Title

Methods for Theoretical Treatment of Local Fields in Proteins and Enzymes

Journal

Chemical Reviews, 125(7)

ISSN

0009-2665

Authors

Eberhart, Mark E
Alexandrova, Anastassia N
Ajmera, Pujan
et al.

Publication Date

2025-02-24

DOI

10.1021/acs.chemrev.4c00471

Peer reviewed

CHEMICAL REVIEWS

pubs.acs.org/CR

Review

Methods for Theoretical Treatment of Local Fields in Proteins and Enzymes

Published as part of Chemical Reviews special issue "Electric Fields in Chemistry and Biology".

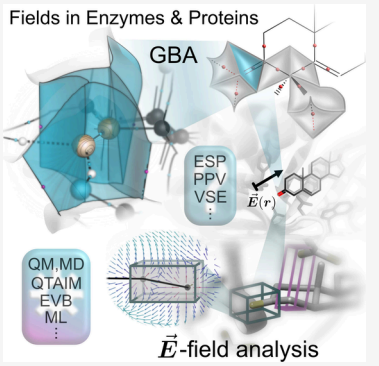
Mark E. Eberhart,* Anastassia N. Alexandrova,* Pujan Ajmera,§ Daniel Bím,§ Shobhit S. Chaturvedi,§ Santiago Vargas,§ and Timothy R. Wilson§

Cite This: *Chem. Rev.* 2025, 125, 3772−3813

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Electric fields generated by protein scaffolds are crucial in enzymatic catalysis. This review surveys theoretical approaches for detecting, analyzing, and comparing electric fields, electrostatic potentials, and their effects on the charge density within enzyme active sites. Pioneering methods like the empirical valence bond approach rely on evaluating ionic and covalent resonance forms influenced by the field. Strategies employing polarizable force fields also facilitate field detection. The vibrational Stark effect connects computational simulations to experimental Stark spectroscopy, enabling direct comparisons. We highlight how protein dynamics induce fluctuations in local fields, influencing enzyme activity. Recent techniques assess electric fields throughout the active site volume rather than only at specific bonds, and machine learning helps relate these global fields to reactivity. Quantum theory of atoms in molecules captures the entire electron density landscape, providing a chemically intuitive perspective on field-driven catalysis. Overall, these methodologies show protein-generated fields are highly dynamic and heterogeneous, and understanding both aspects is critical for elucidating enzyme mechanisms. This holistic view empowers rational enzyme engineering by tuning electric fields, promising new avenues in drug design, biocatalysis, and industrial applications. Future directions include incorporating electric fields as explicit design targets to enhance catalytic performance and biochemical functionalities.



CONTENTS

Received: June 20, 2024
Revised: February 5, 2025
Accepted: February 10, 2025
Published: February 24, 2025

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

1. INTRODUCTION

The journey toward artificial enzyme design, while fraught with challenges, unfolds through well-defined steps. Foremost among these is the unraveling of natural enzymes' operational secrets, including the pivotal role of extended structure surrounding the enzyme's active site. Beyond merely insulating the active site from the cellular environment, the size, structural intricacy, and diversity of these regions hint at a more profound purpose. One hypothesized function is the generation of an electric field that modulates the reaction environment and guides the charge redistribution critical to catalysis. Thus, understanding how electric fields influence enzyme structure and function is critical for advancing the field of enzyme design.[1−4]

This understanding is particularly crucial in the context of *de novo* enzyme design, a cutting-edge approach that starts without a template, making it a more challenging but potentially revolutionary approach. While natural enzymes have evolved complex structural regions that contribute to their catalytic efficiency, *de novo* design aims to engineer these principles from scratch. This involves the strategic arrangement of catalytically active residues and the precise tuning of charge distributions to achieve optimal electrostatic preorganization—an essential factor in determining the rate and selectivity of enzymatic reactions. Advances in computational techniques, including aspects of both quantum and molecular mechanics worlds, are enabling more accurate predictions of electric field effects, driving progress in *de novo* enzyme design.

Yet, a deeper comprehension of how electric fields influence enzyme structures and properties remains a vital precursor to fully realizing the potential of *de novo* enzyme design. This review explores the current understanding of these electric field effects, focusing predominantly on computational research methods, and anticipates some of the developing investigative approaches. This review is structured as follows: Initially, we review the fundamental aspects of static electric fields, focusing on their emergence and calculation within molecular systems. We then explore both the theoretical predictions and experimental evidence of electric fields' critical influence on enzyme catalysis. Following this, the review discusses the seminal role of Empirical Valence Bond Theory in deepening our understanding of enzyme electric fields. Attention then shifts to cutting-edge advancements, including enhancements in interatomic force fields (FFs), which promise more precise enzyme structure predictions and, consequently, more accurate insights into the electric fields they generate. Lastly, we explore innovative methods designed to elucidate the interplay between electric field effects and their overarching structure within enzymes, employing Quantum Theory of Atoms in Molecules (QTAIM) and techniques traditionally applied in 3D field analysis, such as fluid dynamics, yet novel in the context of enzymatic electric field studies. This structure aims to provide a comprehensive overview of the field, setting the stage for future research in *de novo* enzyme design.

1.1. Electric Fields: An Overview

1.1.1. Electric Fields Due to Point Charges. The electric field, $\boldsymbol{E}$, is the negative of the gradient of the electrostatic potential, $V(\boldsymbol{r})$, i.e. $\boldsymbol{E}(\boldsymbol{r}) = -\nabla V(\boldsymbol{r})$; it is the force felt by a charged test particle as it moves through a varying $V(\boldsymbol{r})$. The force, $\boldsymbol{F}$, felt by a point charge q_1 at position $\boldsymbol{r}_1$ due to another "source" point charge q_2 at position $\boldsymbol{r}_2$ is given by Coulomb's law:[5]

$$\boldsymbol{F} = \frac{1}{4\pi\varepsilon_0}\frac{q_2 q_1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|^2}\hat{\boldsymbol{r}} = k\frac{q_2 q_1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|^2}\hat{\boldsymbol{r}} \tag{1}$$

where ε_0 is the permittivity of the medium, k is Coulomb's constant (equal to $\frac{1}{4\pi\varepsilon_0} = 8.99 \times 10^9 N{\cdot}m^2{\cdot}C^{-2}$), and $\hat{\boldsymbol{r}}$ is the unit vector $\frac{\boldsymbol{r}_2 - \boldsymbol{r}_1}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|}$, pointing from $\boldsymbol{r}_1$ to $\boldsymbol{r}_2$. The direction of the force is determined by the signs of the two point charges, attractive for opposite signs and repulsive for same signs.

The electric field due solely to the source point charge, q_2 at the position $\boldsymbol{r}_2$, felt at the point $\boldsymbol{r}_1$, is given by dividing $\boldsymbol{F}$ by q_1 to obtain,

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{\boldsymbol{F}}{q_1} = k\frac{q_2}{|\boldsymbol{r}_2 - \boldsymbol{r}_1|^2}\hat{\boldsymbol{r}} \tag{2}$$

In the presence of several source point charges, say q_1, q_2, ..., q_n, at respective positions $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, ..., $\boldsymbol{r}_n$, the total electric field at an arbitrary point $\boldsymbol{r}$ is the vector sum of the fields due to each charge:

$$\boldsymbol{E}(\boldsymbol{r}) = k\sum_{i=1}^{N}\frac{q_i}{|\boldsymbol{r}_i - \boldsymbol{r}|^2}\hat{\boldsymbol{r}}_i \tag{3}$$

where $\hat{\boldsymbol{r}}_i$ now points from $\boldsymbol{r}$ to $\boldsymbol{r}_i$.

1.1.2. Electric Fields in Continuous Charge Densities. The electric field due to a continuous charge density (i.e., a conceptually infinite number of point charges), as for chemical systems, can be calculated in a similar fashion, in terms of the electron charge density $\rho(\boldsymbol{r})$, nuclear charges Z, and any externally applied fields $\vec{E}_{ext}(\boldsymbol{r})$:

$$\boldsymbol{E}(\boldsymbol{r}) = -\int\frac{\rho(\boldsymbol{r}')(\boldsymbol{r}' - \boldsymbol{r})}{|\boldsymbol{r}' - \boldsymbol{r}|^3}d^3\boldsymbol{r}' + \sum_{i=1}^{N}\frac{Z_i(\boldsymbol{R}_i - \boldsymbol{r})}{|\boldsymbol{R}_i - \boldsymbol{r}|^3} + \boldsymbol{E}_{ext}(\boldsymbol{r}) \tag{4}$$

Note that eqs 1−4 are derived under idealized vacuum conditions and do not explicitly incorporate electronic polarization. In standard nonpolarizable molecular dynamics (MD) and many quantum mechanical (QM) approaches, the protein and solvent environment are often simplified to fixed charges and implicit dielectric backgrounds. Such an approximation can introduce errors, as the electronic redistribution in response to the local field is neglected. Introducing a modest background dielectric constant (e.g., $\varepsilon \approx 2$) has been suggested as a strategy to partially compensate for these limitations, reducing the need for fully polarizable force fields.[6]

In larger chemical systems, such as proteins, calculating the full $\rho(\boldsymbol{r})$ can be computationally expensive. Here, the charges of system subdomains, such as amino acids composing the protein scaffolding, can be used with eq 3 to approximate the electric field over large regions.

1.1.3. Polarization and Dipoles. For a neutral atom in an electric field its positively charged nucleus and the negatively charged electron cloud experience forces in opposing directions. In a sufficiently strong field, the atom, will undergo field-induced ionization, where an electron tunnels through the coulomb barrier and leaves behind a positive ion. However, for less intense fields an equilibrium is reached where the force of

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

the field $\boldsymbol{E}$ pulling the electrons and nucleus apart is balanced by the force of the nuclear-electron attraction, and an induced dipole is formed with a dipole moment $\boldsymbol{p}$ that points in the same direction as $\boldsymbol{E}$ and that is proportional to the field strength according to the atom's polarizability, α:

$$\boldsymbol{p} = \alpha \boldsymbol{E} \tag{5}$$

Placing neutral systems, such as nonpolar molecules, within an electric field can lead to more pronounced polarization compared to individual atoms due to the increased mobility of charges between atoms in chemical bonds. The extent of a molecule's polarizability is also influenced by its orientation relative to the electric field. For instance, CO_2 exhibits a polarizability approximately 2.5 times greater when the field aligns with the molecular axis than when the field is perpendicular.[7,8]

For linear molecules this difference gives rise to two distinct polarizability values, $\alpha_{\parallel}$ and $\alpha_{\perp}$. Consequently, when the field aligns at an angle intermediate to these orientations, the molecule's polarizability is derived from both parallel and perpendicular components, resulting in an induced dipole that may not align with the field direction:

$$\boldsymbol{p} = \alpha_{\parallel}\boldsymbol{E}_{\parallel} + \alpha_{\perp}\boldsymbol{E}_{\perp} \tag{6}$$

For completely asymmetric molecules, a general form for the x, y, and z components of $\boldsymbol{p}$ is calculated using a set of nine constants α_{ij} that define its polarizability tensor:

$$\begin{aligned} p_x &= \alpha_{xx}E_x + \alpha_{xy}E_y + \alpha_{xz}E_z \\ p_y &= \alpha_{yx}E_x + \alpha_{yy}E_y + \alpha_{yz}E_z \\ p_z &= \alpha_{zx}E_x + \alpha_{zy}E_y + \alpha_{zz}E_z \end{aligned} \tag{7}$$

Numerous molecules have permanent dipoles, such as polar molecules, with water serving as a prime example, where the dipole moment extends from the oxygen atom to the midpoint between the two hydrogen atoms. When subjected to an electric field, such a molecule will align its dipole moment with the field.

In enzyme electric field modeling, it is common to quantum mechanically treat an active site model, with the enzyme's extended structure contributing to an electric field that affects the energy of the quantum mechanical region. Usually, the enzyme's polar residues are modeled as point charges, and the active site is treated in an electrostatic embedding scheme.

It is also crucial to note that accurately determining protonation states of ionizable residues and internal groups is essential. Incorrectly assigned protonation states can significantly distort computed fields and energetics. Recent simulations indicate that proteins may sample multiple protonation microstates at equilibrium,[9] underscoring the importance of carefully validating protonation states, potentially through pKa calculations or enhanced sampling methods.

eq 3 can be utilized to calculate the resultant electric field at specific points, if this quantity is of interest, leading to a heterogeneous field distribution across the active site. Occasionally, for specific systems, it can be fruitful to approximate the extended protein structure's influence as a uniform field over the active site, akin to situating the active site within an infinite parallel plate capacitor. However, usually this latter approach lacks chemical detail needed to understand the enzyme function, and as such can be utilized for purely theoretical investigations.

1.2. Electric Fields in Proteins

Warshel first introduced the concept of electrostatic catalysis as a key driver of enzymatic evolution, in the last century,[10−12] emphasizing the enzyme's role in arranging the substrate to align with beneficial electric fields for specific chemical transformations. Recent advancements have highlighted electric fields' critical role in enzymatic catalysis,[13−21] with these insights increasingly incorporated into *de novo* computational enzyme design.[22−30] Previously, the integration of such considerations was limited by the capabilities of quantum chemical calculations, which have now evolved to adequately model these complex interactions.[31] Current experimental methods also allow for the direct measurement of electric fields within enzyme active sites, showcasing their contribution to catalytic rate enhancement.[32−41]

Beyond enzymology, other chemical processes are labile to electrostatic interactions, for example, redox potentials and chemical equilibrium. Warshel and co-workers provide a slightly dated but informative review of electrostatic evaluation methods as well as case studies for how electrostatics influence catalysis, redox, and pKas in protein systems.[42] It is important to distinguish that Ems and pKas reflect equilibrium properties associated with well-defined reactant and product states, while reaction rates are governed by the stabilization of the transition state—often a more elusive entity. Both equilibrium and kinetic properties are influenced by the protein's electrostatic environment. While Ems and pKas are shifted by changes in the electrostatic potential around stable states, reaction rates mostly depend on how the local electric field stabilizes or destabilizes the transition state. Thus, understanding and accurately modeling electrostatic effects is paramount for elucidating and predicting both equilibrium and kinetic aspects of enzyme function. Experiments tying electrostatics to pKa and redox potentials date back decades. For pKas, these methods span a host of methods and include predictions on surface and interior residues.[43−50] Here, we will discuss a few interesting examples.

Adam et al. studied conformational changes in Channelrhodopsins (ChR) to understand proton transfer dynamics and comprehensively account for protein dynamics.[51] They leverage QM/MM and MM calculations to determine the electrostatic effects of K132, a sequence present in many such ChR sequences, is critical. This residue was found to greatly influence proton transfer dynamics via changes in charges on E162 and D292 residues. Another study weighted ensembles of unfolded staphylococcal nuclease structures via electrostatic energies.[49] They computed energies via both Coulomb and Poisson−Boltzmann (PB) formalisms across different pHs. Interestingly, they found that electrostatic contributions to pKas were weak but, at extreme pHs, still contribute significantly to pKas. Yet, another study, compared isoform differences in Monoamine Oxidases, including pKas of titrable residues via linear response approximation version (PDLD/S-LRA).[48] They determined that the alignment in pKa values between the different isoforms could be explained by similar electrostatic environments.

Naturally, many studies aim to explain pKas as the culminating value of a set of different factors.[50−52] For example, Lindman et al. use experimental NMR assays of PGB1 fragments, coupled to Monte Carlo/Gaussian Chain Models, to predict pKa values of the unfolded state of the protein.[52] Effectively, this Monte Carlo approach corrects for whole protein electrostatics to create a complete prediction of

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

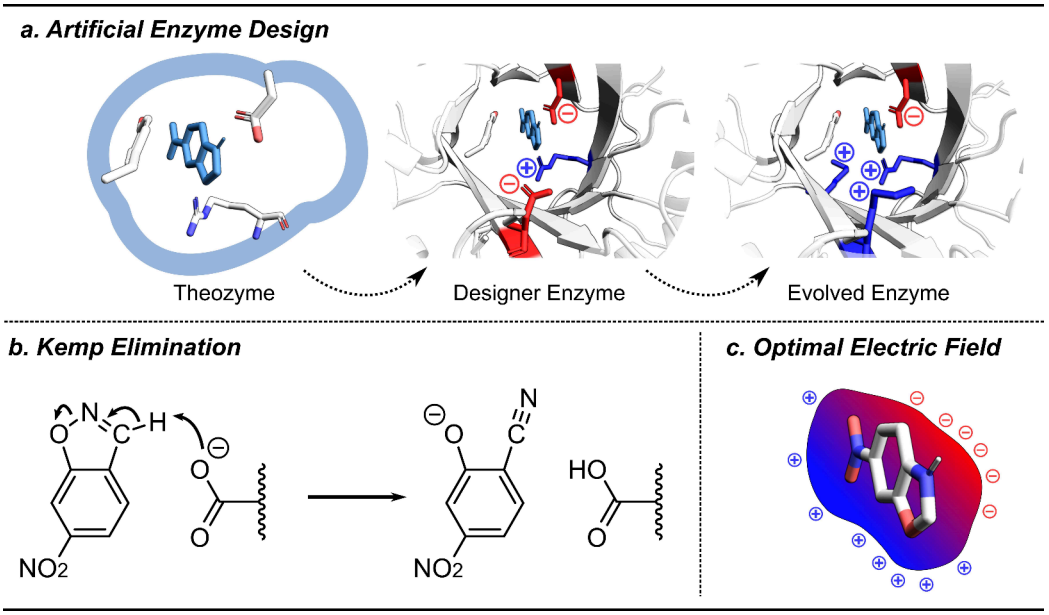


Figure 1. Theoretical search for a *de novo* enzyme of Kemp elimination reaction. (a) Theoretical approach toward an optimized enzyme: substrate positioning between the catalytic residues, incorporating the theozyme inside the supporting enzyme scaffold, and optimization of the internal enzyme electric field through mutagenesis. (b) Mechanism of the Kemp elimination reaction with carboxylate as a base initiating the C−H deprotonation. (c) Optimal electric field surrounding the substrate for the Kemp elimination rate enhancement.

residue pKas in the unfolded protein. They also integrate MD simulations to provide rich insight into the interplay between dynamics, electrostatics, and burial effects.

Recent advances in PB theory and methods have significantly improved the accuracy and utility of pKa predictions in enzymes and related proteins by integrating enhanced conformational sampling, refined dielectric models, and more rigorous validation protocols. Coskun recently reviewed computational methods for pKa calculations, where the conceptual simplicity of PB-based methods is offset by potential issues regarding uncertainty of the optimal dielectric constant, limitation to treating static protein structures, and accurate representation of protein response to ionization,[53] however these limitations are being addressed by recent advances in the field. For instance, toward overcoming dielectric constant uncertainty and treatment of dynamic protein structures, machine learning methods can be used to rapidly predict PB solutions for a range of dielectric constants, accelerating their optimization for specific systems, and allow rapid sampling of protein structures.[54,55]

Pivoting to redox potentials, studies using electrostatic methods such as continuum models to quantify effects on redox potentials date back over 4 decades.[56−59] One such study used a model for the disulfide active site in thioredoxin and DsbA coupled to Poisson−Boltzmann electrostatic calculations to justify experimentally observed redox potential differences between the two systems.[60] Another study, on Heme systems across different cytochrome environments, uses electrostatic calculations to break down energy contributions from solvent, protein-heme residue interactions, and structural changes between electrochemical states.[61] The review by Chen et al. on theoretical biomolecular methods for modeling redox potentials provides more information on methods and applications in this area.[62]

In 2010, Kuznetsov et al. found substantial differences in the redox potentials of Rieske iron−sulfur proteins, as calculated by semicontinuum methods, with and without the larger protein scaffold. Notably, only calculations with the protein electrostatic scaffold agreed well with experiment.[63] More recently, Kanda and Ishikita calculated redox potentials of two Fe_4S_4 clusters (F_A and F_B) in photosystem I (PSI) and green sulfur bacteria reaction centers by solving the linear Poisson−Boltzmann equation using protein structure data.[64] Their calculations revealed that electron transfer from F_A to F_B is energetically downhill in cyanobacterial PSI but isoenergetic in plant PSI, with this difference attributed to varying electrostatic influences from conserved residues like PsaC-Lys51 and PsaC-Arg52. A comprehensive study by Gamiz-Hernandez et al. analyzes rubredoxin redox potentials via continuum models.[65] They analyze different structural parameters including hydrogen bonding, side chain conformations, dielectric environment and charges residues and determine that electrostatic contributions toward redox potentials are notable.

