

Advances in Optimizing Enzyme Electrostatic Preorganization

Matthew R. Hennefarth[†] and Anastassia N. Alexandrova^{*,†,‡}

[†]*Department of Chemistry and Biochemistry, University of California, Los Angeles, 607
Charles E. Young Drive East, Los Angeles, CA 90095-1569, USA*

[‡]*California NanoSystems Institute, University of California, Los Angeles, 570 Westwood
Plaza, Los Angeles, California 90095-1569, USA*

E-mail: *ana@chem.ucla.edu

Phone: +1 310 825-3769

Abstract

Utilizing electric fields to catalyze chemical reactions is not a new idea, but in enzymology it undergoes a renaissance, inspired by Warhsel's concept of electrostatic preorganization. According to this concept, the source of the immense catalytic efficiency of enzymes is the intramolecular electric field that permanently favors the reaction transition state over the reactants. Within enzyme design, computational efforts have fallen short in designing enzymes with natural-like efficacy. The outcome could improve if long-range electrostatics (often omitted in current protocols) would be optimized. Here, we highlight the major developments in methods for analyzing and designing electric fields generated by the protein scaffolds, in order to both better understand how natural enzymes function, and aid artificial enzyme design.

Introduction

Understanding the role and utilization of electric field to catalyze reactions has become a major topic of research. Electric field mediated catalysis breaks the canonical idea that a catalyst must be a reaction-specific chemical entity, such as a molecule or a surface. In general, any reaction that includes repolarization of the reacting system during the transition state (TS) could be influenced by an external electric field toward improvement of the forward reaction rate. Pioneering research by Warshel suggested that enzymes are efficient due to their ability to better stabilize the charge redistribution during the reaction than in solution [1–3]. A protein's ability to stabilize the charge redistribution arises from the optimal orientation of permanent dipoles within the enzyme scaffold. Warshel then summarized his ideas into his theory of protein electrostatic preorganization in 1998 [4] – the notion that enzyme-catalyzed reactions avoid an entropic penalty associated with the reorganization of the environment during TS crossing, since the scaffold (with its electric field) remains relatively fixed (Figure 1). At the same time, this field differentially favors the TS over the reactants. Thus, both the enthalpy and the entropy components of the free energy barrier are lowered. Electrostatic preorganization indeed has been shown to be present in several enzymes, including the most famously studied enzyme, Ketosteroid Isomerase [5–9]. While the generality of electrostatic preorganization in enzymes (and the role activation entropy plays in enzyme kinetics more generally) has been fiercely debated [10–16], a properly chosen electric field has been irrefutably shown (theoretically [17–23] and experimentally [24]) to increase the forward rate of various reactions. Thus, understanding enzymatic catalysis requires including intramolecular fields as one of the likely reactivity regulators.

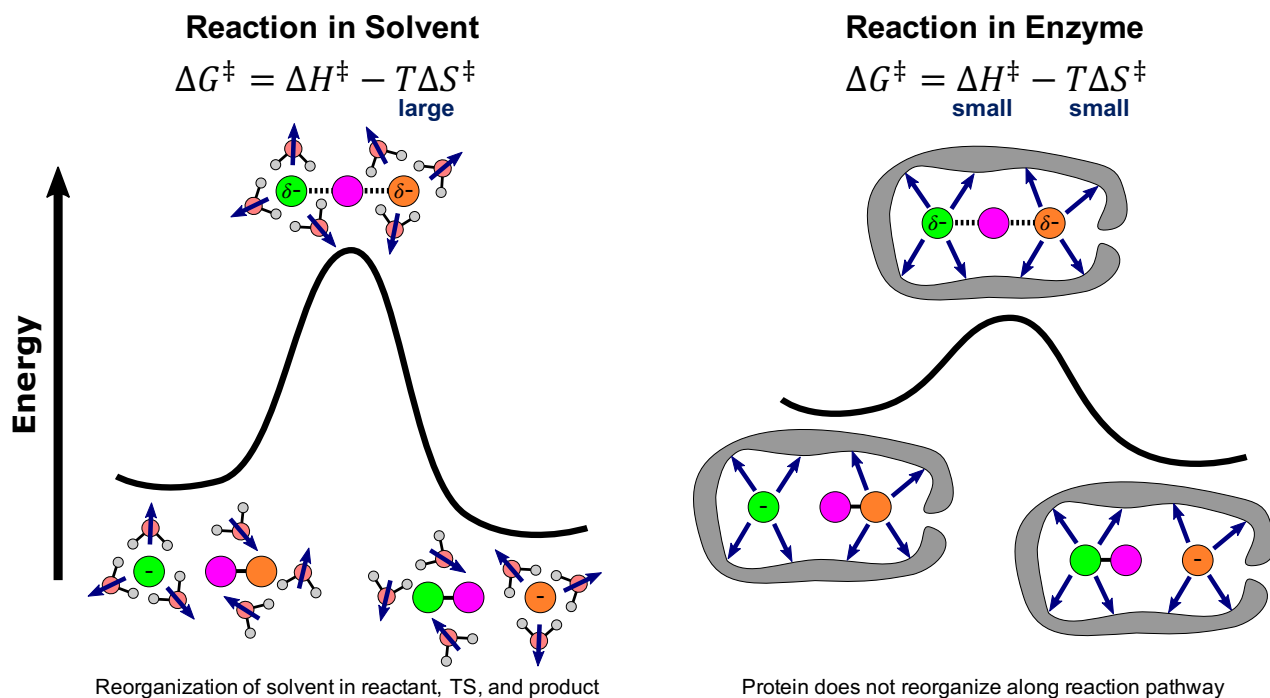


Figure 1. Schematics of a reaction coordinate for a model system in solvent (left) and an enzyme (right). For the reaction in solvent, there is a re-alignment of the solvent dipoles upon going from the reactant to TS, which is reflected in a large entropic contribution to the free energy barrier. Within the protein, the surrounding dipoles (and the field that they produce) are fixed along the reaction coordinate, and thus, no re-alignment is needed, nor possible, and the free energy barrier is thus lowered. This constitutes the effect of protein's electrostatic preorganization.

