Title

PyCPET ▫ Computing Heterogeneous 3D Protein Electric Fields and Their Dynamics

Journal

Journal of Chemical Theory and Computation, 21(8)

ISSN

1549-9618

Authors

Ajmera, Pujan
Vargas, Santiago
Chaturvedi, Shobhit S
et al.

Publication Date

2025-04-22

DOI

10.1021/acs.jctc.5c00138

Peer reviewed

PyCPET - Computing Heterogeneous 3D Protein Electric Fields and Their Dynamics

Pujan Ajmera [a], Santiago Vargas [b], Shobhit S. Chaturvedi [a], Matthew Hennefarth [c], Anastassia N. Alexandrova [a,*]

[a] Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States.

[b] Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

[c] Department of Chemistry, University of Chicago, Chicago, Illinois 60637, United States.

* Corresponding author email: ana@chem.ucla.edu

Abstract

Electrostatic preorganization is an exciting mode to understand the catalytic function of enzymes, yet limited tools exist to computationally analyze it. In particular, no methods exist to interpret the geometry, dynamics, and fundamental components of 3-D electric fields, $\vec{E}(r)$, in protein active sites. Here we present *PyCPET* (Python Computation of Electric Field Topologies), a comprehensive, open-source toolbox to analyze $\vec{E}(r)$, in enzymes. We designed it to be computationally efficient and user friendly with both CPU and GPU accelerated codes. Our aim is to provide a set of functions for rich, descriptive analysis of enzyme systems including dynamics, benchmarking, distribution of streamlines analysis in 3-D $\vec{E}(r)$, computation of point $\vec{E}(r)$, principal component analysis, and 3-D field visualization. Finally, we demonstrate its versatility by exploring the nature of electrostatic preorganization and dynamics in three cases: Cytchrome C, Co-substituted Liver Alcohol Dehydrogenase, and HIV Protease. These test systems, along with previous work, establish *PyCPET* as an essential toolkit for the in-depth analysis and visualization of electric fields in enzymatic systems, unlocking new avenues for understanding electrostatic contributions to enzyme catalysis.

Introduction

Since Warshel introduced the concept of electrostatic preorganization[1], biologists have been at odds to whether electrostatics or dynamics drive enzyme activity. Electrostatics, often measured through electric fields, $\vec{E}(r)$, are able to drive chemical reactivity, both in biology and beyond [2–5]. Enzymes are particularly interesting as they construct intrinsic electric fields that drive reactions over free energy barriers via transition state (TS) stabilization and/or reactant destabilization, and, in contrast to reactions in solution, avoid the entropic penalty for the reorganization of the electrostatic environment – hence they are preorganized. For example, in ketosteroid isomerase (KSI), studies have shown that the protein produces an $\vec{E}(r)$ that reduces the free energy barrier towards enolate formation[6–8]. Also, and perhaps most straightforwardly, fields can assist in processes like charge transfer[9]. These discoveries highlight the importance of analyzing electrostatic preorganization in enzymes and offer an exciting, additional avenue for rational enzyme design – focusing on the protein scaffold and remote electrostatic interactions, in addition to construction of the active site itself[10–12]. Mechanistic studies of enzymatic reactions, computational design of novel enzymes, and experimental approaches focused on the active site $\vec{E}(r)$ can all benefit from theoretical tools for electrostatic preorganization. An alternative school of thought posits that enzyme dynamics have a role in reactivity[13–15]. These competing ideas have

grown, in tandem, not without their share of debate[13,16–19]. Multiple groups noticed that $\vec{E}(r)$ can fluctuate over dynamical trajectories, and have shown the utility of averaged fields over a trajectory to explain fields measured in spectroscopy – so the combination of these two phenomena is not new[8,20,21]. Our group has introduced that the dynamics can be decomposed into individual $\vec{E}(r)$ states, that can differentially drive enzyme reactivities[22,23]. Our approaches demonstrated that the two perspectives can be merged, to construct analytic techniques for unearthing the effects of electrostatic preorganization as a dynamic framework for protein activity.

Here we present an open-source tool, *PyCPET*, or Python-based Computation of Protein Electric field Topologies, that can calculate properties of $\vec{E}(r)$, whether point $\vec{E}(r)$, volume $\vec{E}(r)$, or $\vec{E}(r)$ topological data. It can do so over the course of a dynamics trajectory, adding temporal and statistical components to analysis. Additionally, it offers tools for the comparison of the fields across a trajectory or across systems. This extends previous works on the analysis of $\vec{E}(r)$ and dynamics in enzymes[7,22–24] and we hope to provide the community with a compact, complete, set of tools to interpret $\vec{E}(r)$ at protein active sites from various perspectives.

To improve generalizability, we built these tools to define the region of interest for $\vec{E}(r)$ calculation through several different geometric motifs including atoms, spatial positions, and average atomic positions. This enables customized and high-throughput analysis of $\vec{E}(r)$ across different active site geometries. We accelerated topology calculations through GPU and parallelized C-shared codes. Along with accelerated calculations, we include benchmarking tools so users can avoid excessive resource use. Additional 3-D PCA utilities allow users to measure global changes in $\vec{E}(r)$ between populations. The package offers a clustering procedure, where users can select a set of topological data, either from MD calculations or many single-frame representations, and find representative centroids and distributions across the entire set of frames. Clustering determines $\vec{E}(r)$ states dynamically visited by the system. Additionally, we provide a utility to visualize individual $\vec{E}(r)$s in ChimeraX including filtering and mesh options.

Finally, we demonstrate the wide applicability of the *PyCPET* method across different types and sizes of enzymes on three different systems: Cytochrome C, a cobalt-substituted Liver Alcohol Dehydrogenase, and HIV-1 Protease. These enzymes span varying functions, sizes, and active site characteristics, and thus emphasize the wide applicability of *PyCPET* to understanding electrostatic preorganization and dynamics across enzymology. Instead of an abstract description, we chose to showcase most of the *PyCPET* features, as well as insights that the method provides, using these concrete systems as examples.

Results

Long range electrostatic interactions are required to converge field streamline distribution

PyCPET analyzes electric fields in 3D volumes of interest, such as active sites, produced by the protein macromolecules outside of that space (Fig. 1). In the volume of interest, $\vec{E}(r)$ perturbs the electron density, and thus, affects the active site reactivity. $\vec{E}(r)$ in this volume is constructed on a 3-D grid, and subjected to a variety of analyses based on its topology, using streamline distributions (see Methods). Streamlines of $\vec{E}(r)$ are computed at arbitrary lengths, and their properties such as streamline length and curvature are used to generate topological data. The topological data, which are discretized as 2-dimensional histograms called streamline distributions, are compared using a χ^2 distance.

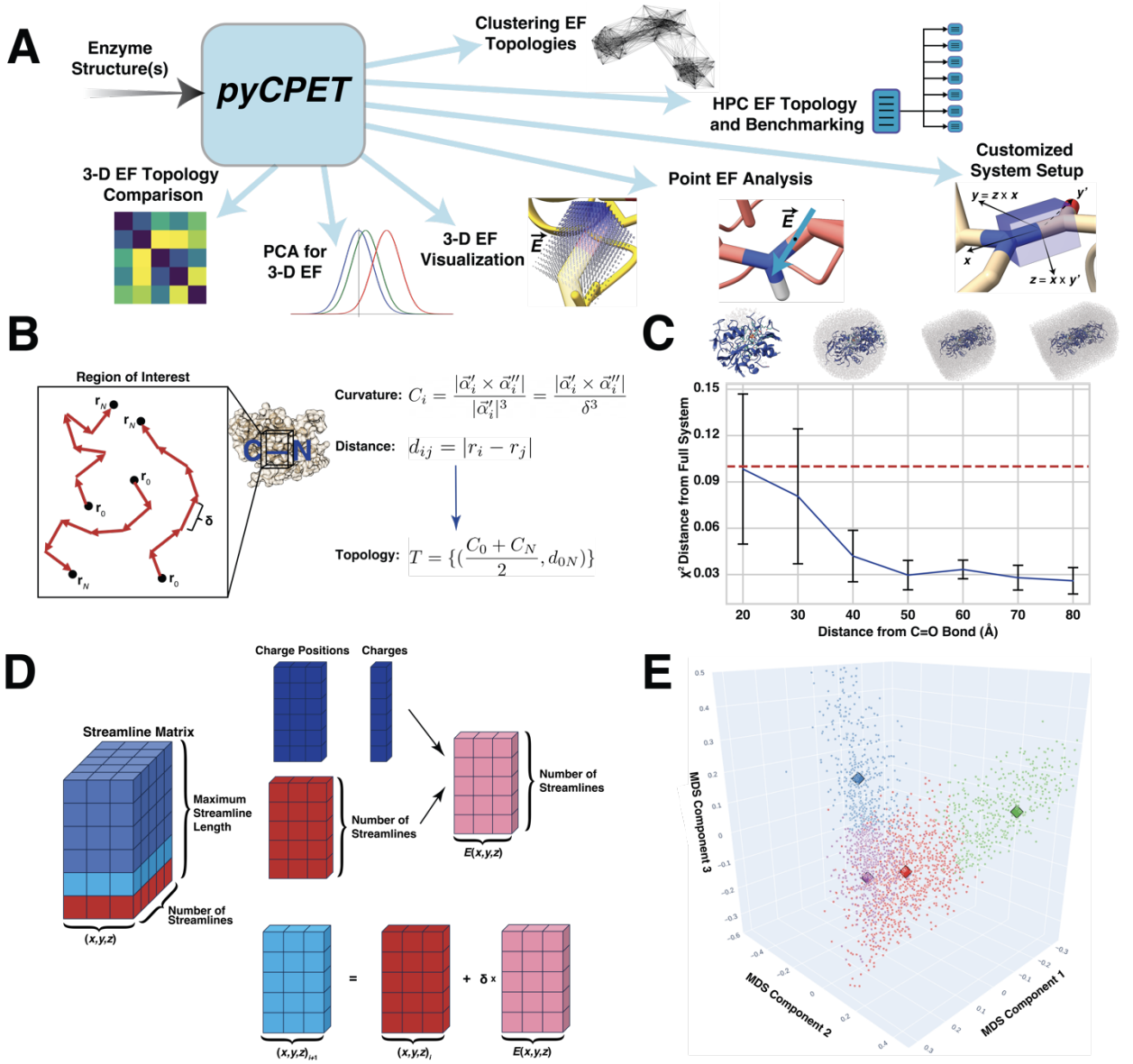


Figure 1: (A) Overall workflow of the *PyCPET* software, including intended use cases like clustering, principal component analysis, and field topology calculations. (B) Distribution of streamlines method to obtain topological characteristics of electric fields. (C) Schematic of GPU acceleration of electric field topology calculations (see Methods for more details). (D) The convergence of the field topology metric when including charges further than 20Å for Cytochrome C heme Fe, showing the residues included at each radius marker. (E) K-medoids partitions the space of electric field topologies, with the dynamics of HIV-1 Proteas shown as an example using multidimensional scaling of the pairwise distance matrix. Bolded points are identified cluster centers. Graphs of convergence and pairwise distance matrix for other two systems studied are in Extended Data Fig. 1

PyCPET's built-in benchmarking protocol allows us to probe the radius from the region of interest required for a converged field topology. For all three proteins chosen for illustration, including 30-40Å of the system from the active site to converge the streamline distribution. This is shown for cobalt-substituted Liver Alcohol Dehydrogenase (CoLADH) in Figure 1C, where we define convergence of this metric as $\chi^2 < 0.1$ relative to the full protein and solvent. It is important to note that this is *larger* than the standard cutoff for long-range Coulomb interactions in MD simulations, emphasizing the relevance of long-range interactions and the sensitivity of the $\vec{E}(r)$ topology[25]. We justified this cutoff by looking at the response of the geometry of the quantum

mechanical charge density, ρ, in the active site, and its convergence with the size of the protein region included in evaluation of electrostatics from previous work[26]. If fewer charges provide convergence relative to the full simulation volume, the user can choose to reduce the cutoff.