Finally, Gaughan et al. systematically probed the redox effects of oriented electric fields by computing vertical excitation energies (VSE).[66] In this study, they selected a set of biologically relevant iron−sulfur clusters and placed point charges around the clusters. They rotated these point charges along a sphere of rotation and demonstrated that external electric fields of around 28.8 MV/cm could change relative stability by as much as 35 kJ/mol depending on their exact orientation. For more relevant studies involving computational prediction of redox potentials in a biological context, interested readers can see refs67−71.

In the realm of *de novo* enzyme design, electrostatic effects have become a pivotal element, shifting the focus from seeking an ideal enzyme structure to identifying an optimal charge distribution or electric field that favors the transition state (TS) over the reactant state.[72,73] This approach involves arranging catalytically active residues, known as theozymes,[74] in a manner that is geometrically integrated into the enzyme's scaffold, as illustrated in Figure 1.[23,75] Through such design, the enzyme's electrostatics are tailored to boost catalysis by lowering the free energy barrier. Building on this concept,

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

Sokalski et al. and Hartke et al. advanced an inverse design strategy, creating abstract optimal catalytic fields (OCFs)[73,76] and globally optimal catalysts (GOCATs),[77] marking a novel starting point in the pursuit of ideal biocatalysts.

Computational enzyme design, leveraging charge complementarity and optimization of internal electric fields, has successfully produced enzymes for novel and non-natural reactions.[78−81] A prime example is Kemp elimination, a benchmark for proton transfer without a corresponding natural enzyme, illustrated in Figure 1.[82,83] This reaction, characterized by substrate deprotonation and the opening of its heterocyclic ring, leads to significant substrate charge accumulation. Employing TS stabilization and Rosetta toolbox for protein backbone design, Röthlisberger et al. developed eight Kemp eliminases, notably KE07, KE59, and KE70, showcasing activity.[80] Note that the focus of the theoretical design at the time was solely on the immediate environment of the substrate, i.e. without considerations of the electric fields from the remote parts of the protein.

This methodology was further applied to create enzymes catalyzing complex reactions, such as multistep retro-aldol reactions with five different protein scaffolds,[79] and two enzymes with stereoselective Diels−Alder reactivity.[84] Modern computational tools play a crucial role in mirroring biochemical experiments, allowing for accurate predictions of catalytic residue positions and internal electric fields within active sites, aligning with X-ray structural data.[85−89] The direct correlation between internal electric fields and enzymatic activity aids in identifying mutagenic targets for scaffold improvement, enabling systematic enhancement of electrostatic catalysis by analyzing the contributions of individual atoms.[72,73,76,87,90,91]

Despite significant advances in enzyme modeling protocols, limitations persist in accurately capturing the complexities of enzymatic catalysis.[28,92,93] These limitations include the omission of factors like configurational entropy changes, second-shell or long-range interactions, and dynamical effects, which are crucial for efficient catalysis and turnover.[29,94−101] Computational constraints often lead to simplifications, such as neglecting enzyme flexibility and using discrete ligand placements, thereby missing the subangstrom precision vital for aligning theoretical predictions with experimental results.[102] Consequently, modeled enzymes frequently show optimistic catalytic predictions, with actual activities significantly lower than those observed in nature. For instance, of 59 computationally designed enzymes, only 8 showed activity in Kemp eliminase experiments as per ref 80, and only 2 out of 84 designs demonstrated activity for Diels−Alder reactions in ref 84. Rational enhancements have focused on the active site and chemically active residues, overlooking the broader protein scaffold's potential influences on function, dynamics, and stability. Additionally, entropy's role in stabilizing transition states and destabilizing reactant states highlights the importance of dynamics in catalytic efficiency.[87,96,97,103−106]

With computational methods struggling to account for dynamical aspects, there's a tendency to overdesign active sites. This challenges further optimization efforts, suggesting experimental directed evolution, as recognized by F. Arnold's 2018 Nobel Prize,[107] as a potential strategy. Here, *de novo* designed enzymes emerge as valuable starting points for *in vitro* evolution.[78,108]

Evolved variants have demonstrated significant improvements in catalytic rates, nearing the efficiencies of natural enzymes, although not yet reaching the diffusion limit.[80,81,102,109−111] *In vitro* evolution applied to the computationally designed Kemp eliminase KE07 led to an over 200-fold increase in k_{cat}/k_M across seven rounds of random mutagenesis.[80,109] These enhancements often involved mutations at the active site's base, introducing polar or charged residues to better position Lys222 (acting as an acid in the reaction) and mitigate its adverse interaction with Glu101 (acting as the base—see Figure 1). This adjustment is supported by the shifted pKa values of Glu101 in evolved variants. However, the exact mechanisms—whether improvements in electrostatic stabilization of the TS, destabilization of the reactant state, or reduction in overall reorganization energy—remain debated.[112−116] Directed evolution of KE70 yielded a greater than 400-fold increase in catalytic efficiency, attributed to enhanced substrate binding, electrostatic refinement, and active site stabilization in conformations favorable for catalysis.[110] Yet, in KE70, electric field optimization was less emphasized, with mutations enhancing substrate affinity through adjustments in its hydrophobic cavity.[115] This underscores a limitation of laboratory directed evolution: electric field optimization tends to be local, constrained by the initial non-naturally optimized enzyme scaffold design. Thus, the concept of electrostatic preorganization may be overlooked, leading to an electrostatic mismatch between the active site and scaffold, inadequately supporting the TS charge distribution. Consequently, the potential for optimizing *de novo* enzymes via *in vitro* evolution is estimated to be limited to roughly 3 orders of magnitude.[115] Future enzyme modeling could benefit from initial computational scaffold optimization, focusing on precise internal electric field adjustments.[115]

Achievements in enzyme optimization have been realized by linking laboratory directed evolution with computational approaches, notably through parallel or iterative combinations of molecular dynamics (MD) simulations and structural analysis of initial, enhanced, and inactive variants.[117−119] This methodology involves integrating computationally identified mutations into experimental scaffolds for empirical evaluation, while discarding mutations found to be ineffective or detrimental. For instance, the evolved Kemp eliminase KE70 variants were crafted using this strategy, with approximately half of the mutations originating from computer optimizations and the other half from random mutagenesis.[110] Another example involves the iterative refinement of the inactive Kemp eliminase scaffold HG-1, employing X-ray crystallography and MD simulations.[81] The initial analysis revealed HG-1's active site was overly solvent-exposed with excessively flexible catalytic residues. These issues were systematically addressed through iterative computational and experimental evolution, culminating in the HG-3 scaffold, which achieved a k_{cat}/k_M of 430 M^{-1} s^{-1}. Subsequent enhancements through 17 rounds of random mutagenesis produced the HG-3.17 variant, boasting a k_{cat}/k_M of 230,000 M^{-1} s^{-1}, nearing the efficiency of natural enzymes. The efficiency gains were primarily linked to the stabilization of the TS's developing negative charge.[102]

De novo enzyme modeling has seen significant advances through laboratory directed evolution, highlighting the need for enhancing theoretical models to reduce dependence on extensive experimental work. The optimization of electric field stabilization for the TS of Kemp eliminase KE15 represents a milestone, achieved entirely through computational methods by Head-Gordon et al.[26] By introducing just four targeted

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

computational mutations, they developed an enzyme variant that demonstrated a 43-fold increase in catalytic efficiency, with an experimental k_{cat}/k_M reaching 403 M^{-1} s^{-1}.

The field of artificial enzyme design stands to gain from enhanced modeling of long-range electrostatic networks among charged and polar residues, alongside the use of MD simulations to evaluate residue couplings and fluctuations. Traditional FFs with additive potentials often fall short in capturing these complex interactions. Although more computationally intensive, the AMOEBA polarizable FF aligns more closely with electric fields obtained from multiscale density functional theory calculations.[89]

Furthermore, explicitly acknowledging the role of electrostatics in guiding force field development can strengthen this connection. Improved force fields—ranging from polarizable models to machine learning-based potentials—are increasingly designed to capture subtle electrostatic features, such as polarization and environment-dependent dielectric responses. By more faithfully reproducing the underlying electrostatic environment, these advanced force fields promise better predictions of both equilibrium properties (e.g., Ems and pKas) and kinetic parameters (e.g., transition state stabilization). This holistic electrostatic modeling is crucial for achieving more accurate enzyme design and understanding how electrostatic preorganization drives catalysis.

However, quantum mechanical descriptions offer even higher precision by accounting for charge penetration and transfer effects—elements not fully captured by polarizable FFs.[100,120,121] Employing an *ab initio* MD approach, Kulik et al. showed the extent of charge distribution variability, noting substantial charge accumulation or depletion around charged or polar residues.[100] Such dynamics challenge the capabilities of fixed-charge additive FFs, underscoring the discovery of strong correlations among residues, including significant long-range couplings beyond adjacent residues.

Our group has employed QTAIM to devise a sensitive yet straightforward quantum mechanical (QM) probe for assessing electrostatic preorganization.[120,122,123] We have demonstrated that both the electrostatic potential and density at bond- and ring-critical points linearly correlate with the applied electric field.[123] Given their correlation with reaction barriers, these metrics serve as robust predictors of reactivity. It is crucial to acknowledge that actual protein structures create complex, nonuniform electric fields, significantly influencing reaction pathways.[124] The intricate link between changes in topological electric fields and reactivity presents a considerable challenge, underscoring the necessity for a holistic view of electric fields. To address this, we have introduced a methodology to associate electric field topologies with reactivity, utilizing a global distribution of field streamlines to quantify reactivity in chemically akin systems.[125] We discuss these developments in detail in subsequent sections.

The modulation of local electric fields (LEFs) over long ranges is essential for protein functionality, potentially acting as a mechanism to initiate or control chemical reactions.[126−130] Such modulations can alter redox potentials and pKa values, induce conformational changes, facilitate charge transfer reactions, align substrates within active sites, or lower barriers in the rate-limiting steps of enzyme catalysis. A notable challenge confronting the simulation of these varying electric fields arises from how the favorability of proton or charge transfer, partial atomic charges, electrostatic density, and electric fields depend on the chosen QM region's size and composition, whether employing cluster models or quantum mechanics/molecular mechanics (QM/MM) strategies.[122]

Theoretical model enhancements can be achieved by moving beyond rigid backbones to explore mutations that trigger new conformations, improve TS entropic stabilization, and fine-tune protein dynamics. Future efforts aim to refine the HG3.17 Kemp eliminase enzyme by embracing backbone flexibility, a strategy that previously facilitated the computational optimization of Kemp eliminase KE70,[110] enhancing TS stabilization. Assessing dynamic effects necessitates the examination of extensive ensembles of reactant and TS structures through detailed and accurate MD simulations. Current MD simulations typically assume static protonation states, overlooking how local charges and electric fields may alter protonation states and affect protein properties, including folding. Optimizing proteins in future studies could also involve minimizing field fluctuations and investigating the interplay between protein stability and functionality.[29] Additionally, designing enzymes with metal cofactors introduces further complexity, necessitating a quantum mechanical approach to metal coordination, comprehensive protein backbone sampling, and active site polarization considerations.[120]

1.3. Stark Spectroscopy

Measuring electric fields at the molecular level presents significant challenges due to the complexity and dynamic nature of biological systems. Traditional methods such as dielectric spectroscopy and electrochemical techniques often lack the spatial resolution required to probe localized fields within proteins. Stark spectroscopy, a technique derived from the Stark effect, is a powerful analytical method to probe electric fields in targeted regions of proteins and enzymes[131] The Stark effect is observed when the energy levels of atoms or molecules are perturbed by an external electric field, leading to shifts in their spectral lines. This effect, discovered simultaneously by Johannes Stark and Anthony Lo Surdo in 1913,[132,133] provides critical insights into the electronic structure of atoms and molecules, as well as the interaction between electromagnetic radiation and matter.

Stark spectroscopy exploits this principle to measure the electric field's influence on molecular vibrations, offering a window into the electrostatic environment within complex systems such as proteins and enzymes. By observing how the frequencies of specific vibrational modes shift in response to changes in electric fields, the magnitude and direction of the fields can be deduced at precise locations within a molecule. This capability is particularly useful for understanding how proteins harness electric fields for their function, as these fields can significantly influence biochemical reactions and molecular interactions. The application of vibrational stark effect (VSE) spectroscopy, which focuses on vibrational transitions, has been especially fruitful in dissecting the roles of electric fields in biological systems, providing a quantitative measure of electrostatic effects that are otherwise challenging to study.[134] To quantify the electric fields using the VSE, a calibration step is first performed in known environments. This involves measuring the vibrational frequency shift ($\bar{\nu}_{obs}$ generally measured in units of cm^{-1}) of a probe molecule in various solvents with known electric field strengths ($|\boldsymbol{F}_{env}|$, in MV/cm). The relationship between the frequency shift and the electric field is characterized by the Stark tuning rate ($|\Delta\boldsymbol{\mu}_{probe}|$, in

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

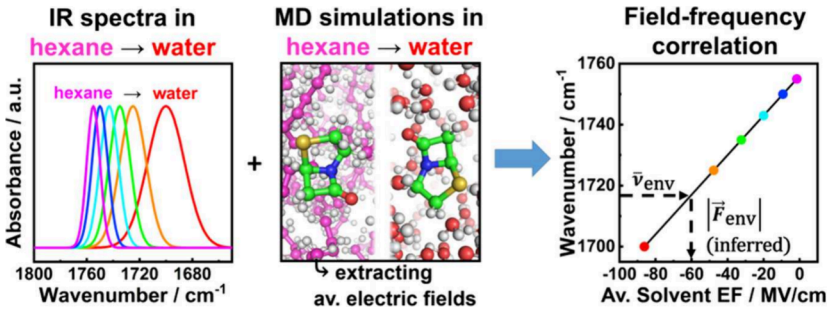


Figure 2. Illustration of the solvatochromic method for deriving electric field-frequency calibrations using VSE. Reproduced with permission from Kozuch et al., J Phys Chem B, 2021, 125 (17), 4415−4427. Copyright 2021 American Chemical Society.[140]

$cm^{-1}/(MV/cm)$, which is determined through this calibration. The fundamental equation governing this relationship is

$$\bar{\nu}_{obs} = \bar{\nu}_{probe} - |\Delta\boldsymbol{\mu}_{probe}|\cdot|\boldsymbol{F}_{env}| \quad (8)$$

where $\bar{\nu}_{probe}$ represents the vibrational frequency in the absence of any electric field, such as in a vacuum. This calibration ensures that when VSE is applied to an unknown environment, such as a protein active site, the measured frequency shift can be accurately translated into the LEF using the previously determined Stark tuning rate.

The experimental applications of Stark spectroscopy in determining the electric fields of proteins and enzymes have shed light on the complex ways in which molecular machinery of life exploits electrostatic forces.[135] For instance, studies using VSE have revealed the presence of strong electric fields within the active sites of enzymes such as ketosteroid isomerase (KSI).[39,136] Calculations based on crystallographic structures indicate that the electrostatic environment in KSI's active site is preorganized to favor the TS geometry. Specifically, electric fields in the range of −127.6 to −141.7 MV/cm have been measured, which drives the substrate-like ligand to adopt a TS-like geometry upon binding. This oriented electric field minimizes the need for dipole reorientation during the reactive event, thereby stabilizing the transition state and lowering the activation energy by approximately 7 kcal/mol. Similarly, research on dihydrofolate reductases has shown how ligand-electrostatic interactions contribute to enzyme catalysis, offering insights into the fundamental principles of molecular recognition and catalytic efficiency.[137] VSE has also been pivotal in understanding the inhibition mechanisms of β-lactamases by avibactam.[40,41] These studies have demonstrated that avibactam binding significantly alters the LEFs within the enzyme's active site, transitioning from high-field environments that stabilize charge-separated transition states to lower-field conditions that reduce the hydrolysis rate of the covalent bond between the inhibitor and the enzyme. Specifically, VSE measurements revealed that while ancestral penicillin-binding proteins exhibit smaller electric fields (−59 MV/cm) that render the ester linkage resistant to hydrolysis, descendant TEM β-lactamases experience larger electric fields (−140 MV/cm) that facilitate bond hydrolysis by stabilizing the transition state. Avibactam counteracts this evolutionary trend by inducing a low-field environment, thereby preventing the rapid hydrolysis of its covalent linkage and effectively inhibiting the enzyme's activity. These applications not only underscore the utility of VSE in deciphering the electrostatic underpinnings of enzyme activity but also in drug design, where manipulating electric fields could lead to new therapeutic strategies. Finally, VSE has been crucial in uncovering transitions of functionally relevant protein conformations and detailing local interactions,[37,138] as well as in measuring changes in enzyme electric fields due to mutations,[33] alterations in ligand coordination,[87] and conformational shifts during the reaction cycle.[139] Through these diverse applications, VSE continues to be a cornerstone technique for exploring the fundamental role of electric fields in biochemical processes.

However, accurately interpreting VSE data is complicated by factors such as solvent effects, probe orientation, and the dynamic fluctuations of the molecular environment, necessitating rigorous calibration and computational support. MD simulations play a vital role in validating key assumptions behind VSE measurements and extending the technique's applicability to complex systems like proteins. By simulating the solvation environment around a VSE probe molecule like a carbonyl (C=O) or nitrile (C≡N) group, MD can provide a detailed picture of the LEFs experienced by the probe within different solvents or biological matrices. Accurately modeling electric fields for VSE calculations using molecular dynamics typically requires polarizable FFs,[140] which enable the dynamic adjustment of atomic charges in response to the local electrostatic environment and accurately capture the fluctuating fields around the probe. These FFs provide a more realistic representation of molecular interactions compared to nonpolarizable counterparts, thereby enhancing the reliability of electric field calculations derived from MD simulations. However, at least one study has demonstrated that when combined with QM/MM methods, nonpolarizable FFs such as Amber ff99SB can also produce results that align well with experimental observations.[88] These simulations account both for the intricate interactions and dynamic fluctuations of the surrounding molecules, enabling more accurate interpretations of VSE spectral data. A common experimental approach called the solvatochromic method is used to calibrate the relationship between vibrational frequencies and electric field strengths for a given VSE probe (Figure 2).[139] This involves measuring the IR spectra of the probe dissolved in a range of solvents, from nonpolar hexanes to highly polar water. The vibrational frequencies obtained from these measurements are then correlated with electric field strengths calculated from MD simulations of the probe in each solvent environment. A field-frequency correlation plot is generated by plotting the observed vibrational frequency shifts ($\bar{\nu}_{obs}$) against the known electric field strengths ($|\boldsymbol{F}_{env}|$) from the calibration step. As described in eq 8, this plot ideally exhibits a linear relationship, as predicted by VSE theory, where the slope of the line corresponds to the Stark tuning rate ($|\Delta\boldsymbol{\mu}_{probe}|$). Confirming

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

this linearity validates a critical assumption—that the observed frequency shifts predominantly arise from the Stark effect rather than changes in covalent bond strengths.

Additionally, MD can model local field effects[141,142] where the field experienced by the probe differs from the bulk electric field. For more complex protein systems where direct electric field measurements are challenging, MD simulations of the protein environment enable applying VSE spectroscopy by computationally modeling how factors like mutations or ligand binding perturb the LEFs around an embedded VSE probe group.[134,143] Finally, MD also allows prescreening of new VSE probe molecules by simulating their behavior in well-characterized environments before applying them to more complex biological systems.[144] This ensures frequency shifts can be reliably mapped back to electric field changes in the system of interest. Together, VSE spectroscopy with MD simulations provides a powerful approach for precisely measuring and interpreting electric fields in biological environments ranging from solvents to enzymes to proteins.

A more accurate method to model electrostatic interactions, obtain localized electric field strengths, and even predict experimental shifts in frequencies is through hybrid quantum mechanics/molecular mechanics (QM/MM) calculations. The pivotal study by Wang and He on KSI exemplifies how QM/MM calculations can unveil the profound influence of electrostatic fields on enzymatic catalysis.[88] Their work demonstrates that the QM-derived electric fields based on snapshots from QM/MM molecular dynamics simulations provided better quantitative agreement with experimental observations by using the nonpolarizable FF Amber ff99SB as a comparison. The study also established a direct link between the electric fields within the KSI active site and the enzyme's catalytic activity, highlighting the potential for rationally designing enzymes with enhanced efficiency by strategically manipulating these electrostatic fields. Further showcasing the precision of QM/MM methods, Layfield and Hammes-Schiffer's study demonstrated that this methodology can accurately reproduce the experimentally measured shifts in vibrational frequencies upon binding of the intermediate analogue equilinen to KSI, for two different nitrile probe positions within the active site while also dissecting the intricate influence of specific residues on the electrostatic environment experienced by the probes.[85] Additionally, the comprehensive QM/MM approach developed by Sandberg, Rudnitskaya, and Gascón for predicting Stark shifts in proteins achieves quantitative agreement with experimental findings.[145] Noteworthy for integrating molecular dynamics with moving-domain QM/MM techniques, their method accurately predicts absorption frequency changes caused by electrostatic perturbations in the protein environment. By modeling the interplay between quantum mechanical and molecular mechanical interactions, their approach precisely captures localized electric fields and their effects on vibrational probes. Furthermore, it addresses challenges in predicting electrostatic field changes from protein site mutations and facilitates determining the protonation states of nearby ionizable residues. Finally, Ringer and MacKerell's QM/MM analysis showed that QM/MM can be used to validate traditional molecular mechanics FFs by accurately replicating the electrostatic environments around VSE probes within macromolecular systems.[146]

VSE spectroscopy, in conjunction with computational methodologies, continues to serve as a cornerstone for understanding the critical influence of electric fields in the realm of biochemistry. This fusion of experimental and computational techniques opens a window into the sophisticated manipulation of electric fields across biological systems, ranging from stabilizing catalytic transition states to covalent inhibition to the nuanced control of ligand binding and specificity. Such comprehensive insights not only validate the foundational principles derived from Stark spectroscopy but also lay the groundwork for the deliberate engineering of biomolecules with tailored properties.

