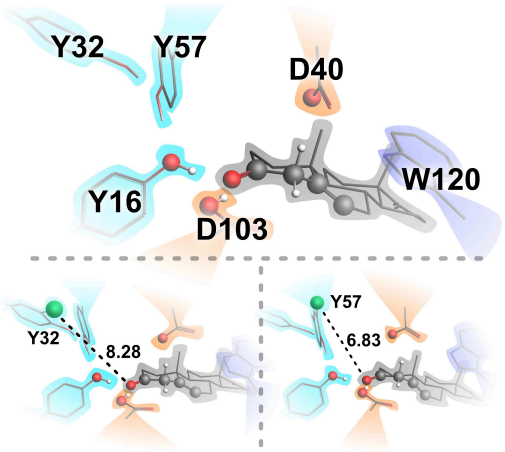


Figure 2: Closer depiction of the arrangement of amino acids forming the "oxyanion hole" about the substrate in KSI (top) and the positioning of the 3-chlorotyrosines in the KSI mutants (bottom) with Cl···O1 distance (Å) indicated.

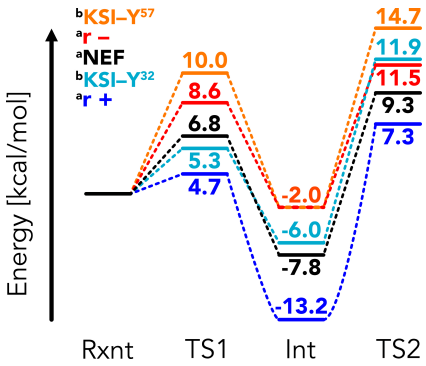


Figure 3: Reaction profiles for the systems in this study. Ordering at top-left is the same as in the TS1 column. [a]Data for the NEF and EEF systems taken from Fuller *et al.*[11] [b]Data for KSI–Y systems taken from Hennefarth and Alexandrova.[6]

ants that had a tyrosine mutated to a 3-chlorotyrosine.[6] Because the tyrosines of interest are involved in the extended hydrogen bonding network around the oxyanion hole (see Figure 2),[23] this affects carbonyl activation, altering electrostatic preorganization. They found that a mutated Tyr_{32} lowered the barrier, and a mutated Tyr_{57} raised the barrier relative to the wild type enzyme.

The present investigation uses the small-scale KSI active site, EEF directions, and 3-chlorotyrosine KSI variants from References 6 and 11 as the starting point for our calculations, and we explicitly use the same calculated reaction barrier energies. The five systems: wild type (WT) KSI (aka NEF: no electric field), KSI^{r+}, KSI^{r-}, KSI-Y^{32} and KSI-Y^{57} provide a minimum example set of KSI enhancement and inhibition via global and local perturbations; oriented EEFs and chlorination of amino acids respectively. The regional changes in charge density and energy underlying the reaction barrier shifts should be accessible using this limited sample.

2.2 Assessing the local charge density origins of KSI catalytic enhancement

A common approach in the search for *local* relationships between $\rho(\boldsymbol{r})$ and energy related catalytic properties (reaction barriers) is to seek correlations between the property of interest and charge density metrics at points, along one-dimensional paths, or within arbitrarily defined volumes. This approach has been employed to understand the enhancement and inhibition mechanisms of KSI using $\rho(\boldsymbol{r})$ from *ab-initio* calculations.[4–6,11]

Fuller *et al.* checked for correlations between KSI-catalyzed reaction barrier shifts due to EEFs and a number of local properties, such as inter-atomic distances and values of $\rho(\boldsymbol{r})$ at bond critical points (CPs).[11] The strongest correlation found was that of the O2–H1 bond length (see Figure 4 for atom numbering), which correlated positively with the change in reaction barrier, while the value of $\rho(\boldsymbol{r})$ at the corresponding bond CP anticorrelated, indicating that direct facilitation of the deprotonation step affects the barrier energy shift. A weaker positive correlation was found with the O1–H2 bond length, indicating, as anticipated, that activation of the carbonyl bond—by decreasing the Asp_{40}–substrate distance—also lowers the reaction barrier.

Given the catalytic role of electrostatics in KSI, Hennefarth and Alexandrova investigated the electric field itself, $\boldsymbol{E}(\boldsymbol{r})$, in the KSI active site using both point and regional properties.[6] Here, the set of systems included those exposed to oriented EEFs, and also two 3-chlorotyrosine KSI mutants

with respectively higher and lower catalytic activity than WT KSI. They found that the electric field magnitude, $|\boldsymbol{E}(\boldsymbol{r})|$, at the O2–H1, C2–H1, and C1=O1 bond CPs correlated strongly with the change in reaction barrier, but only for WT KSI structures; KSI-Y^{32} and KSI-Y^{57} were outliers to this trend. We take these observations to indicate that the mechanism, through $\rho(\boldsymbol{r})$, by which the KSI mutants enhance/hinder catalysis may be different than that due to EEFs. Furthermore, the correlation at the carbonyl bond CP showed a nearly constant relationship between $|\boldsymbol{E}(\boldsymbol{r})|$ and reaction barrier energy.

Hennefarth and Alexandrova also conducted a regional electric field curvature analysis within two separate rectilinear volumes, one containing the carbonyl C1=O1 atoms, and the other containing the C2–H1$\cdots$O2 atoms of the reaction site. By evaluating the total curvature along $\boldsymbol{E}(\boldsymbol{r})$ streamlines within each volume, they generated histograms that reflect the relative occupations of high and low curvature regions—a regional $\boldsymbol{E}(\boldsymbol{r})$ fingerprint. The same relative regions of multiple systems could then be compared by computing their corresponding histogram distances, thus providing a scalar similarity metric also useful for statistical evaluation.

Using this similarity metric, Hennefarth and Alexandrova found that the $\boldsymbol{E}(\boldsymbol{r})$ curvature about the carbonyl bond had a stronger correlation with reaction barrier than that of the reaction site, with 3-chlorotyrosine mutants included in the analysis and which were no longer outliers. This result is counter to the results from point properties (and those of Fuller *et al.*) that the strongest correlations occur in the reaction site rather than within the substrate. The regional results indicate that activation of the carbonyl enhances the reaction rate, where point-based analysis results emphasize changes at the deprotonation site. That is, a regional approach seems to highlight the underlying chemistry at work common to both the EEF and mutant KSI systems, resulting from changes to the electrostatic preorganization in the KSI oxyanion hole.[10,24]

2.3 The chemical bonding structure of the charge density

Here we extend The Quantum Theory of Atoms in Molecules (QTAIM), considering only volumes over which energy is well-defined.[25,26] In its standard form, QTAIM is used to locate the boundaries of the "atoms in molecules," aka atomic basins, which possess unambiguous energies, volumes, and charges. That atomic basins have a well-defined energy results from their satisfaction of a zero flux (of the charge density gradient) boundary condition. For an arbitrary region in $\rho(\boldsymbol{r})$, one may calculate the average regional kinetic energy using the gradient or Laplacian forms of the quantum mechanical kinetic energy operator, but in general these values will not match one another. Over a region bounded by zero-flux surfaces, however, these values will agree, and hence the regional energy of an atomic basin is unambiguous.[25]

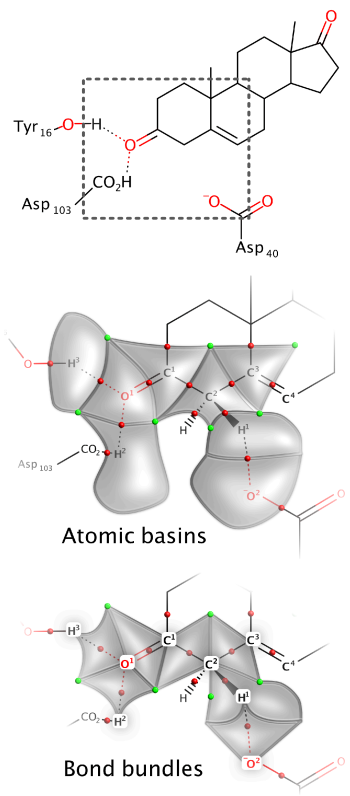


Figure 4: Cartoon depictions of the atomic basins and bond bundles analyzed in this study. Red and green spheres represent bond and ring CPs respectively.