Distinct field signatures emerge from dynamical clustering

Clustering is a popular technique for partitioning MD simulations, it has been used to find representative structures based on the protein geometry[27]. We demonstrate an alternative approach: using the $\vec{E}(r)$ streamline distribution to find representative clusters of $\vec{E}(r)$ configurations, rather than atomic positions. *PyCPET* can compute the $\vec{E}(r)$ streamline distribution at each frame of an MD trajectory (Figure 1B), and then uses the χ^2 distance (Eq. 3 in Methods) to construct a pairwise distance matrix, and perform K-Medoids clustering to obtain representative structures.

The K-medoids algorithm finds the optimum number of clusters for $\vec{E}(r)$ streamline distributions, as the number of clusters past which the within-cluster sum of squares does not improve. To quantify the success of the algorithm, we compute the silhouette score – a metric that measures how well-matched each point is to its assigned cluster, where a higher score is better. The algorithm optimally yields (Extended Data Fig. 2) four to five unique $\vec{E}(r)$ clusters for each system, with silhouette scores of 0.49, 0.34, 0.33, and 0.51 for Cytochrome C (Cyt C), wild-type (WT) CoLADH, mutated CoLADH, and HIV protease (HIV-P), respectively. These scores indicate distinguishable, well-separated clusters. We then visualized the identified clusters using multidimensional scaling of the pairwise distance matrix, with an example of WT Co-substituted LADH in Fig. 1E. To ensure that the MD simulations captured a variety of $\vec{E}(r)$ conformations and that clustering did not simply sort frames by their simulation replica, we also show multidimensional scaling graphs of individual MD runs, colored by cluster identity for Cyt C as an example (Extended Data Fig. 3). We see here that creating a pairwise distance matrix, and clustering allows users to understand the underlying dynamics of $\vec{E}(r)$, and partition MD trajectories by field for deeper analysis (Fig. 1E, Extended Data Fig. 1C-E).

3-Dimensional dynamic electric field analysis provides insights into reactivity

Electric fields and the reducing ability of Cytochrome C Hemes

Cytochrome C (Cyt C) is important as a redox partner in biological systems where it transfers electrons to Complex IV in mitochondria[28]. It contains an iron heme center that acts as a reductant in its ferrous (Fe^{2+}) form. Prior studies underscored the role of electrostatic preorganization in heme function and catalysis[23,29,30], and here we use *PyCPET* to probe 3-D electric field at the heme Fe in a human Cyt C (PDB 2N9I). *PyCPET* allows us to specify the orientation, location, and volume where we computed the electric field (Fig. 2A). At each representative frame obtained from clustering the MD trajectory of ferrous Cyt C, we visualized the 3-D $\vec{E}(r)$ using the visualization tool in *PyCPET*. This utility generates files for the *ChimeraX* software[31] (Fig. 2B).

Each representative $\vec{E}(r)$ has geometrically unique features captured through the clustering algorithm, as can be seen upon comparison of the four cluster centers, denoted CYTC-1, CYTC-2, CYTC-3, and CYTC-4, ordered by their frequency of appearance during dynamics (Figure 2B). All four $\vec{E}(r)$ conformations for Cyt C have a dominant homogeneous direction toward the same

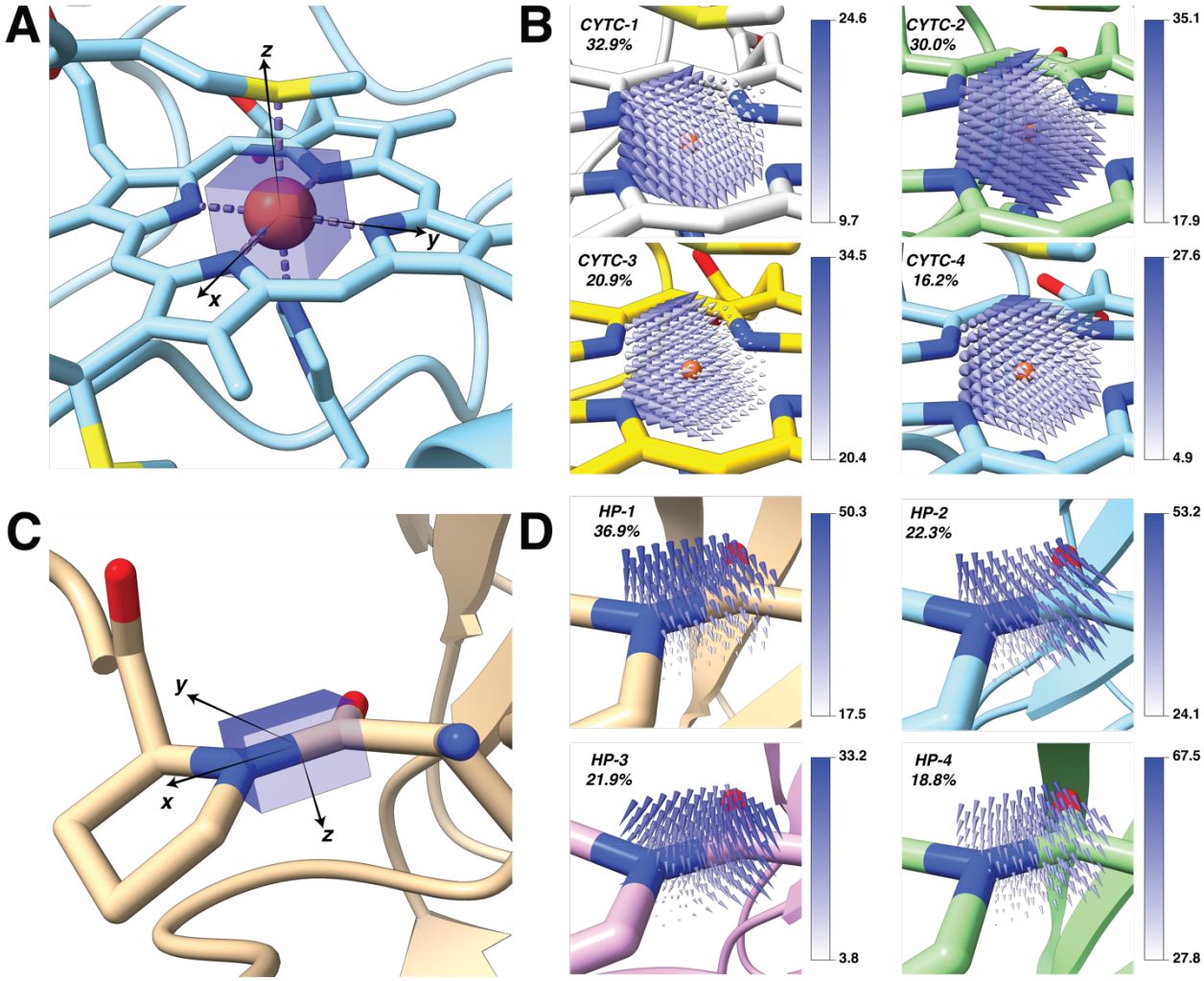


Figure 2: (A) Orientation of the electric field box studied for the ferrous Cyt C. The volume axes are oriented by the heme Fe coordinates, and two adjacent porphyrin nitrogen atoms. (B) Representative electric fields observed in Cyt C from topology-based clustering, in order of population. (C) Orientation of the electric field box studied for HIV-P. The axes are defined by the peptide C-N bond center, the proline nitrogen, and the carboxyl oxygen (*PyCPET* automatically reorients axes for an orthonormal basis). (D) Representative electric fields observed in HIV-P from topology based clustering, in order of population. Scales next to (B) or below (D) each representative electric field show the range of magnitudes present in calculations, in MV/cm.

porphyrin nitrogen, but the topology-based clustering method captures subtle differences between them. CYTC-1, which appears in highest frequency, has contributions in the z direction from Fe toward the coordinating His residue, with a small component pointing toward the coordinating Met. CYTC-2 is a rotated field toward the y-axis, and has a weaker field contribution in the -z direction compared to CYTC-1. CYTC-3 is highly similar to CYTC-2, further evidenced by similarities in their magnitude ranges. CYTC-4, the least-visited $\vec{E}(r)$ conformation, is the most heterogeneous, showing components facing both coordinating ligands, and in both the x and y-axis directions. All these 3-D electric fields have the same characteristic of facing away from the solvent face of the protein where interaction with a Cytochrome C oxidase, for example, would take place – an electric field facing in this direction would favor oxidation of this ferrous form. Our analysis provides a way to capture major and subtle fluctuations in the field geometry, and by taking a perturbative approach of computing the exerted $\vec{E}(r)$ only from the surrounding environment, removes any sort of bias.

To link this quantity to reactivity, we computed oxidation potentials at each representative configuration using QM/MM calculations. The four representative conformations, CYTC-1-4, yielded absolute reduction potentials of 4.79eV, 4.73eV, 5.36eV, and 3.54eV. For comparison, experimental values range from 4.37 to 4.68eV[32,33], and an *ab initio* study by Chen and Hayashi reported a value of 4.34eV[34]. The lowest computed reduction potential was for the least frequent electric field configuration, CYTC-4. The *PyCPET* method of partitioning molecular dynamics by $\vec{E}(r)$ configurations reveals the dynamic nature of the reduction potential, which cannot be captured by experimental methods (which are slow and only measure averages), or QM calculations on single structures. Clustering allows us to weight and focus on a few physically informed frames. To ensure that the QM/MM differences in redox potential were not due to merely changes in active site geometry, we obtained QM-region only reduction energies at the QM/MM optimized geometries of 0.94eV, 0.92eV, 0.88eV, and 0.74eV, i.e. in a range near DFT error and not following the same trend as the full QM/MM, confirming this methods ability to capture $\vec{E}(r)$ configurations with distinct chemical relevance.