Since enzymes can catalyze reactions with incredible efficiency and stereo/enantioselectivity under physiological conditions, it is of interest to be able to design artificial enzymes that can catalyze any reaction of human interest with the same prowess. The most successful computational efforts produced the HG3 Kemp Eliminase [25], featuring k_{cat}/K_M on the order of $430 \text{ M}^{-1} \text{ s}^{-1}$. This performance, while top for computational design, falls several orders of magnitude short relative to that of natural enzymes (k_{cat}/K_M on the order of $10^5 \text{ M}^{-1} \text{ s}^{-1}$ [26]). In 2017, the Nobel prize in chemistry was awarded to Arnold and her pioneering work on laboratory directed evolution, which has been extremely successful in improving catalytic efficiency of designed enzymes and altering the functionality of natural enzymes. For example, the HG3 protein underwent 17 rounds of directed evolution to reach HG317 [27] with many orders of magnitude gain in efficiency (a final k_{cat}/K_M of about $230,000 \text{ M}^{-1} \text{ s}^{-1}$). Further, the family of natural

cytochrome P450 enzymes have been evolved to catalyze a wide number of reactions, from C-H insertion [28,29] to lactam synthesis [30]. Clearly, the objective function that computational enzyme design is currently optimizing is missing some key factors, and directed evolution is picking up the task. One of the factors which is missing from current design protocols is the incorporation of the favorable protein's electrostatic preorganization (or the electric field created by distant residues). In this review, we discuss the current work in trying to understand optimal electric fields for enzymatic catalysis, properly model proteins with respect to their electric fields, as well as methods developed to tackle the inverse design problem (how to create an active site that produces the best electrostatic preorganization). Being able to properly account for the protein's electric field, and to design enzymes with optimal electrostatic preorganization should augment computational design efforts.

Modeling Electrostatics in Proteins

Proteins often comprise several hundred to thousands of atoms, and that is too expensive to be modeled using quantum mechanics, especially in view of the sampling needed for converged free energies. Instead, classical or QM/MM molecular dynamics (MD) is used, based on solving Newton equation of motions on the approximate potential energy surfaces. In determining the forces on atoms, long-range electrostatic interactions are included via the classical Coulomb's law (though are routinely approximated using the particle mesh Ewald summation with an appropriate cutoff to increase the speed of calculations [31]), and using partial charges assigned to each atom type within the force field. Thus, the accurate representation of electrostatics depends on the quality of the force field. Recently, Essex et al. showed that both Amber ff14SB [32] and Charmm C36m [33] fall short in reproducing the electric field within the active site of the CypA bound to HIV capsid protein [34]. Instead, the AMOEBA polarizable force field [35], which adds correcting terms to the partial charges to account for the asymmetric electronic distribution around the nuclei,

was shown to accurately reproduce the quantum mechanical picture of the electric field within the active site along an MD trajectory.

Another critical yet non-trivial aspect is the protonation states of residues within the protein, which can be non-standard (fortunately rarely). They can be important structurally, and, both through structure, and directly, impact the electric field within the protein. In turn, fields and local dielectrics influence the protonation states, and the resultant protein fold. Many current implementations of MD have protonation states fixed during the course of the simulations. However, recent efforts have been made towards implementing titration within MD allowing the protonation states of residues to change [36–40]. Because equilibrating protonation states requires particularly lengthy simulations (especially with explicit solvent, as is usually the case), our group has recently developed a titratable feature within the pi-DMD software [41,42] which takes advantage of implicit solvation and simplified discrete potential energy surfaces, to simulate proteins at much longer time scales [43].

Finally, ions in solution, and posttranslational protein modifications may impact the electric fields within active sites. These effects are rarely modeled, but it is also likely that they are too remote to exert a great influence on the catalysis.

In summary, in order to evaluate and understanding electric fields in proteins, and to eventually enable their incorporation into computational enzyme design protocols, accurate modeling is essential. The key aspects of such accurate modeling include proper force fields (polarizable, if possible), treatment of protonation states concurrent with the dynamics, and possibly accounting for co-present ions.

Understanding the Optimal Electric Field

Our understanding of how an electric field couples to a reaction of choice to modulate the reactivity is incomplete. Recent works by Shaik have emphasized that the electric field parallel

with the reaction axis (direction of electron reorganization) controls the forward rate of reaction for single-step reactions [17,44]. For example, for the Diels-Alder reactions, it has been shown that the electric field parallel to the bond-forming axis controls the rate of reaction [19,20,24]. Additionally, we showed that in the family of natural heme-iron oxidoreductases, the direction of the field at the Fe is highly conserved and parallel to the Fe-O bond. The strength of this field is characteristic of the protein function, i.e. oxygenases, catalases, and peroxidases all have distinct fields (Figure 2) [45]. However, for certain reactions, it is impossible to choose a simple external field to align with a single dipole change during the reaction, because more than one dipole changes and/or more than one bond forms/breaks. In such cases, the reaction might still benefit from a field, but the field might need to be heterogeneous and more intricate.