1.4. Empirical Valence Bond Theory

Empirical Valence Bond (EVB) theory, crafted by Warshel and Weiss,[147] offers a framework for comparing the energy dynamics of chemical reactions inside enzymes with those in solutions. At the heart of the theory is the idea that environmental effects are mainly electrostatic, affecting the energy of ionic states while keeping the covalent state consistent across different settings. The EVB method offers several advantages over other QM/MM approaches, particularly in its ability to provide conceptual clarity and intuitive understanding of enzymatic reactions.

Indeed, Warshel and Weiss noted that chemical intuition is often sufficient to simplify the reaction coordinate to a few key valence bond states, highlighting the critical electrostatic interactions that contribute to catalysis and allowing for direct insights into electrostatic preorganization within enzymes. This simplification enables the decomposition of activation energies into distinct contributions—such as electrostatics and solvation effects—enhancing interpretive power when analyzing results. Furthermore, EVB is computationally efficient, facilitating extensive sampling and the simulation of larger systems over longer time scales, which is essential for capturing dynamic enzyme behavior. Its flexibility and transferability, owing to empirical parametrization, make it adaptable to various reactions and systems. These advantages, especially the interpretive benefits, make EVB an invaluable tool for exploring electric fields in enzymes. A recent robust implementation of EVB within GROMACS makes the method widely accessible.[148]

The EVB theory utilizes a secular eq (eq 9) to calculate the energies of resonance forms based on their spatial configurations and off-diagonal Hamiltonian matrix elements:

$$\Delta F^{\ddagger} \approx \frac{(\Delta F^{0} + \lambda)^{2}}{4\lambda} - \langle H_{12}\rangle_{TS} + \frac{\langle H_{12}\rangle_{RS}^{2}}{\Delta F^{0} + \lambda} \tag{9}$$

Here, $\Delta F^{\ddagger}$ is the activation free energy; ΔF^{0} represents the standard free energy change of the reaction; and λ is the solvent reorganization energy, which mainly reflects the changes in the solvent−solvent interaction during the reaction. The term $\langle H_{12}\rangle_{TS}$ refers to the average electronic coupling (off-diagonal Hamiltonian matrix element) between the reactant and product states at the transition state, while $\langle H_{12}\rangle_{RS}$ represents the average electronic coupling between these states at the reactant state (RS).

A pivotal aspect of the EVB method is calculating the free energy changes associated with solvation and electrostatic interactions, both in solution and within the enzyme's active site. Notably, the first term on the right side of eq 9 corresponds to the classical Marcus expression for the activation free energy of electron-transfer reactions.[149] This term represents the reorganization free energy, accounting for the energy required to reorganize the solvent and enzyme environment during the reaction without considering elec-

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

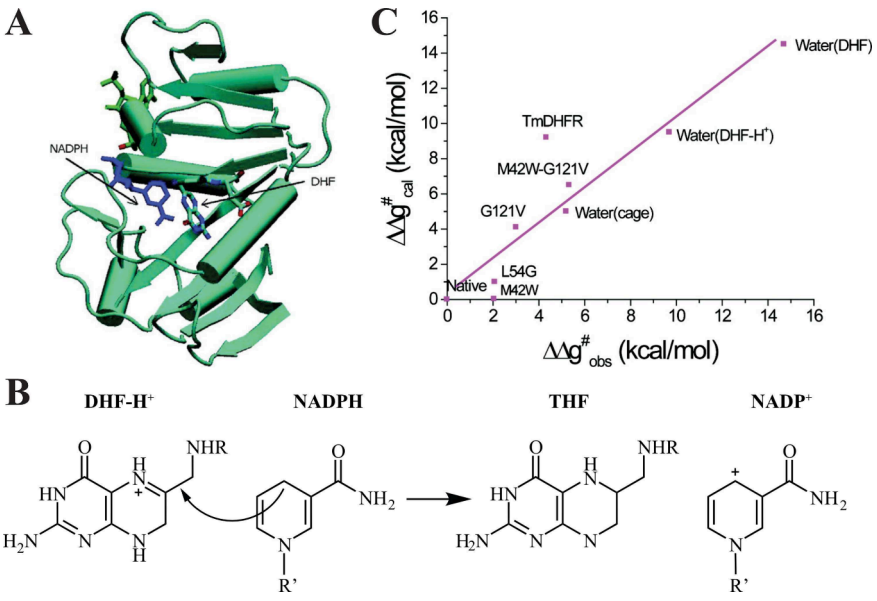


Figure 3. (A) The dihydrofolate reductase protein. (B) The reacting system in the catalytic reaction of DHFR, where R indicates the benzoyl glutamic acid moiety of DHF-H+, and R' indicates the remaining adenine, phosphate, and ribose groups of NADPH. (C) Correlation between the calculated and experimentally observed changes in activation energies for dihydrofolate reductase (DHFR) catalysis in the native enzyme (labeled "Native" in the plot), the designated mutants (M42W, L54G, G121 V, and M42W-G121 V), the thermophilic TmDHFR (labeled "TmDHFR"), and the reference solution reaction in water. The water-based references—Water (cage), Water (DHF-H+), and Water (DHF)—indicate (i) the reaction carried out in a solvent cage, and (ii) dihydrofolate in its protonated or neutral forms, respectively.

tronic coupling between the reactant and product states. However, the EVB method extends beyond Marcus theory by including the additional two terms on the right. As noted by Villa and Warshel,[150] these terms allow for the treatment of regular chemical reactions, not just electron-transfer processes. The second term on the right accounts for the stabilization of the transition state due to electronic coupling between the reactant and product states, while the final term corrects for the mixing of electronic states at the reactant state. This inclusion of electronic coupling effects enables the EVB method to model reactions involving significant changes in electronic structure, providing a more comprehensive framework than Marcus theory alone. See ref 12 for a more complete discussion.

In solutions, the model accounts for interactions with solvent dipoles, drawing on empirical evidence to accurately represent solvation effects. Within the enzyme, solvation of the ionic form involves analyzing charge and dipolar interactions, emphasizing the significance of electrostatic preorganization in enzymatic catalysis. By incorporating these additional terms, the EVB method effectively bridges the gap between electron-transfer reactions and more complex chemical transformations, justifying its application in studying enzymatic reactions and electrostatic effects.

The EVB method's flexibility allows for adaptation to various reactions, although it may become more complex when considering complicated metalloenzyme-catalyzed transformations. Nonetheless, its conceptual clarity and interpretive advantages make EVB particularly valuable for exploring electric fields in enzymes. In fact, the method has been used quite widely for the assessment of the field's role in enzymatic catalysis, and in governing such properties as pKa and reduction potentials.[42] An extensive, albeit slightly dated, compilation of these studies can be found in Table 2 of ref 12, which emphasizes the electrostatic foundations of enzymatic catalysis. Here, for the sake of completeness, we provide a succinct summary and concentrate on the most current research findings.

The EVB approach was first applied to the study of the catalytic reaction mechanism of lysozyme.[147] In this context, it proved useful in elucidating the critical influence of electrostatic interactions between the enzyme's active site and the ionic resonance forms of the reactive state. This offered a quantitative explanation for the enzyme's rate enhancement. This process notably involves the proton transfer from glutamic acid 35 to the polysaccharide's O4. It also involves the cleavage of the protonated C−O4 bond. Finally, the stabilization of the carbonium ion transition state by ionized aspartic acid 52 is also a part of this process.

Further extending its application, EVB was utilized to explore orotidine 5′-phosphate decarboxylase, aiming to uncover the origins of its catalytic function.[151] The findings underscored the importance of electrostatic transition state stabilization, alongside the structural induced fit of the protein, in minimizing necessary reorganization for catalysis. EVB simulations have also provided insights into the mutational effects within purine nucleoside phosphorylase, where a specific mutation from Asn243 to Asp significantly altered substrate specificity by modifying the electrostatic preorganization of the active site.[17]

Beyond these specific cases, EVB's utility in elucidating the role of electrostatics has been demonstrated across a diverse array of enzymes, including monoamine oxidase A.[152] Methyltransferases are another example.[153] Enzymes dependent on coenzyme B12 have also been studied.[154,155] Candida Antarctica Lipase A is another case.[156] GTPases have been examined as well.[157] Such studies have consistently revealed electrostatics as a key determinant in enzymatic function. Notably, the methodology has been applied to deconstruct the minimal influence of flexibility and dynamics in enzymatic catalysis, as illustrated by studies on dihydrofolate reductase.[158,159]

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

EVB has also proved applicable to metalloenzymes. This is exemplified by the catalysis study of diethyl 7-hydroxycoumarinyl by a custom-designed zinc metalloenzyme,[160] which serves as an example system to illustrate the general workflow. Figure 3 depicts the protein, its mechanism, and reaction profiles across various environments. Two ionic configurations were selected for EVB, D-H A^+ and D^+ H-A, represented as NADPH DHF-H^+ and $NADP^+$ HTF, respectively. The potential energy surfaces (PES) of these states (H_{11} and H_{22}), along with the mixing term (H_{12}, which is considered constant in the gas phase, solution, and within the protein), are defined by the elements of the Hamiltonian matrix:

$$H_{ii} = \varepsilon_i = \alpha^i_{\text{gas}} + U^i_{\text{intra}}(\boldsymbol{R}) + U^i_{\text{inter}}(\boldsymbol{R}, \boldsymbol{r}) + U^i_{\text{solvent}}(\boldsymbol{r})$$

$$H_{ij} = A\left(\frac{1}{1 + e^{-B(\varepsilon_j - \varepsilon_i)+C}} + \frac{1}{1 + e^{B((\varepsilon_j - \varepsilon_i)-C)}} + 1\right)e^{-\mu R} \quad (10)$$

Here, the atomic coordinates of the reactants or products in the diabatic states are given by $\boldsymbol{R}$, and coordinates of the surrounding water or protein by $\boldsymbol{r}$. a_{gas} is the energy of a given diabatic state in the gas phase, where all the fragments are at infinite separation. $U^i_{\text{intra}}(\boldsymbol{R})$ is the intramolecular potential of the solute system (relative to its minimum) in this state. $U^i_{\text{inter}}(\boldsymbol{R}, \boldsymbol{r})$ gives the interaction between the solute and the surrounding solvent atoms. $U^i_{\text{solvent}}(\boldsymbol{r})$ is the potential energy of the solvent.

Solving the secular equation for this description yields the adiabatic ground-state energy (Eg) and the corresponding eigenvector (Cg). Using these EVB potential energy surfaces, umbrella sampling MD trajectories and free energy perturbation simulations allow for the calculation of reaction free energies. The reaction barrier variations due to mutations reveal distinct differences aligning with experimental data (Figure 3c).

The KSI system was demonstrated to operate via electrostatic preorganization using EVB in 2010.[14] This preceded the notable experimental investigation in 2014.[136] Kamerlin et al. elucidated the impact of protein electrostatics on both the substrate, equilenin, binding and the initial reaction barriers within KSI.[159] They leveraged the substrate's pKa shift relative to its value in aqueous solution as a metric for the electrostatic influence upon binding. This shift, found to exceed 5.5 units, closely correlates with binding affinity, underscoring the predominance of electrostatic effects in binding. Moreover, the protein's electrostatics were revealed to furnish even more substantial stabilization to the transition state, according to EVB analysis. This research underscored the necessity of incorporating the full substrate rather than reduced models to ensure the substrate's correct orientation within the binding pocket, optimizing its alignment with the protein's electric field. Additionally, it cautioned against employing transition state analogues (TSA) for evaluating electrostatic contributions to catalysis.[161]

Roca et al. undertook an analysis of chorismite mutase (CM) proteins using the EVB method,[162] aiming to capture their experimentally observed catalytic performance. CMs are notable for catalyzing the Claisen rearrangement (Figure 4), a pericyclic reaction that does not straightforwardly benefit from a uniform electric field that might activate a specific bond. EVB, by not assuming nor requiring field directionality, assesses the influence of the protein's complete electric field

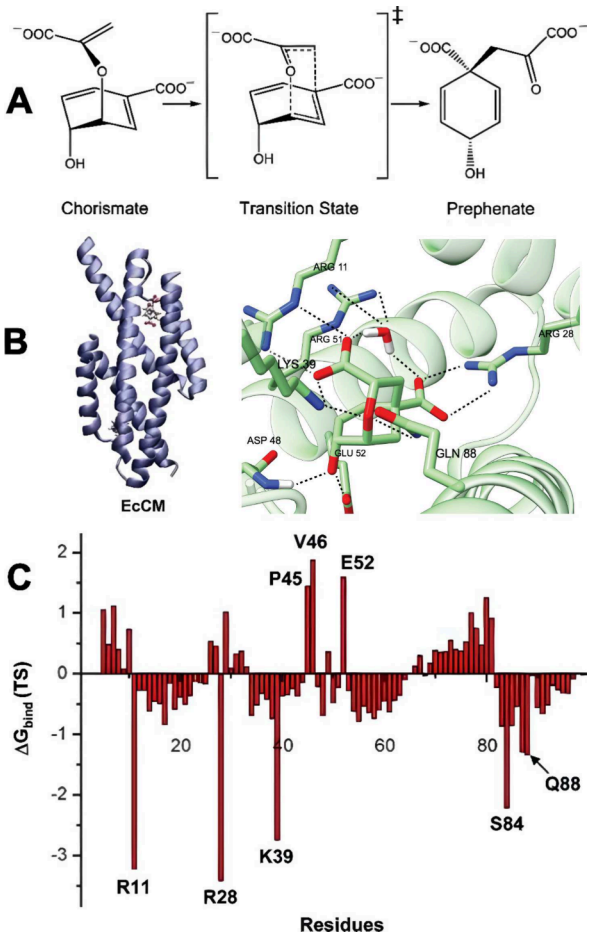


Figure 4. (A) Rearrangement of chorismate to prephenate. (B) Structure of the dimeric EcCM protein and the active site. (C) Electrostatic group contributions for the TS binding in the native EcCM in kcal/mol; these data are derived from EVB simulations, where the reference state for energies is the uncatalyzed reaction in solution. The contributions are shown for the residues of both subunits, which are close to the active site. Reproduced with permission from Roca et al., Biochemistry, 2009, 48 (14), 3046−3056. Copyright 2024 American Chemical Society.[162]

on the active site. Long-standing debates around CMs have questioned the source of their catalytic proficiency, with one theory suggesting the importance of bringing substrates into a near-attack conformation (NAC). Yet, further studies have demonstrated that the electrostatic stabilization of the transition state is the more significant catalytic mechanism. Nature presents several CM classes, alongside artificial CMs. Figure 4B illustrates the structure of EcCM and its active site, while Figure 4C highlights the electrostatic contributions of various residues throughout the protein scaffold to CM catalysis, noting that not all influential residues are charged, thus presumably affecting electrostatics with their backbone dipole.

Researchers have proposed EVB as a strategic tool for enzyme design, focusing on the influence of the scaffold's electric field.[163] They advocate EVB for evaluating the catalytic efficiency of designed enzymes prior to experimental validation.[162] In computational design, initially predicted proteins often show modest catalytic effects compared to their natural counterparts. These nascently active proteins undergo several rounds of directed evolution, incorporating

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

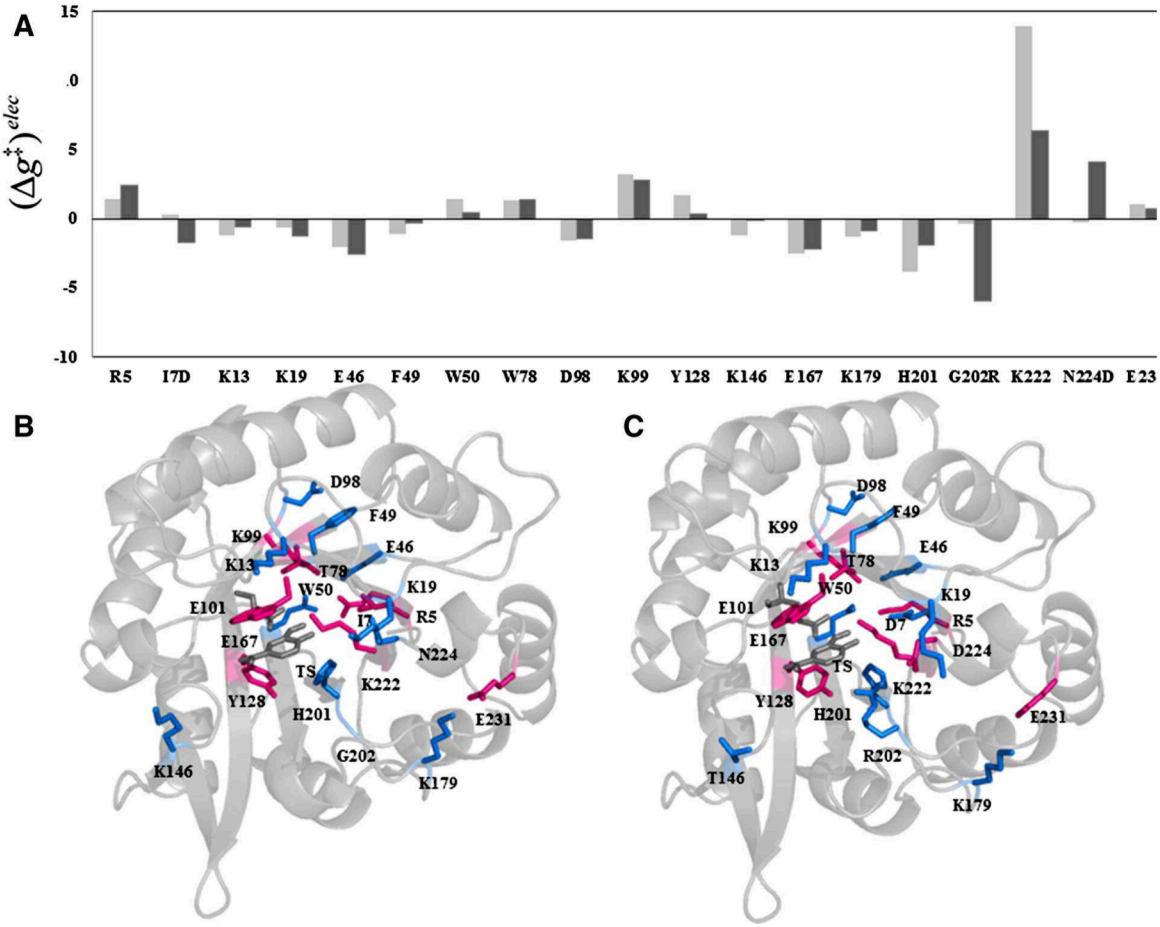


Figure 5. Residues providing significant contributions to the catalytic effect. (A) Kemp eliminase residues with electrostatic contributions exceeding 1 kcal/mol in magnitude in the KE07 design (light gray) and the R7 1/3H variant (dark gray). Residues with favorable (negative, blue) and unfavorable contributions (positive, magenta) to catalysis in KE07 design (B; PDB ID 2RKX) and the evolved R7 1/3H variant (C; PDB ID 3IIV) are displayed in the corresponding transition state structures. EVB method was also applied to examine the designed and evolved kemp eliminases by Frushicheva et al.[112] The authors found that for this particular reaction it is not very easy to incorporate TS stabilization via electrostatics, and that the evolution instead brings additional active site dissolvation. Reproduced from Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, Vol 1834, Labas et al., Optimization of the reorganization energy of the Kemp eliminase KE07, 908−917, Copyright (2013), with permission from Elsevier.[114]

targeted mutations throughout the scaffold to enhance catalysis. This iterative evolution underscores missing elements in the initial theoretical design, emphasizing the protein's role in catalysis. Many believe that directed evolution primarily enhances mid- and long-range electrostatics, filling the gap in electrostatic preorganization. To factor in reorganization energy early in the design process, a linear response strategy considering only the configurational space of reactants and products has been introduced.