HIV-1 Protease active site electric fields

HIV-1 protease (HIV-P) is critical to the life cycle of HIV, facilitating the processing of precursor peptides to final, mature proteins[35]. Although the specific mechanism is debated between concerted, oxyanion, and gem-diol mechanisms (Extended Data Fig. 4A) [36–41], we examine $\vec{E}(r)$ that the protein scaffold exerts on the peptide bond through *PyCPET*. A study by Krzemińska and coworkers[41], on electrostatic preorganization in HIV proteases, emphasized two $\vec{E}(r)$ components, one pointing toward the reactive aspartate residues and one along the direction of the peptide bond (Figure 2C). Another study demonstrated the importance of electron flow toward the C-N bond for two possible mechanisms, and the role of charged amino acids in reducing the barrier for both mechanisms[42].

Utilizing tools in *PyCPET*, we examine the dynamics of $\vec{E}(r)$ and contextualize them across these studies. We computed streamline distributions of 3-D $\vec{E}(r)$ around the peptide C-N bond (Fig. 2C) along MD simulations of an HIV-1 protease (PDB 1KJ4). The four representative $\vec{E}(r)$s are shown in Figure 2D, labeled HP-1, HP-2, HP-3, and HP-4 in the order of population. All four fields have a dominant direction perpendicular to the C-N bond, but *PyCPET* teases out changes in the field curvature. In HP-3 and HP-4, the field component along the C-N bond is not ubiquitous in the studied volume, and this component is not present at all in HP-1 and HP-2. Krzemińska and Robiero both showed that there is an overall $\vec{E}(r)$ present in both directions observed in our 3-D analysis: the perpendicular one favoring electron transfer from the catalytic Asp residue to the C-N bond through a water molecule required for both gem-diol and oxyanion-based mechanisms, and the field along the C-N bond favoring the charge separation required for a concerted mechnism[41,42]. Here we not only confirm this, but also show that $\vec{E}(r)$ is not static across protein dynamics – a reality which traditional methods fail to capture. The change in $\vec{E}(r)$ structure along dynamics supports the possibility of multiple mechanisms being favored in the enzyme, where HP-1 and HP-2 solely support charge transfer to the bond and likely a gem-diol/oxyanion based mechanism, while HP-3 and HP-4 sustain an electric field required for a cyclic concerted mechanism, where both charge transfer and C-N bond breaking are simultaneous[38]. In addition,

we find that these fields are heterogeneous across an entire 3-D field volume, highlighting the importance of a multidimensional approach to understanding electrostatic preorganization.

Liver Alcohol Dehydrogenase exhibits a dipole-stabilizing 3D electric field

The Boxer group used Stark spectroscopy to demonstrate that cobalt-substituted liver alcohol dehydrogenases (CoLADH) exert $\vec{E}(r)$ favorable for aldehyde hydrogenation, which can be enhanced by introducing the S48T mutation[43]. They observed $\vec{E}(r)$ in the same direction as the reactive C=O bond, which corresponds to reducing the reaction barrier by stabilizing the charge-separated TS (Extended Data Fig. 4B). *PyCPET's* tools allow us to observe this 3-D $\vec{E}(r)$ along the C=O bond (Fig. 4A) and understand what residues contribute toward reactivity. We compute $\vec{E}(r)$ streamline distributions across MD simulations of WT Co-substituted LADH and S48T Co-substituted LADH (Fig. 4B), capturing changes in $\vec{E}(r)$ throughout dynamics, with an emphasis on the S48 residue influence. However, the representative electric fields in the S48T mutant are similar in geometry to representative WT electric fields (Extended Data Fig. 5), a finding that does not clarify the observations that the S48T mutant enhances $\vec{E}(r)$ along that direction.

To amend this outcome, we utilized PCA within *PyCPET*, to decompose variations in $\vec{E}(r)$ between WT and the mutant LADH, including magnitudes and geometry. Among the PCA components, we observed only two $\vec{E}(r)$ components that clearly distinguished the mutants (Fig. 3C). PC0 is an $\vec{E}(r)$ pointing generally in the direction opposite to the C=O bond dipole (Fig. 3C, left). The WT LADH, projected onto PC0, fluctuates around 0, having no alignment to it, while the S48T mutant is bimodal with a slight positive bias. Along PC0, these two variants have no clear separation in their mean projection, and the component points to residue 48 – the bimodal nature of S48T is likely due to varying conformations of the residue's hydroxo-group, not present in WT (Extended Data Fig. 6). On the other hand, PC1 shows that there is a dominant direction from the oxygen toward the carbon in the C=O bond (Fig. 3C, right). S48T is negatively aligned with this field component, while the WT Co-LADH is positively aligned with this field. Therefore, S48T exerts a field throughout dynamics more strongly aligned with the C=O dipole compared to the WT, confirming the experimental result by Zheng, et. al[43]. This brings up an important consideration with using PCA as an analytical tool – components with the most variability are not always most physically informative, chemical intuition required for interpretation, and users should consider a larger set of components in analysis to not miss important elements. Thus, *PyCPET* and its 3-D electric field analysis features can tease out intricacies in the electric field geometry and compare mutants along dynamic trajectories. Overall, we see *PyCPET* as a rich tool kit that allows us and other users multiple avenues for analysis of active site electric fields and their dynamics.

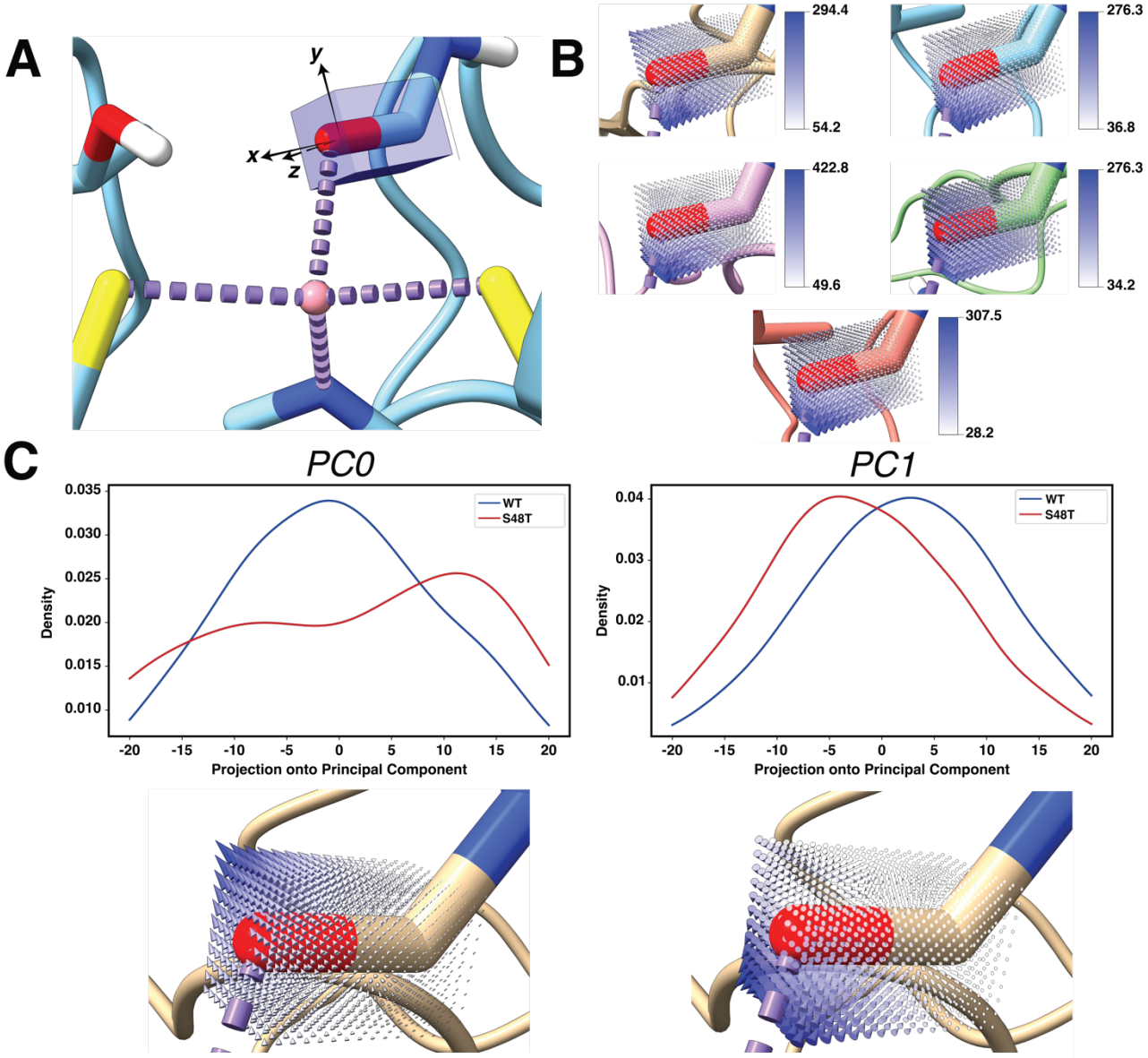


Figure 3: (A) The volume of interest in Co-LADH, based on measured active site electric fields from Zheng (44). The carboxyl bond center, as well as the carboxyl oxygen and amido nitrogen, orient the axes. (B) The representative electric fields present in S48T Co-LADH, with a scale (in MV/cm). Representative electric fields for WT Co-LADH are in Extended Data Fig. 3. (C) The projections of 3-D fields from S48T and WT Co-LADH onto two principal components, PC0 and PC1. The projection is simply a dot product of the 3-D field from each frame onto the principal component. (D) The principal components PC0 and PC1 themselves, displayed on the region of interest, are either opposing or in the same direction of the C=O bond dipole, with curvature toward the key S48T residue for PC0.

mber

of streamline samples, as well as a large radius of protein charges. This leads to an increase in computational cost; therefore, we have developed CPU and GPU-accelerated codes for *PyCPET* that reduce the cost of computing streamline distributions. In our CPU-accelerated implementation, we have parsed the problem of computing properties of field streamlines by

assigning each streamline to a thread (see Methods), and therefore achieving roughly linear accelerations in computation with number of threads provided. Our GPU-accelerated implementation computes all streamlines simultaneously on a grid, and therefore outperforms CPU-accelerated codes on several metrics.