Bond bundle analysis is an extension of QTAIM that recognizes a further partitioning of atomic basins into the smallest regions bounded by zero-flux surfaces, called differential gradient bundles (dGBs).[27,28] To each dGB there corresponds an unambiguous energy, and an atomic basin decomposition into dGBs produces a continuous and "well-defined energy space." The topology of this space reveals precise boundaries between intra-atomic regions of charge accumulation called bond wedges. Bond wedges of adjacent atoms then combine to form bond bundles.[14,29]

Figure 4 illustrates the eight atomic basins and seven bond bundles considered in this study, as well as the bond and ring CPs that lie along or interior to their boundaries. See Figure S3 in the supplemental information for a more realistic, three-dimensional representation of bond bundles. We will appeal to this sort of abstract representation of bond bundles for the remainder of this manuscript. Each bond bundle has a set of properties commonly associated with a chemical bond, such as an energy and a number of electrons, which, when taken over all bonds in a system, recover the system energy and electron count. Like atomic basins, bond bundles have precise, non-overlapping boundaries that combine to fill all space, giving rise to bond volume—one of a number

of newly quantifiable bond properties. Of course, the integrated value of electron density within a bond bundle gives its electron count.

Methods

All results are from *ab-initio* calculations including the residues shown experimentally and theoretically to control the electrostatic preorganization in KSI.[30] The initial placement of these residues is that determined in References 6 and 11, starting from the experimentally determined KSI structure.[22] While the KSI active site $|\boldsymbol{E}(\boldsymbol{r})|$ can fluctuate between 135 and 150 $^{\mathrm{MV}}/_{\mathrm{cm}}$ at physiological temperatures,[31] here, as in References 6 and 11, we are interested in elucidating reactivity using information from the reactant state, so we exclusively use relaxed, reactant-state conformations where $|\boldsymbol{E}(\boldsymbol{r})|$ is closer to $144 \pm 6 ^{\mathrm{MV}}/_{\mathrm{cm}}$.[20] Subsequent bond bundle investigations will focus on the significance of these fluctuations due to large thermal movements in biological molecules, as well as analogous effects due to molecular vibrations in simple molecular systems.

All *ab-initio* calculations were performed using the ADF package of the Amsterdam Modeling Suite.[32–34] Relaxed system geometries were obtained with initial coordinates from References 6 and 11 as mentioned above. Optimization of all five systems was performed using a triple-ζ STO all-electron basis set with one polarization function,[35] with the Minnesota'06-2X XC energy density functional[36,37] and "good" numerical integration quality. The NEF system relaxation also included implicit COSMO solvation[38,39] using Allinger solvent radii and a dielectric constant of $\varepsilon = 4.0$. Subsequent single-point calculations were run with the same basis set and functional, with the same COSMO settings now used for all systems. All applied electric fields were of magnitude 10 $^{\mathrm{MV}}/_{\mathrm{cm}}$, a value on the order of that determined for the field produced by the protein beyond the active site by Fuller *et al.* using AMBER charges and classical multipole expansion,[11] and an order of magnitude less than that produced within the KSI active site.[20]

Topological analysis and gradient bundle decomposition was performed on static $\rho(\boldsymbol{r})$ grids with the Bondalyzer software suite of the Molecular Theory Group at Colorado School of Mines,[40] an add-on to the Tecplot360 visualization package.[41] The error associated with these calculations arises mostly from the use of regular volumetric grid field data, the resolution of which governs the error. Given a "correct" $\rho(\boldsymbol{r})$, the accuracy of the method increases with the number of differential gradient bundles used, and with increased discretization of the underlying gradient paths. When a grid spac-

ing of 0.025 bohr is used, and each atomic basin is decomposed into at least five-thousand differential gradient bundles, as in this investigation, error in atomic basin electron populations is generally less than 1%, *i.e.* more than 99% of atomic basin charge density is recovered. Additionally, the current method of condensed maximum basin identification results in ±0.2% error in the integrated bond wedge properties.

Comparison and statistical analysis of gradient bundle integrated properties was performed in Python, and the MatPlotLib[42] and SeaBorn[43] libraries were used for plotting line/scatter and bar charts respectively.[44] One-dimensional scatter plots (Figure 7 and Section 5 of the SI) show the KSI-catalyzed reaction barrier height as a function of integrated regional properties, with a linear fit produced with `sklearn.linear_model.LinearRegression` in SciPy.[45] R^2, R^2_{adj}, and p-values included with the linear fits (and the other statistical measures included for each fit in the SI) were produced using the ordinary least squares `summary` in StatsModels.[46] Horizontal bar charts (Figure 8 and Section 4 of the SI) are a compact, graphical depiction of many such scatter plots, showing the sign of the correlation (fit line slope) and the correlation coefficient of the linear fit, but not the slope of the fit line, and were constructed using MatPlotLib. Correlation diagrams (Figure 6 and Section 7 of the SI) are typical correlation diagrams showing whether the properties within a particular region correlate with each other (*e.g.* how volume and energy correlate within some particular bond bundle), produced using the SeaBorn `heatmap`. Section 6 of the SI contains other correlation diagrams showing whether particular regional properties correlate across similar regions (*e.g.* how bond bundle volume correlates across all bond bundles).

Chemical diagrams were composed in MarvinSketch.[47] The left two images in Figure 1 were compiled with VMD.[48] Final figures were compiled in Affinity Designer.[49]

3 Results and discussion

3.1 KSI charge density response to a catalyzing EEF

To illustrate the concrete nature of gradient bundle properties, we begin by inspecting $\rho(\boldsymbol{r})$ redistribution in the KSI active site due to a catalyzing uniform EEF of magnitude $10^{\mathrm{MV/cm}}$, which is an order of magnitude less than that produced by the first and second coordination sphere residues, and on the order of the field produced by the protein beyond the active site.[11] The KSI active site

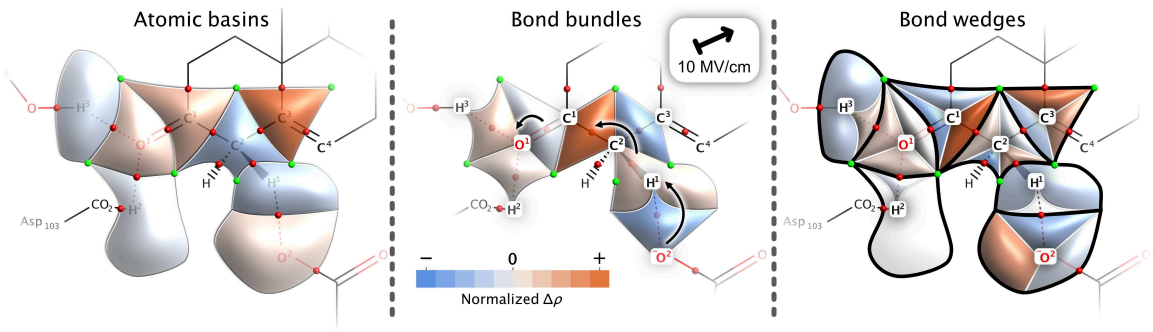


Figure 5: The atomic basins, bond bundles, and bond wedges of KSI (left, middle, and right respectively) shaded according to the changes in their regional electron count due to a $10^{\mathrm{MV}/\mathrm{cm}}$ oriented EEF (the r+ direction), which are listed in Table 1. The center image includes the electron-pushing arrows of the deprotonation reaction step.

EEF response is captured in the gradient bundle decomposition.

Table 1 contains regional electron counts (ρ) for atomic basins, bond bundles, and bond wedges in the KSI active site, as well as their change due to the catalyzing r+ EEF. These regional property changes are graphically depicted in Figure 5, where regions are shaded orange or blue to indicate EEF-induced $\rho(\boldsymbol{r})$ accumulation or depletion respectively. See Section 3 in the supplemental information for similar tables for all five systems, for ten different regional properties.

The field activates the carbonyl bond, which is assumed to occur by redistributing electron charge density to the carbonyl O atom, and is evidenced here by the accumulation of charge in the O1 atomic basin (Figure 5; left). However, the charge redistribution *within* the atomic basin (*i.e.* between bond wedges), recovers a more detailed description of the activation. The charge accumulation in the O1$\cdots$H2 and O1$\cdots$H3 bond bundles (Figure 5; center) is a consequence of charge depletion in the O1=C1 bond bundle. More pronounced is the charge redistribution occurring at the other end of the conjugated system, interior to the substrate. The C1–C2 bond bundle, which is believed to increase its bond order from single to double in the deprotonation step, accumulates nearly 0.09 electrons in response to the field, offset by decreases in the carbonyl bond and more so in the C2–C3 bond.

At the reaction site, the C2–H1 bond, which is broken in the deprotonation step, accumulates charge due to the EEF, while the incipient O2$\cdots$H1 bond loses charge, seemingly the reverse of the anticipated electron motion for a forming bond. Inspection of the corresponding bond wedge values (Figure 5; right), however, reveals the expected behavior. The increase in the C2–H1 bond bundle electron count is due to the contribution of the C2 atom. The H1 bond wedge component of the C2–

Table 1: Regional electron counts in the KSI active site with (EEF) and without (NEF) an applied external electric field of $10^{\mathrm{MV}/\mathrm{cm}}$ pointing from the O nuclear position to the C; the r+ direction. Unnumbered atoms were not included in the study. All regions are truncated at the $\rho = 0.001$ isosurface. Complete gradient bundle integration tables are available in the SI, Section 3.