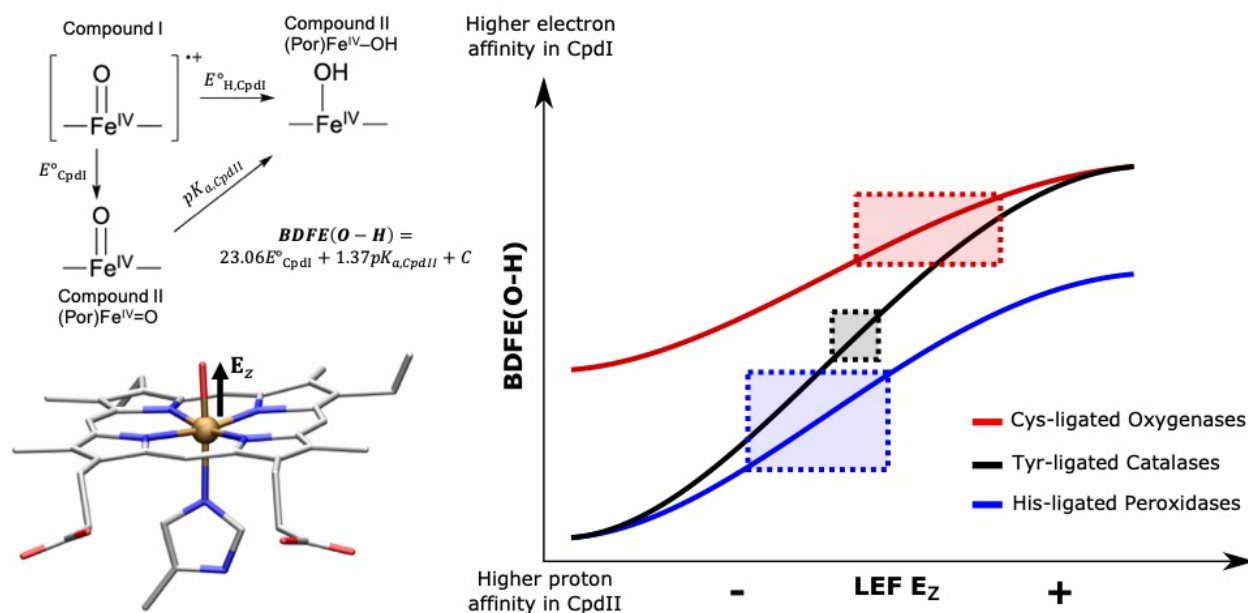


Figure 2. Electric fields generated from the protein scaffold in Heme-iron oxidoreductases are indicative of their function. Top left: Thermodynamic cycle for calculating the BDFE(O-H) between Compound I (CpdI) and II (CpdII). Bottom left: direction of the local electric field (LEF) generated by the protein scaffold (along the Fe-O bond). Positive direction of this field goes from iron to oxygen. Right: Illustrative graph of the BDFE(O-H) as a function of the LEF. Shaded regions illustrate field strengths where canonical function of a protein occurs. For more-negative LEF, there is not enough oxidative power for CpdI, whereas for more positive LEF, there is off-pathway oxidation. Hence, the proteins have tuned their electric fields to fall within the canonical function regions.

Indeed, complexity arises in analyzing the electric field within enzyme active sites, since it in general is a 3D vector field, and it is far from uniform in any chemical system (Figure 3). Thus, the relationships between topological changes of the field in chemically relevant regions and the reactivity of a protein can be non-trivial and call for a more global view on the field than its assessment at a single point or projected on a single dipole. We have proposed two metrics by which to quantify the electrostatic preorganization within a protein. The first is through the changes in the topological parameters of the full, 3-D reactant state electron density. We showed a subset of these features (e.g., the electrostatic potential and electron density at bond and ring critical points) to correlate linearly with the applied electric field, and also correlate with the reaction barrier, thus establishing a rigorous quantum mechanical link between electrostatic preorganization and reactivity. This was shown for Histone Deacetylase 8 (HDAC8) [46], KSI [18], and the Diels-Alder reactions in solution and catalyzed by artificial enzymes [19]. As such, the topology of the electron density contains a signature of the electric field which is applied to the system.

More recently, we have proposed a way to directly compare the electric field topologies through the global distribution of field lines, as a metric of the differences in reactivity of chemically related systems (e.g. in several variants of an enzyme). In both the KSI and Diels-Alder reactions, we have shown that topologically similar electric fields correspond to similar barriers of the corresponding reaction [19,47]. Furthermore, we showed how fields only at discrete points (which is what is often computed/measured) do not fully capture a relationship to the reaction barrier [47]. Instead, a more global, 3-D geometry of the electric field around relevant bonds are predictive of the reactivity. The charge density and field topologies computed at the reactant states are rigorous and straightforward enough and allow for an accurate and quick estimation of electrostatic preorganization and resultant reactivity in enzyme. For example, the impact of point-

mutations on the electric field that the scaffold generates, and on the reaction barrier, could be quickly screened via ground state MD. Note that these descriptors, based on charge density and fields in the reactant (1) permit avoiding explicit (and tedious) barrier calculations, and (2) are continuous and information-rich, which makes them suitable for machine learning approaches (S. Vargas *et al.*, unpublished).