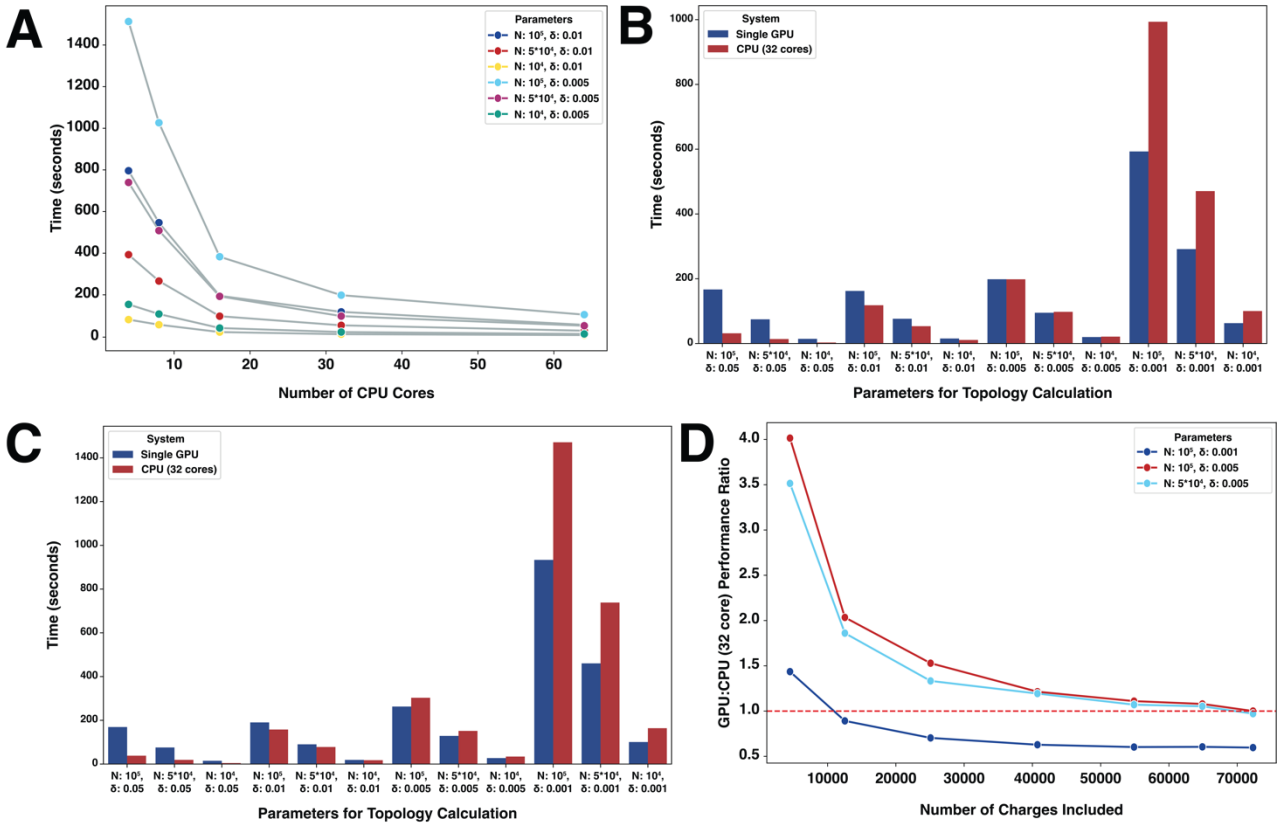


Figure 4: (A) Scaling of CPU-accelerated code to compute electric field streamline distributions across 4,8,16,32, and 64 cores. (B) Performance of CPU (32 cores) and GPU-accelerated streamline distribution calculations across several numbers of streamlines and stepsizes, of a box of size 1.5 x 1.0 x 1.0 and a size of (C) 3.0 x 2.0 x 2.0 (Å). (D) . N represents the number of streamlines computed, and δ represents the stepsize of the streamline (in Å).

with the number of cores provided, as the problem is grossly parallelizable (Fig. 4A). We have implemented C-shared libraries to boost performance on computation of $\vec{E}(r)$. We compare the performance of both CPU (with 32 cores) and GPU-accelerated codes in Fig. 4B and Fig. 4C, for various number of streamlines and step sizes, at smaller (Fig. 4B) and larger (Fig. 4C) volumes. Since a larger volume means longer computed streamlines, the GPU's ability to store and propagate a matrix of streamlines simultaneously provides faster performance times (see Supplementary Notes). The same effect is seen at smaller step sizes, as more steps are required for a streamline of the same size. Therefore, we prescribe the GPU-accelerated code for systems where larger volumes and/or finer stepsizes are required for a well-converged $\vec{E}(r)$ streamline distribution. Lastly, we compared GPU to CPU (with 32 cores) performance based on system size, proxied by number of charges included (Figure 4D), and implies further utility of the GPU code for systems with >100,000 point charges. Furthermore, for individual workstations with a single GPU, but limited number of dedicated CPUs, the GPU code would be preferred for performance,

as based on the scaling relationships shown in Figure 4, the GPU code would outperform 16-core CPU acceleration on several topology parameters and system sizes.

Discussion

Here we demonstrate that *PyCPET* is a widely-applicable tool for computing 3-D electric fields, generating topological features, and analyzing their dynamics. We built the tool with flexibility in mind as users can uniquely specify relevant active sites in their own protein systems. This approach bridges the gap between electrostatic preorganization and a 3-D perspective on electronic structure with important extensions toward sampling dynamics. In addition, *PyCPET* features clustering algorithms, PCA, and baseline methods to compute 1-dimensional $\vec{E}(r)$ at points for cursory analysis. We have provided a compelling set of use cases for our toolkit, that when combined with previous studies[7,22,23], highlight how users can functionalize these methods in their own studies. To facilitate future work, we implemented several benchmarking tools to ensure that parameters for clustering, streamline distribution analysis, and 3-D electric fields yield desirable accuracy for studied systems. Benchmarking on a few representative structures of new systems can save a considerable amount of compute while equivalent or better signal. This project is open-source source, modular, and can be configured for several modes of analysis and high-throughput studies. We provide several example notebooks in our repository (https://github.com/pujanajmera/pycpet). Finally, the tool is compatible with the popular biomolecular software *ChimeraX* for easy visualization of 3D fields.

PyCPET builds on several studies in our group. We first showed the utility of computing $\vec{E}(r)$ topological data through a study KSI mutants and the effects of oriented external electric fields. Here we found the χ^2 distances of $\vec{E}(r)$ streamline distributions to explain reaction barriers for enolate formation[7]. Additionally, Vargas, et. al[23] showed that 3-D fields were more useful than point $\vec{E}(r)$ for the classification of enzyme function across a dataset of hemes, and was further enhanced by the consideration of dynamic trajectories. Chaturvedi, et. al.[22] also demonstrated the utility of dynamics for interpreting how $\vec{E}(r)$ impact catalysis beyond a static perspective, in a study of the causes of new-to-nature chemistry seen in a lab-evolved protoglobin.

Analyzing three additional enzyme systems in this study, we demonstrate the interpretive power and versatility of *PyCPET* to understand the nature of active site electric fields. For example, looking at the electric field configurations in Cyt C showed the dynamic sampling of reduction potentials, all while still reproducing averaged experimental results. The 3-D electric field and its dynamics in the HIV-1 protease studied shed light on the nature of preorganized $\vec{E}(r)$ along and perpendicular to the peptide bond. Lastly, we utilized both 3-D field PCAs and topological calculations to tease out the impact of a mutation on the dynamics of $\vec{E}(r)$ in liver alcohol dehydrogenases. Our discussion of the three example systems serve as a template for future uses of *PyCPET*. The power of *PyCPET* over other methods is two-fold: dynamics and high-dimensional analysis. These two properties allow for much richer studies across molecular dynamics while also allowing users to focus on representative structures for higher-order analysis such as QM/MM mechanistic calculations.

Methods

Theory of enzymatic electric fields

Traditional analyses of electric fields in enzymes have focused on determining the electric field, in direction and/or magnitude, at a single point of interest at the active site or substrate:

$$\vec{E}(r) = \sum_{i=1}^{N} q_i \frac{r - r_i}{(|r - r_i|)^3} \quad Eq.1$$

However, this is a limited approach as ρ in the active site is delocalized, while $\vec{E}(r)$ exerted by the enzyme is heterogeneous in the active site volume. This motivates us to consider a 3-D approach for electric field analysis and comparison. Quantifying similarities and differences in 3-D vector fields is no meek task. Several algorithms have been developed, some focusing on detecting local features, and others focusing on global measures of similarity[44–46]. We focus on the latter, as we wish to consider all features of $\vec{E}(r)$ that could perturb ρ.

PyCPET is based on our method to extract topological data of a 3-D $\vec{E}(r)$, inspired by fluid dynamics [44]. In this approach, we compute a distribution of streamlines of $\vec{E}(r)$ in a volume. Each streamline is computed by propagating a point in the volume along the field it experiences by steps of size δ, up to an arbitrary length. A streamline's topological data is encoded in it's mean of start and end curvatures (Eq. 2) and the Euclidean distance of the start and end points. This process is repeated until information about *N* streamlines in $\vec{E}(r)$ are computed (the procedure is highlighted in Fig. 1B). The whole field, therefore, can be represented as a 2-D histogram of streamline mean curvatures and Euclidean distances. These histograms are clearly invariant under rotations of the field and scaling of field magnitudes. This makes a streamline distribution approach a robust metric for generating fingerprints of 3-D electric fields, and focuses truly on the underlying geometry.

$$\kappa(\alpha) = \frac{|\alpha' \times \alpha''|}{|\alpha'^3|} \quad Eq.2$$

Two histograms, say H and H', can be compared using a χ^2 distance metric (Eq. 3). This permits analysis of the histograms as distributions, compared to others like Euclidean or taxi-cab distances which would treat the histograms as vectors. This value, assuming H and H' are normalized, range between 0 and 1. The approach has been shown to be more robust than comparison of local vector field features[44]. Randomly generated distributions H and H' produce a χ^2 value of around 0.18 (Supplementary Notes). This provides a useful reference point from which to compare the similarities of any two distributions.

$$\chi^2(\mathrm{H}, \mathrm{H}') = \sum_{i}^{N} \frac{(H[i] - H'[i])^2}{H[i] + H'[i]} \quad Eq.3$$

Robust parameter determination for computing field streamline distributions

To extract topological data of the global electric field, several parameters need to be determined for every system: the volume to calculate the 3-D $\vec{E}(r)$, the number of streamlines, and the step size. First, the volume and location where the streamlines are computed should be based on a *chemically relevant space*, and is reaction specific. We define this space as one in which ρ undergoes significant deformation for a reaction to proceed[47]. For example, in a bond breaking reaction, this volume constitutes the bond itself. However, if a conjugated system is involved, the reaction may involve perturbation of ρ over the entire region of the conjugation. For the sake of

understanding the isolated effect of $\vec{E}(r)$ that the protein scaffold exerts, we zero charges in the chemically relevant region. For Cytochrome C, we zeroed the charges on the heme and coordinating residues. In Alcohol Dehydrogenase, we zeroed only the substrate, as we wished to compare to an experimental study that probed the effect of the entire protein scaffold and coordinating metal on the substrate[43]. For HIV-1 Protease, we zeroed the catalytic aspartate residues as well as the peptide substrate.