The groundbreaking endeavor of computationally designing a series of enzymes unveiled Kemp elimination, transforming 5-nitrobenzisoxazole into cyanophenol, a reaction unfamiliar to existing enzymes (Figure 5).[15,80] Among these, KE59, leveraging an α/β barrel structure (PDB ID 1A53), significantly enhanced its catalytic efficacy by roughly 2,000 times following directed evolution.[111] This evolution notably fine-tuned the electrostatic interactions, achieving a remarkable reduction in reorganization energy by 27.4 kcal mol^{-1} from the wild type to its evolved R7 1/3H counterpart, illustrating the potency of directed evolution in refining electrostatic preorganization. Accordingly, the meticulous examination of the Kemp variants underscored that *in vitro* evolution precisely tailors reorganization energy.[112,114]

Further exploration led to the directed evolution of KE70, a design based on a TIM-barrel, which saw a nearly 400-fold boost in catalytic efficiency (k_{cat}/k_M).[110] Employing the EVB method to dissect the evolutionary path revealed an impressive enhancement in rate by a factor of 10^3.[114] This version capitalized on mutations that uniformly maintained or enhanced reorganization energy, validated by experimental congruence. Here, E101 functioned as the general base, K222 provided hydrogen bond support, and W50 engaged in a π-stacking interaction directly with the substrate. Computational findings, consistent with experimental observations, indicate that mutations arising during *in vitro* evolution either had a neutral impact or positively affected reorganization energy.

The evolution of Kemp eliminases, KE07 and HG3, presents a compelling narrative of catalytic diversification as unveiled by EVB calculations. Directed evolution of KE07 predominantly destabilized the ground state, while HG3's evolution stabilized the TS.[116] Further EVB investigations into KE07's evolution revealed a shift toward an alternative substrate binding mechanism, ultimately dominating its evolutionary end-

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

game.[164] This nuanced development across different variants demonstrates the transformative power of directed evolution, guided by strategic considerations of reorganization energy revealed by EVB methods.

The EVB literature frequently emphasizes the primacy of electrostatic contributions from the protein scaffold to enzymatic catalysis, as opposed to dynamics. It is essential to recognize that conformational dynamics influence the protein's generated electric fields and electrostatics, which in turn alter potential energy surfaces, affecting flexibility and dynamics. Hence, completely separating dynamic and electrostatic effects might not be entirely feasible, although possible under certain conditions. Despite general views, protein movements are likely disconnected from reaction processes, as convincingly argued by Warshel.[165] Yet, Schwartz identified specific protein promoting vibrations (PPV) that do couple with reaction dynamics.[166,167] Warshel and colleagues argue that vibrations, including PPVs, if adhering to a Boltzmann distribution, do not modify reaction barriers. Nonetheless, proteins experiencing thermal fluctuations generate varying electric fields.[98] Our observations revealed noticeable variances in the 3-D electric field geometry within the active site, depending on molecular dynamics trajectories. We discovered that certain field configurations, encountered at varying degrees of catalytic efficiency, could be significantly catalytic,[168] with some infrequent fields (manifesting briefly) exhibiting the most catalytic activity. Crucially, diverse fields may facilitate different TSs, i.e. altering the reaction mechanism. This implies that a single enzyme might exhibit multiple catalytic mechanisms induced by distinct electric fields produced by the protein, without direct momentum exchange (or dynamic coupling) between the protein and the reaction coordinate. In addition, different TSs achieved by different fields will break down the Boltzmann-based argument presented above. Consequently, it is vital to examine electric fields within configurations most pertinent to catalysis (those leading to the highest turnover frequencies), possibly derived from the tail of the Boltzmann distribution of protein movements. These configurations and their associated fields should be the focus when analyzing their impact on the catalyzed reaction's transition state. In essence, the predominant protein conformation and its typical electric field best represent the reactant state, while the transition state might be more accurately described by the infrequent yet highly active protein conformations and their unique fields, differing from those in the reactant state. Although EVB may not capture this field variability, it remains a potent theoretical framework for analyzing protein electrostatics.

2. DEVELOPING METHODS

2.1. Force Fields

2.1.1. Treatment of Electrostatics in Enzymes and Force Field Parameterization. As QM/MM simulations serve as the computational foundation for exploring electric fields in enzymes, their quest for accuracy naturally drives the development of more sophisticated FFs, crucial for precise electrostatic representation. This pursuit highlights the paramount importance of electrostatic characterizations in computational enzymology, a fact supported by extensive research.[169−173] The significance of this endeavor is 2-fold. On one hand, electric fields within enzyme active sites are known to accelerate rate-limiting reaction steps, and influence the spatial arrangement of substrates and adjacent residues, thereby potentially increasing field intensities.[89] On the other hand, while approximations such as fixed point charges for atoms or amino acid residues might suffice for modeling the protein scaffold's long-range electrostatic effects, a detailed portrayal of the protein's larger structure demands a more intricate electrostatic treatment. This becomes increasingly crucial in long-term simulations accounting for protein thermal fluctuations, where the precise modeling of the protein's expansive structure is vital for understanding how long-range electrostatic fields affect the active site's dynamics.

Our progress toward improved FFs may be categorized into three areas: (i) electronic polarization models, (ii) parametrized polarization models, and (iii) machine-learned interatomic potentials. This discussion will explore these advancements, complemented by insights from several comprehensive reviews on the breakthroughs of computational electrostatics.[169−171,174−178]

2.1.2. Electric Field Polarization. At the foundation of electrostatic modeling lies the concept of interactions between fixed point charges, which may represent individual atoms or groups of atoms, such as amino acid residues. This fundamental approach, however, grapples with two significant limitations. Initially, it falls short of depicting anisotropic charge density distributions—an issue that plagues both atom-centered and atom-group point charges. Moreover, it inadequately captures charge penetration effects that emerge when spherical charge density distributions intersect. These challenges manifest in the polarization of charge density around a point charge, necessitating direction-specific adjustments.

In exploring electronic polarization within chemical systems, two primary methodologies emerge: one attributes polarization to charge redistribution within individual atoms, employing either the induced dipole model[179−182] or the Drude oscillator model—alternatively known as the charge-on-spring or shell model.[183] The second methodology considers polarization as stemming from charge redistribution among atoms, utilizing frameworks like the fluctuating charge model, also referred to as charge equilibrium or chemical potential equilibrium.[184]

Using the induced dipole model, each point charge has a polarizability α in response to a field, resulting in a dipole with some direction and magnitude. This approach also extends to include atomic multipoles to obtain a more robust representation of atomic polarizability. Because the total energy is determined by the potential energy of the multipole expansion in response to those of nearby atoms, and in response to the field, the induced dipoles are solved in a self-consistent fashion such that total energy is minimized for the system.

In the Drude oscillator model, some amount of charge is taken from each point charge and given to a so-called Drude particle that is placed separate from the point charge, creating a dipole in the direction between the two points, with a magnitude proportional to their distance and relative charges. The Drude particle is connected to the point charge by a harmonic spring, and the positions of the Drude particles for all point charges in the system are again determined self-consistently (or approximated using an extended Lagrangian technique) such that total energy is minimized. The fluctuating charge model represents the charge for each point charge at multiple sites at and around the point, and uses the electronegativity equalization principle to minimize system

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

energy by balancing the electronegativity, chemical potential, and chemical hardness for each point charge.[185,186]

These three approaches differ in how they allow the charge density to deviate from a radial distribution around each point charge. The induced dipole and Drude oscillator models allow a dipole to form, and in fact the two were recently shown to be numerically equivalent,[187] since the Drude particle displacement is generally so slight that it approximates to a point dipole. The fluctuating charge model allows charge to move farther from an atomic site, from one atom to another, or even from one molecule to another, recovering a representation of polarization and charge transfer, and it can be expanded to include out-of-plane polarization at increased computational cost by including more virtual sites for each atom.[188] Additionally, fluctuating charge and induced dipole can be combined to achieve higher agreement with quantum-mechanical calculations,[189] again at a higher computational cost. However, it has been shown that inclusion of higher order (multipole) electrostatics produces results that more closely agree with those of DFT methods.[89]

2.1.3. Parameterization and Assessment of Polarizable Force Fields. The polarization models discussed above incorporate atom-specific parameters that account for changes in atomic polarization based on the types and positions of neighboring atoms. The derivation of these parameters, contingent on the model's nature, can be conducted through local or global methodologies, aiming to minimize discrepancies in energy or other attributes against high-fidelity benchmarks, whether experimental or derived from quantum mechanics. Local parametrization aligns parameters with individual atomic characteristics—charges, electronegativities, polarizabilities, and multipole moments, along with interatomic distances—based on molecular polarizabilities. While conceptually direct, this approach risks overfitting and may compromise transferability due to the selection and quality of reference data employed. Conversely, global parametrization seeks to refine parameters by minimizing deviations across entire systems against experimental or QM benchmarks, across broad chemical system sets. This method's complexity varies with the polarization model and the parameters it incorporates, as well as the choice of reference data.

Notable polarizable FFs utilized in molecular dynamics simulations of proteins and enzymes, such as CHARMM-Drude,[184,190−194] AMBER polarizable model[195] (FF02pol, FF02EP,[196] FF02r1,[197] ff12pol[198]), AMOEBA,[199] SIBFA,[200,201] ReaxFF polarizable model,[202] and LAMMPS[203,204] (using QEQ,[205−207] CORESHELL,[208,209] or DRUDE[210] packages), among others,[176] benefit from such optimized parameter sets for enhanced accuracy and transferability. The CHARMM Drude FFs employ the Drude oscillator model to represent anisotropic charge distributions and cover a wide range of biomolecules, including proteins, DNA, and lipids.[183,211,212] AMBER ff02pol is one of the earliest polarizable FFs, used in the study of proteins and nucleic acids, employing an induced dipole method,[196] while the ff12pol model includes damping functions to improve the treatment of long-range interactions.[198] The AMOEBA polarizable FFs use atomic induced dipoles for polarization and atomic multipoles up to quadrupole for permanent electrostatics,[199] and have been applied to various biological problems, demonstrating accurate structural predictions for systems including DNA, RNA, and proteins.[213,214] SIBFA is an *ab initio* polarizable FF that includes electrostatic multipole, short-range repulsion, polarization, charge transfer, and dispersion contributions,[200] initially developed for divalent-cation metalloproteins and later extended to halogen compounds and nucleic acids.[215−217] It has been implemented into Tinker-HP for parallel MD simulations, with recent advancements incorporating GEM for a more accurate representation of electron density, and combining with SIBFA and AMOEBA. ReaxFF is a bond-order-dependent FF parametrized using a large training set of QM data including thermodynamic and kinetic information,[202] useful for studying reactions which may include multiple intermediates along the reaction path,[218] including those in biological contexts.[219] Like other popular FF codes, it benefits from GPU-acceleration,[220] and using the ReaxFF implementations in LAMMPS,[221] AMS/ReaxFF,[222] or PureMD,[223,224] simulations of 10k+ atoms are feasible. LAMMPS is a versatile MD code that supports a number of polarizable FFs, including fluctuating charge, core−shell, Drude models, and a recently added polarizable embedding scheme for QM/MM calculations,[225] and is widely used for simulations of biomolecules and materials.[203,204]

Enhancements by Lin and MacKerell improved the CHARMM-Drude FF for halogenated molecules,[226] while refinements to the DNA model addressed issues like weak base stacking and Z-DNA unwinding.[211] QM energy profile fitting has also been used to adjust ion and water models for better compatibility with DNA models.[183] The CHARMM Drude-2013 polarizable protein FF was recently further optimized by Lin et al. to address limitations observed in β-sheet stability and structural fidelity over long molecular dynamics simulations.[227] The updated FF, termed Drude-2019, includes reoptimization of backbone parameters targeting the $(Ala)_5$ peptide's conformational properties in solution and gas-phase properties of the alanine dipeptide. Side-chain atomic polarizabilities and Thole screening factors for selected C_β, C_γ, and C_δ atoms were refined using QM dipole moments and molecular polarizabilities. Side-chain χ_1 and χ_2 dihedral parameters were optimized against QM data and PDB survey data. Nonbonded interactions between charged residues were also improved to better match QM interaction energies and experimental osmotic pressures. Validation through MD simulations of various peptides and proteins, including β-sheet structures and transmembrane ion channels, demonstrated that Drude-2019 offers smaller root-mean-square deviations and better agreement with experimental NMR data compared to CHARMM36m and Drude-2013, enhancing the FF's applicability for longer and more complex biomolecular simulations.

Ongoing development of AMOEBA aims to enhance accuracy and transferability by calibrating energy components to high-level QM energy decomposition and by utilizing automated optimization.[228] Improvements include better capturing electrostatic interactions with empirical damping functions,[229] refining the polarization model,[230] and parametrizing van der Waals interactions with the buffered-14−7 potential.[231]

For more information on improvements to these FFs, a recent review by Polêto and Lemkul discusses the advances in developing protein FFs, focusing on polarizable FFs and the limitations of additive FFs.[232] The article covers the parametrization strategies for FFs such as CHARMM, AMBER, OPLS, and GROMOS, which use experimental data, QM calculations, and automated fitting methods like ForceBalance.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

The inclusion of experimental solution data and QM calculations has improved the accuracy of FFs in capturing structural, dynamic, and thermodynamic properties. Polarizable FFs like AMOEBA and Drude models are shown to offer better representation of interaction energies in various chemical environments compared to nonpolarizable FFs. The authors highlight the need for more diverse experimental data, including structural and thermodynamic properties, to further refine FFs. Examples provided include the refinement of torsional parameters and nonbonded interactions, such as the development of the IPolQ model and the CHARMM22/CMAP FF. The review emphasizes the potential of machine learning and automated methods to enhance FF development, reducing biases and improving parameter accuracy, which we expand on below.

Assessment of the accuracy of FF-predicted electrostatics, and the resulting electric fields, is a necessary and ongoing process involving benchmarking of results against QM calculations and experimental observations. Kirsh et al. evaluated the accuracy of polarizable (AMOEBA) and nonpolarizable (AMBER) FFs in modeling electric fields in proteins using nitrile-containing photoactive yellow protein (PYP) variants.[233] Transition dipole moment (TDM) measurements were used to overcome hydrogen bonding issues in nitrile frequency measurements. Simulations showed that AMOEBA, which includes higher-order multipole parameters, more accurately reproduced experimentally derived TDM electric fields than AMBER, especially in hydrogen-bonding environments, thanks to AMOEBA's better modeling of negative charge density along the nitrile bond axis, leading to more head-on hydrogen bonds. Low-temperature IR spectra confirmed multiple nitrile populations, supporting the reliability of AMOEBA for accurate protein electrostatics modeling. This is particularly important as nitriles are commonly found on drugs, and accurate modeling is essential for understanding protein function and for computational screening of nitrile-containing compounds, potentially improving drug design and efficacy.

Amin et al. benchmarked polarizable and nonpolarizable FFs for Ca^{2+}−peptide interactions against a QM data set.[234] The systems analyzed included dipeptides bound to Ca^{2+}, with a particular focus on interactions involving carboxylate groups from Asp and Glu residues. The authors found that the Drude polarizable FF, prior to any parametrization, better approximated QM interaction energies than nonpolarizable FFs but suffered from polarization catastrophes at short Ca^{2+}−carboxylate distances. To mitigate this, the Drude FF was optimized using Boltzmann-weighted fitting, which improved its accuracy in MD simulations of calmodulin's N-lobe. Additionally, the CTPOL FF, which incorporates charge-transfer and polarization effects,[235,236] was evaluated. The optimized Drude FF showed improved performance, reducing discrepancies in ion-ligand interactions. This benchmark process led to parameter optimization and resulted in specific FF improvements for accurately capturing the structure and dynamics of ion-protein interactions, which is crucial for applications involving metalloproteins.

Proper treatment of classical electrostatics via polarizable FFs is also important in QM/MM approaches. In a study on the catalytic mechanism of reductive dehalogenase PceA, Zhang et al. used MD simulations and QM/MM calculations to explore the role of the proximal [4Fe−4S] cluster in proton-coupled electron transfer (PCET).[237] The QM/MM scheme utilized ChemShell to integrate QM calculations with Turbomole (B3LYP/Def2-SVP) for the QM region and DL_POLY (AMBER ff14SB FF) for the MM region, using an electronic embedding approach to account for the enzyme environment's polarizing effects on the QM region. The authors found that the $[4Fe{-}4S]^{1+}$ cluster facilitates the reduction of organohalide compounds by mediating a PCET process, with Arg305 acting as the proton donor, while the deprotonated Tyr246 stabilizes Arg305's conformation and proton donation ability. This novel mechanism helps rationalize the selective dechlorination of trichloroethene to cis-1,2-dichloroethylene. The QM/MM scheme revealed that the $[4Fe{-}4S]^{1+}$ cluster's participation in PCET enhances both exchange and superexchange interactions, overcoming previous uncertainties related to electron and proton sources in the dechlorination process. The findings have significant implications for the bioremediation of toxic chloroalkenes, commonly found in industrial and agricultural pollutants, highlighting the potential for engineering more efficient biocatalysts for environmental detoxification.

Yan et al. used QM/MM to evaluate TS (de)stabilization effects of electric fields generated by scaffold residues in proteins, specifically focusing on the O−O heterolysis reaction in tyrosine hydroxylase.[238] The QM region included the heme with the Fe−OOH moiety, ligand imidazole ring, and substrate, while the MM region was represented using AMBER ff14SB. The authors found that electric fields due to residues far from the active site could be approximated as uniform, while the heterogeneity of electric fields due to residues near the active site necessitated a direct QM/MM calculation for accurate assessment. The results demonstrated that scaffold residues contribute to the preorganized electric field, influencing the catalytic efficiency of enzymes.

Song and Wang combined state-averaged CASSCF with the AMOEBA FF to study photoreactions in proteins.[239] The model incorporated a variational treatment of intramolecular polarization and link atom schemes to handle QM/MM boundaries through covalent bonds. Single-link and double-link atom schemes were evaluated, with the double-link scheme providing more accurate results for both ground and excited states. The model was applied to the O−O bond dissociation in NanoLuc, showing significant differences in reaction pathways and conical intersections compared to gas phase and polarizable continuum model (PCM) calculations. This work demonstrates the combination of multireference quantum chemistry with polarizable FFs in QM/MM models, paving the way for more accurate studies of photoreactions in complex biological systems.

Demonstrating the catalytic potential of designed LEFs, Siddiqui et al. investigated the enzymatic degradation of polyethylene terephthalate (PET) using engineered cytochrome P450 (CYP450) enzymes. Combining MD simulations, molecular docking, QM/MM, and binding free energy calculations, the authors focused on three variants: CYP450GcoA, CYP450OleT, and CYP450BSβ, finding that only CYP450GcoA, with a preorganized LEF favorably aligned along the O−C1 bond, could efficiently catalyze PET degradation. The comprehensive study demonstrated that PET degradation occurs in two steps: C−H hydroxylation and O-dealkylation, with the latter driven by the enzyme's LEF, engineered via delibrate mutation based on alignment with the reaction axis. The significance of this investigation lies in demonstrating that LEFs can be used to guide enzyme

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

engineering, leading to new strategies for enzymatic catalysis of non-native reactions, and in this case offering a greener solution for plastic degradation.

More information on QM/MM methods for enzyme modeling can be found in the tutorial review by Ahmadi et al., which covers QM, QM/MM, and QM/MM/MD modeling approaches to enzyme modeling.[240] Also see the tutorial review by de la Lande et al., which focuses on QM/MM methods using the deMon2k software package.[241]

2.1.4. Coarse-Grained Force Fields. Coarse-grained (CG) FFs have emerged as powerful tools in biomolecular simulations, offering a balance between computational efficiency and accuracy.[242] Unlike atomistic FFs, which model every atom explicitly, CG FFs simplify the system by grouping atoms into larger units or "beads." This reduction in complexity significantly enhances computational performance, enabling the simulation of larger systems and longer time scales that would be prohibitive with atomistic approaches.

Some example CG FFs for biochemical systems are SIRAH,[243] MARTINI,[244,245] UNRES,[246] CABS,[247] PRIMO,[248] PRIME,[249] and OPEP.[250] CG FFs offer substantial speed-ups while retaining essential physical properties, making them ideal for studying phenomena such as protein folding,[251] membrane fusion,[252] and the self-assembly of complex biomolecular structures, where capturing long-range interactions and large conformational changes is crucial.

Due to the simplified nature of CG FFs, their application to individual biological macromolecules can be more nuanced. For instance, while they excel at capturing the general behavior and dynamics of macromolecular assemblies, they may lack the detailed resolution needed for investigating fine structural features or specific atomic-level interactions within a single protein or nucleic acid. This is because the coarse-graining process inherently involves a loss of detailed atomic information, which can be critical for understanding precise mechanisms such as enzyme catalysis, ligand binding, or allosteric regulation.