	Electron count (ρ) [e]			
Atomic basin decomposition	NEF	EEF	Δ	$\%\Delta$
C1	5.151	5.185	0.035	0.67
C2	6.207	6.187	-0.021	-0.33
C3	6.182	6.255	0.073	1.18
H1	0.862	0.846	-0.017	-1.92
H2	0.370	0.369	-0.001	-0.22
H3	0.380	0.367	-0.013	-3.46
O1	9.245	9.258	0.013	0.14
O2	9.324	9.329	0.005	0.05
Total	**37.721**	**37.795**	**0.074**	**0.20**
Bond bundle decomposition	NEF	EEF	Δ	$\%\Delta$
C1 — C bond wedge	2.049	2.010	-0.040	-1.93
C1 — C2 bond bundle	3.435	3.524	0.089	2.58
↪ *C1 bond wedge*	*2.073*	*2.157*	*0.084*	*4.07*
↪ *C2 bond wedge*	*1.363*	*1.367*	*0.004*	*0.31*
C1 — O1 bond bundle	3.362	3.353	-0.009	-0.28
↪ *C1 bond wedge*	*1.029*	*1.019*	*-0.010*	*-0.99*
↪ *O1 bond wedge*	*2.334*	*2.334*	*0.001*	*0.04*
C2 — H bond wedge	1.625	1.574	-0.051	-3.16
C2 — C3 bond bundle	3.214	3.171	-0.042	-1.31
↪ *C2 bond wedge*	*1.567*	*1.563*	*-0.004*	*-0.25*
↪ *C3 bond wedge*	*1.647*	*1.609*	*-0.038*	*-2.32*
C2 — H1 bond bundle	2.465	2.478	0.013	0.52
↪ *C2 bond wedge*	*1.652*	*1.683*	*0.030*	*1.84*
↪ *H1 bond wedge*	*0.813*	*0.795*	*-0.018*	*-2.17*
C3 — C bond wedge	1.769	1.842	0.073	4.15
C3 — C bond wedge	2.766	2.804	0.038	1.37
H1 — O2 bond bundle	3.662	3.617	-0.046	-1.25
↪ *H1 bond wedge*	*0.049*	*0.050*	*0.001*	*2.15*
↪ *O2 bond wedge*	*3.613*	*3.566*	*-0.047*	*-1.30*
H2 — Asp_{103} bond wedge	0.335	0.335	0.000	0.12
H2 — O1 bond bundle	3.539	3.543	0.005	0.13
↪ *H2 bond wedge*	*0.035*	*0.034*	*-0.001*	*-3.46*
↪ *O1 bond wedge*	*3.503*	*3.509*	*0.006*	*0.17*
H3 — Tyr_{16} bond wedge	0.321	0.300	-0.022	-6.77
H3 — O1 bond bundle	3.467	3.482	0.014	0.42
↪ *H3 bond wedge*	*0.059*	*0.067*	*0.009*	*14.66*
↪ *O1 bond wedge*	*3.408*	*3.414*	*0.006*	*0.17*
O2 — Asp_{40} bond wedge	2.114	2.109	-0.005	-0.24
O2 lone pair wedge	3.597	3.654	0.057	1.58
Total	**37.721**	**37.795**	**0.074**	**0.20**

H1 bond decreases by 0.018 electrons. Because the H1 atom's other bond wedge, corresponding to the O2$\cdots$H1 bond, only increases by 0.001 electrons, we conclude that the remaining 0.017 electrons—the amount lost by the H1 atomic basin—is transferred to the C2 atom rather than redistributed within the H1 atom. That is, shared charge density in the C2–H1 bond was lost by the H1 atom, responsible for at least half of the charge density gained in the corresponding C2 bond wedge; a precursor to its subsequent gains when the bond is broken. Furthermore, given the much lower electron counts of the O-bonded H atoms—around $0.37e$ in this case—we posit that the H1 atom, at $0.86e$, loses roughly half an electron during its abstraction to Asp_{40}. The EEF-induced decrease of $0.017e$ constitutes around 4% of the necessary H1 atomic charge depletion as dictated by the reaction, so here too the $\rho(\boldsymbol{r})$ response to the EEF appears to facilitate deprotonation. It's tempting to discuss this in terms of electron flow across boundaries. However, a more correct approach is to recognize that the boundaries themselves move.

3.2 Correlated gradient bundle property redistribution

The specification of real-space bonding regions manifests a number of new (or newly quantified) bond properties, some of which are geometric. To assess the contribution of boundary motion to regional property responses, we can compare gradient bundle field properties against gradient bundle geometric properties. We will include in our analysis two field properties—ρ and kinetic energy (T)[1]—and two geometric properties—volume (V) and bond wedge solid angle (α). The former represent fundamental chemical bonding concepts—bond electron count and bond energy respectively—that are quantified and generalized via their counterpart gradient bundle properties. The latter represent two of a number of regional geometric charge density descriptors, which are necessary components of a real-space chemical bond.[15] Using this approach, we will see that gradient bundle properties can vary independently in response to applied EEFs or enzyme mutation, and that these two types of perturbations affect gradient bundles in qualitatively different ways.

Although these geometric descriptors are newly computed $\rho(\boldsymbol{r})$ properties, they too, like the field properties, can have straightforward chemical interpretations. In addition to bond wedge V, we can describe a bond wedge by α. Conceptually, one can think of α as the percent of near-nuclear ("core") $\rho(\boldsymbol{r})$ contained within a bond wedge. In a methane molecule, each bond wedge on the sp^3

[1]Due to the virial theorem, gradient bundle kinetic energy is equal to minus the total energy,[14,50,51] and so relative changes to T give us insight into changes in total bond energy.

C atom would have α of ¼. In a benzene molecule, each sp^2 C atom is found to have a C–H bond wedge and two C–C bond wedges with respective α of 0.30, 0.35, and 0.35. For the sp^2 C3 atom from this study, which, in the WT KSI system, has α of 0.28, 0.30, and 0.43 (from Section 3.1 in the supplemental information). Bond wedge V and α both describe atomic partitioning and depend on bond wedge bounding surfaces, though α specifically describes the surfaces close to the nucleus.

As a system changes, these surfaces move, and this motion primarily determines redistribution of properties between bond wedges.[16,52] Numerically, this means that changes to geometric bond wedge properties (*e.g.* V and α) often predict changes to other bond wedge properties. However, bond wedge surface motion close to the nucleus can be different from the motion in the inter-atomic ("valence") region. For example, a decrease in α accompanied by an increase in V would indicate that the bond wedge surfaces contracted in the near-nuclear region and expanded in the inter-atomic region. Moreover, bond wedge property responses to a perturbation are not equal, and can be compared to assess whether a bond wedge merely expanded into more space with the same average property distribution (*e.g.* V, T, and ρ increase proportionately), or if the property content of the region actually changed (*e.g.* V decreases while T and ρ increase). Rather than inspect relative gradient bundle geometric and field property redistribution resulting from the r+ EEF directly, we will simultaneously compare the redistribution resulting from both types of perturbations.

Figure 6 shows gradient bundle property correlation matrices for a selection of atomic basins, bond bundles, and bond wedges, generated using the data from all five systems: wild type KSI (a.k.a. NEF), EEF-enhanced KSI^{r+}, EEF-inhibited KSI^{r-}, enhanced mutant KSI-Y,[32] and inhibited mutant KSI-Y.[57] With both types of catalytically enhancing and inhibiting perturbations considered, we see that the correlation among gradient bundle property responses varies from region to region. Within the H1 atomic basin at the reaction site (top-left), we see that V correlates weakly with ρ and T, indicating that atomic basin boundary motion occurs primarily in space with little charge or energy content. That is, the boundary motion responsible for this V change occurs in regions that contribute minimally to ρ and T. Contrast to the H2 atomic basin, part of the oxyanion hole, in which V, T, and ρ all strongly correlate. Similar contrast exists between the C1 and C3 atomic basins (left column, rows three and four), demonstrating that property correlation varies from atom to atom regardless of species. Note that α is always equal to 1 for an atomic basin, so it provides no information here.