To make the streamline distribution method widely applicable, as well as efficient, we developed a benchmarking method for the number of streamlines N and step size δ parameters illustrated in Fig. 1B. We perform MD simulations of each enzyme and sample three frames, at random, to compute field streamline distributions with several values of N and δ. We chose N = 100,000 and δ = 0.001 as the high sampling limit, as both parameters combined imply a sub 0.1 Å resolution of the field, and thus precise sampling. By generating a χ^2 pairwise distance matrix, we obtain values of N and δ at which the distributions converge. All three systems demonstrated converged field streamline distributions where we define convergence as a χ^2 distance less than 0.1 relative to the high sampling limit, at N = 100,000 and δ = 0.01. Users can perform similar parameter benchmarking on other systems with scripts we have provided.

Principal Component Analysis

PyCPET includes the PCA functionality to compare 3-D, multicomponent fields across different sets of proteins, or between different dynamics frames of the same protein. PCA is a popular dimensionality reduction tool that constructs eigenvectors of the covariance matrix with maximum eigenvalues. This corresponds to producing the direction along which there is maximum variability in the data. Our implementation takes in a set of 3-D $\vec{E}(r)$ vectors of total size (3, N_x, N_y, N_z, T), where N_x, N_y, and N_z are the number of points in each dimension of the electric field volume, and T is the number of electric field frames provided. From PCA, we obtain eigenvectors that maximally explain the variance in the data of shape (3, N_x, N_y, N_z), which we can then visualize using PyCPET's built-in tools.

MD Simulations

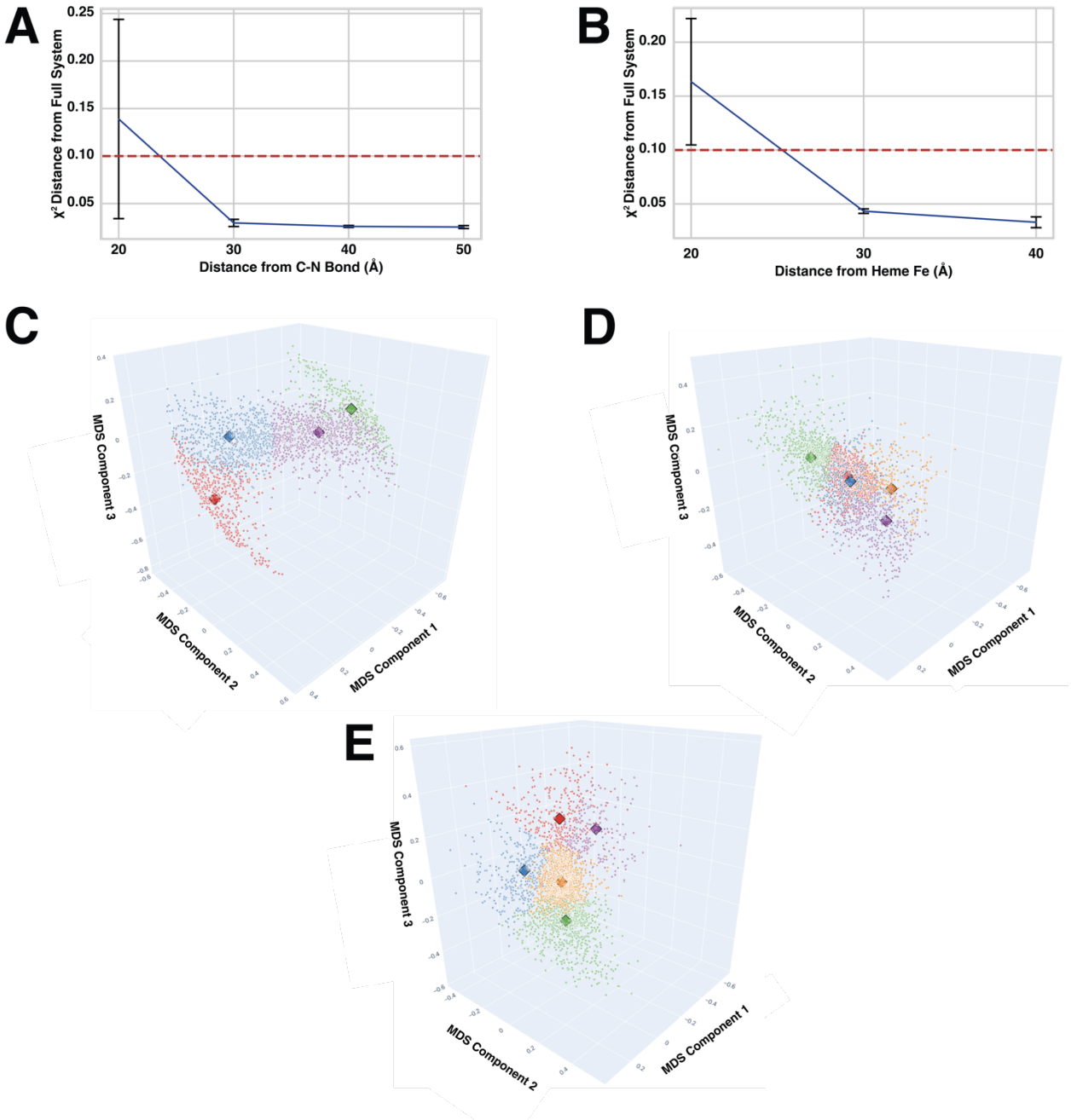
Molecular Dynamics (MD) simulations were performed for all three enzyme systems using Amber22[48]. All metal sites (Cytochrome C, Liver Alcohol Dehydrogenase) were parametrized using MCPB.py[49]. To obtain equilibrated structures for production runs, we ran a solvent-only minimization, full system minimization, heating simulation to 300K, NPT restrained simulation for 1ns to achieve uniform density, and an NPT equilibration simulation for 5ns. We then ran five 100ns NPT production run simulations for each enzyme to provide dynamic sampling.

QM/MM Calculations

For the Cytochrome C system, we performed QM/MM calculations using Py-Chemshell[50,51] with TURBOMOLE[52]. We defined the QM region as the Fe-containing heme, axial ligand, and the Cys residues ligated to the heme propionate. The QM region was treated with unrestricted DFT, using the B3-LYP D3-BJ functional, with def2-TZVP basis set for Fe and def2-SVP for all other atoms for optimizations. For single-point energy calculations, we used the def2-TZVP basis set for all atoms in the QM region. The MM region was defined as 8Å beyond the QM region, and was treated with the AMBER force field. All atoms past the MM region were frozen, and both the MM and frozen regions were treated with electrostatic embedding.

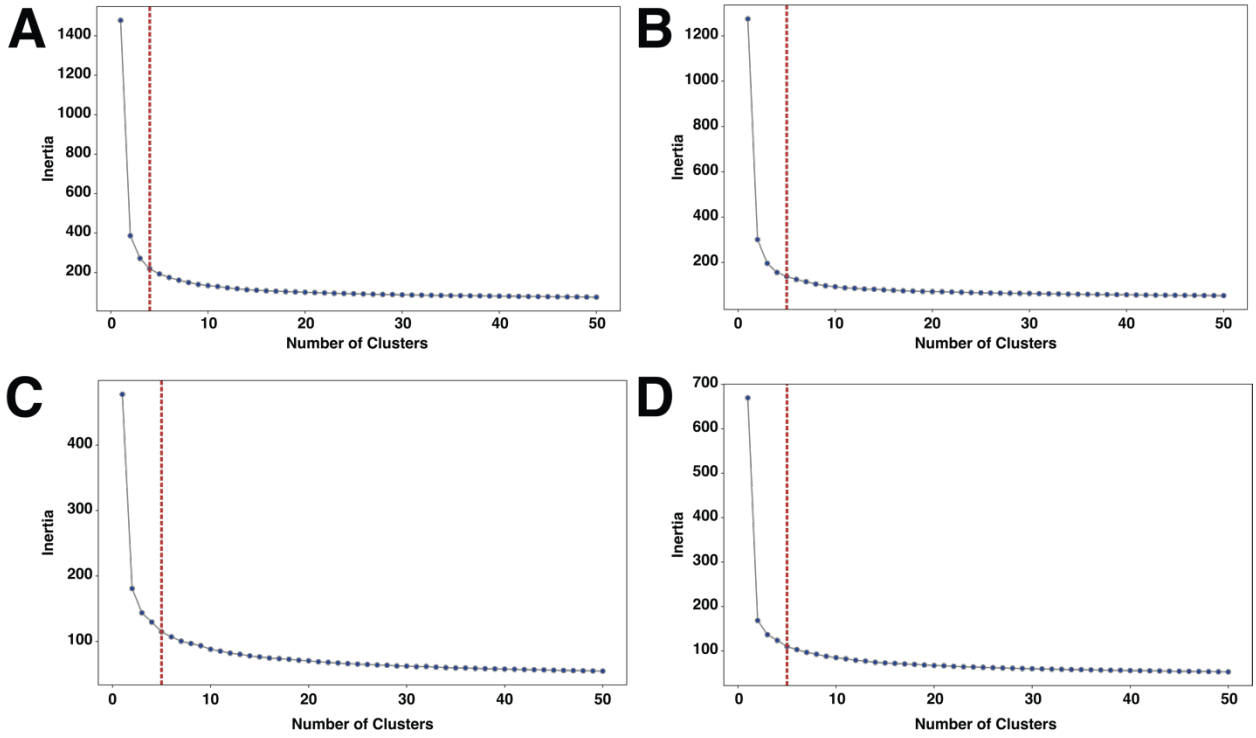
Extended Data

Extended Data Figure 1: Convergence of electric field streamline distribution and electric field dynamics for three enzyme systems