To address this shortcoming, researchers have developed alternative CG approaches such as shape-based coarse-graining (SBCG) allowing higher bead granularity, hybrid CG/atomistic schemes in order to embed finer resolution atomistic regions within surroundings represented with CG FFs, and employ advanced parametrization and machine learning to improve the ability of the CG beads to capture the behavior of the groups of atoms they represent.

Bryer et al. presented advancements in the SBCG method, termed SBCG2, with an improved topology representing network, a charge density Fourier Shell Correlation method for granularity selection, and an iterative refinement protocol for bond and angle parameter optimization.[253] The systems investigated include the HIV-1 capsid and heteromultimeric cofilin-2 bound actin filaments. The SBCG2 models maintain assembly characteristics with subnanometer resolution and achieve high simulation performance using the GPU-resident NAMD3 engine, with the HIV-1 capsid simulations exceeding 1 μs per day without particle mesh ewald (PME) and nearly 300 ns per day with PME, and actin filaments achieving up to 4 μs per day. The SBCG2 methodology is implemented in VMD and facilitates efficient simulation of large-scale macromolecular complexes.

Demonstrating an advancement in algorithmic CG FF parametrization, Caceres-Delpiano et al. employed an automated approach using ForceBalance to optimize the SIRAH CG FF by reproducing hydration free energy gradients derived from atomistic molecular simulations.[254] The method specifically targeted uncharged side-chains and the protein backbone, with MD and QM calculations, incorporating multiple sources of experimental or simulated reference data. The optimized parameters significantly improved agreement with experimental hydration free energies, achieving an R^2 of 0.985, compared to the original SIRAH 1.0 and 2.0 models which had R^2 values of 0.104 and 0.404 respectively. Despite challenges in optimizing parameters for charged side-chains, the new SIRAH-OBAFE FF showed enhanced protein stability in MD simulations, with reduced RMSD values.

In the area of machine-learning-based improvements to CG FFs, Kanada et al. presented an advanced coarse-grained molecular dynamics (CGMD) simulation method that combines a smoothed hybrid potential with a neural network model to enhance the structural transition dynamics of proteins.[255] The hybrid potential integrates an AI-based potential with minimal CG potentials, targeting bond lengths and excluded volume interactions. The AI potential is trained using a diverse structural ensemble from multicanonical MD simulations, with energy profiles smoothed by energy minimization. This methodology was applied to chignolin and TrpCage, achieving a high correlation coefficient (R^2 > 0.89) between predicted and true energies. CGMD simulations using the hybrid potential showed significantly improved transition dynamics between metastable states compared to conventional CGMD and all-atom MD, while maintaining protein properties.

Another recent machine-learning-based approach to improve CG FFs was presented by Navarro et al.[256] Using a top-down approach via MD, neural network potentials (NNPs) were trained via differentiable trajectory reweighting using only the native conformations of proteins, eliminating the need for extensive labeled data or memory-intensive simulations. Validation was conducted using Markov state models to predict native-like conformations from CG simulations. The theoretical transferability of the method and its potential for new protein force fields are highlighted in the study, where trained NNPs are shown to generalize to new proteins and accurately fold proteins outside the training set, achieving results comparable to other CG methods and demonstrating the efficiency and potential of this machine learning approach in protein modeling.

Lastly, a cutting-edge investigation by Carrer et al. introduces ∂-HylleraasMD (∂-HyMD), a fully end-to-end differentiable molecular dynamics software based on the Hamiltonian hybrid particle-field (HhPF) formalism.[257] Leveraging the JAX autodiff framework for automatic differentiation, the software performs automated optimization of CG FF parameters. The optimization process employs a trivially parallel algorithm where independent simulations are run in parallel, and their trajectories are processed using reverse mode automatic differentiation to calculate the gradient of the loss function, used iteratively to optimize the FF parameters, ensuring convergence and avoiding memory and numerical stability issues. The authors demonstrate the effectiveness of ∂-HyMD by optimizing FF parameters for standard phospholipids, including those with zwitterionic or anionic heads and saturated or unsaturated tails. The optimized FF parameters yield improved density profiles compared to those derived from gradient-free optimization methods and accurately predict properties like lateral pressure

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

profiles, and their transferability to other systems such as triglycerides was demonstrated.

For more information on advancements in CG-based biomolecular modeling, see recent reviews[258,259] including that of Roel-Touris and Bonvin,[260] which covers the nuance of the coarse-graining process, new parametrization strategies for CG models including machine-learning-based parametrization, extensions of multiscale hybrid methods, and use of sophisticated sampling and scoring schemes to enable longer simulation time scales. The review also covers the development of integrative modeling, fusing experimental data with computational models to generate more accurate and reliable structural models. The Protein Data Bank (PDB) has since 2014 established a new integrative model database (PDB-Dev) for archiving such models. Examples include structures determined using CG/hybrid computational approaches with various experimental techniques. Chemical cross-linking data have been extensively used, leading to models of complexes such as the heptameric module of NPC, exosome complex, and Complement C3(H2O). Other examples involve combining cross-linking with 2DEM/3DEM, SAS, mutagenesis, NMR, and single molecule FRET data. Specific cases like the ATP synthase membrane motor and ISWI ATPase complex showcase the integration of CG modeling with experimental data for detailed structural insights. The Nuclear Pore Complex (NPC), modeled using the Integrative Modeling Platform (IMP), represents a significant achievement, where multiscale modeling combined with extensive experimental data resulted in a subnanometer precision structure of a large protein assembly.

2.1.5. Machine Learning Interatomic Potentials. Harnessing artificial intelligence (AI) to develop machine learning-based interatomic potentials represents a groundbreaking stride toward crafting more accurate and comprehensive models of enzyme behaviors. These machine learning potentials, trained on data sets of atomic configurations and their quantum mechanically computed energies/forces, seek to blend quantum accuracy with the computational speed of traditional FFs.

Machine learning interatomic potentials (MLIPs) historically face challenges in scaling and accuracy when modeling complex proteins. Despite this, advancements in charge descriptors, high-performance algorithms, and the advent of fourth-generation models promise significant breakthroughs.[261] The pioneering work by Behler-Parrinello on atom-centered potentials marked the beginning of this field.[262] These early models, however, were hampered by reliance on descriptors that captured the local atomic environment in a generalized or oversimplified manner and did not include sufficient detail to capture all relevant physical and chemical interactions—such as multicenter bonding. In addition, capturing long-range electrostatic and dispersion forces proved challenging due to computational cost and the fact that local descriptors inherently neglect these forces.

Fourth-generation frameworks, such as CENT[263] and BpopNN,[264] rely on more sophisticated descriptors that capture the local atomic environment with greater detail and accuracy. These descriptors are designed to encode, not only the immediate surroundings of an atom, but the broader context of its position within the molecule or material, allowing the model to account for both short-range and long-range interactions. Some fourth-generation models incorporate global optimization techniques or charge equilibration mechanisms that allow for a dynamic adjustment of charge distribution based on the overall molecular or material structure. This approach uses methods akin to those found in DFTB[265] and ReaxFF[224] to mimic the way real charges redistribute in response to electrostatic forces, improving the accuracy of predictions related to electrostatic interactions. Message-passing neural networks represent another innovative approach to FF development. These networks update atom representations iteratively, effectively "passing" information between atoms across the molecule or material and allowing the model to capture the influence of distant atoms, overcoming the limitation of earlier models that struggled with long-range interactions.[266]

These advancements have significantly improved the efficiency and accuracy of computational methods. Recent explorations into protein modeling with MLIPs have been promising, including the Allegro algorithm's application to simulate the entire HIV Capsid,[267] using Graph-Based Embedding of Molecules (GEMS) which combined a bottom-up and top-down methodology that improved the modeling of short-range and long-range interactions in systems of over 10,000 atoms (Figure 6).[268] This latter study,

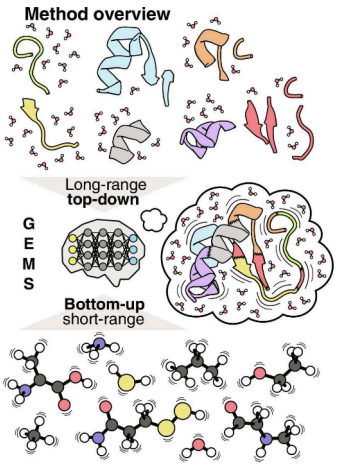


Figure 6. Biomolecular dynamics with machine-learned quantum-mechanical FFs trained on diverse chemical fragments. Graph-Based Embedding of Molecules (GEMS) uses both bottom-up and top-down approaches by creating fragments from large and small molecules which are used in training machine learning models to predict molecular behavior or properties. Adapted from ref 268.

comparing *ab initio* calculations with AmberFF and a MLIP, Spookynet, revealed that MLIPs provide more accurate representations of protein dynamics over extended time scales, uncovering new dynamics in crambin and poly alanine systems.

Despite these algorithmic advances, their application to systems influenced by induced or intrinsic electric fields remained limited until the advent of alternative MLIP algorithms designed for electric field operations and electrostatic potential energy terms. Notably, Christensen et al. introduced a kernel method utilizing fictitious effective charges, coupled to electric fields as system energetic descriptors.[269] Similarly, FieldSchNet treats solvent effects through an effective scalar field sensitive to electrostatics,[270] while Gao and Remsing devised a dual approach for long-range and short-range interactions, with the former incorporating a perturbative correction to atomic descriptors influenced by electric fields.[271] Jiang et al.'s method leverages Gaussian-type orbitals, informed

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

by local chemical environments and virtual fields, to embed atomic densities, offering direction-equivariant descriptors.[272] These innovative techniques, promising for protein system applications, have primarily been tested on smaller systems, not yet extending to protein dynamics governed by (and governing) electrostatics.

Shifting focus to protein studies involving charge-embedding schemes, Zinovjev introduced an ML/MM embedding strategy that updates atomic energies within a machine learning-treated "quantum region" based on electrostatic interactions with the surrounding MM region,[273] adaptable to various MLIPs. This approach demonstrated superior performance in embedding the PF-00835231 inhibitor within the SARS-CoV-2 Mpro Complex compared to semiempirical methods, all while reducing computational demands. Another method applies a deep potential range correction (DPRc) to adjust MM potentials with a distance-dissipating interaction term near the QM region, predicting RNA cleavage reactions' free energy barriers and kinetic isotope effects.[274] Shao and colleagues adopted Δ-learning to refine semiempirical energy estimates through a bespoke machine-learned embedding scheme in CM simulations,[274] achieving DFT-comparable accuracy at a fraction of the cost. An intriguing study combined hybrid MLIP/AMOEBA MD approaches, integrating machine learning's short-range precision with AMOEBA's efficient long-range, polarizable FF handling.[275] The approach was tested across various solvated proteins, such as DHFR, COX, and SARS-CoV2Mpro, albeit with mixed accuracy results. These advancements herald a new frontier in electrostatically informed protein system studies.

2.2. Other Computational Methods for Biomolecular Simulations

In addition to the discussed FFs and MLIPs, several other methods have been developed to approximate electrostatic interactions and other properties in biomolecular systems, which we will briefly outline here.

Graph theory-based methods model molecules such as proteins, nucleic acids, and their complexes as graphs, where nodes correspond to atoms, residues, or functional groups, and edges represent chemical bonds or interactions.[276] This abstraction facilitates the analysis of intricate networks of interactions, enabling the identification of key structural motifs, functional domains, and critical interaction pathways.[277] Graph theory-based methods are particularly useful for studying protein folding, where the network of intramolecular interactions can be analyzed to predict folding pathways and identify stable intermediates. Additionally, these methods are employed to explore protein−protein interaction networks, elucidating the functional connectivity and signaling pathways within cellular systems.[278] By leveraging algorithms from graph theory, such as shortest path, clustering, and centrality analysis, researchers can gain insights into the robustness, modularity, and dynamics of biomolecular networks. Graph theory-based methods have also been integrated with other computational techniques, such as MD and ML.[279] Though not related to the treatment of local fields directly, the structural information derived from these methods can be used to inform FF development and improve the accuracy of electrostatic models in biomolecular simulations.

Monte Carlo (MC) methods are a class of computational techniques that utilize random sampling to explore the configurational space of biomolecular systems and calculate their thermodynamic properties.[280−283] In biomolecular modeling, MC methods are continually employed to simulate the behavior of proteins, nucleic acids, and other biological molecules under various conditions.[284,285] By generating random configurations and using importance sampling algorithms like Metropolis-Hastings, MC methods can efficiently sample the most probable states of a system, providing insights into equilibrium properties such as energy, entropy, and free energy.[286−289] Variants of MC methods, such as Grand Canonical Monte Carlo (GCMC)[290−296] and Replica Exchange Monte Carlo (REMC),[297−300] extend these capabilities to systems with fluctuating particle numbers and enhanced sampling of high-energy barriers, respectively. These techniques help elucidate protein folding, ligand binding, solvation effects, and other complex phenomena in biological systems, offering a robust framework for enzyme and protein research and design.[301] Like graph theory-based methods, MC methods do not directly related to the prediction of local fields in enzymes and proteins, but they are essential tools for exploring the conformational space and thermodynamic properties of biomolecules, which can inform the development of FFs and other computational models.

Continuum solvent models, including Poisson−Boltzmann (discussed below) and Generalized Born methods,[302−304] are widely used to estimate the electrostatic solvation free energy and binding affinities of biomolecules.[305−307] Polarizable Continuum Models (PCM) are commonly used in computational enzyme modeling to simulate the solvent environment by treating it as a polarizable dielectric continuum, thereby providing a more accurate representation of solvation effects on the local electrostatics of proteins and enzymes.[308−312]

Conceptually adjacent to coarse-grained FFs, Fragment-Based Quantum Chemistry (FBQC) methods similarly decompose large systems into fragments but often incorporate more detailed quantum mechanical calculations for each fragment and their interactions.[313−316] These methods facilitate the study of large biomolecular assemblies, enabling the exploration of their structural and functional properties with high accuracy and reduced computational cost, providing a bridge between detailed electronic structure methods and large-scale biomolecular simulations. Discussed below is the Effective Fragment Potential method, which is a type of FBQC method.

2.2.1. Effective Fragment Potential Method. The Effective Fragment Potential (EFP) method is designed to simulate large molecular systems with high accuracy and efficiency.[317−319] Unlike traditional MD or FF methods, which often rely on empirical parameters to describe atomic interactions, the EFP method is rooted in QM principles, treating the system as a collection of fragments, each represented by a set of potentials derived from quantum mechanical calculations. These fragments interact through electrostatic, polarization, dispersion, and exchange-repulsion forces, capturing the essential physics of molecular interactions without the need for extensive QM calculations on the entire system.

One of the primary advantages of the EFP method over conventional MD simulations is its ability to accurately model electronic polarization effects, including polarization explicitly by allowing the fragments to polarize in response to the local electrostatic field. This leads to a more realistic representation of molecular interactions compared to that provided by many MD approaches, particularly in systems where polarization

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

plays a significant role, such as in solvated biomolecules or in the presence of strong electric fields. Another key distinction between the EFP and FF simulations is in the treatment of dispersion and exchange-repulsion interactions.[320] FFs often use empirical van der Waals parameters to approximate these interactions, while EFP derives these interactions from QM calculations on the fragments, ensuring that they are more physically accurate and system-specific. This quantum-derived approach makes EFP method well-suited for large-scale simulations where empirical parameters may fail. EFP can also be used to parametrize other FFs, such as CG FFs.[321]

Since its introduction, EFP has been extended to include a DFT-based parametrization,[322] and a number of improvements to the method have been implemented including a parallel version of the method.[323] Charge-transfer interactions have become treatable in EFP through the use of perturbation theory, canonical MOs, and Fock matrices from SCF calculations.[324] Short-range behavior of the Coulomb, induction, and dispersion interactions have been improved by incorporating damping functions,[325] and work has been done to improve the derivation of multipoles used in EFP.[326] Improvements have also been made to its treatment of polarizability[327] and of dispersion[328,329] and exchange-repulsion interactions.[320]

Tazhigulov et al. employed a polarizable embedding scheme to simulate the redox potentials of biomolecules, particularly focusing on the FAD cofactor reduction in cryptochrome 1 from *Arabidopsis thaliana*.[330] The investigation used DFT (oB97X-D/6-31G(d)) with the BioEFP (the extension of the effective fragment potential[319,320,331] method to biological macromolecules)[318] representing the classical surroundings. The study highlighted the critical role of environment polarization and long-range electrostatic interactions, showing that ignoring these factors can lead to significant errors in computed redox potentials, with differences as large as 1.4 V. The results demonstrate that accurate estimates of redox potentials require proper treatment of polarization and long-range interactions, which were essential in achieving good agreement with experimental data. The study helps researchers' understanding of redox processes in biological systems, which are crucial in areas like energy storage, photovoltaics, and medical applications such as DNA repair and circadian rhythm regulation.

Recently, Slipchenko presented a hybrid QM/EFP approach for analyzing solvatochromic shifts in electronically excited states by decomposing these shifts into contributions from individual solvent molecules using.[331] Using two model systems, a water pentamer and hydrated uracil, Slipchenko demonstrated that the decomposition into individual solvent contributions highlights significant nonadditivity in solute–solvent interactions and emphasizes the importance of solute polarization in determining the total solvatochromic shifts, concluding that these insights can guide the design of materials with targeted optical properties. The results underline the necessity of considering many-body effects and long-range interactions for accurately predicting solvatochromic shifts.

Merging fragment molecular orbital (FMO) and EFP, the Effective Fragment Molecular Orbital (EFMO) method extends EFP to include the simulation of reactive systems. Unlike the traditional EFP method, which primarily computes intermolecular interaction energies, EFMO handles both covalently bonded systems and noncovalently bound molecular clusters. Sattasathuchana et al. recently outlined improvements to the EFMO method, including a new memory-based implementation for solving the coupled perturbed Hartree–Fock (CPHF) and time-dependent Hartree–Fock (TDHF) equations.[332] The improved EFMO method, parallelized using hybrid MPI/OpenMP, achieves nearly ideal strong scaling, enabling efficient calculations on massively parallel supercomputers. Demonstrating excellent accuracy (<1 kcal/mol error per fragment), the authors show that EFMO can now perform calculations that include long-range polarization and dispersion interactions on systems as large as hydrated mesoporous silica nanoparticles with explicit water solvent molecules, showcasing its feasibility for exascale computing architectures.

2.2.2. Poisson–Boltzmann Equation. Solving the Poisson–Boltzmann (PB) equation is the standard approach in situations where detailed and accurate modeling of electrostatic interactions is crucial, particularly when these interactions significantly influence the system's behavior, such as in pKa calculations of ionizable residues in proteins. The PB equation has its roots in the fields of electrostatics and statistical mechanics,[333−335] and was originally formulated to describe the electrostatic potential in a medium containing mobile charge carriers. In contrast to more traditional force field (FF) methods that rely on empirical parameters, the PB equation provides a continuum approach that directly captures the essential physics of electrostatics, making it especially valuable for complex biological systems.

The underlying Poisson equation for the electrostatic potential $\phi(\mathbf{r})$ takes the form

$$\nabla\cdot[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -\rho(\mathbf{r}) \tag{11}$$

where $\rho(\mathbf{r})$ is the charge density and $\varepsilon(\mathbf{r})$ the spatially dependent absolute permittivity. The ionic charge density following a Boltzmann distribution, $\rho_{\text{ion}}(\mathbf{r})$, is

$$\rho_{\text{ion}}(\mathbf{r}) = \sum_i c_i^0 \, Z_i \, e \exp\left(\frac{-Z_i \, e \, \phi(\mathbf{r})}{k_B T}\right)$$

for species i with valence Z_i and number density c_i^0 (m^{-3}), where e is the elementary charge, k_B the Boltzmann constant, and T the temperature. Hence the Poisson–Boltzmann equation becomes

$$\nabla\cdot[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -\rho_{\text{ion}}(\mathbf{r}) = -\sum_i c_i^0 \, Z_i \, e \exp\left(\frac{-Z_i \, e \, \phi(\mathbf{r})}{k_B T}\right) \tag{12}$$

which is nonlinear in ϕ.