Bond bundle property correlation is more dynamic. In the substrate carbonyl bond bundle (C1–

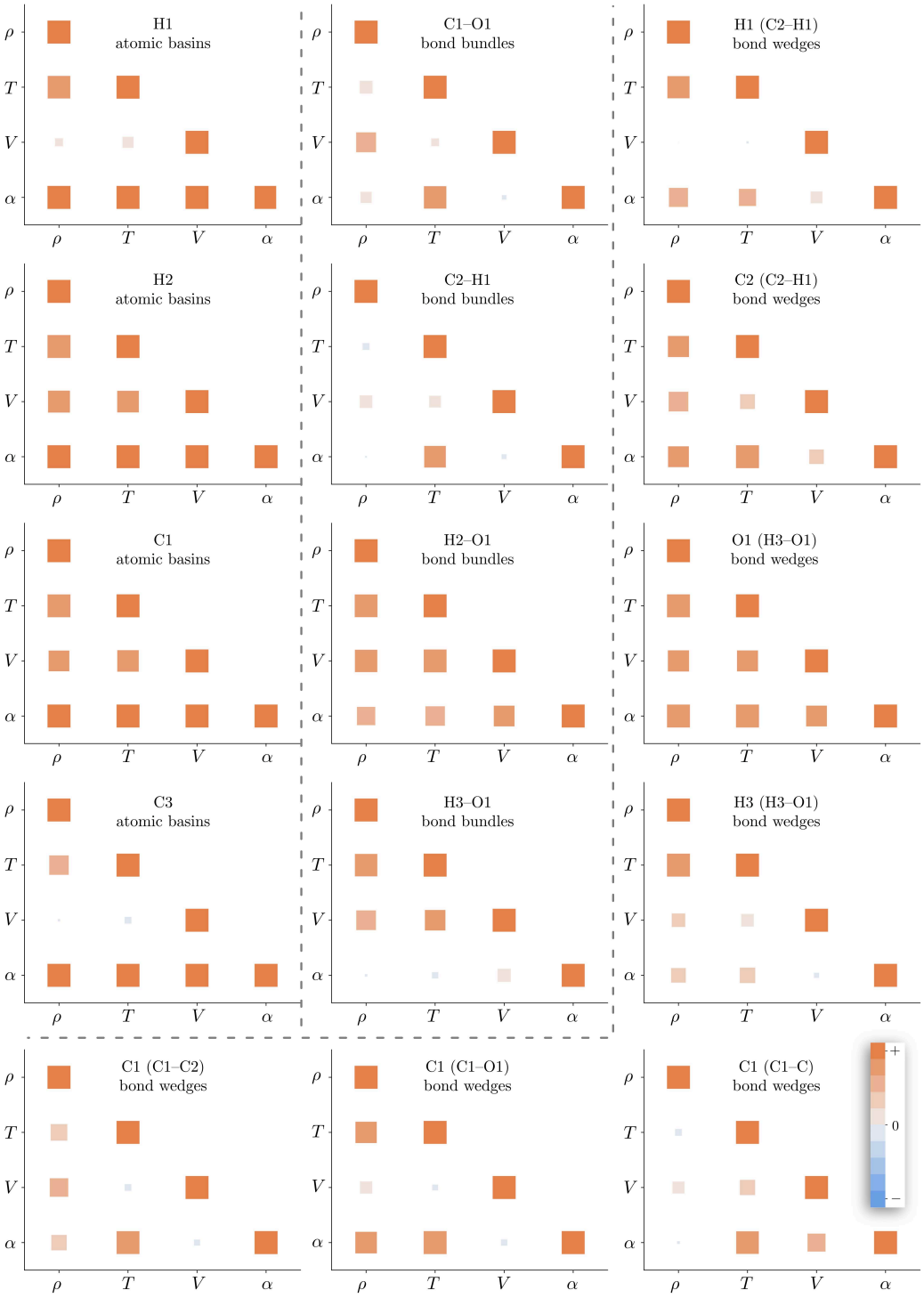


Figure 6: Correlation matrices showing property correlations within a selection of gradient bundles. Electron density (ρ), kinetic energy (T), volume (V), and solid angle (α) are included. Full sets of correlation matrices are provided in the SI, Section 7.

Table 2: Regional property changes for each system relative to NEF KSI for the C1 bond wedge corresponding to the C1 bond with the unnumbered C atom.

C1 (C1–C)	ρ (NEF : 2.049)		T (NEF : 13.899)		V (NEF : 21.068)		α (NEF : 0.3688)	
bond wedges	Δ	%Δ	Δ	%Δ	Δ	%Δ	Δ	%Δ
r+	-0.040	-1.93	-0.306	-2.21	-0.183	-0.87	-0.0065	-1.75
r-	-0.030	-1.45	-0.204	-1.47	-0.360	-1.71	-0.0047	-1.28
Y^{32}	-0.001	-0.05	-0.339	-2.44	5.608	26.62	-0.0091	-2.48
Y^{57}	-0.014	-0.67	-0.382	-2.75	5.433	25.79	-0.0101	-2.73

O1; top-center of Figure 6), α and V weakly anticorrelate, suggesting that inter-atomic region bond bundle surface motion is independent (and somewhat opposite) from near-nuclear region surface motion. Also in the C1–O1 bond bundle, ρ and T weakly correlate with each other, but strongly correlate with V and α respectively, indicating that changes in T result primarily from near-nuclear region bond bundle surface motion (where electronic kinetic energy is greatest), while inter-atomic surface motion is responsible for changing ρ. The ability of T and ρ to change independently is most stark in the C2–H1 bond bundle at the reaction site (center column, second row), where they anticorrelate, and again (as with the C1–O1 bond bundle) this results from differences in bond bundle surface motion in the near-nuclear versus inter-atomic region, as evidenced by anticorrelation between V and α. Not all bond bundles present such dynamic property correlation, however. Considering the carbonyl O atom's lone pair coordinated hydrogen bonds (center column, rows three and four), within the H2–O1 bond bundle, all four properties strongly correlate. Contrast to the other O atom lone pair hydrogen bond, H3–O1, where α more weakly correlates with V and anticorrelates with both T and ρ, so again the bond bundle surface motion near the nuclear CP is quite different from that far from the nuclear CP.

Varying bond bundle property correlations are partially explained as the combination of their constituent bond wedges. For example, the H3–O1 bond bundle property correlation appears to combine that of the O1 bond wedge component (right column, third row), in which all four properties correlate, and that of the H3 bond wedge component (right column, fourth row), which seems responsible for the dynamic correlation of the resulting bond bundle.

While uncorrelated behavior between T and ρ can be accompanied by a lack of correlation between α and V, this is not always the case, as in the H3 component of the H3–O1 bond bundle where ρ and T strongly correlate *despite* anticorrelation between V and α. In this case it appears that the H3 bond wedge surface motion differs in the near-nuclear versus inter-atomic regions, but

ρ and T occupy both regions in the same proportions, such that boundary motion does not affect one more than the other. The same behavior is evident in the C1 bond wedge components of the C1–O1 and C1–C2 bond bundles (bottom row, left two columns). More interesting is the opposite case, where α and V do correlate, while T and ρ do not, as in the C1 atom's third bond wedge to another (unnumbered) C atom (bottom row, right column). Here α and V correlate, indicating similar motion of near-nuclear and inter-atomic bond wedge surfaces, but regardless, ρ and T weakly anticorrelate. More ambiguous than the previous cases, we turn to the tabulated data in Table 2 to find that this behavior stems from disproportionate changes to bond wedge properties resulting from applied EEFs compared to enzyme mutation. Specifically, T, ρ, and α are lowered in response to all four perturbations, though the response of T is proportionate with respect to both types, but not the response of ρ, which is much less responsive to chorination of amino acids than to applied EEFs. Meanwhile, bond wedge V dramatically increases in response to atomic substitution, while slightly decreasing in response to applied EEFs. When the C1 atomic basin is considered as a whole (Figure 6; left column, third row) this uncorrelated behavior vanishes. We conclude that the lack of correlation between ρ and the other three properties in the C1–C bond wedge in Figure 6 results from the two mutant systems for which bond wedge ρ barely changes relative to the other properties, in marked contrast to the effects of applied EEFs.

These observations show that gradient bundle properties respond differently to chorination of amino acids than to applied EEFs. Although T, V, and α are all functionals of ρ, they can respond independently of ρ, as in Table 2 where enzyme mutations and applied EEFs had the same relative effect on T, but dramatically different effects on ρ and V.