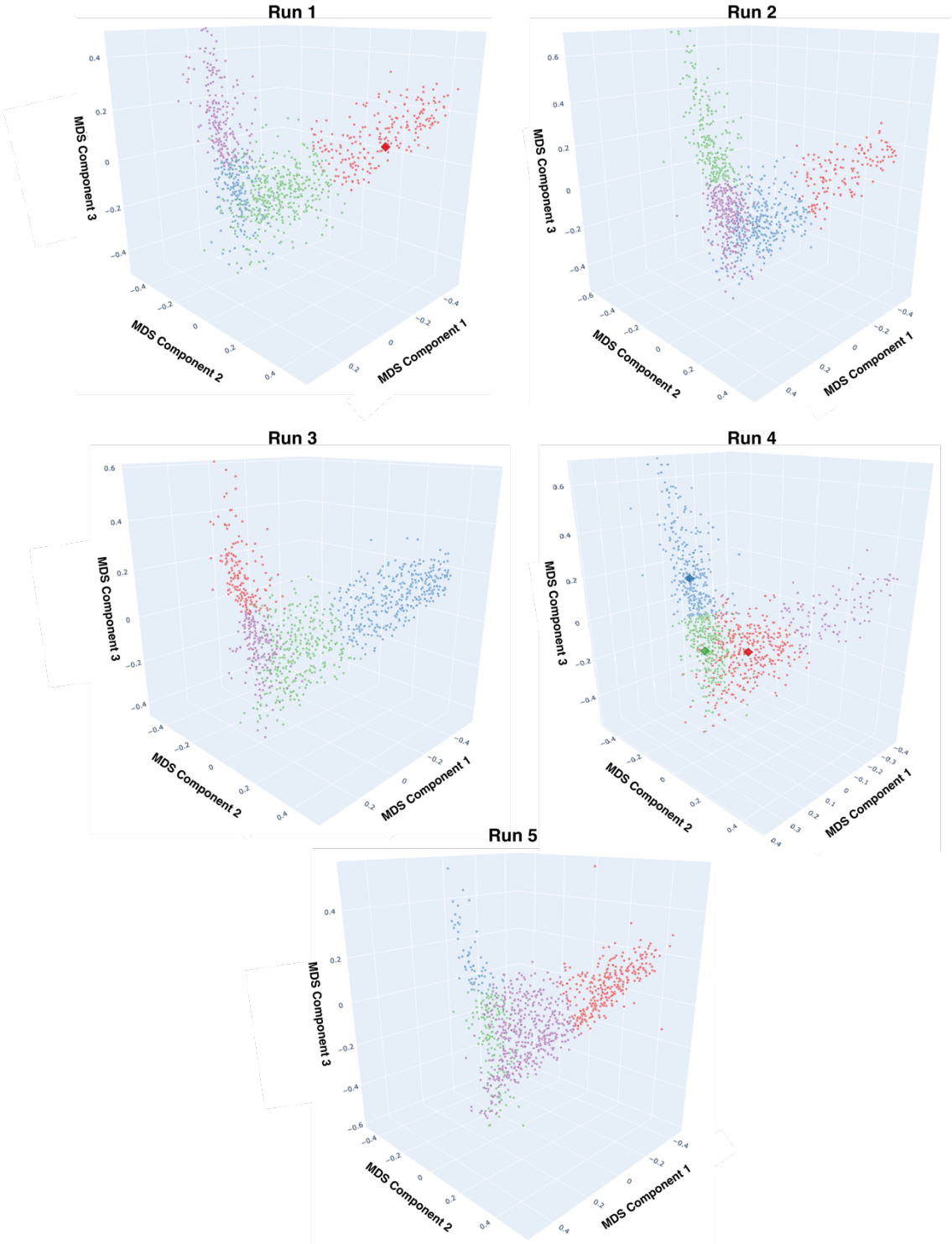
Extended Data Figure 1: Convergence of the electric field streamline distribution along radii from the HIV-1 Protease active site (A) and the Cytochrome C active site (B). Clustering of electric field dynamics for Cytochrome C (C), S48T Co-substituted LADH (D), and WT Co-substituted LADH (E), represented through multidimensional scaling.

Extended Data Figure 2: Determining ideal number of clusters for partitioning electric field topology dynamics



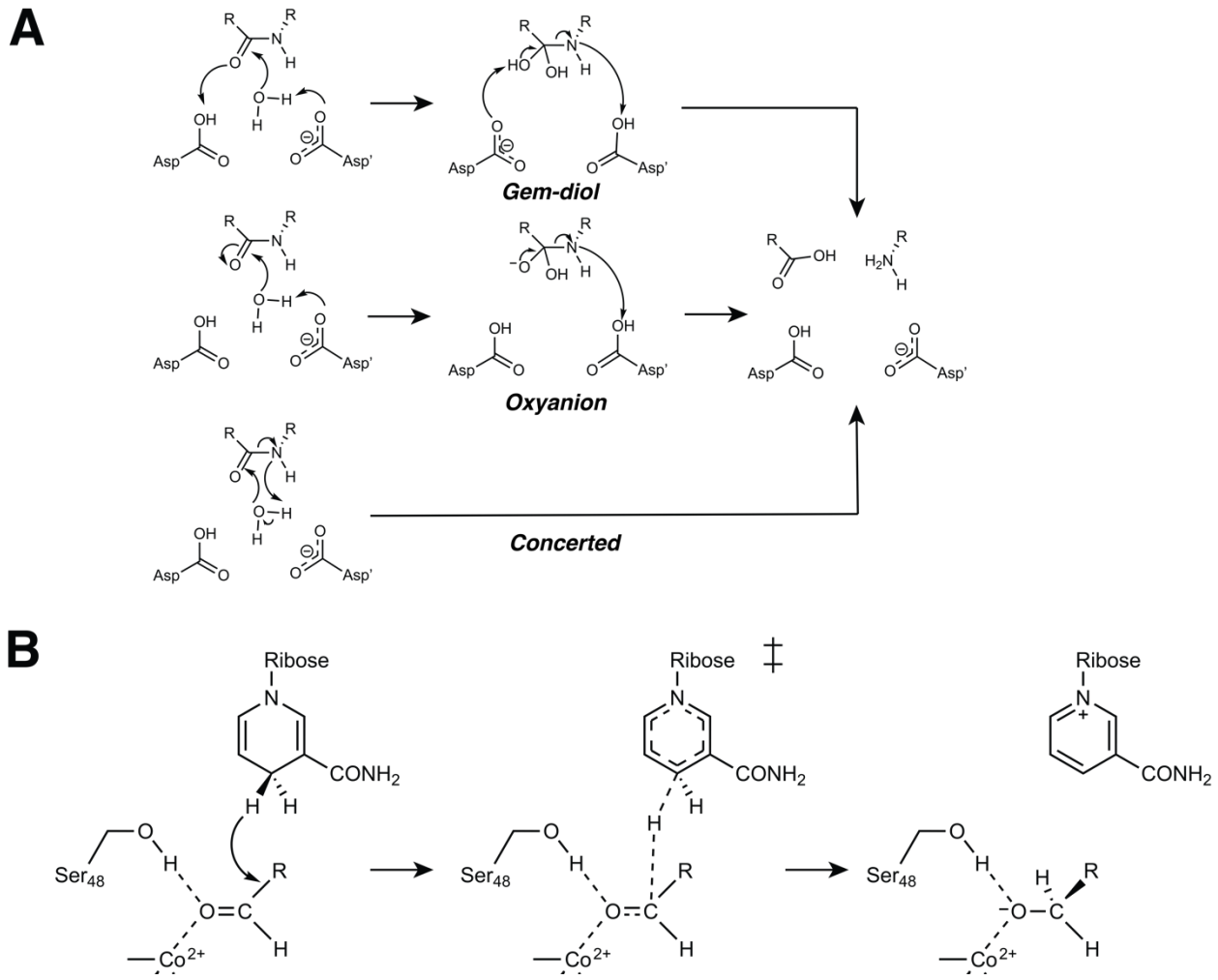
Extended Data Figure 2: Inertia, or sum of the squares of points to their cluster centers, as a function of the number of clusters provides a method to determine the correct number of clusters. The algorithm determines the optimal number of clusters, denoted by the red line, for (A) Cytochrome C, (B) HIV-1 Protease, (C) S48T Co-substituted LADH, and (D) WT Co-substituted LADH.

Extended Data Figure 3: Electric field topology-based clustering yields clusters that are replicate-independent



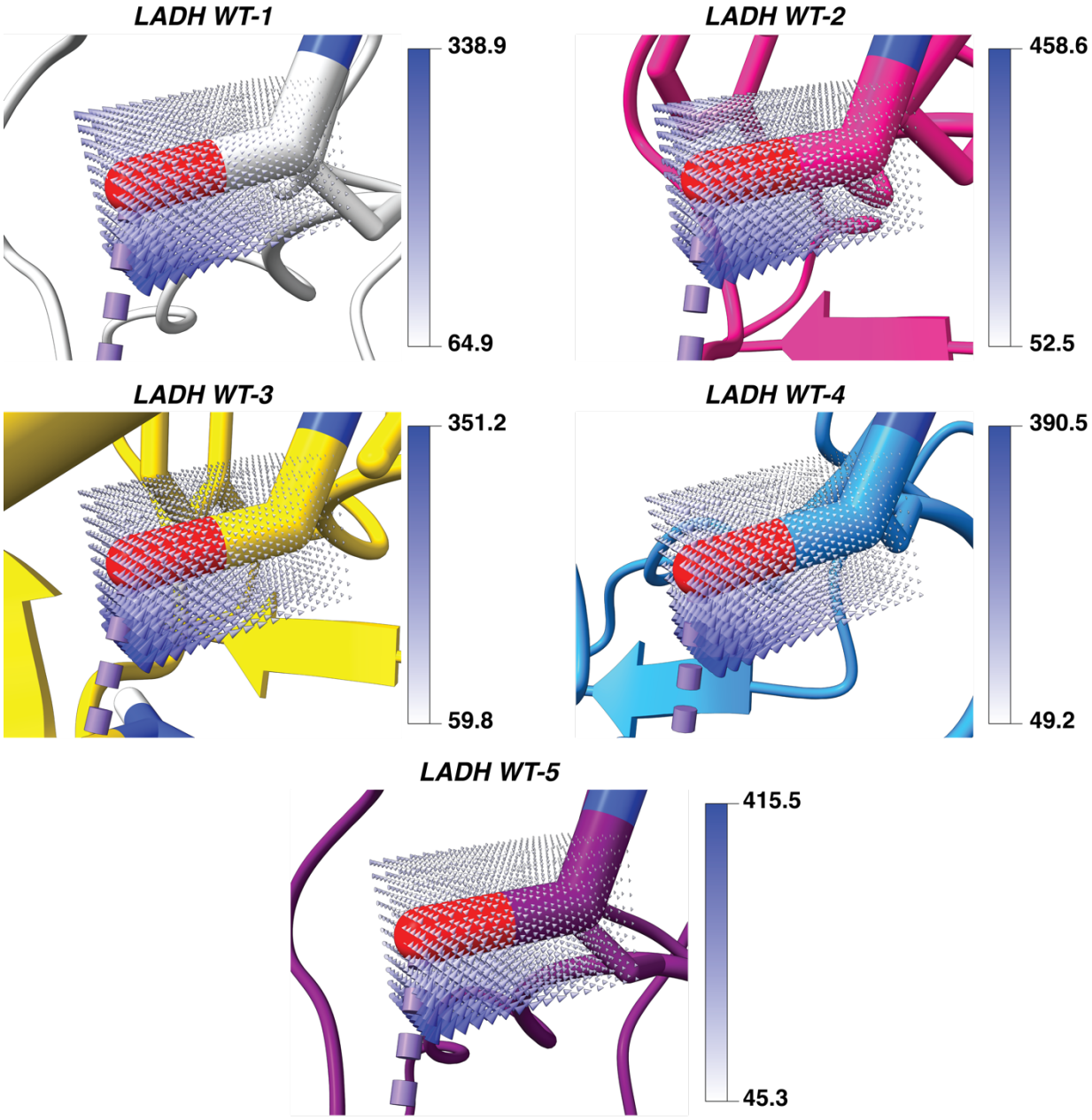
Extended Data Figure 3: Multidimensional scaling graph represented in Fig. 1E in Main, separated by trajectory run, demonstrating the independence of the electric field clustering from simulation runs. Cluster centers identified in Fig. 1E are bolded here as well.