It is not uncommon to separate the charge density into an ionic contribution and a fixed contribution that does not obey Boltzmann statistics, such as the contribution from a charged surface.[336] In such cases

$$\rho(\mathbf{r}) = \rho_{\text{fixed}}(\mathbf{r}) + \rho_{\text{ion}}(\mathbf{r})$$

giving the PB equation as

$$\nabla\cdot[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -\rho_{\text{fixed}}(\mathbf{r}) \underbrace{- \sum_i c_i^0 \, Z_i \, e \exp\left(\frac{-Z_i \, e \, \phi(\mathbf{r})}{k_B T}\right)}_{\rho_{\text{ion}}(\mathbf{r})} \tag{13}$$

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

Where $\phi(\mathbf{r}) \ll k_bT$ a linear approximation can be applied, yielding:

$$\nabla\cdot[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -\rho_{\text{fixed}}(\mathbf{r}) + \kappa^2\phi(\mathbf{r}) \quad (14)$$

where κ, the Debye screening parameter, is defined in terms of the ionic strength I of the solution:

$$I = \frac{1}{2}\sum_i c_i^0 Z_i^2, \quad \text{and} \quad \kappa^2(\mathbf{r}) = \frac{2e^2 I(\mathbf{r})}{k_B \quad T}$$

By employing a continuum description of solvent and ionic species, the PB equation allows for efficient computation of electrostatic potentials in systems where traditional FF-based methods are either too costly or less straightforward. In the context of proteins and enzymes, where electrostatic interactions critically influence stability, substrate binding, and catalysis, PB calculations can yield insights into pKa values, solvation energies, and binding affinities. This capability complements MD simulations and QM-based methods, offering a balance of accuracy and computational efficiency well-suited to large biomolecular assemblies. Within enzyme and protein modeling, the PB equation has been used to analyze the electrostatic environment of active sites, quantify the influence of surface charges, and assess how mutations alter protein stability and reaction mechanisms.

Recent work has leveraged PB-based approaches for more accurate pKa predictions. Meyer and Knapp presented a method that combines electrostatic energy computations with MD simulations using different protonation patterns to predict pKa by solving the PB equation.[337] Their approach reduced the root-mean-square deviation between computed and measured pKa values from 1.17 to 0.96 pH units compared to previous methods, with further improvement to 0.79 pH units when conformations were energy-minimized with a dielectric constant of $\varepsilon = 4$. More recently, Aleksandrov et al. introduced a method that integrates Monte Carlo simulations, a Drude polarizable FF, and an implicit PB solvation model.[338] This approach achieved excellent convergence and accuracy in pKa calculations for lysozyme and other proteins, outperforming the additive CHARMM36 FF and demonstrating reduced sensitivity to assumptions about the internal dielectric constant.

Beyond pKa estimation, PB theory has been incorporated into process modeling. Briskot et al.[339] applied PB-based considerations combined with a basic Stern model to predict retentate and permeate compositions in ultrafiltration/diafiltration (UF/DF) processes. Their model accurately described low to moderate Donnan potentials and improved predictions under high concentration conditions, providing a valuable tool for understanding and controlling protein processing steps in biotechnological applications. Similarly, Gama et al.[340] employed a modified PB equation (PBEm) to investigate the binding of lysozyme onto a mesoporous silica surface, demonstrating that both electrostatic and van der Waals forces govern adsorption under varying pH and ionic strength conditions.

Methodological assessments have further highlighted the strengths and limitations of various PB-based approaches, offering benchmark data and best practices for improved predictions.[341,342] Additionally, PB techniques have been employed to analyze how pH and electrostatic environments influence protein−protein and enzyme−substrate interactions, thereby offering a window into complex pH-dependent processes relevant to enzyme catalysis.[343,344] Further improvements in modeling the dielectric environment, including spatially varying dielectric constants, have led to more precise pKa estimates and a better understanding of the factors governing enzyme activity.[345,346] Tools like PypKa provide advanced, optimized, parallel PB-based pKa predictions and can be incorporated into existing computational research workflows with ease.[347]

For a more in-depth exploration of PB-related methods, advancements, and applications—ranging from fundamental electrostatic principles to complex biomolecular and industrial systems—readers are referred to comprehensive reviews available in the literature.[307,348−351]

2.3. Representing Charge Density and Electric Field Geometry

In any open chemical system, such as an enzyme active site, an electric field, whether internally generated by the surrounding chemical environment or an applied externally, will alter the electron distribution, necessarily leading to changes in nuclear positions. A comprehensive approach to studying electronic structure involves analysis of the full quantum mechanically mediated electron density.

According to Kohn's theorem, this density determines all ground state system properties, for example the anharmonicity of the reaction coordinate mode, which in turn affects the barrier height. The subtle shifts in active site charge density, influenced by the long-range effects of electric fields, can significantly alter reaction rates, sometimes by orders of magnitude. These long-range effects stem from the enzyme's entire extended structure, underscoring the importance of understanding the relationships between enzyme activity and the global structure of the electron density. Once discovered these relationships will serve as a foundational element of enzyme design—allowing for the tailoring of electron density to elicit preferred responses to electric fields. However, a necessary step toward this end is to devise methods to describe and quantify the global structure of electric fields and charge densities. Thankfully, robust formalisms exist for detailing the geometric structure of 3D scalar (electron density) and vector (electric fields themselves) fields, offering a foundation for this advanced exploration.

2.3.1. The Topological Character of Electric Fields and the Electron Charge Density. Topological analysis offers a significant method for examining how systems respond to perturbations and is applicable to both scalar fields like charge density, $\rho(\boldsymbol{r})$, or the electrostatic potential, $V(\boldsymbol{r})$, and vector fields such as $\nabla\boldsymbol{\rho}(\boldsymbol{r})$ and electric fields, i.e. $-\nabla\boldsymbol{V}(\boldsymbol{r})$. The essence of a 3D scalar field's topology is captured by its critical points—locations where the gradient of the field vanishes. Defined by the field's three principal curvatures (the principal components of the field's Hessian matrix) at these points, there are four primary types: maxima, where all curvatures are negative, minima, with all curvatures positive, and two forms of saddle points, distinguished by the sign of the two curvatures being either positive or negative, with the third curvature of opposite sign. A common notation is to distinguish critical points with two indices corresponding to the rank and signature of the Hessian matrix at the point in question. The first number indication the spatial dimension and the second indicates the signature (the sum of the signs of the eigenvalues, with negative values for a maximum). Thus, a maximum is

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

denoted (3,-3) a minimum as (3,3) and the two saddle points (3,1) and (3,-1) (Figure 7A).

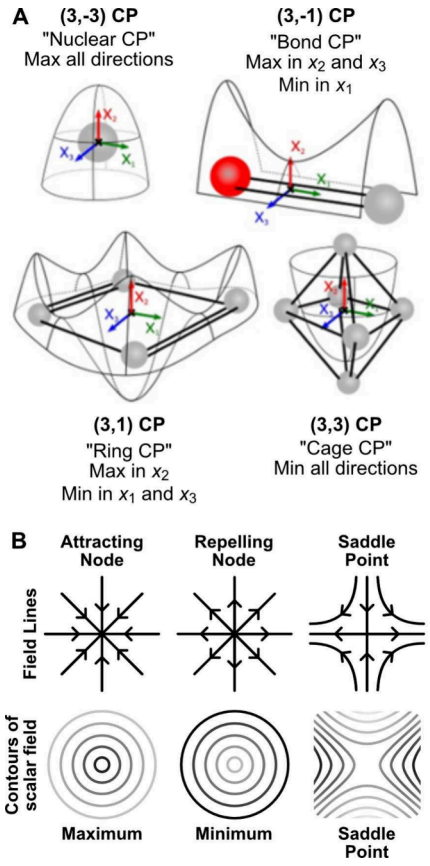


Figure 7. (A) Schematic showing the various critical points of a scalar field and their associated designation in QTAIM formalism where the scalar field is $\rho(r)$. There are four types of critical points: maxima, minima, and two kinds of saddle points. (B) Various topological features present within a two-dimensional gradient field reflecting the underlying shape of the scalar field about a critical point. A maximum acts to attract the field, a minimum acts to repel the field and a saddle point has mixed character. The field may be represented also by its isosurfaces (contours) about a critical point, which is the dual representation of the gradient vector field.

The distinction between a field's topology and its geometry, often not fully appreciated, plays a crucial role in the study of scalar fields. Topology concerns itself with the types and numbers of critical points present in the field, a foundational aspect in defining the field's basic structure. On the other hand, geometry quantifies field attributes, not only those of critical points—such as their locations, the scalar field's intensity, and principal curvatures at these points—but also including global measures that capture the field's overall shape and distribution. These geometric properties, especially when observing their variations due to perturbations, offer a comprehensive framework for analyzing a molecule's response to environmental changes. The geometry of a scalar field extends beyond the vicinity of critical points, incorporating both local and global characteristics to provide a fuller understanding of the field's relationship to molecular behavior.

In a related way, the topology of a 3D vector field, including gradients of charge density or the electric field is characterized by specific equilibrium points where the vector magnitude diminishes to zero. For electric fields, these points represent locations where a charged "test particle" would experience no net force. Within the broader categorization of equilibrium points in general 3D vector fields, six types are recognized: attracting nodes, repelling nodes, saddle points, attracting focuses, repelling focuses, and centers. However, given the electric field is also a gradient field, attention narrows to the initial trio, as the latter types are not observed for gradient fields (Figure 7B).

Attracting nodes, located at the local maxima of the underlying scalar field, draw the gradient vector field inward from all directions, suggesting a converging force landscape. In contrast, repelling nodes, representing local minima, emit a diverging vector field, indicating an outward force in all directions. Saddle points, on the other hand, exist where the scalar field exhibits a dichotomy—maximizing in one direction while minimizing perpendicularly, creating a directional flux that converges toward and diverges away from the equilibrium point. This stability dichotomy, distinguishing between stable and unstable equilibrium points, underscores the nuanced force balance at these fixed points, dictating the potential for a particle's deviation from equilibrium upon perturbation.

Gradient vector fields are typically depicted through gradient paths or streamlines, which start at minima of the core scalar field and end at its maxima. An alternative but equivalent depiction uses isosurfaces or level surfaces, such as equipotential surfaces and charge density isosurfaces, to represent the field variable (bottom row Figure 7B). Since a streamline or gradient path is always perpendicular to its corresponding set of isosurfaces, these two representations are dual to each other, conveying the same information.

The topology of vector fields, akin to scalar fields, hinges on the presence and types of fixed points. Yet, it is the geometric attributes—details like the exact locations and strengths of these fixed points, the dimensions and curvatures of streamlines or gradient paths, and the curvatures of isosurfaces—that offer deeper insights into how molecules react to external influences. This intertwined analysis of topology and geometry is foundational to QTAIM, providing a model for uncovering the complex links between molecular properties and the structure of their charge density.

2.3.2. QTAIM. QTAIM's core tenet rectifies the ambiguity inherent in defining local kinetic energy. The two commonly used kinetic energy operators—the gradient and the Laplacian forms—give different results for the kinetic energy of arbitrarily defined regions.[352,353] Bader noted, however, that for regions bounded by a surface through which the flux of the gradient of the charge density is everywhere zero, these differences vanish.[352,354] Thus, these zero flux surfaces (ZFSs) bounded regions possess definite kinetic and hence total energies. Significantly, in a molecule, every nucleus is fully contained in only one such region, delineating a Bader atom or sometimes termed a topological atom or atomic basin. With well-defined boundaries, each atom may be characterized not only with a definite energy but also with a rigorous electron count and volume. Qualitative chemical concepts such as electron flow between atoms and their associated energy change, become rigorous within the QTAIM formalism.

Each Bader atom is categorized based on the topology of its charge density, which, as mentioned, is determined by the number and kind of critical points (CPs). The QTAIM framework uses specific nomenclature for the four types of CPs, highlighting their chemical roles (Figure 7A). A local maximum, a (3,-3) CP, aligns with an atomic nucleus and is

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

termed a nuclear CP (nCP). Charge density minima, (3,3) CPs, are identified as cage CPs (cCP) and typically reside at the heart of cage-like structures. Saddle points, either (3,1) or (3,-1) CPs, are referred to as ring CPs (rCP) and bond CPs (bCP), respectively. Notably, bCPs are often found along charge density ridges that extend from one nCP to another, mirroring the atomic connections depicted in the Lewis structure model and thus are called bond paths.

Given its association with chemical bonds, much research has been directed toward establishing correlations between geometric measures of the charge density about bCPs and the properties of bonds such as their strength, stiffness, ionicity/covalency, etc. Notably, the Laplacian of the density at a bCP plays a central role in many of these correlations.

The Laplacian of the charge density at a bCP, $\nabla^2\rho_b$, is the trace of the Hessian matrix evaluated at this point. In a diagonal form, where the z-direction is taken as parallel to the bond path and the x- and y-directions are those of principal curvature perpendicular to the bond path, $\nabla^2\rho = \rho_{xx} + \rho_{yy} + \rho_{zz}$ and by definition, $\rho_{zz} > 0$ while ρ_{xx} and $\rho_{yy} < 0$. The sign and magnitude of $\nabla^2\rho_b$ has been used as an indicator of ionicity. For ionic interactions the magnitude of curvature parallel to the bond path is large, while in the perpendicular direction it is small. For covalent bonds the opposite is found. Hence, $\nabla^2\rho_b > 0$ is often said to be indicative of ionic bonding, while $\nabla^2\rho_b < 0$ is consistent with covalent interactions. Metallic bonds are intermediate with $\nabla^2\rho_b \approx 0$.[352]

Another set of geometric features associated with the bCP is its directionalities, which are defined as $\sqrt{\left|\frac{\rho_{xx}}{\rho_{zz}}\right|}$ and $\sqrt{\left|\frac{\rho_{yy}}{\rho_{zz}}\right|}$.[355−358] Geometrically these values give the tangents of the characteristic angles θ and ϕ of Figure 8 and recover the shape of the charge density isosurface passing through a bCP.

In addition to these purely geometric parameters there are several calculated parameters at the bCP that have been used to characterize bonding. Of these the local bCP kinetic (G_b) and potential (V_b) energies figure prominently. For a stable nuclear configuration via the virial theorem one can show:[352]

$$\frac{1}{4}\nabla^2\rho_b = 2G_b + V_b \tag{15}$$

and naturally one can define a local total energy $H_b = G_b + V_b$. As an aside, defining a local kinetic energy and hence total energy at a single point e.g. the bCP is antithetical to the QTAIM canons as a point is not bonded by a ZFS. Nonetheless, various functional forms for G_b in terms of ρ_b have been proposed.[359] These quantities have been argued to be indicative of the degree of covalency,[360] particularly as related to hydrogen bonds.[361−366] When both $\nabla^2\rho_b$ and $H_b < 0$, bonds are considered covalent, while when both are positive the bond is noncovalent. The mixed case where $\nabla^2\rho_b > 0$ but $H_b < 0$ is considered partially covalent. Thus, a negative value of H_b has been implicated as an indicator covalency. However, working with a more diverse set of calculation, this greater covalency[366] was argued to be the result of increased electrostatic contribution to the interaction energy—in a loose sense echoing the EVB paradigm.

It is worth mentioning that quantities such as $\nabla^2\rho_b$ offer a precise description of the local charge density geometry. Efforts to link this exact measure with the more abstract concepts of chemical bonding—such as covalency, ionicity, and metallicity—inevitably encounter ambiguity. Research grounded in QTAIM benefits from identifying correlations between specific charge density geometrical descriptors and quantifiable molecular properties.

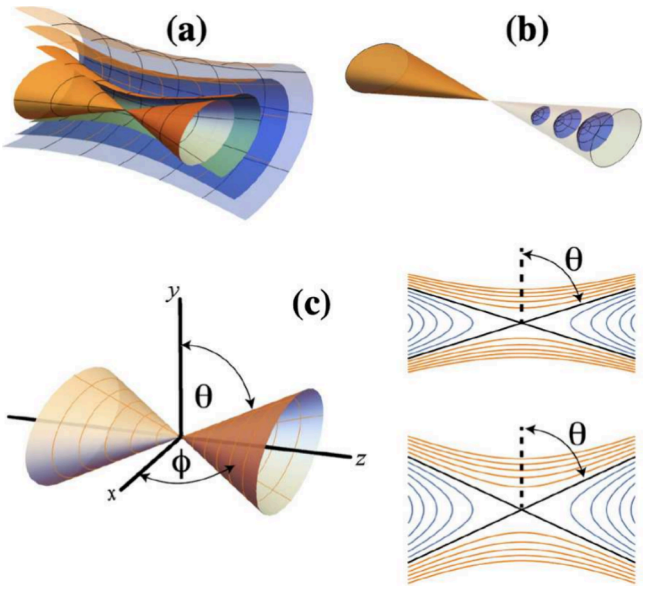


Figure 8. Isosurfaces near a bCP. The isosurface passing through the bCP will have the form of an elliptic cone, with the bond path coincident with its axis. This cone is the asymptotic boundary separating the exterior isosurfaces (a) from those interior to the cone (b). (c) The cone is fully characterized by the characteristic angles θ and ϕ. As the angle θ decreases, the hyperbolic region (orange) becomes more curved, while the convex region contours (blue) become less curved. Reproduced with permission from Wilson et al., J Phys Chem A, 2021, 125 (50), 10622−10631. Copyright 2021 American Chemical Society.[357]

In addition to the properties inferred from bCPs are those characterizing Bader atoms, particularly its electron count. Taking Ω_i to be the atomic basin (volume bounded by a ZFS) of Bader atom i with atomic number Z_i, then the number of electrons in Ω_i is given by the integral of the density over the atomic basin, i.e.,

$$N(\Omega_i) = \int_{\Omega_i} \rho(\boldsymbol{r})d\boldsymbol{r} \tag{16}$$

and the charge of the i^{th} atomic basin is $q(\Omega_i) = Z_i - N(\Omega_i)$.

Other useful atomic properties are its localization and delocalization indices, $\Lambda(\Omega)$ and $\delta(\Omega, \Omega')$ respectively. A discussion and derivation of these indices is given by Fradera et al.[367] Conceptually, $\delta(\Omega, \Omega')$ gives the number of electrons shared between atom Ω and Ω', one-half this quantity summed over all atoms sharing an interatomic (ZFS) surface with atom Ω gives its number of shared electrons. The number of unshared (localized) electrons is given by $\Lambda(\Omega)$. Obviously the sum of these two quantities gives $N(\Omega)$.

Recent extensions to QTAIM have introduced a further partitioning of Bader atoms into regions enclosed by zero-flux surfaces (ZFSs), known as gradient bundles (GBs).[370,371] Like individual Bader atoms, each GB may be characterized by its distinct electron count, energy, and related properties. Notably, certain gradient bundles have been identified as the volumes occupied by chemical bonds.[371−373] Bond bundles can be further decomposed into two bond wedges, which are the volumes given by the intersection of the bond bundle and the Bader atoms of the bound pair (Figure 9). Bond bundles and

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

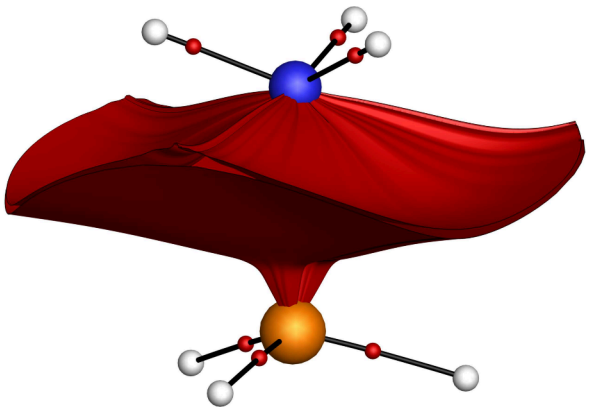


Figure 9. Bounding surfaces of the N−B bond bundle in borane-ammonia; the union of one N bond wedge and one B bond wedge, and containing the shared portion of the N−B interatomic surface.[354,368,369]

bond wedges may be unambiguously described by their electron count, energy, volume, and other properties. For every Bader atom property there is a corresponding bond bundle/bond wedge property. Bond bundle energies and electron counts have been shown to be consistent with expectations derived from intuitive chemical concepts.[374] Thus, while traditional QTAIM lays the groundwork for quantifying charge and energy redistributions due to any perturbation, the enhanced QTAIM framework provides a more detailed account of this charge redistribution and its energy changes between bonds.

2.3.3. Applications of QTAIM to Enzymes. A set of papers reported the use of QTAIM to gauge the effect of fields produced by an enzyme's extended structure on active site charge density and to correlate these effects with properties. These studies serve to emphasize that QTAIM parameters provide sensitive quantitative probes of electronic structure.

The first of these studies was conducted in the context of computational metalloenzyme redesign with carboxypeptidase A (CPA) and its several mutants serving as a test case.[120] Specifically, a different sequence of the peptide was the target substrate (terminal Phe was replaced with Asp), and the pocket had to be redesigned to accommodate the change of the peptide charge, which was achieved with the V243R mutation of CPA (Figure 10). Valdez et al. used QM/DMD to determine the structures and charge densities of native and mutant CPA active sites with the bound substrates, and then compared the QTAIM parameters of the resulting charge densities and correlated their similarity to the transition state charge density, which they argued should correlate with enzyme efficiencies.