3.3 KSI catalytic enhancement and inhibition

Having demonstrated that enzyme mutation affects gradient bundle properties differently than applied EEFs, we now investigate the relationship between this redistribution and changes to catalytic activity. As in References 6 and 11, we now check for linear correlations between reaction barrier energy and regional properties across the five systems: wild type KSI (a.k.a. NEF), EEF-enhanced KSI;$^{r+}$ EEF-inhibited KSI;$^{r-}$ enhanced mutant KSI-Y;[32] and inhibited mutant KSI-Y.[57] We start by inspecting some particularly relevant individual correlations, then move to a method of viewing correlations of all regions and properties simultaneously. We will find that the bond bundle and bond wedge $\rho(\boldsymbol{r})$ redistribution accompanying the r+ EEF is well represented in the reaction barrier energy

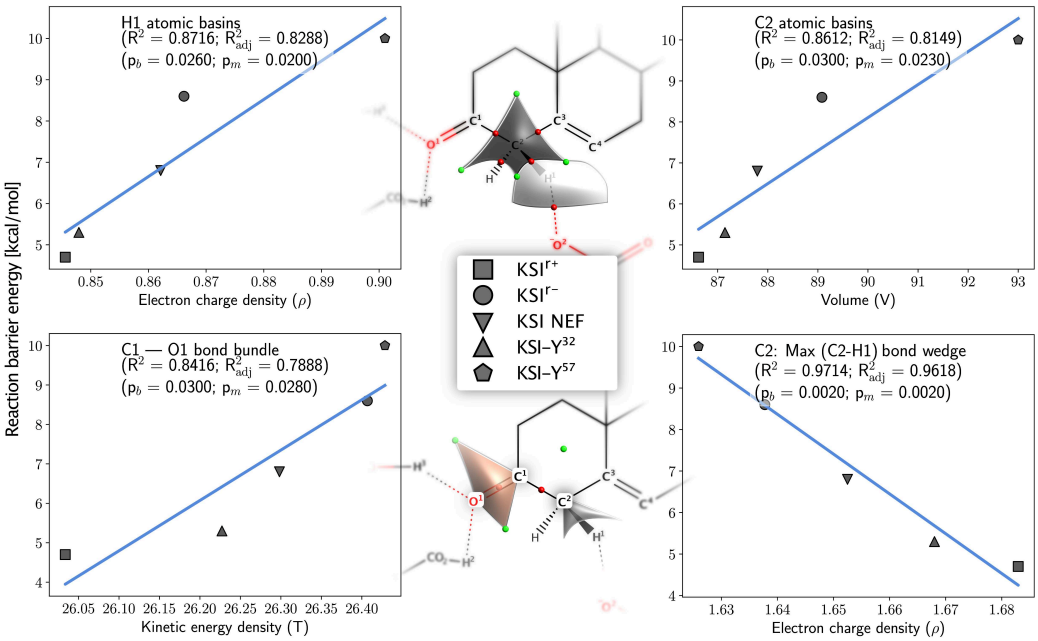


Figure 7: Sampling of atomic basin and bond bundle properties correlated against KSI-catalyzed deprotonation reaction barrier energy. Center: Schematic representation of the atomic basins (top) and bond bundles (bottom) shown. Sides: Plots of regional properties vs barrier energy. Full sets of plots used to generate correlations presented in this manuscript are available in the SI, Section 5.

correlations, but that this is not the case for atomic basins, for which there are notable exceptions. By considering correlations between system energy and gradient bundle properties, we will see that the mechanism by which amino acid mutation affects catalytic proficiency appears different from that of applied EEFs, and that this difference stems from the ability of gradient bundle properties to independently vary.

Beginning with inspection of individual correlations, Figure 7 shows a selection of some of the fits of reaction barrier height as a function of different gradient bundle condensed properties (see Section 5 in the supplemental information for the complete set of plots for all systems, and section SI-3 for the corresponding tabulated gradient bundle integration data).[2] Atomic basin correlations were strongest for the C2 and H1 atoms. Specifically, the H1 atomic ρ correlates positively with reaction

[2]To aid in the statistical interpretation of the fits, in addition to the correlation coefficient (R^2) we have included the adjusted correlation coefficient (R^2_{adj}) and p-values for the constant and linear fit coefficients (p_b and p_m respectively). R^2_{adj} is always $\leq R^2$, and a large difference between the two indicates that there are unhelpful independent variables in the regression model. For example, in the two top plots in Figure 7, the points are skewed to the left, that is, the left half of the x-axis is more thoroughly sampled than is the right half, resulting in R^2_{adj} being about 0.054 less than R^2. Meanwhile, the C2 bond wedge fit (bottom-right) has more even sampling of the x-axis, and less of a resulting penalty in its R^2_{adj}. P-values indicate the likelihood that a given fit can be explained by the null hypothesis; that it is explained by chance. A p-value of less than 0.05 conventionally indicates that the null hypothesis can be rejected and that the fit is statistically significant. Section 5 of the supplementary information contains additional measures of statistical significance, independence, and separation.

barrier, as does the C2 atomic V. Intra-atomic property redistribution additionally indicates that property shifts within the conjugated substrate carbon system—*e.g.* carbon atom rehybridization—play an important role. In this case, positive correlation of O1=C1 bond T and negative correlation of the C2-to-H1 bond wedge ρ together indicate that, to lower the reaction barrier, charge should increase in the C2 bond wedge (gained from the heterolytic cleavage of the C2–H1 bond), and kinetic energy should decrease in the O1=C1 bond (thus increasing total energy, weakening the carbonyl bond). This is in agreement with results in the previous section, where the catalyzing EEF did in fact produce these property shifts.

Moving on to the simultaneous inspection of correlations of all regions and properties, it is helpful to simplify and plot multiple correlation coefficients at once to see how *e.g.* volume and energy each correlate, and to access a more immediate chemical interpretation. Inspired graphically by Reference 53, we have included correlations of regional ρ, T, V, and α as bar charts in Figure 8. Regions are sorted according to the reaction barrier correlations of ρ, which is also used to graphically shade representative regions as in Figure 5. Note that the shading has opposite meaning between the two figures, in regards to energetic significance. In Figure 8 a blue-shaded region anticorrelates with reaction barrier energy, so an increase in regional properties should lower the reaction barrier. Unlike in Figure 5, where an orange region was one in which ρ increased in response to a catalyzing EEF. Regions and properties that resulted in low correlation coefficients are still included in the bar charts in order to contrast the full set of regions. Redistribution of $\rho(\boldsymbol{r})$ and other gradient bundle integrated properties is more correlated to reaction barrier energy for some regions/properties than others.

Looking at ρ redistribution between bond bundles, the C1–C2 bond most anticorrelates with reaction barrier energy, indicating that promoting (or hindering) its transition from single to double bond generally lowers (or raises) the reaction barrier. The O1=C1 and O1$\cdots$H2 bond bundles respectively correlate and anticorrelate with barrier height, indicating that activation of the carbonyl bond—by increasing the O1 atom lone pair density and decreasing the carbonyl bond density—lowers the barrier. At the deprotonation site, the C2–H1 bond properties anticorrelate with barrier height, paradoxically suggesting that the electron count (and other properties, especially T) of the breaking bond should increase rather than decrease, which is again resolved by inspecting the bond's constituent bond wedges. Altogether, the bond bundle property correlations seem to indicate the same underlying catalytic charge density shifts as those resulting from an applied EEF, including