Extended Data Figure 4: HIV-1 Protease and Liver Alcohol Dehydrogenase Mechanisms



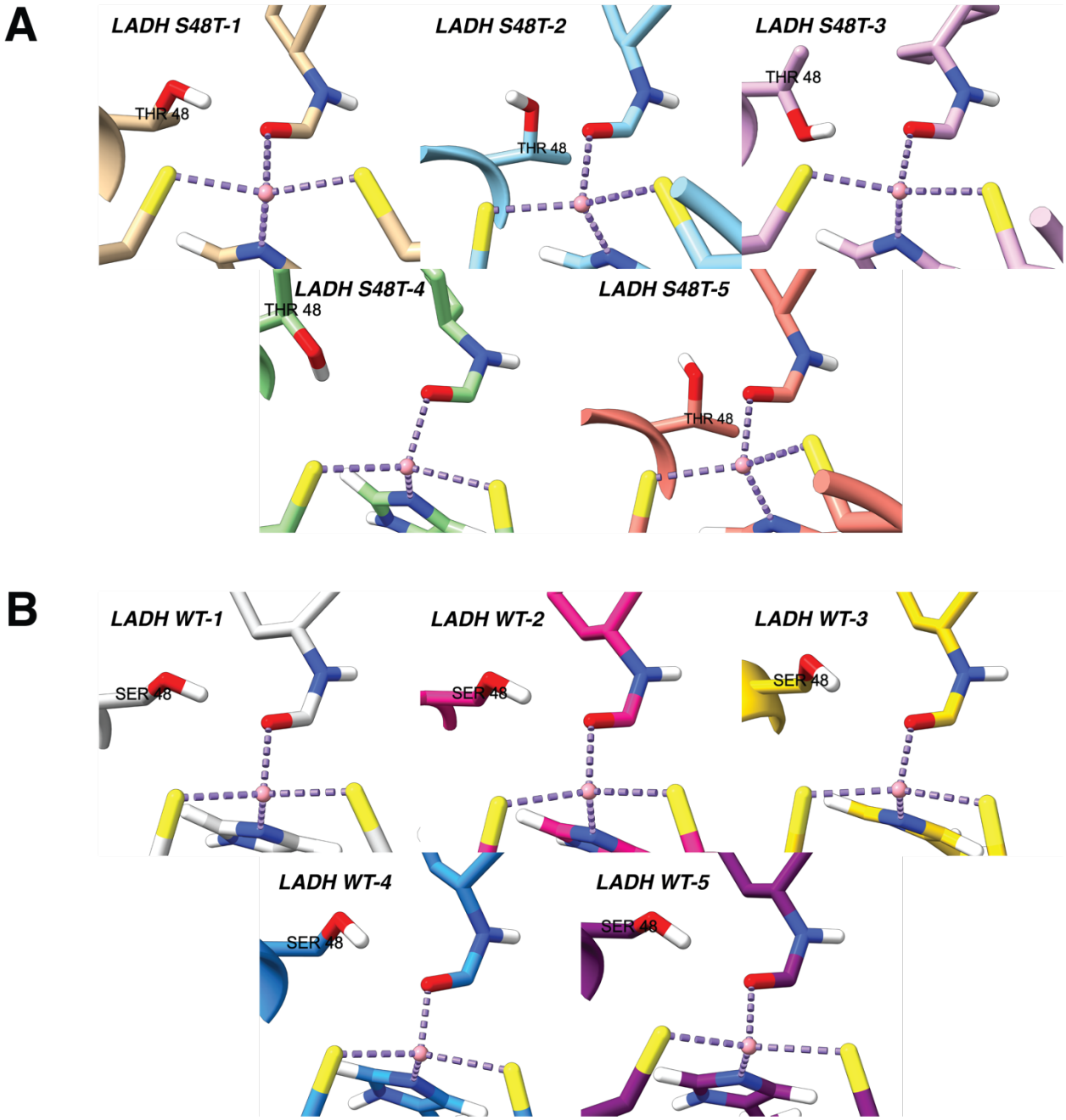
Extended Data Figure 4: (A) 3 proposed mechanisms for HIV-1 Protease, including a gem-diol intermediate (top), oxyanion intermediate (middle), and a concerted mechanism (bottom). (B) Mechanism for the rate-limiting step of cobalt-substituted Liver Alcohol Dehydrogenase

Extended Data Figure 5: Representative electric fields for wild-type LADH



Extended Data Figure 5: Five representative fields obtained from electric field topology-based clustering for WT Co-substituted LADH. In order from LADH WT-1 to LADH WT-5, these clusters have populations of 34.6%, 30.4%, 13.6%, 12.6%, and 8.8%. Field values next to pictures are in MV/cm.

Extended Data Figure 6: Fluctuations of residue 48 across LADH dynamics



Extended Data Figure 6: Representative frames for S48T (A) and WT (B) Co-substitued LADH, zoomed out to show the residue 48, labeled in each image. Note the structure diversity of this residue in the S48T mutant to the wild-type, an explanation for PC 0's behavior.

References

(1) Warshel, A. Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites. *J. Biol. Chem.* **1998**, *273* (42), 27035–27038. https://doi.org/10.1074/jbc.273.42.27035.

(2) Ramanan, R.; Danovich, D.; Mandal, D.; Shaik, S. Catalysis of Methyl Transfer Reactions by Oriented External Electric Fields: Are Gold–Thiolate Linkers Innocent? *J. Am. Chem. Soc.* **2018**, *140* (12), 4354–4362. https://doi.org/10.1021/jacs.8b00192.

(3) Pacchioni, G.; Lomas, J. R.; Illas, F. Electric Field Effects in Heterogeneous Catalysis. *J. Mol. Catal. Chem.* **1997**, *119* (1–3), 263–273. https://doi.org/10.1016/S1381-1169(96)00490-6.

(4) Lee, V.-J. Heterogeneous Catalysis: Effect of an Alternating Electric Field. *Science* **1966**, *152* (3721), 514–514. https://doi.org/10.1126/science.152.3721.514.

(5) Aragonès, A. C.; Haworth, N. L.; Darwish, N.; Ciampi, S.; Mannix, E. J.; Wallace, G. G.; Diez-Perez, I.; Coote, M. L. Electrostatic Catalysis of a Diels–Alder Reaction. *Nature* **2016**, *531* (7592), 88–91. https://doi.org/10.1038/nature16989.

(6) Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme Electric Fields Power Catalysis in the Active Site of Ketosteroid Isomerase. *Science* **2014**, *346* (6216), 1510–1514. https://doi.org/10.1126/science.1259802.

(7) Hennefarth, M. R.; Alexandrova, A. N. Direct Look at the Electric Field in Ketosteroid Isomerase and Its Variants. *ACS Catal.* **2020**, *10* (17), 9915–9924. https://doi.org/10.1021/acscatal.0c02795.

(8) 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* (32), 12487–12492. https://doi.org/10.1021/jacs.9b05323.

(9) Li, N.; Yan, S.; Wu, P.; Li, J.; Wang, B. Local Electric Fields Drives the Proton-Coupled Electron Transfer within Cytochrome P450 Reductase. *ACS Catal.* **2024**, *14* (10), 7893–7900. https://doi.org/10.1021/acscatal.4c02215.

(10) Chaturvedi, S. S.; Bím, D.; Christov, C. Z.; Alexandrova, A. N. From Random to Rational: Improving Enzyme Design through Electric Fields, Second Coordination Sphere Interactions, and Conformational Dynamics. *Chem. Sci.* **2023**, *14* (40), 10997–11011. https://doi.org/10.1039/D3SC02982D.

(11) Hennefarth, M. R.; Alexandrova, A. N. Advances in Optimizing Enzyme Electrostatic Preorganization. *Curr. Opin. Struct. Biol.* **2022**, *72*, 1–8. https://doi.org/10.1016/j.sbi.2021.06.006.

(12) Shaik, S. My Vision of Electric-Field-Aided Chemistry in 2050. *ACS Phys. Chem. Au* **2024**, *4* (3), 191–201. https://doi.org/10.1021/acsphyschemau.3c00064.

(13) Schwartz, S. D. Protein Dynamics and Enzymatic Catalysis. *J. Phys. Chem. B* **2023**, *127* (12), 2649–2660. https://doi.org/10.1021/acs.jpcb.3c00477.

(14) Callender, R.; Dyer, R. B. The Dynamical Nature of Enzymatic Catalysis. *Acc. Chem. Res.* **2015**, *48* (2), 407–413. https://doi.org/10.1021/ar5002928.

(15) Agarwal, P. K. A Biophysical Perspective on Enzyme Catalysis. *Biochemistry* **2019**, *58* (6), 438–449. https://doi.org/10.1021/acs.biochem.8b01004.

(16) Zoi, I.; Antoniou, D.; Schwartz, S. D. Electric Fields and Fast Protein Dynamics in Enzymes. *J. Phys. Chem. Lett.* **2017**, *8* (24), 6165–6170. https://doi.org/10.1021/acs.jpclett.7b02989.