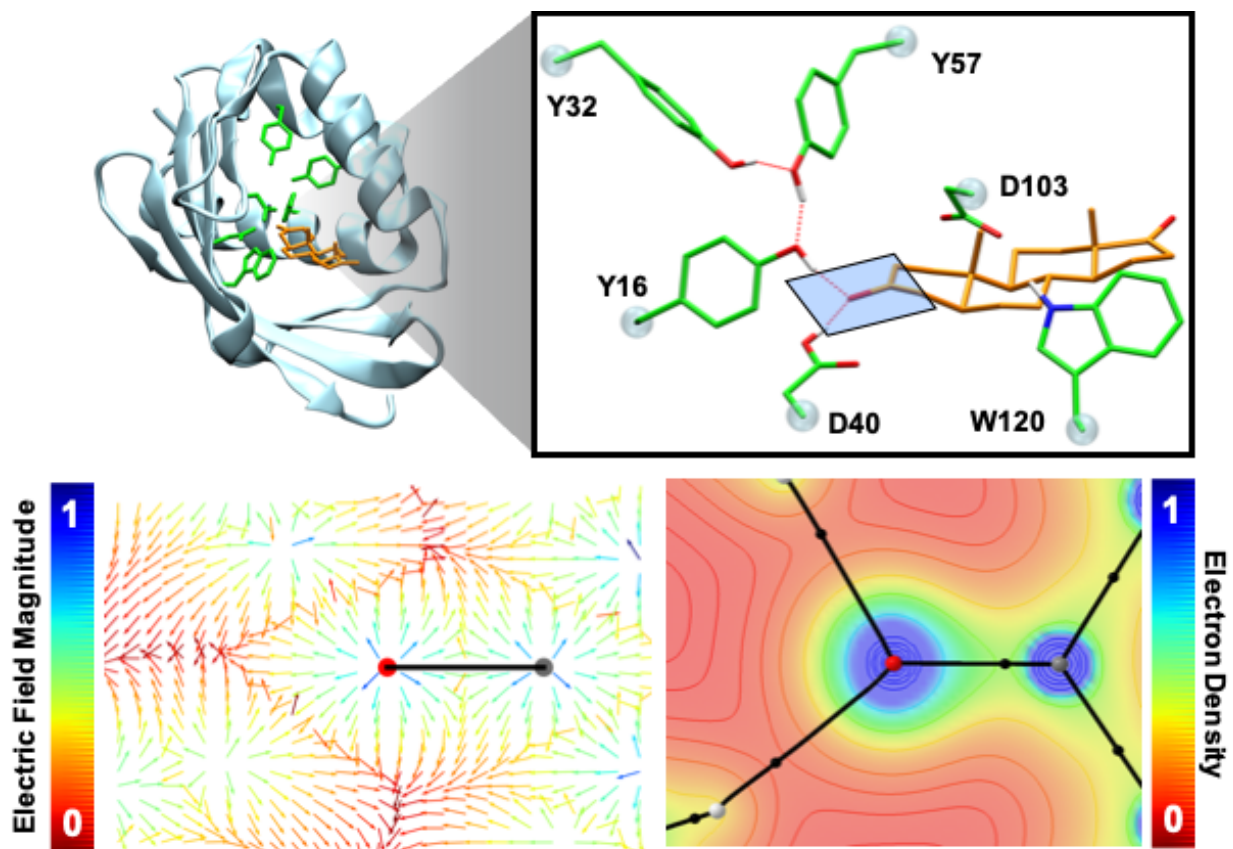


Figure 3. One of the most studied proteins for electrostatic preorganization, KSI (PDB Code 1OH0 [48]): it has been shown that the electric field around the substrate carbonyl (outlined in top right) gives the protein the large catalytic efficiency. Our recent work has shown that topological considerations of both the internally defined electric field (bottom left) through the global distribution of streamlines [49] and the electron density (bottom right) through the Quantum Theory of Atoms in Molecules formalism [50] can be used as rigorous and accurate metrics of protein's electrostatic preorganization.

Inverse Design Problem

Since proteins produce heterogeneous fields in active sites, and the optimal field benefitting a reaction is also often heterogeneous, the design of a novel, optimal field, or a protein that would produce it for a given reaction is in general a complicated problem. A field may be organized by

specifically placed charged groups around the active site. Hence, first, the entire space around the active site needs to be sampled with various distributions of charge embeddings, and the desired distribution would have to be mapped on a chemical space obtainable in a protein (Figure 4). Very recently, advances have been made on the inverse design problem that greatly reduced the dimensionality of the space needed to investigate, as well as utilize machine learning algorithms to speed up the search.