Bond paths, bond bundles, location and charges of CPs, and Bader atom charges were calculated for the active site region of both the native and the V243R_FpepD mutant enzyme. Some of this detail is shown in Figure 11. The variation in several of these parameters was deemed to be indicative of greater efficiency of the native enzyme, while both the native CPA and the mutant featured geometrically similar, reactive conformations with the respective bound substrates. For example: the more negative Bader charge on the water oxygen was argued to make the water molecule a better nucleophile in the native enzyme compared to V243R_FpepD; and the increased charge density at the Zn−O1 bond CP and the size of the Zn−O1 bond bundle of the native enzyme was argued to be indicative of the greater stabilization of the carbonyl oxygen.

As part of this research Valdez et al. performed a mechanistic study identifying the structure and charge density along the reaction path, again for the native and a mutant enzyme. They identified a topological change inherent to the reaction in which a ring opens through the merging of a rCP with a bCP. The charge difference between these two critical points serves as a measure of the transition state energy, which was substantially less for the native enzyme. Essentially this is an appeal to the Hammond postulate,[375] asserting that for the native enzyme the charge density is closer to the transition state than it is in the mutant. An important takeaway from this study is that through QTAIM a quantitative measure can be associated with the historically qualitative Hammond postulate.

The other two studies in this series focused on Histone Deacetylase 8 (HDAC 8). Traditionally a Zn^{2+} ion was thought to be essential to its activity. However, there is experimental evidence suggesting HDAC8 is catalytically active with a variety of divalent metal ions. In a theoretical study intended to shed light on this possibility, Nechay et al.[376] used mixed quantum-classical QM/DMD methods to construct a comparatively large cluster models of the active site with Fe^{2+}, Co^{2+}, Mn^{2+}, Ni^{2+}, Mg^{2+}, or Zn^{2+} as the divalent ion. The charge density was determined as part of these calculations and subsequently its evolution through the reaction was followed.

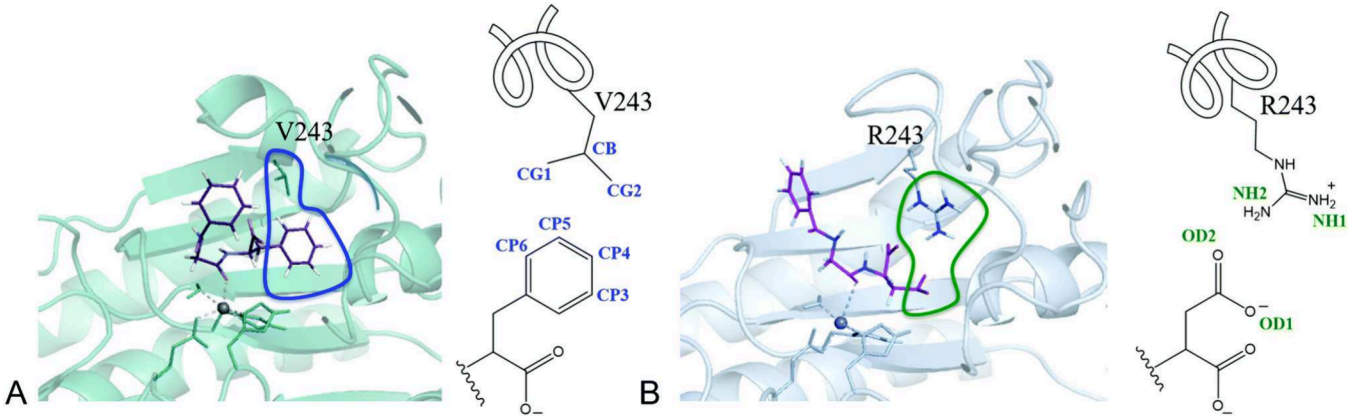


Figure 10. (A) Native CPA and (B) the V243R mutant with their bound substrates. Outlines indicate the regions of the binding site defining the specificity of the CPA to a given substrate. Reproduced from ref 120 with permission from the Royal Society of Chemistry.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

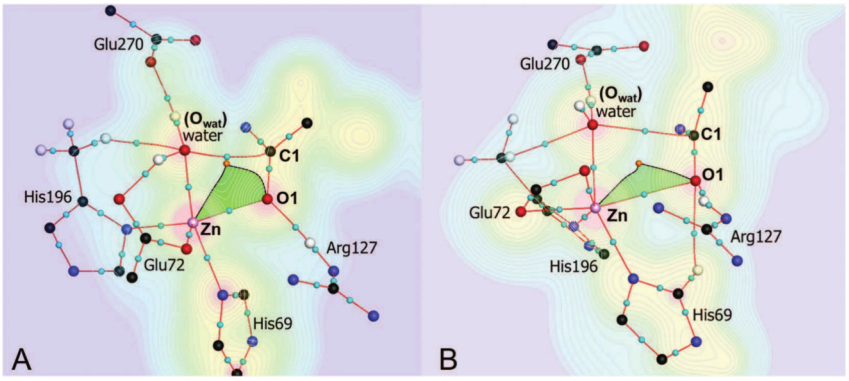


Figure 11. Bond paths of interest in the native (A) and V243R$_{FpepD}$ mutant enzyme (B), in the reactant state. Contours in $\rho(r)$ are drawn on a cut plane on a logarithmic scale from 10^{-3} to 1e ·Bohr −3. Red lines indicate bond paths. The pictured portion of the Zn−O1 bond bundle is shaded green with black lines showing approximate edges. The following coloring scheme is used: Zn-purple, O-red, C-black, H-white, N-blue, bond CP-cyan, and ring CP-orange. Reproduced from ref 120 with permission from the Royal Society of Chemistry.

Calculated QTAIM parameters were then used to understand the different binding affinities for each metal as well as their abilities to bind and orient the substrate for deacetylation. And once again, QTAIM made it possible to quantitatively assess the nearness of the reactant and transition state charge densities.

A crucial step in one of the investigated mechanisms involves proton transfer between a water molecule and H143, see Figure 12. For this mechanism the activation energy was found to be smaller when the active site contained Zn^{2+} compared to Mn^{2+}. A rationale can be found in the relative topologies of the active site models depicted in Figure ME3. Nechay et al. found that when Zn^{2+} is present, a bCP and corresponding bond path forms between the water oxygen and the carbon atom of the substrate carbonyl. This requires a topologically necessary rCP. When Mn is present, there is no water to substrate carbonyl bCP or necessary bond path and rCP. Ni^{2+}, Fe^{2+}, and Co^{2+}, give the same topology as Zn^{2+}, while Mg^{2+}, has the same topology as Mn^{2+}. As the reaction involves proton transfer and accompanying formation of a bond path, the Zn^{2+} charge density is closer to that of the transition state than the Mn^{2+} charge density. Hence, one would expect a lower activation energy for this reaction step.

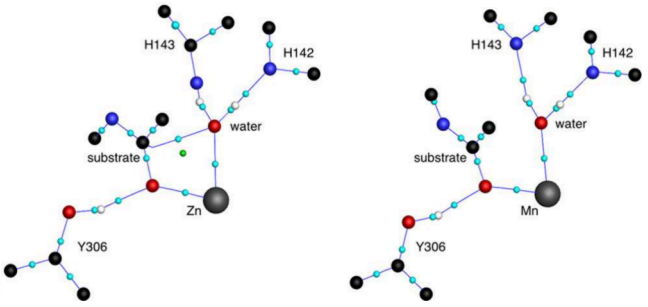


Figure 12. Critical points and bond paths of interest in the active site of HDAC8 with Zn (left) and Mn (right). Sphere coloring is as follows: C-black, N-blue, O-red, metal-gray, bond CP-cyan, ring CP-green. Reproduced with permission from Nechay et al., J Phys Chem B, 2016, 120 (26), 5884−5895. Copyright 2016 American Chemical Society.[376]

To assess the effects of an enzyme's extend structure on active site LEFs Morgenstern et al. followed the convergence of QTAIM parameters as the number of surrounding amino acid residues was increased.[122] They performed DFT calculations of progressively larger active site models of HDAC8 consisting of a central Zn^{2+} ion and successively more of the surrounding environment as modeled with cluster of approximately 3, 4, 5, 6, and 7 Å (Figure 13).

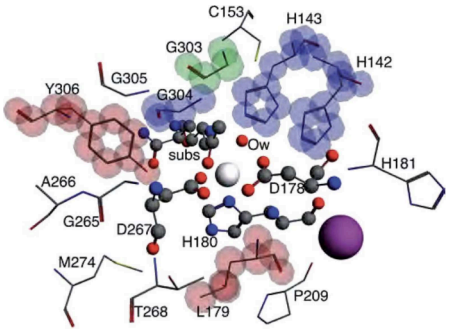


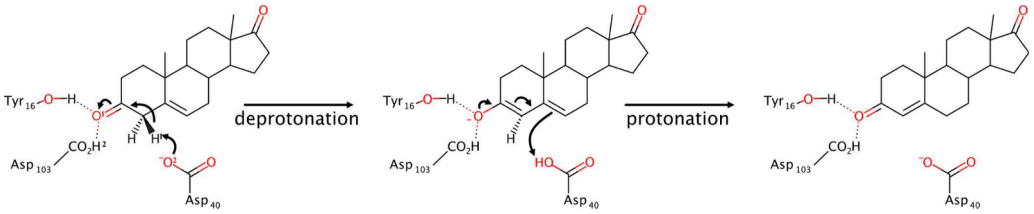
Figure 13. Structure about the active site of HDAC8 as modeled with clusters of 3−7 Å. The 3 Å system consists of the central Zn^{2+} ion (white sphere), substrate (subs), water molecule (Ow), D178, H180, and D267, all shown in ball and stick model. The 4 Å system adds L179 and Y306, shown with stick model highlighted in red. Five Å adds H142, H143, and G304, stick model highlighted in blue. Six Å includes G303, highlighted in green. The largest 7 Å system includes a K^+ ion (purple sphere) and C153, H181, P209, G265, A266, T268, M274, and G305, shown with nonhighlighted stick model. Reproduced from ref 122 with permission from the Royal Society of Chemistry.

They found that the magnitude of the charge density at critical points and Bader atom charges converged once the immediately adjacent residues around the point or atom of interest were included in the modeled environment. Further including a dielectric constant or point charges to the calculations had only a small effect on atomic charges but did not change the converged value of the critical point charge densities. In contrast, the locations of critical points were affected by the extended environment of the protein at all cluster sizes and converged only after virtually the full structure of the protein and solvent were included in the model.

These findings indicate that the locations of critical points are influenced by dipole moments from even distant residues. Since critical point positions had been demonstrated to correlate with reactivity and reaction barriers,[120,376] the authors

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

Scheme 1. KSI Catalyzed Reaction



argued that it is the full, extended structure of an enzyme that mediates its reactivity. Thus, the positions and curvatures of the charge density at CPs could be used to optimize computationally designed enzymes.

The work of Yang et al. provides a vivid example as to ways in which QTAIM parameters can be useful in elucidating enzymatic reaction mechanisms.[377] The specific interest was to assess the role of hydrogen bonding in Methyl transferases (MTases) where enzymatic efficiency has been attributed to electrostatic- and charge-transfer-driven stabilization of the transition state or alternatively to hydrogen bonding induced changes to the active site that enhance substrate binding and catalysis. Using QM/MM steered-molecular dynamics (SMD) simulations the authors investigated four distinct MTases: Cap Methyltransferase 1 (CMTr1), Protein L-Isoaspartate (D-Aspartate) O-Methyltransferase (PIMT), *Plasmodium falciparum* Phospho-ethanolamine Methyltransferase (PfPMT), and C-Methyltransferase (HcgC). These four enzymes have diverse regulatory and synthetic functions.

Making the questionable but nonetheless common assumption that bCP values of local energy could be used to determine bond energies, the energetic contributions from the hydrogen bonds, CH ···X (X = N, O), along the reaction coordinate were obtained by estimating the hydrogen bond energy as one-half bCP potential energy density. While the hydrogen bond contributions to the stabilization of the transition state were not constant across the four MTases, in some cases and at some points in the reaction path they were significant, for individual hydrogen bonds over a narrow reaction coordinate interval reaching values as large as 8 kcal/mol. Summed hydrogen bond energies reached maximum values as high at 18 kcal/mol, though the values obtained when averaging over the reaction coordinate were less than 5 kcal/mol.

Notably, this investigation did not reveal a universal stabilizing effect of hydrogen bonds across all MTases. While in N-type PfPMT and C-type HcgC hydrogen bonding was assessed as stabilizing the transition state by 2−5 kcal/mol, no or limited stabilization was found for CMTr1 and PIMT. These findings motivated further calculations exploring the role of charge transfer and electrostatics in mediating the reactivity of the MTases investigated. The authors ultimately concluded that the intrinsic reactivity of MTases was due to these latter factors and not hydrogen bonding per se.

Sowlati-Hashjin et al. took a significant step in expanding the number of QTAIM parameters that are useful in describing the response of the charge density to an external electric field (EEF) in a study of diatomic molecules under the influence of an EEF.[378] EEFs were found to stabilize homodiatomic molecules, though heteronuclear diatomic molecules were found to be stabilized or destabilized depending on the direction of the field relative to the molecule's dipole moment.

For both homonuclear and heteronuclear diatomic molecules, EEF were found to alter the indicators of ionic/covalent character at the bCP as well as the electron localization and delocalization indices of the associated atomic basins. For homonuclear diatomic molecules the EEF was observed to decrease the electron density at the bCP, and, as expected, induced polar character (indicated by changes to the localization and delocalization indices) with accompanying induced curvature of the interatomic surface and associated changes to the atomic volumes, atomic energies, and atomic populations. Field induced changes to ionic character seem to be a general effect across molecular systems. Shaik et al. have noted the same phenomenon in enzymes using measures of ionic/covalent character that are distinct from those used in QTAIM studies.[379] These finding are consistent with the induced ionic resonance forms of the EVB theory and serves to emphasize the general applicability of QTAIM approaches.

Sowlati-Hashjin et al. also demonstrated that most atomic properties, e.g. volumes, energies, and electron counts varied linearly with field strength, while bCP and bond path properties, such as electron density at the bCP and bond length, varied nonlinearly. Delocalization index showed mixed behavior, correlating linearly with the magnitude of the EEF for homonuclear diatomic molecules and nonlinearly for heteronuclear diatomic molecules.

The field effects found by Sowlati-Hashjin et al. for diatomic molecules are applicable generally as a series of investigation of KSI demonstrates. KSI is a well-studied enzyme that catalyzes the repositioning of a double C=C bond in the steroid substrate as shown in Scheme 1.

Freindorf et al. performed a comprehensive investigation of steroid isomerization by KSI. Among the many questions of interest was the role of hydrogen bonding in transition state stabilization.[380] As part of this investigation the energy density, H_b, at crucial bCPs along the reaction coordinate was calculated and compared with other methods of estimating bond covalent character and bond strength. They found that H_b did not provide a reliable measure of bond strength where the environment, such as solvent effects, is significant. This observation was based on poor correlations between H_b and the calculated value of the stretching mode force constant, k^a, for O−H interactions. On the other hand, the correlation was found to be quite good for the less environmentally sensitive C−H interactions. The authors correctly point out that H_b is a property evaluated at a single point along the bond path, whereas the local mode force constant, as a second-order property, sensitively captures the environment between the two atoms forming the bond or interaction under consideration.[380]

However, the assumption that H_b, if providing a reliable measure of "bond strength," would correlate with the stretching mode force constant is questionable. At best, one would expect H_b to correlate with bond energy, related to the

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

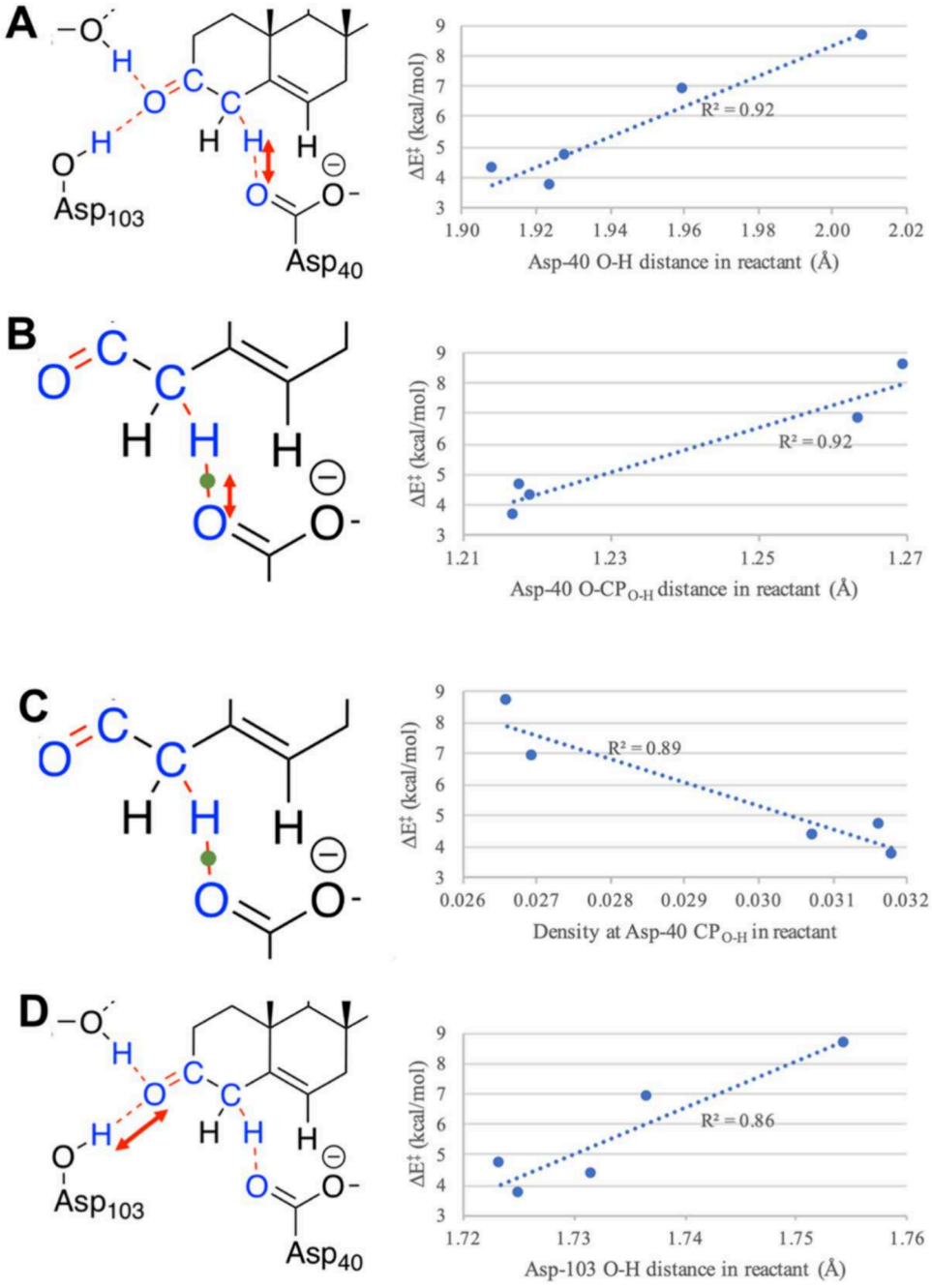


Figure 14. Geometric and QTAIM parameters found to exhibit the best correlations with the computed reaction barriers for the first step of the reaction at varying external electric fields: (A) the Asp-40 O−H distance, (B) the distance between the Asp-40 CPO−H and the Asp-40 O atom, (C) the charge density at the Asp-40 CPO−H, and (D) the Asp-103 O−H distance. Reproduced with permission from Fuller et al., J Chem Inf Model 2019, 59 (5), 2367−2373. Copyright 2019 American Chemical Society.[123]

potential well depth, not with a force constant, related to the potential well curvature. Others have shown that force constants are related to the bond directionality (Figure 8).[356]

In another investigation of KSI, Fuller et al. determined that geometric features of the enzyme active site charge density served as quantum mechanically rigorous probes of LEFs.[123] Field effects were calculated by subjecting a large, isolated cluster representing KSI's active site with the first coordination sphere residues, to external electric fields of 10 MV/cm with varying directions. These fields were assumed to, at least locally, model the electric fields produced by the enzyme's structure beyond the active site. A field applied parallel to the substrate carbonyl bond, pointing from O to C, was calculated to lower the reaction barrier to deprotonation and raise the barrier to protonation, while a field in the opposite direction had the reverse effect.