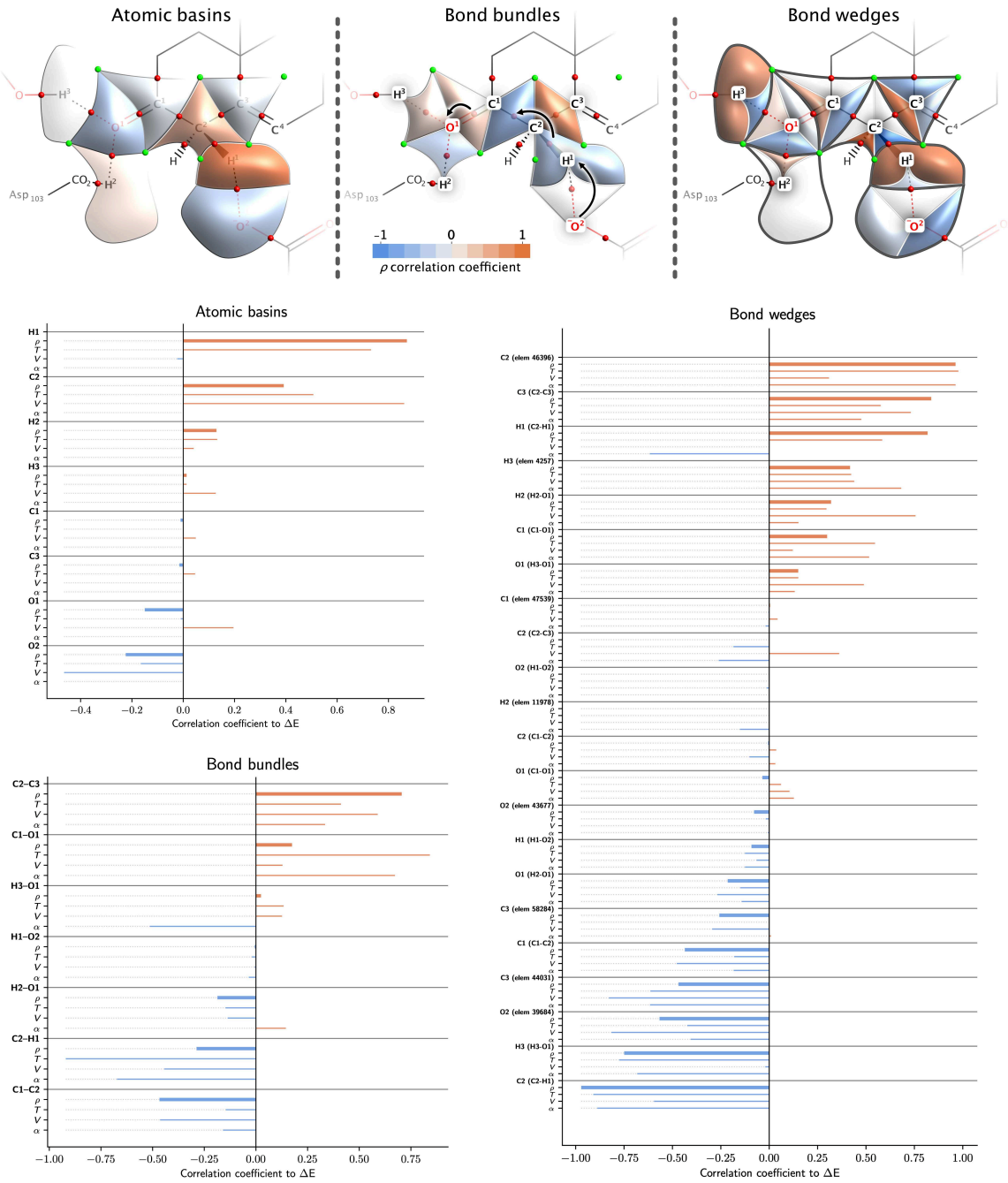


Figure 8: Atomic basin, bond bundle, and bond wedge (top; left to right) property correlations with reaction barrier energy, signed as positive or negative to indicate correlation and anti-correlation. Above are cartoon, schematic depictions of the overlaid on the Lewis representation of the active site, with electron pushing arrows in the center pane denoting the deprotonation reaction step from Scheme 1. Regions are shaded above, and sorted in the plots below, according to the sign and magnitude of their electron population (ρ) correlations with reaction barrier. The regional kinetic energy (T), volume (V), and normalized solid angle (α) are plotted as well. Complete multi-variable correlation bar charts are available in the SI, Section 4.

the reversed behavior at the deprotonation site. Here the results indicate a structure property relationship between reactant state substrate bond bundle properties and the barrier height of the KSI-catalyzed reaction, regardless of the perturbation causing the bond bundle properties to change.

Considering the correlations of atomic basin properties, the C2 and H1 atoms at the deprotonation site correlate with barrier height, indicating their combined atomic electron count should be lowered in order to facilitate deprotonation. The O1 atom weakly anticorrelates, suggesting that it should be activated via accumulating charge density. Both of these match the redistribution that resulted from the applied r+ field.

However, while the conjugated carbon system experienced the greatest amount of inter-atomic ρ redistribution due to the EEF (Figure 5), those (C1 and C3) atomic properties show almost no correlation with reaction barrier. Instead, the H1 and C2 atoms most strongly correlate. This suggests that accumulation of $\rho(\boldsymbol{r})$ in the C1 and C3 atomic basin due to the r+ EEF is not a general requirement for catalytic enhancement. On the other hand, the r+ EEF-induced bond bundle redistribution does seem indicative of the general behavior shown in Figure 8.

Indeed, the (anti)similarity between bond bundle correlations and EEF-induced bond bundle property shifts is stark. With very few exceptions, regions of $\rho(\boldsymbol{r})$ accumulation due to the r+ field (Figure 5) are those that anticorrelate with reaction barrier (Figure 8). Likewise, regions of $\rho(\boldsymbol{r})$ depletion due to the field correlate positively with reaction barrier. For example, the C1–C2 bond bundle, which experienced the greatest increase in $\rho(\boldsymbol{r})$ due to the EEF, anticorrelates with barrier height. Given that the r+ system was included in the regional correlations, and that it is the most rate-enhancing system in the (small) sample, it is no surprise that it should be represented in the resulting correlations. However, if KSI$^{\mathrm{r+}}$ was simply dominating the correlations, one would expect all regions to match, but as noted above, atomic basin regional correlations share less similarity with their r+ field-induced redistribution in KSI. This interplay can only be recovered via bond wedge and bond bundle analysis.

We now return to the observation that the H1 and C2 atomic basins most strongly correlate with reaction barrier energy, and how it relates to the underlying mechanism of KSI catalysis. In general, the catalytic effect of an oriented EEF depends on the magnitude of the field, its angle relative to the dipole of the reaction coordinate (*e.g.* parallel to a bond that is broken or formed in the reaction), and the magnitude of that dipole.[54–66] The local electrostatic preorganization in KSI most underlying its catalytic proficiency is that within the oxyanion hole, which is aligned along the substrate carbonyl

bond (the r+ direction).[20,66] One might ask how nature determined that the best way to achieve this atomic basin redistribution is via a carbonyl-oriented field, rather than a field oriented along the reaction coordinate (C2–H1 bond). We previously observed that a catalyzing EEF, applied to a simple Diels-Alder reaction, primarily shifted electron density between bonds bundles *within* rather than *between* reactant molecules, even though the field was applied *along the reaction coordinate.*[16] That is, the field-aligned bond was not the most affected. Here too we observe, in Figure 5, that the r+ field causes bond bundle redistribution *within* the substrate molecule through its conjugated carbon system, and that this results in the same H1 and C2 atomic basin electron count changes that most correlate with reaction barrier energy. Additionally, the carbonyl is a better candidate for electric field-induced transition state stabilization because its dipole is strong and does not reorient during the reaction, while the C–H (reaction coordinate) dipole is weak and does reorient. Thus the carbonyl provides a more consistent EEF "grip" on the molecule (dipole orientation), and more "leverage" (dipole magnitude) for inducing intra-molecular bond bundle redistribution of charge and other properties.

A final and somewhat unexpected similarity to the results of the previous section is the additivity of bond wedge property correlations to predict those of bond bundles and atomic basins. For example, the two H3 bond wedges strongly correlate and anticorrelate respectively, and the H3 atomic basin has nearly zero correlation, as if its bond wedges combined to give the whole. The same behavior is apparent in the C1 and C3 atoms. Bond wedge property correlations with reaction barrier height combine to predict those of bond bundles in a similar way, and in this case without exception; the correlation of each bond bundle appears to be the sum of its bond wedge correlations. While gradient bundle properties are definitionally additive, it cannot be said that gradient bundle energetic significance is additive, in this case in relation to the catalyzed reaction barrier. Regardless, bond wedge regional energetic significance does seem to combine to predict bond bundle and atomic basin significance. How far this energetic additivity extends has yet to be investigated, but it is a welcome exception to the lack of such additivity that has been noted as a major obstacle in enzymology.[19]

Regarding the nearly uniform agreement between the correlations of different regional properties with reaction barrier energy, we observed in Figure 6 that regional V, T, and α often correlate strongly with ρ, as we have shown previously for organic systems.[14] Hence, it is unsurprising that these properties should yield similar correlations with respect to reaction barrier height, as in KSI.

However, there are exceptions. For example, within the C1–O1 bond bundle in Figure 8, T and α correlate more strongly with reaction barrier ($R^2 \approx 0.7$) than do ρ and V ($R^2 \approx 0.2$). Disagreements of this type typically indicate weak correlation between the properties themselves, in this case indicating that within the C1–O1 bond bundle, T and α do not correlate with ρ and V, just as we observed in Figure 6. The connection between regional property correlation and reaction barrier height correlation is observable for almost all regions in the study. The H1 atomic basin, for example, has strong ρ and T barrier correlation ($R^2 \approx 0.8$) but weak V anti-correlation ($R^2 \approx 0.02$), and we saw in Figure 6 that ρ and T correlate strongly with each other and weakly with V. The C2–H1 bond wedge is particularly relevant because it yields some of the strongest overall correlations to reaction barrier height, though V correlates more weakly ($R^2 \approx 0.55$) than ρ, T, or α ($R^2 \approx 0.89$ to 0.98), and in Figure 6, ρ, T, and α correlate strongly with each other and more weakly with V.

Because the correlation among regional properties appears to be closely tied to their relative correlations with reaction barrier height, it appears that different types of system changes (*e.g.* an applied EEF vs. amino acid mutation) enhance catalysis through different mechanisms that affect and utilize some gradient bundle properties more than others. We conclude that weak correlation between a region's properties indicates the region is affected by different perturbations in different ways. Conversely, strong correlation between a region's properties, *together with strong correlation to reaction barrier height* indicates that the region has similar energetic significance and catalytic functionality regardless of the type of system change, *i.e.* that it more fundamentally underlies the catalysis. In future investigations involving larger samples, *e.g.* many mutant and EEF-exposed systems, we can further test this conclusion by comparing results of similar analyses on EEF-only, mutant-only, and mixed subsamples. Additionally, by including the local electric field in the analysis, we can directly address the pivotal role of electrostatic preorganization, *e.g.* the strong observed link between $|\boldsymbol{E}(\boldsymbol{r})|$ at the substrate carbonyl and KSI catalytic activity (see Figure 4d of Reference 66).