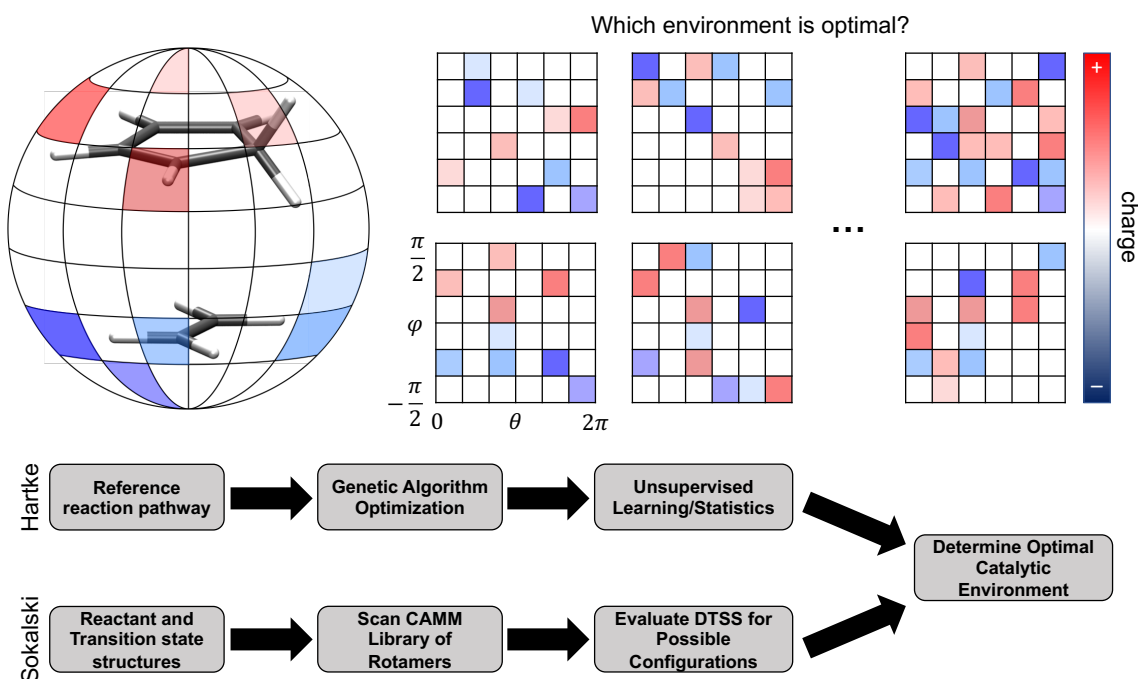


Figure 4. Electrostatic inverse design problem consists of finding the optimal placement of charged residues or point charges around the reaction that decreases the reaction barrier. Considering a partitioned spherical surface around, e.g., a Diels-Alder reaction, with each patch of the partition assigned a charge density (blue - negatively charged; red - positively charged), the total number of possible environments is theoretically infinite, since the charge at any point can be any real-valued number, and the partitioning can be infinitely fine. This yields an enormous search space. In addition, solutions will not be unique with various configurations of the electrostatic environment yielding similar reaction barriers, as well as possible solutions completely altering the reaction pathway, which is no longer feasible within a protein. Efforts by Hartke and Sokalski have attempted to reduce this search space by using machine learning or by minimizing the E_{DTSS} for the given reaction to determine an optimal catalytic environment.

Head-Gordon et al. implemented an iterative procedure, similar to directed evolution, with the objective being that each mutation brings an increase in the electric field along the reaction

coordinate or substrate positioning in the active site [51]. The four mutations introduced in this way improved the k_{cat}/K_M by ca. 20-fold for the KE15 Kemp Eliminate. Specifically, they considered the electric field along each bond that is either formed or broken during the course of the reaction. While this method shows promise, they only consider the electric field change from mutating a residue from a positive to negative or negative to positive charge, whereas in reality, upon mutation, there is a change in the electric field due to slight (or large) geometrical changes in the entire protein fold. Improvement of the method could also take into account these geometrical impacts on the electric field in the active site.

Sokalski et al. recently introduced a bottom-up approach for optimizing the placements of charged amino acids around a reaction toward speeding it up. The idea stems from differential transition state stabilization (DTSS) energy [52–54], which is a measure of the change in reaction barriers between the WT and “mutated” system by placing point charges around the system ($q_i(\mathbf{r}_i)$). They observed that the long-range multipolar electrostatic component reflects the charge-redistribution along the reaction pathway. Hence, they define the DTSS energy as a function of the catalytic field (Δ_s), which is the difference in electrostatic potentials between the transition state and reactant state.

$$E_{DTSS} \approx \sum_i q_i(\mathbf{r}_i)(V^{TS}(\mathbf{r}_i) - V^S(\mathbf{r}_i)) = - \sum_i q_i(\mathbf{r}_i)\Delta_s(\mathbf{r}_i)$$

This catalytic field, when mapped on a surface surrounding the reaction, defines regions where positively and negatively charged residues should be placed in order to decrease the reaction barrier. Using this catalytic field in conjunction with a rotamer library for sampling amino acid conformations, Sokalski was able to reproduce directed evolution mutations in the second-coordination sphere of the KE07 Kemp Eliminate protein, indicating that these mutations represent an optimization of the protein’s electrostatic preorganization [55].

In a similar vein, Dittner and Hartke, have utilized a genetic algorithm to discover the optimal placement of ‘globally optimized catalysts’ (GOCATs) (in the particular case of their studies, point charges) around a reaction, by minimizing the reaction barrier through a series of nudged elastic band (NEB) calculations [56,57]. Several constraints are imposed on the system to ensure that the reactant and product remain stationary states. The first iteration of the algorithm was applied to a simple SN2 Menshutkin reaction and a second to a Diels-Alder reaction. Interestingly, their method was able to reproduce the results of experimental and theoretical work on these system, by regenerating the optimal field that was determined by Shaik and Coote [20,24]. However, the approach has only worked with optimizing partial charge distributions. The realization of these charge distributions within a larger catalyst such as an enzyme would require proper placement of the charged residues attached to the protein scaffold, or placement of charged ions near the reactive site [58]. The generalization of the method to GOCATs allows them to change point charges for charged amino acid residues or charged ions, though the placement of these entities within a protein still poses a difficult task, since the protein sequence and fold are interdependent. Additionally, for each point charge distribution, a costly NEB calculation is performed, which might become prohibitively expensive especially within a full protein. However, these are important steps forward in designing for optimal intramolecular fields in enzymes.