(17) Pisliakov, A. V.; Cao, J.; Kamerlin, S. C. L.; Warshel, A. Enzyme Millisecond Conformational Dynamics Do Not Catalyze the Chemical Step. *Proc. Natl. Acad. Sci.* **2009**, *106* (41), 17359–17364. https://doi.org/10.1073/pnas.0909150106.
(18) Warshel, A.; Bora, R. P. Perspective: Defining and Quantifying the Role of Dynamics in Enzyme Catalysis. *J. Chem. Phys.* **2016**, *144* (18), 180901. https://doi.org/10.1063/1.4947037.
(19) Kohen, A. Role of Dynamics in Enzyme Catalysis: Substantial versus Semantic Controversies. *Acc. Chem. Res.* **2015**, *48* (2), 466–473. https://doi.org/10.1021/ar500322s.
(20) Suydam, I. T.; Snow, C. D.; Pande, V. S.; Boxer, S. G. Electric Fields at the Active Site of an Enzyme: Direct Comparison of Experiment with Theory. *Science* **2006**, *313* (5784), 200–204. https://doi.org/10.1126/science.1127159.
(21) Zheng, C.; Mao, Y.; Kozuch, J.; Atsango, A. O.; Ji, Z.; Markland, T. E.; Boxer, S. G. A Two-Directional Vibrational Probe Reveals Different Electric Field Orientations in Solution and an Enzyme Active Site. *Nat. Chem.* **2022**, *14* (8), 891–897. https://doi.org/10.1038/s41557-022-00937-w.
(22) Chaturvedi, S. S.; Vargas, S.; Ajmera, P.; Alexandrova, A. N. Directed Evolution of Protoglobin Optimizes the Enzyme Electric Field. *J. Am. Chem. Soc.* **2024**, *146* (24), 16670–16680. https://doi.org/10.1021/jacs.4c03914.
(23) Vargas, S.; Chaturvedi, S. S.; Alexandrova, A. N. Machine-Learning Prediction of Protein Function from the Portrait of Its Intramolecular Electric Field. *J. Am. Chem. Soc.* **2024**, jacs.4c09549. https://doi.org/10.1021/jacs.4c09549.
(24) Hennefarth, M. R.; Alexandrova, A. N. Heterogeneous Intramolecular Electric Field as a Descriptor of Diels–Alder Reactivity. *J. Phys. Chem. A* **2021**, *125* (5), 1289–1298. https://doi.org/10.1021/acs.jpca.1c00181.
(25) Piana, S.; Lindorff-Larsen, K.; Dirks, R. M.; Salmon, J. K.; Dror, R. O.; Shaw, D. E. Evaluating the Effects of Cutoffs and Treatment of Long-Range Electrostatics in Protein Folding Simulations. *PLoS ONE* **2012**, *7* (6), e39918. https://doi.org/10.1371/journal.pone.0039918.
(26) Morgenstern, A.; Jaszai, M.; Eberhart, M. E.; Alexandrova, A. N. Quantified Electrostatic Preorganization in Enzymes Using the Geometry of the Electron Charge Density. *Chem. Sci.* **2017**, *8* (7), 5010–5018. https://doi.org/10.1039/C7SC01301A.
(27) Karpen, M. E.; Tobias, D. J.; Brooks, C. L. Statistical Clustering Techniques for the Analysis of Long Molecular Dynamics Trajectories: Analysis of 2.2-Ns Trajectories of YPGDV. *Biochemistry* **1993**, *32* (2), 412–420. https://doi.org/10.1021/bi00053a005.
(28) Bertini, I.; Cavallaro, G.; Rosato, A. Cytochrome *c* : Occurrence and Functions. *Chem. Rev.* **2006**, *106* (1), 90–115. https://doi.org/10.1021/cr050241v.
(29) Bím, D.; Alexandrova, A. N. Local Electric Fields As a Natural Switch of Heme-Iron Protein Reactivity. *ACS Catal.* **2021**, *11* (11), 6534–6546. https://doi.org/10.1021/acscatal.1c00687.
(30) De Biase, P. M.; Doctorovich, F.; Murgida, D. H.; Estrin, D. A. Electric Field Effects on the Reactivity of Heme Model Systems. *Chem. Phys. Lett.* **2007**, *434* (1–3), 121–126. https://doi.org/10.1016/j.cplett.2006.11.104.
(31) Meng, E. C.; Goddard, T. D.; Pettersen, E. F.; Couch, G. S.; Pearson, Z. J.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX : Tools for Structure Building and Analysis. *Protein Sci.* **2023**, *32* (11), e4792. https://doi.org/10.1002/pro.4792.

(32) Donald, W. A.; Leib, R. D.; Demireva, M.; O'Brien, J. T.; Prell, J. S.; Williams, E. R. Directly Relating Reduction Energies of Gaseous $Eu(H_2O)_n^{3+}$, $n = 55-140$, to Aqueous Solution: The Absolute SHE Potential and Real Proton Solvation Energy. *J. Am. Chem. Soc.* **2009**, *131* (37), 13328–13337. https://doi.org/10.1021/ja902815v.
(33) Fawcett, W. R. The Ionic Work Function and Its Role in Estimating Absolute Electrode Potentials. *Langmuir* **2008**, *24* (17), 9868–9875. https://doi.org/10.1021/la7038976.
(34) Cheng, C.; Hayashi, S. Ab Initio Evaluation of the Redox Potential of Cytochrome c. *J. Chem. Theory Comput.* **2021**, *17* (2), 1194–1207. https://doi.org/10.1021/acs.jctc.0c00889.
(35) Brik, A.; Wong, C.-H. HIV-1 Protease: Mechanism and Drug Discovery. *Org. Biomol. Chem.* **2003**, *1* (1), 5–14. https://doi.org/10.1039/b208248a.
(36) Suguna, K.; Padlan, E. A.; Smith, C. W.; Carlson, W. D.; Davies, D. R. Binding of a Reduced Peptide Inhibitor to the Aspartic Proteinase from Rhizopus Chinensis: Implications for a Mechanism of Action. *Proc. Natl. Acad. Sci.* **1987**, *84* (20), 7009–7013. https://doi.org/10.1073/pnas.84.20.7009.
(37) Jaskolski, M.; Tomasselli, A. G.; Sawyer, T. K.; Staples, D. G.; Heinrikson, R. L.; Schneider, J.; Kent, S. B. H.; Wlodawer, A. Structure at 2.5-.ANG. Resolution of Chemically Synthesized Human Immunodeficiency Virus Type 1 Protease Complexed with a Hydroxyethylene-Based Inhibitor. *Biochemistry* **1991**, *30* (6), 1600–1609. https://doi.org/10.1021/bi00220a023.
(38) Lawal, M. M.; Sanusi, Z. K.; Govender, T.; Tolufashe, G. F.; Maguire, G. E. M.; Honarparvar, B.; Kruger, H. G. Unraveling the Concerted Catalytic Mechanism of the Human Immunodeficiency Virus Type 1 (HIV-1) Protease: A Hybrid QM/MM Study. *Struct. Chem.* **2019**, *30* (1), 409–417. https://doi.org/10.1007/s11224-018-1251-9.
(39) Sanusi, Z. K.; Lawal, M. M.; Gupta, P. Lal.; Govender, T.; Baijnath, S.; Naicker, T.; Maguire, G. E. M.; Honarparvar, B.; Roitberg, A. E.; Kruger, H. G. Exploring the Concerted Mechanistic Pathway for HIV-1 PR—Substrate Revealed by Umbrella Sampling Simulation. *J. Biomol. Struct. Dyn.* **2022**, *40* (4), 1736–1747. https://doi.org/10.1080/07391102.2020.1832578.
(40) Kumar, M.; Mandal, K.; Blakeley, M. P.; Wymore, T.; Kent, S. B. H.; Louis, J. M.; Das, A.; Kovalevsky, A. Visualizing Tetrahedral Oxyanion Bound in HIV-1 Protease Using Neutrons: Implications for the Catalytic Mechanism and Drug Design. *ACS Omega* **2020**, *5* (20), 11605–11617. https://doi.org/10.1021/acsomega.0c00835.
(41) Krzemińska, A.; Moliner, V.; Świderek, K. Dynamic and Electrostatic Effects on the Reaction Catalyzed by HIV-1 Protease. *J. Am. Chem. Soc.* **2016**, *138* (50), 16283–16298. https://doi.org/10.1021/jacs.6b06856.
(42) Ribeiro, A. J. M.; Santos-Martins, D.; Russo, N.; Ramos, M. J.; Fernandes, P. A. Enzymatic Flexibility and Reaction Rate: A QM/MM Study of HIV-1 Protease. *ACS Catal.* **2015**, *5* (9), 5617–5626. https://doi.org/10.1021/acscatal.5b00759.
(43) Zheng, C.; Ji, Z.; Mathews, I. I.; Boxer, S. G. Enhanced Active-Site Electric Field Accelerates Enzyme Catalysis. *Nat. Chem.* **2023**, *15* (12), 1715–1721. https://doi.org/10.1038/s41557-023-01287-x.
(44) Dinh, H. Q.; Xu, L. Measuring the Similarity of Vector Fields Using Global Distributions. In *Structural, Syntactic, and Statistical Pattern Recognition*; Da Vitoria Lobo, N., Kasparis, T., Roli, F., Kwok, J. T., Georgiopoulos, M., Anagnostopoulos, G. C., Loog, M., Eds.; Lecture Notes in Computer Science; Springer Berlin Heidelberg: Berlin, Heidelberg, 2008; Vol. 5342, pp 187–196. https://doi.org/10.1007/978-3-540-89689-0_23.

(45) Chiao-Fe Shu; Jain, R. C. Vector Field Analysis for Oriented Patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **1994**, *16* (9), 946–950. https://doi.org/10.1109/34.310692.
(46) Lavin, Y.; Batra, R.; Hesselink, L. Feature Comparisons of Vector Fields Using Earth Mover's Distance. In *Proceedings Visualization '98 (Cat. No.98CB36276)*; IEEE: Research Triangle Park, NC, USA, 1998; pp 103-109,. https://doi.org/10.1109/VISUAL.1998.745291.
(47) 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* (5), 2367–2373. https://doi.org/10.1021/acs.jcim.8b00958.
(48) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9* (9), 3878–3888. https://doi.org/10.1021/ct400314y.
(49) Li, P.; Merz, K. M. MCPB.Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016**, *56* (4), 599–604. https://doi.org/10.1021/acs.jcim.5b00674.
(50) Kästner, J.; Carr, J. M.; Keal, T. W.; Thiel, W.; Wander, A.; Sherwood, P. DL-FIND: An Open-Source Geometry Optimizer for Atomistic Simulations. *J. Phys. Chem. A* **2009**, *113* (43), 11856–11865. https://doi.org/10.1021/jp9028968.
(51) Lu, Y.; Farrow, M. R.; Fayon, P.; Logsdail, A. J.; Sokol, A. A.; Catlow, C. R. A.; Sherwood, P.; Keal, T. W. Open-Source, Python-Based Redevelopment of the ChemShell Multiscale QM/MM Environment. *J. Chem. Theory Comput.* **2019**, *15* (2), 1317–1328. https://doi.org/10.1021/acs.jctc.8b01036.
(52) Franzke, Y. J.; Holzer, C.; Andersen, J. H.; Begušić, T.; Bruder, F.; Coriani, S.; Della Sala, F.; Fabiano, E.; Fedotov, D. A.; Fürst, S.; Gillhuber, S.; Grotjahn, R.; Kaupp, M.; Kehry, M.; Krstić, M.; Mack, F.; Majumdar, S.; Nguyen, B. D.; Parker, S. M.; Pauly, F.; Pausch, A.; Perlt, E.; Phun, G. S.; Rajabi, A.; Rappoport, D.; Samal, B.; Schrader, T.; Sharma, M.; Tapavicza, E.; Treß, R. S.; Voora, V.; Wodyński, A.; Yu, J. M.; Zerulla, B.; Furche, F.; Hättig, C.; Sierka, M.; Tew, D. P.; Weigend, F. TURBOMOLE: Today and Tomorrow. *J. Chem. Theory Comput.* **2023**, *19* (20), 6859–6890. https://doi.org/10.1021/acs.jctc.3c00347.