Overall, in regards to the mechanism of KSI catalytic augmentation, this approach leads us to chemically similar conclusions to those of direct inspection of EEF-induced bond bundle property redistribution from the previous section. Activation of the substrate carbonyl bond, and charge accumulation in the C1–C2 bond bundle, as result from the catalyzing r+ EEF, here suggest a more general structure-property relationship between KSI active site gradient bundle properties and reaction barrier height. Correlations of reaction barrier energy to bond bundle regional properties recover a picture of property redistribution that resembles the expected qualitative electron redistribution of

the catalyzed chemical reaction, as implied with electron-pushing formalisms, but in unambiguous, quantitative terms, absent the presuppositions of organic chemistry. Surprisingly, we also found that bond wedge correlations to reaction barrier energy appear to combine to predict those of bond bundles and atomic basins. That is, typically the energetic relevance of bond wedges appears to be additive, contrary to the conventional notion of lacking energetic additivity in enzymes, such that the sum of local energies cannot be used to deduce resulting changes to reactivity.[19] We also expect that this approach can be used to investigate reactant state destabilization in addition to transition state stabilization. Lastly, by extending this analysis to the many computable properties of gradient bundles, we saw that not all properties correlate with reaction barrier to the same degree, and that these differences typically stem from weak correlation between the regional properties themselves.

4 Conclusion

Here we have presented a method of computing the energies and energy-mediated properties of real-space bonding regions in any chemical system, applied to the specific problem of KSI catalysis. We inspected the redistribution of $\rho(\boldsymbol{r})$ due to a catalyzing EEF in the KSI active site, and directly observed, without the use of chemical presuppositions, that the field catalyzes by shifting charge in the forward reaction direction. Extending the analysis to include multiple perturbed KSI systems, we recovered a similar overall picture of property redistribution, generally descriptive of KSI augmentation via either chlorination of amino acids or applied EEFs. Moreover, even though KSI can be enhanced or inhibited by either type of perturbation, they affect gradient bundle properties differently, and hence augment KSI through different mechanisms.

The surfaces separating bond wedges, along with concepts like bond wedge solid angle that describe their relative atomic occupation, are necessary physical components of a real-space chemical bond. In the same way, the ability of a bond wedge to change one integrated property independent from or in different proportion to its other properties (Figure 6) is something expected of a well-defined, real-space chemical system. These concepts are put to abundant qualitative and quantitative use in other chemical systems, *e.g.* via equations of state. Gradient bundle analysis provides the same spatial and energetic grounding at the atomic scale, where regional additivity can be assumed and leveraged. The gradient bundle properties presented here, together with the larger integration tables contained in the supplementary information, are but a small sample of what can currently

be calculated. We have only begun to investigate the statistical and conceptual significance of this extensive set of properties.

Constructing a complete, descriptive network of enzyme $\rho(\boldsymbol{r})$ structure-property relationships, such that local changes in $\rho(\boldsymbol{r})$ could be used to accurately predict catalytic rate enhancement, will be a formidable task owing to the lack of underlying energetic additivity between the independent properties of enzymes.[19] Bond bundles do possess spatial and energetic additivity, are uniquely defined in any chemical system, and their analysis appears to naturally leverage our hard-won understanding of chemical bonding. Furthermore, in this case the correlation of bond wedge properties with rate enhancement does, in fact, appear to be additive, so there is a readily quantifiable sense in which catalyzing features may have underlying energetic additivity.

Supporting information

Supporting information is provided, containing:

- 3D depictions of some of the regions discussed in the manuscript
- Coordinates and energies of *ab-initio* calculations
- Atomic basin, bond bundle, and bond wedge tabulated integration values *with many more properties than are included in this manuscript*
- Multi-variable bar charts of simplified regional property correlations
- Plots of fits of each region and property with reaction barrier energy
- Correlations among regional properties

Acknowledgments

This work was supported by the National Science Foundation grant CHE-1903808, and by the Office of Naval Research grant N00014-05-C-0241.

References

[1] Levitt, M.; Warshel, A. Computer Simulation of Protein Folding. *Ann. Soc. R. Sci. Med. Nat. Brux.* **1975**, *253*, 694.

[2] Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley-Interscience: New York, 1997.

[3] Messer, B. M.; Roca, M.; Chu, Z. T.; Vicatos, S.; Kilshtain, A. V.; Warshel, A. Multiscale Simulations of Protein Landscapes: Using Coarse-Grained Models as Reference Potentials to Full Explicit Models. *Biochim. Biophys. Acta, Mucoproteins Mucopolysaccharides* **2010**, *78*, 1212–1227.

[4] Kim, K. S.; Oh, K. S.; Lee, J. Y. Catalytic Role of Enzymes: Short Strong H-bond-induced Partial Proton Shuttles and Charge Redistributions. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6373–6378.

[5] Oh, K. S.; Cha, S.-S.; Kim, D.-H.; Cho, H.-S.; Ha, N.-C.; Choi, G.; Lee, J. Y.; Tarakeshwar, P.; Son, H. S.; Choi, K. Y. *et al.* Role of Catalytic Residues in Enzymatic Mechanisms of Homologous Ketosteroid Isomerases,. *Acta Chem. Scand. Ser. B* **2000**, *39*, 13891–13896.

[6] Hennefarth, M. R.; Alexandrova, A. N. Direct Look at the Electric Field in Ketosteroid Isomerase and Its Variants. *ACS Catal.* **2020**, *10*, 9915–9924.

[7] Warshel, A. Electrostatic Basis of Structure–Function Correlation in Proteins. *Acc. Chem. Res.* **1981**, *14*, 284–290.

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

[9] 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.

[10] Hennefarth, M. R.; Alexandrova, A. N. Advances in Optimizing Enzyme Electrostatic Preorganization. *Curr. Opin. Struct. Biol.* **2022**, *72*, 1–8.

[11] Fuller, J.; Wilson, T. R.; Eberhart, M. E.; Alexandrova, A. N. Charge Density in Enzyme Active Site as a Descriptor of Electrostatic Preorganization. *J. Chem. Inf. Model.* **2019**, *59*, 2367–2373.

[12] Vargas, S.; Hennefarth, M. R.; Liu, Z.; Alexandrova, A. N. Machine Learning to Predict Diels–Alder Reaction Barriers from the Reactant State Electron Density. *J. Chem. Theory Comput.* **2021**, *17*, 6203–6213.

[13] Jones, T. E.; Eberhart, M. E. The Bond Bundle in Open Systems. *Int. J. Quantum Chem.* **2010**, *110*, 1500–1505.

[14] Wilson, T. R.; Rajivmoorthy, M.; Goss, J.; Riddle, S.; Eberhart, M. E. Observing the 3D Chemical Bond and Its Energy Distribution in a Projected Space. *ChemPhysChem* **2019**, *20*, 3289–3305.

[15] Wilson, T. R.; Alexandrova, A. N.; Eberhart, M. E. Electron Density Geometry and the Quantum Theory of Atoms in Molecules. *J. Phys. Chem. A* **2021**, *125*, 10622–10631.

[16] Wilson, T.; Eberhart, M. In *Advances in Quantum Chemical Topology Beyond QTAIM, Chapter 17*, 1st ed.; Rodríguez, J. I., Guzman, F., Anderson, J. S. M., Eds.; Elsevier: Netherlands, 2022.

[17] Hoffmann, R. Qualitative Thinking in the Age of Modern Computational Chemistry—or What Lionel Salem Knows. *J. Mol. Struct. THEOCHEM* **1998**, *424*, 1–6.

[18] Kraut, D. A.; Sigala, P. A.; Pybus, B.; Liu, C. W.; Ringe, D.; Petsko, G. A.; Herschlag, D. Testing Electrostatic Complementarity in Enzyme Catalysis: Hydrogen Bonding in the Ketosteroid Isomerase Oxyanion Hole. *PLoS Biol.* **2006**, *4*, e99.

[19] Herschlag, D.; Natarajan, A. Fundamental Challenges in Mechanistic Enzymology: Progress toward Understanding the Rate Enhancements of Enzymes. *Acta Chem. Scand. Ser. B* **2013**, *52*, 2050–2067.

[20] Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme Electric Fields Power Catalysis in the Active Site of Ketosteroid Isomerase. *IEEE Comput. Sci. Eng. Mag.* **2014**, *346*, 1510–1514.