Conclusions

Being able to utilize enzymes to catalyze any reaction of human interest has been a long-standing goal within the catalysis community. Current enzyme design protocols omit long-range electrostatic interactions (or the protein’s electrostatic preorganization) due to the difficulty in quantifying and understanding the optimal environment for a given reaction. Advances in proper modeling of these long-range interactions firstly rely on obtaining accurate thermodynamic structural ensembles of the protein scaffold, which, within classical MD simulations, depends on

the accuracy of the force field, and in some cases the treatment of protonation states of amino acids and placement of ions. Further, recent work has analyzed how electric fields in general can be utilized as catalysts, and more specifically, how protein electric fields influence their function. To this end, our group has shown that the enzyme active site electron density topology as well as the electric field topology report on the reactivity of the system. These findings equip researchers with tools to rigorously assess the protein's electrostatic preorganization, based on the properties of the reactant state. New methods are being developed for the design of enzyme active site environments that improve electrostatic preorganization. Several foci of this sort can be highlighted: mimicking directed evolution by hand, utilizing machine learning, and optimizing for the DTSS energy to reduce the effective search space. While these works have focused on the primary and secondary coordination spheres for the reaction, ultimately, the entire protein scaffold will have to be considered when designing an artificial enzyme with the long-range electrostatics in mind. Even though remote parts of the protein are unlikely to exert strong enough fields on the active site, they impact the overall protein structure and the placement of the more proximal residues and affect the fields in this indirect way. We also note that much of the work on optimizing electric fields is within the context of a static protein, whereas in reality protein dynamics can alter the field, and ideally for catalysis one would wish for a narrow field distribution around the optimum for the reaction. Thus, future work should be motivated in reducing electric field fluctuations as well as optimizing the average field.

Acknowledgement

This work was supported by the NSF CHE-1903808 grant.

Annotations

* of special interest, and ** of outstanding interest

* Hennefarth MR, Alexandrova AN: **Direct Look at the Electric Field in Ketosteroid Isomerase and Its Variants.** *ACS Catal* 2020, **10**:9915–9924.

The electric field within the wildtype Ketosteroid Isomerase protein with and without external electric fields along with several mutants is probed at several locations. It is shown that evaluating the electric field at discrete points, such as at the geometrical center of bond, does not correlate with the changes in reaction barrier. Instead, full consideration of the electric field topology in a region around chemical bonds of interest yields correlations with the changing reaction barrier due to the external fields as well as point mutations.

*Bradshaw RT, Dziedzic J, Skylaris CK, Essex JW: **The Role of Electrostatics in Enzymes: Do Biomolecular Force Fields Reflect Protein Electric Fields?** *J Chem Inf Model* 2020, **60**:3131–3144.

Using the peptidyl-prolyl isomerase cyclophilin A, it is shown that the AMOEBA polarizable force field is more accurate than the Amber ff14SB and Charmm C36m force fields for calculation the electric field within the protein's active site. Thus, the choice of force fields is important when computationally evaluating the electric field generated by the protein scaffold.

*Vaissier V, Sharma SC, Schaettle K, Zhang T, Head-Gordon T: **Computational Optimization of Electric Fields for Improving Catalysis of a Designed Kemp Eliminase.** *ACS Catal* 2018, **8**:219–227.

Residues within the KE15 Kemp Eliminase protein are screened for their electric field contributions within bond breaking/forming locations of the active site. Those which contribute electric fields that hinder the reaction are mutated such that they instead improve the electric field, similar to how laboratory directed evolution is performed. Subsequently, the resulting protein showed a 20-fold increase in catalytic efficiency by tuning the active site electric field.

Beker W, Sokalski WA: **Bottom-Up Nonempirical Approach to Reducing Search Space in Enzyme Design Guided by Catalytic Fields. *J Chem Theory Comput* 2020, **16**:3420–3429.

Using the KE07 Kemp Eliminase, protein mutations are guided using a library of atomic multipoles for side-chain rotamers in conjunction with the catalytic field (which optimizes the differential transition state stabilization energy). Mutations were in qualitative agreement with laboratory directed evolution for the same protein, indicating that directed evolution optimized the protein's electrostatic preorganization.

Dittner M, Hartke B: **Globally optimal catalytic fields for a Diels-Alder reaction. *J Chem Phys* 2020, **152**.

Globally optimal catalysts (point charges in this case) are optimally placed around a simple Diels-Alder reaction using a genetic algorithm without any *a priori* information. The resulting charge distributions regenerates the optimal electric field for the reaction from prior computational research. Future extensions include optimizing placement of protein residues, rather than point charges, around the reaction.

References

1. Warshel A: **Energetics of enzyme catalysis (lysozyme/enzyme mechanism/dielectric effects in enzymes and in solutions/relationship between protein folding and catalysis).** *Proc Natl Acad Sci* 1978, **75**:5250–5254.
2. Warshel A: **Calculations of Enzymatic Reactions: Calculations of pKa, Proton Transfer Reactions, and General Acid Catalysis Reactions in Enzymes.** *Biochemistry* 1981, **20**:3167–3177.
3. Hwang JK, Warshel A: **Semiquantitative Calculations of Catalytic Free Energies in Genetically Modified Enzymes.** *Biochemistry* 1987, **26**:2669–2673.
4. Warshel A: **Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites.** *J Biol Chem* 1998, **273**:27035–27038.
5. Wang L, Fried SD, Markland TE: **Proton network flexibility enables robustness and large electric fields in the ketosteroid isomerase active site.** *J Phys Chem B* 2017, **121**:9807–9815.
6. Welborn VV, Head-Gordon T: **Fluctuations of Electric Fields in the Active Site of the Enzyme Ketosteroid Isomerase.** *J Am Chem Soc* 2019, **141**:12487–12492.
7. Kamerlin SCL, Sharma PK, Chu ZT, Warshel A: **Ketosteroid isomerase provides further support for the idea that enzymes work by electrostatic preorganization.** *Proc Natl Acad Sci* 2010, **107**:4075–4080.
8. Fried SD, Bagchi S, Boxer SG: **Extreme electric fields power catalysis in the active site of ketosteroid isomerase.** *Science (80-)* 2014, **346**:1510–1514.
9. Wu Y, Fried SD, Boxer SG: **A Preorganized Electric Field Leads to Minimal Geometrical Reorientation in the Catalytic Reaction of Ketosteroid Isomerase.** *J Am Chem Soc* 2020, **142**:9993–9998.
10. Åqvist J, Kazemi M, Isaksen GV, Brandsdal BO: **Entropy and Enzyme Catalysis.** *Acc Chem Res* 2017, **50**:199–207.
11. Jencks WP: **Binding Energy, Specificity, and Enzymic Catalysis: The Circe Effect.** 2006:219–410.
12. Jencks WP: **From Chemistry to Biochemistry to Catalysis to Movement.** *Annu Rev Biochem* 1997, **66**:1–18.
13. Wolfenden R, Snider MJ: **The Depth of Chemical Time and the Power of Enzymes as Catalysts.** *Acc Chem Res* 2001, **34**:938–945.
14. Bruice TC, Lightstone FC: **Ground State and Transition State Contributions to the Rates of Intramolecular and Enzymatic Reactions.** *Acc Chem Res* 1999, **32**:127–136.
15. Shurki A, Štrajbl M, Villà J, Warshel A: **How Much Do Enzymes Really Gain by Restraining Their Reacting Fragments?** *J Am Chem Soc* 2002, **124**:4097–4107.
16. Villa J, Strajbl M, Glennon TM, Sham YY, Chu ZT, Warshel A: **How important are entropic contributions to enzyme catalysis?** *Proc Natl Acad Sci* 2000, **97**:11899–11904.
17. Shaik S, Mandal D, Ramanan R: **Oriented electric fields as future smart reagents in chemistry.** *Nat Chem* 2016, **8**:1091–1098.
18. Fuller III J, Wilson TR, Eberhart ME, Alexandrova AN, Fuller J, Wilson TR, Eberhart ME, Alexandrova AN: **Charge Density in Enzyme Active Site as a Descriptor of Electrostatic Preorganization.** *J Chem Inf Model* 2019, **59**:2367–2373.
19. Hennefarth MR, Alexandrova AN: **Heterogeneous Intramolecular Electric Field as a Descriptor of Diels–Alder Reactivity.** *J Phys Chem A* 2021, doi:10.1021/acs.jpca.1c00181.
20. 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.
21. Bhattacharyya K, Karmakar S, Datta A: **External electric field control: driving the reactivity of metal-free azide–alkyne click reactions.** *Phys Chem Chem Phys* 2017, **19**:22482–22486.
 22. Che F, Gray JT, Ha S, Kruse N, Scott SL, McEwen JS: **Elucidating the Roles of Electric Fields in Catalysis: A Perspective.** *ACS Catal* 2018, **8**:5153–5174.
 23. Welborn VV, Ruiz Pestana L, Head-Gordon T: **Computational optimization of electric fields for better catalysis design.** *Nat Catal* 2018, **1**:649–655.
 24. Aragonès AC, Haworth NL, Darwish N, Ciampi S, Bloomfield NJ, Wallace GG, Diez-Perez I, Coote ML: **Electrostatic catalysis of a Diels–Alder reaction.** *Nature* 2016, **531**:88–91.
 25. Privett HK, Kiss G, Lee TM, Blomberg R, Chica RA, Thomas LM, Hilvert D, Houk KN, Mayo SL: **Iterative approach to computational enzyme design.** *Proc Natl Acad Sci* 2012, **109**:3790–3795.
 26. Bar-Even A, Noor E, Savir Y, Liebermeister W, Davidi D, Tawfik DS, Milo R: **The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters.** *Biochemistry* 2011, **50**:4402–4410.
 27. Blomberg R, Kries H, Pinkas DM, Mittl PRE, Grütter MG, Privett HK, Mayo SL, Hilvert D: **Precision is essential for efficient catalysis in an evolved Kemp eliminase.** *Nature* 2013, **503**:418–421.
 28. Zhang RK, Chen K, Huang X, Wohlschlager L, Renata H, Arnold FH: **Enzymatic assembly of carbon–carbon bonds via iron-catalysed sp³ C–H functionalization.** *Nature* 2019, **565**:67–72.
 29. Zhang RK, Huang X, Arnold FH: **Selective C H bond functionalization with engineered heme proteins: new tools to generate complexity.** *Curr Opin Chem Biol* 2019, **49**:67–75.
 30. Cho I, Jia Z, Arnold FH: **Genetically Tunable Enzymatic C – H Amidation for Lactam Synthesis.** *Chem Rxivs* 2019, doi:10.26434/chemrxiv.7711418.v1.
 31. Darden T, York D, Pedersen L: **Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems.** *J Chem Phys* 1993, **98**:10089–10092.
 32. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C: **ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB.** *J Chem Theory Comput* 2015, **11**:3696–3713.
 33. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, Grubmüller H, MacKerell AD: **CHARMM36m: an improved force field for folded and intrinsically disordered proteins.** *Nat Methods* 2017, **14**:71–73.
 34. Bradshaw RT, Dziedzic J, Skylaris CK, Essex JW: **The Role of Electrostatics in Enzymes: Do Biomolecular Force Fields Reflect Protein Electric Fields?** *J Chem Inf Model* 2020, **60**:3131–3144.
 35. Ponder JW, Wu C, Ren P, Pande VS, Chodera JD, Schnieders MJ, Haque I, Mobley DL, Lambrecht DS, DiStasio RA, et al.: **Current Status of the AMOEBA Polarizable Force Field.** *J Phys Chem B* 2010, **114**:2549–2564.
 36. Lee MS, Salsbury FR, Brooks CL: **Constant-pH molecular dynamics using continuous titration coordinates.** *Proteins Struct Funct Bioinforma* 2004, **56**:738–752.
 37. Goh GB, Hulbert BS, Zhou H, Brooks CL: **Constant pH molecular dynamics of proteins in explicit solvent with proton tautomerism.** *Proteins Struct Funct Bioinforma* 2014, **82**:1319–1331.