[21] Warshel, A.; Sharma, P. K.; Chu, Z. T.; Åqvist, J. Electrostatic Contributions to Binding of Transition State Analogues Can Be Very Different from the Corresponding Contributions to

Catalysis: Phenolates Binding to the Oxyanion Hole of Ketosteroid Isomerase. *Acta Chem. Scand. Ser. B* **2007**, *46*, 1466–1476.

[22] 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.; Oh, B.-H. High-Resolution Crystal Structures of Δ^5 -3-Ketosteroid Isomerase with and without a Reaction Intermediate Analogue. *Acta Chem. Scand. Ser. B* **1997**, *36*, 14030–14036.

[23] Sigala, P. A.; Fafarman, A. T.; Schwans, J. P.; Fried, S. D.; Fenn, T. D.; Caaveiro, J. M. M.; Pybus, B.; Ringe, D.; Petsko, G. A.; Boxer, S. G. *et al.* Quantitative Dissection of Hydrogen Bond-Mediated Proton Transfer in the Ketosteroid Isomerase Active Site. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E2552–E2561.

[24] Hennefarth, M. R.; Alexandrova, A. N. Heterogeneous Intramolecular Electric Field as a Descriptor of Diels–Alder Reactivity. *J. Phys. Chem. A* **2021**, *125*, 1289–1298.

[25] Bader, R. F. W.; Nguyen-Dang, T. T.; Löwdin, P.-O. *Advances in Quantum Chemistry*; Academic Press, 1981; Vol. Volume 14; pp 63–124.

[26] Matta, C. F.; Boyd, R. J. *The Quantum Theory of Atoms in Molecules: From Solid State to DNA and Drug Design*; Weinheim: Wiley-VCH, 2007.

[27] Morgenstern, A.; Wilson, T.; Miorelli, J.; Jones, T.; Eberhart, M. E. In Search of an Intrinsic Chemical Bond. *Comput. Theor. Chem.* **2015**, *1053*, 31–37.

[28] Morgenstern, A.; Eberhart, M. Bond Dissociation Energies from the Topology of the Charge Density Using Gradient Bundle Analysis. *Phys. Scr.* **2016**, *91*, 23012.

[29] Wilson, T. R.; Eberhart, M. Quantum Theory of Atoms in Molecules in Condensed Charge Density Space. *Can. J. Chem.* **2019**, *97*, 757–762.

[30] 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.

[31] 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.

[32] te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. Chemistry with ADF. *J. Comput. Chem.* **2001**, *22*, 931–967.

[33] Fonseca Guerra, C.; Snijders, J. G.; te Velde, G.; Baerends, E. J. Towards an Order-N DFT Method. *Theor. Chem. Acc.* **1998**, *99*, 391–403.

[34] ADF 2019.3, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands. http://www.scm.com, 2019.

[35] Van Lenthe, E.; Baerends, E. J. Optimized Slater-type Basis Sets for the Elements 1–118. *J. Comput. Chem.* **2003**, *24*, 1142–1156.

[36] Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125*, 194101.

[37] Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[38] Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, *99*, 2224–2235.

[39] Pye, C. C.; Ziegler, T. An Implementation of the Conductor-like Screening Model of Solvation within the Amsterdam Density Functional Package. *Theor. Chem. Acc.* **1999**, *101*, 396–408.

[40] Wilson, T. R.; Eberhart, M. E. Bondalyzer. https://github.com/moltheorygroup/BondalyzerTecplotAddon, 2022.

[41] Inc., T. Tecplot 360 2013R1. https://www.tecplot.com, 2013.

[42] Hunter, J. D. Matplotlib v3.5.2. *Comput. Sci. Eng.* **2007**, *9*, 90–95.

[43] Waskom, M. L. Seaborn: Statistical Data Visualization. *J. Open Source Softw.* **2021**, *6*, 3021.

[44] Wilson, T. R. Gradient Bundle Analysis Post-Processing: Direct and Statistical Comparisons. https://github.com/moltheorygroup/gbapost, 2022.

[45] Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272.

[46] Seabold, S.; Perktold, J. Statsmodels: Econometric and Statistical Modeling with Python. 9th Python in Science Conference. 2010.

[47] MarvinSketch V21.19. ChemAxon, 2022.

[48] Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics v1.9.4. *J. Mol. Graphics* **1996**, *14*, 33–38.

[49] Affinity Designer v1.10.5. Serif, 2022.

[50] Ruedenberg, K.; Schmidt, M. W. Physical Understanding through Variational Reasoning: Electron Sharing and Covalent Bonding. *J. Phys. Chem. A* **2009**, *113*, 1954–1968.

[51] Slater, J. C. The Virial and Molecular Structure. *J. Chem. Phys.* **1933**, *1*, 687–691.

[52] Morgenstern, A.; Morgenstern, C.; Miorelli, J.; Eberhart, M. E.; Wilson, T. The Influence of Zero-Flux Surface Motion on Chemical Reactivity. *Phys. Scr.* **2016**, *18*, 5638–5646.

[53] Luchi, A. M.; Villafañe, R. N.; Gómez Chávez, J. L.; Bogado, M. L.; Angelina, E. L.; Peruchena, N. M. Combining Charge Density Analysis with Machine Learning Tools To Investigate the Cruzain Inhibition Mechanism. *ACS Omega* **2019**, *4*, 19582–19594.

[54] Meir, R.; Chen, H.; Lai, W.; Shaik, S. Oriented Electric Fields Accelerate Diels–Alder Reactions and Control the Endo/Exo Selectivity. *ChemPhysChem* **2010**, *11*, 301–310.

[55] Aragonès, A. C.; Haworth, N. L.; Darwish, N.; Ciampi, S.; Bloomfield, N. J.; Wallace, G. G.; Diez-Perez, I.; Coote, M. L. Electrostatic Catalysis of a Diels–Alder Reaction. *Ann. Soc. R. Sci. Med. Nat. Brux.* **2016**, *531*, 88–91.

[56] Shaik, S.; Mandal, D.; Ramanan, R. Oriented Electric Fields as Future Smart Reagents in Chemistry. *Nat. Chem.* **2016**, *8*, 1091–1098.

[57] Wang, Z.; Danovich, D.; Ramanan, R.; Shaik, S. Oriented-External Electric Fields Create Absolute Enantioselectivity in Diels–Alder Reactions: Importance of the Molecular Dipole Moment. *J. Am. Chem. Soc.* **2018**, *140*, 13350–13359.

[58] Wang, C.; Danovich, D.; Chen, H.; Shaik, S. Oriented External Electric Fields: Tweezers and Catalysts for Reactivity in Halogen-Bond Complexes. *J. Am. Chem. Soc.* **2019**, *141*, 7122–7136.

[59] Mattioli, E. J.; Bottoni, A.; Zerbetto, F.; Calvaresi, M. Oriented External Electric Fields Affect Rate and Stereoselectivity of Electrocyclic Reactions. *J. Phys. Chem. C* **2019**, *123*, 26370–26378.

[60] Shaik, S.; Danovich, D.; Joy, J.; Wang, Z.; Stuyver, T. Electric-Field Mediated Chemistry: Uncovering and Exploiting the Potential of (Oriented) Electric Fields to Exert Chemical Catalysis and Reaction Control. *J. Am. Chem. Soc.* **2020**, *142*, 12551–12562.

[61] Liu, F.; Liu, M. An Intuitive Electric-field Contribution Decomposition Model for Chemical Processes and Its Applications on Diels-Alder Reactions. *Chem. Res. Chin. Univ.* **2020**, *36*, 1241–1248.

[62] Yu, S.; Vermeeren, P.; Hamlin, T. A.; Bickelhaupt, F. M. How Oriented External Electric Fields Modulate Reactivity. *Chem. Eur. J.* **2021**, *27*, 5683–5693.

[63] Shaik, S.; Danovich, D.; Dubey, K. D.; Stuyver, T. *Effects of Electric Fields on Structure and Reactivity*; The Royal Society of Chemistry, 2021; pp 12–70.

[64] Léonard, N. G.; Dhaoui, R.; Chantarojsiri, T.; Yang, J. Y. Electric Fields in Catalysis: From Enzymes to Molecular Catalysts. *ACS Catal.* **2021**, *11*, 10923–10932.

[65] Sowlati-Hashjin, S.; Karttunen, M.; Matta, C. F. Manipulation of Diatomic Molecules with Oriented External Electric Fields: Linear Correlations in Atomic Properties Lead to Nonlinear Molecular Responses. *J. Phys. Chem. A* **2020**, *124*, 4720–4731.

[66] Fried, S. D.; Boxer, S. G. Electric Fields and Enzyme Catalysis. *Annu. Rev. Biochem.* **2017**, *86*, 387–415.