38. Chen Y, Roux B: **Constant-pH Hybrid Nonequilibrium Molecular Dynamics–Monte Carlo Simulation Method.** *J Chem Theory Comput* 2015, **11**:3919–3931.
39. Harris RC, Shen J: **GPU-Accelerated Implementation of Continuous Constant pH Molecular Dynamics in Amber: p K a Predictions with Single-pH Simulations.** *J Chem Inf Model* 2019, **59**:4821–4832.
40. Grünewald F, Souza PCT, Abdizadeh H, Barnoud J, de Vries AH, Marrink SJ: **Titrateable Martini model for constant pH simulations.** *J Chem Phys* 2020, **153**:024118.
41. Proctor EA, Ding F, Dokholyan N V.: **Discrete molecular dynamics.** *WIREs Comput Mol Sci* 2011, **1**:80–92.
42. Shirvanyants D, Ding F, Tsao D, Ramachandran S, Dokholyan N V.: **Discrete Molecular Dynamics: An Efficient And Versatile Simulation Method For Fine Protein Characterization.** *J Phys Chem B* 2012, **116**:8375–8382.
43. Reilley D, Alexandrova AN, Wang J, Dokholyan N: **Tit-DMD – A Rapid, Coarse-Grained Quasi-All-Atom Constant pH Molecular Dynamics Framework.** 2021, doi:10.26434/chemrxiv.13717870.v2.
44. Shaik S, de Visser SP, Kumar D: **External Electric Field Will Control the Selectivity of Enzymatic-Like Bond Activations.** *J Am Chem Soc* 2004, **126**:11746–11749.
45. Bím D, Alexandrova AN: **Local Electric Fields as a Natural Switch of Heme-Iron Protein Reactivity.** 2021, doi:10.26434/chemrxiv.13615853.v1.
46. Morgenstern A, Jaszai M, Eberhart ME, Alexandrova AN: **Quantified electrostatic preorganization in enzymes using the geometry of the electron charge density.** *Chem Sci* 2017, **8**:5010–5018.
47. Hennefarth MR, Alexandrova AN: **Direct Look at the Electric Field in Ketosteroid Isomerase and Its Variants.** *ACS Catal* 2020, **10**:9915–9924.
48. Kim SW, Cha SS, Cho HS, Kim JS, Ha NC, Cho MJ, Joo S, Kim KK, Choi KY, Oh BH: **High-resolution crystal structures of delta5-3-ketosteroid isomerase with and without a reaction intermediate analogue.** *Biochemistry* 1997, **36**:14030–6.
49. Dinh HQ, Xu L: **Measuring the Similarity of Vector Fields Using Global Distributions.** In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. . Springer, Berlin, Heidelberg; 2008:187–196.
50. Bader RFW: **Atoms in molecules.** *Acc Chem Res* 1985, **18**:9–15.
51. Vaissier V, Sharma SC, Schaettle K, Zhang T, Head-Gordon T: **Computational Optimization of Electric Fields for Improving Catalysis of a Designed Kemp Eliminase.** *ACS Catal* 2018, **8**:219–227.
52. Sokalski WA: **Theoretical model for exploration of catalytic activity of enzymes and design of new catalysts: CO2 hydration reaction.** *Int J Quantum Chem* 1981, **20**:231–240.
53. Sokalski WA: **The physical nature of catalytic activity due to the molecular environment in terms of intermolecular interaction theory: derivation of simplified models.** *J Mol Catal* 1985, **30**:395–410.
54. Szeftczyk B, Mulholland AJ, Ranaghan KE, Sokalski WA: **Differential Transition-State Stabilization in Enzyme Catalysis: Quantum Chemical Analysis of Interactions in the Chorismate Mutase Reaction and Prediction of the Optimal Catalytic Field.** *J Am Chem Soc* 2004, **126**:16148–16159.
55. Beker W, Sokalski WA: **Bottom-Up Nonempirical Approach to Reducing Search Space in Enzyme Design Guided by Catalytic Fields.** *J Chem Theory Comput* 2020,

- 16:3420–3429.
56. Dittner M, Hartke B: **Globally Optimal Catalytic Fields - Inverse Design of Abstract Embeddings for Maximum Reaction Rate Acceleration.** *J Chem Theory Comput* 2018, **14**:3547–3564.
57. Dittner M, Hartke B: **Globally optimal catalytic fields for a Diels-Alder reaction.** *J Chem Phys* 2020, **152**.
58. Nechay MR, Gallup NM, Morgenstern A, Smith QA, Eberhart ME, Alexandrova AN: **Histone Deacetylase 8: Characterization of Physiological Divalent Metal Catalysis.** *J Phys Chem B* 2016, **120**:5884–5895.