A systematic search was conducted to find geometric features of the charge density that correlated with the calculated changes to the reaction barriers under the various applied fields. It was noted that the strongest correlations were not with the charge density geometry around the activated carbonyl group but around Asp40, where the H-transfer part of the reaction takes place, Figure 14. The authors noted that, DFT calculations indicate that the charge density in the active

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

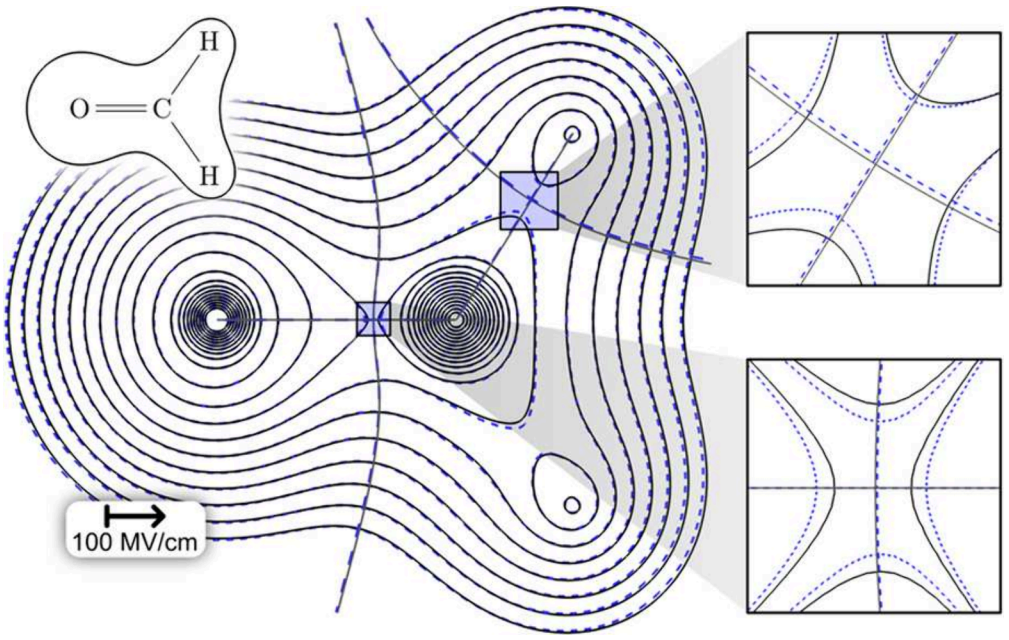


Figure 15. Contours of $\rho(r)$ in the molecular plane of formaldehyde with and without a 100 MV/cm uniform electric field applied along the C=O internuclear axis, pointing from O to C. The C−H and C=O bond saddle points (top and bottom right, respectively.) are shown in more detail with bold lines designating the interatomic surface and the lighter line designating the internuclear axis (bond path). The unperturbed charge density is shown with black contours, and the field-induced density is shown with dashed blue contours. Reproduced with permission from Wilson et al., J Phys Chem A, 2021, 125 (50), 10622−10631. Copyright 2021 American Chemical Society.[357]

site of KSI is highly responsive to minor changes in the external electric field, and this responsiveness both reflects and predicts the reaction barrier. This finding invited a series of questions, importantly: How are distinct and sometimes distant charge density critical points correlated? Building on the results reported in ref 123, Wilson et al. embarked on a more detailed analysis of the global response of the charge density to external electric fields.[357] Rather than focus on the charge density's critical points, the authors considered the shape of the density, as reflected by its isosurfaces about critical points.

Figure 8 illustrates that the shape of isosurfaces about a bCP is controlled by its characteristic angles. To study EEF effects on these angles, Wilson et al. initially used formaldehyde as a surrogate system for the field induced carbonyl activation of KSI. The intersection of formaldehyde's isosurfaces with the molecular plane are depicted in Figure 15, with and without an EEF along the C=O internuclear axis.

The shape of the contours about the C−H and C=O bCPs are as represented in Figure 8. However, it is important to keep in mind that charge density contours are automatically closed loops and their corresponding isosurfaces are closed as well. As such, the integral of the contour (isosurface) curvature around the loop (surface) must be 2π (4π). Thus, it is mathematically required that perturbations altering the curvature of a contour at one point must be offset by changes at other points along the contour. The same principle holds for isosurfaces.

In the case of formaldehyde, the effect is visible in Figure 15. The EEF alters the bCP isosurface curvatures—measured by its characteristic angles. These changes are offset by compensatory changes to contour and isosurface curvatures along the distal segments of the contours.

Regions where a perturbation leads to an increase in contour curvature correspond to the same regions where there is a decrease in charge density, and the opposite is true for regions where curvature decreases. This relationship suggests that areas of curvature increase and decrease are interconnected both mathematically and physically. Mathematically, they are linked by the principle of conserved isosurface curvature, and physically, they are connected through the conserved number of electrons within a molecule.

Building on this concept, a novel parameter was introduced to the QTAIM framework—the total isosurface curvature enclosed within a region bounded by zero-flux surfaces (ZFSs). This parameter is proportional to the region's volume. Specifically, a region will have a larger volume if it encompasses isosurfaces that are more positively curved.

Field induced carbonyl activation of formaldehyde was found to be associated with reduced negative isosurface curvature and hence a reduction in volume of the C=O bond bundle, which is offset, as required by geometric principles, by an increase in positive isosurface curvature and thus larger volumes available to the O lone pairs.

With these findings in hand, Wilson et al. calculated the volume distributions around the carbonyl of KSI and found that the complex geometry around the Tyr16 residue and the substrate carbonyl O atom (See Scheme 1) to be responsible for the curvature distribution activating the carbonyl. Specifically, the hydrogen bond between the carbonyl O and Tyr16 is part of a hexagonal ring as is evidenced by a rCP. Isosurfaces above and below a ring are concave and necessarily negatively curved. Thus, for KSI, allowing for a greater positive curvature and greater volume for the carbonyl O to Tyr16 hydrogen bond bundle, in turn decreases the volume of the C=O bond bundle. Crucial to this curvature distribution is the position of the Tyr16 H atom. Changes on the order of few hundredths of an angstrom to its location alter the curvature distribution substantially. This study illustrates the complex, interconnected structure of the charge density—changes in one part of the molecule are propagated to other molecular

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

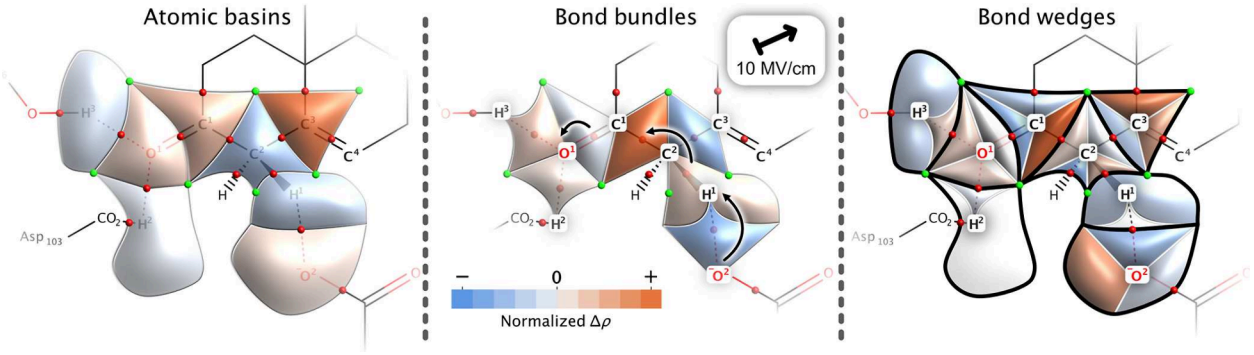


Figure 16. Atomic basins, bond bundles, and bond wedges of KSI (left, middle, and right, respectively) shaded according to the changes in their electron count due to a 10 MV/cm EEF oriented parallel to the O^1-C^1 internuclear axis. The center image includes the electron-pushing arrows of the deprotonation reaction step. Reproduced with permission from Wilson et al., J Phys Chem B 2022, 126 (46), 9443−9456. Copyright 2022 American Chemical Society.[381]

regions. This fact is neither surprising nor new. What is novel, is that this propagation is governed by mathematical principles that may be quantified with an extended set of QTAIM parameters.

The relationship between isosurface curvature (approximated by bond bundle volumes) and field/structure-produced activation of carbonyl motivated a fuller bond bundle analysis of five KSI enzymes of known and varying catalytic activity. The five enzymes including two mutants and the native enzyme were subjected to strategically directed external electric fields.[381]

The volumes, total energies, electron counts, and a shape parameter were calculated for the Bader atoms, bond bundles, and bond wedges for each of the five active site models and compared to the zero-field native enzyme. Graphical representations of the perturbation produced changes proved particularly informative. An example is reproduced in Figure 16.

This figure shows the electron redistribution due to an EEF of 10 MV/cm oriented parallel to the O^1-C^1 internuclear axis. The additional information afforded by bond bundle analysis can be illustrated by considering, as an example, the change to the electron density of C^1. The EEF increases the electron count of this atom. However, this total change results from a large increase in the C^1-C^2 bond bundle and a smaller to negligible decrease in the density contained in the O^1-C^1 bond bundle. In turn, the density increase in the C^1-C^2 bond bundle is due to a redistribution within the C^1 Bader atom, as is clear from an inspection bond wedge resolved electron redistribution. The changes within the Bader atoms, bond bundles, bond wedges and the other shape parameters were then correlated with catalytic activity.

The researchers discovered that catalytic enhancement resulted from promoting both inter- and intra-atomic electron density redistribution in the forward direction of the catalyzed reaction. Though the redistribution applies to both types of perturbed systems (mutants and EEFs) the authors observed that bond properties (e.g., volume, energy, electron count) can respond independently and disproportionately depending on the type of perturbation. The findings suggest that catalytic enhancement or inhibition occurs through distinct pathways, with certain bond properties being more significantly affected by one perturbation type over another.

The exploration of gradient bundle analysis has led to the discovery of numerous geometric characteristics that could potentially function as vectors for machine learning, with the aim of predicting activation energies. This concept is bolstered by the findings of Vargas et al.,[392] who demonstrated that QTAIM parameters—such as density, ellipticity, electrostatic potential, and localization indices for various CPs and atomic basins of the reactant state—could be leveraged in supervised machine learning to forecast reaction barrier energies. These QTAIM parameters, which are derived from the observable total charge density, align with Bader's original vision of being both computationally and experimentally attainable.

In a practical application of this theory, Vargas et al. conducted a study on the reaction barriers of a wide range of Diels−Alder reactions, compiling a comprehensive data set of electron density and related mathematical descriptors for the reactants. This data set was then streamlined using feature selection techniques to isolate a set of critical variables that reflect underlying physical principles. Utilizing these variables, the team successfully developed several regression models with strong predictive capabilities based on physical descriptors. Moreover, they were able to qualitatively forecast the activity sequence for three Diels−Alderase enzymes, demonstrating that, by focusing solely on the reactant state, there was no need to identify the transition state geometry to estimate relative TS energy. This suggests that the extensive array of QTAIM parameters uncovered through gradient bundle analysis holds promising potential for enhancing machine learning methodologies.

2.4. 3D Electric Fields

The aforementioned findings through multiple points of analysis like QTAIM, bond/gradient bundles, and EVB all suggest that different electric fields might promote or inhibit catalysis through different mechanisms. Analysis of reactivity through the lens of electrostatic preorganization necessitates a shift beyond electric field analysis at a single point, emphasizing the need to compare the global structure of electric fields. Previous analysis techniques with QTAIM have demonstrated the importance of the electron density and its topology as a descriptor of reactivity, and have emphasized the 3-dimensional nature of the chemical bond.[123,354,372,382] When looking at chemistry that enzymes employ, methods must assess the degree to which electrostatic preorganization influences the 3D charge density that drives chemical reactivity. Prior studies compute electric fields at individual points or even project electric field components on points along interatomic axes,[115,130,383] but further evidence has shown that this metric

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

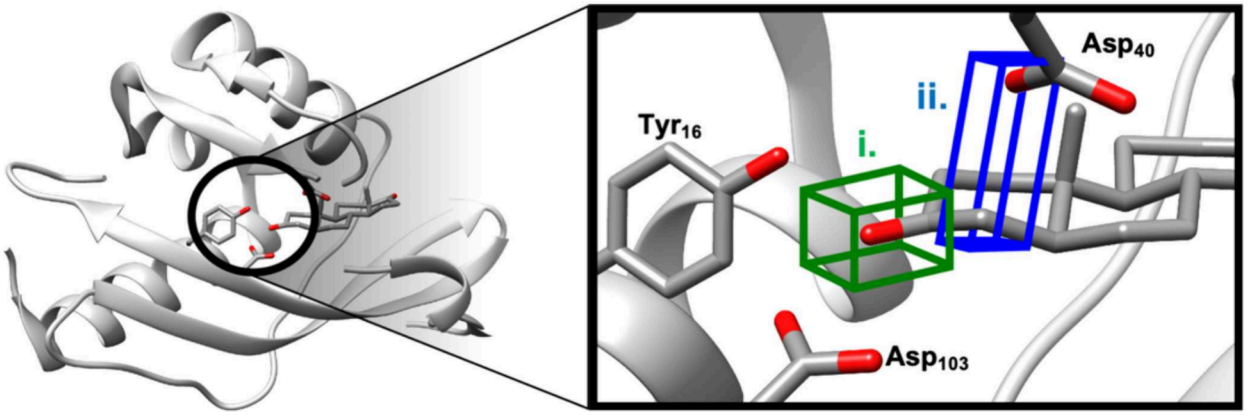


Figure 17. Regions (i) and (ii) within KSI that were analyzed via the global distribution of streamlines. Reproduced with permission from Hennefarth and Alexandrova, ACS Catal, 2020, 10 (17), 9915−9924. Copyright 2020 American Chemical Society.[384]

may not be an adequate descriptor for the heterogeneous effects that protein scaffolds can exert on their active sites. We have demonstrated the utility of considering the 3-dimensional geometry of electric fields when studying the reactivity of KSI, diels-alderase, evolved protoglobin, and across natural heme protein families.[124,125,168,384,385]

The heterogeneity of local 3D fields at enzyme active sites necessitates the ability to compare vector fields and evaluate their similarity. Methods to compare and analyze vector fields can generally be categorized into those that use single or a set of local descriptors, and those that use global descriptors for vector fields. Local descriptors employ analysis of singularities. Singularities can be readily generalized to three dimensions, and allows for structures like spiral saddles and 3D orbits. The direct classification of critical points has been explored by analysis of Jacobian eigenvalues and eigenvectors, and employs techniques like polynomial interpolation, vector field convolutions, and extending attachment/detachment nodes, to name a few.[386−389] Local descriptors and connections between them, however, are not always sufficient to uniquely identify and compare 3D vector fields, due to limitations in their construction from and around singularities rather than knowledge of the entire 3D field. Vector field flows within the active site are, on the other hand, rich, complex structures that include, but are not limited to, singularities. These complex objects determine chemical reactivity, and thus, a global descriptor that can take in an entire vector field as an input is more desirable for analyzing enzyme electric fields.

The primary method we use for describing global topological characteristics of 3D fields is a distribution of streamlines method based on the work from H. Quynh Dinh and Liefei Xu.[390] The process starts by randomly selecting points within these regions and generating the intersecting streamlines. Next, two points along the same streamline are chosen at random, and both the Euclidean distance and mean curvature (eq 17) between them are calculated. Here, $\alpha(t)$ is a parametric representation of a streamline, and the curvature κ employs the first and second derivatives of this representation. With a sufficiently thorough sampling, these measurements provide insights into the curvature distribution of the vector field. This distribution is statistically analyzed through a 2D histogram plotting Euclidean distance against curvature, using a bin width tailored to the data's distribution. These 2D histograms can be compared using a normalized χ^2 distance metric (eq 18). The χ^2 distance metric D compares any two histograms f and g, each with N bins, and if f and g are normalized, $D(f, g)$ ranges between $[0,1]$. $D(f, g) = 0$ indicates perfectly matched histograms, and $D(f, g) = 1$ indicates maximally dissimilar histograms (no overlapping bins).

$$\kappa = \frac{|\alpha'(t) \times \alpha''(t)|}{|\alpha'(t)|^3} \tag{17}$$

$$\chi 2\text{: } D(f, g) = \frac{1}{2}\sum_{i=1}^{N} \frac{(f[i] - g[i])^2}{f[i] + g[i]} \tag{18}$$

A benefit of this method over others for comparing fields is its ability to differentiate between and within sets of simple and complex fields. Dinh and Xu demonstrated this by showing this for sets of fluid flow data.[390] Although simpler methods were able to differentiate between a set of simple and a set of complex fields, within each set, there was no clear distinction. Upon employing a distance-curvature distribution, they found that each histogram was a more unique descriptor of the field.

This method was first employed to examine the effect of the electric field's global structure on the properties of KSI and its variants.[384] This technique entails sampling characteristics of the electric field across regions of chemical interest, with an initial focus on areas of KSI shown in Figure 17, the regions denoted i and ii corresponding to a carbonyl moiety and a hydrogen bonded region involving Asp40 O−H−C interaction respectively. The choice of the volumes was motivated by the fact that activation of the carbonyl is implicated in the reactivity, and H-transfer to and from Asp40 is involved in the two reaction steps. For the KSI analysis, the method involved sampling 100,000 streamlines per region and binning these points in a 200 × 200 histogram.

The measure's effectiveness is demonstrated in Figure 18, comparing the electric field characteristics of the wild type KSI under six varied electric fields and two KSI variants across the regions i and ii. The intensity of the gray shade corresponds to the level of similarity: the darker it is, the higher the similarity. For instance, wild type KSI shows different electric field characteristics in region i, particularly around the carbonyl, compared to its behavior under the g+ field, yet exhibits closer similarity in region ii, around the Asp40 O−H−C interaction. Additionally, Figure 18's right side plots the reaction barrier differences, $\Delta\Delta E^{\ddagger}$, against the histogram distances, $D(f, g)$, revealing a notable correlation between the electric field influence in region i and the reaction barrier sensitivity.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

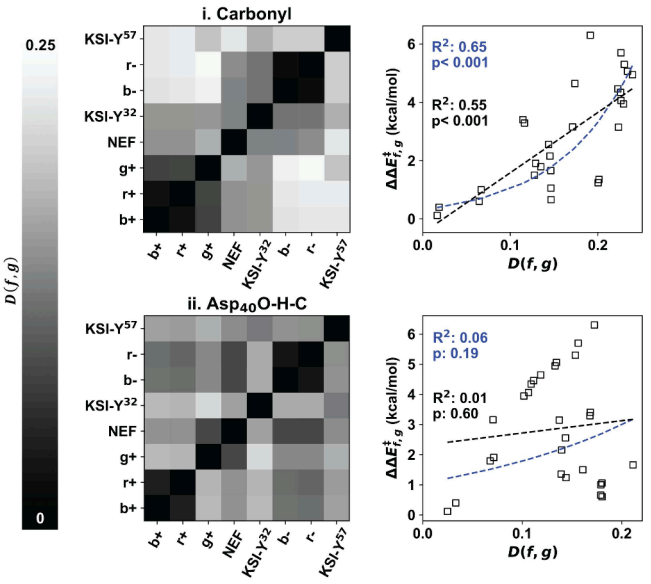


Figure 18. Dissimilarity measurement between systems with different applied external electric fields. (i) analysis of the region around the carbonyl and (ii) region around the Asp40 O−H−C region. In the dissimilarity matrices lighter shades of gray indicate greater dissimilarity. Graphs on the right compare the change in the reaction barrier between two systems ($\Delta\Delta E^{\ddagger}$) to the distance between their histograms D(f, g). Exponential and linear fits are shown in blue and black, respectively. R and p values are shown with each graph. Reproduced with permission from Hennefarth and Alexandrova, ACS Catal, 2020, 10 (17), 9915−9924. Copyright 2020 American Chemical Society.[384]

The streamline distribution method demonstrates resilience against variations in the bounding box size for chemically relevant areas in this case. Its versatility allows for sampling a diverse set of vector field attributes, including but not limited to streamline length, local field strength, torsion and curvature, as well as their separation and convergence patterns. The method's effectiveness in capturing the global geometry of the field is attributed to the extensive area covered during sampling. This technique has been efficiently parallelized to analyze the distributions of well over 10,000 streamlines, facilitating detailed studies of vector field characteristics.

The generation of distance matrices as shown in Figure 18 makes this metric a starting point for clustering molecular dynamics trajectories. This method was extended to molecular dynamics trajectories for two systems. First, it was employed for the directed evolution of protoglobin by Chaturvedi et al.[168] The 3D electric field dissimilarity matrices were used for clustering and obtaining structures with representative fields out of trajectories. Representative structures were treated with QM/MM, to demonstrate the link between the active site electric field and experimentally observed reactivity (Figure 19A). Here, the bounding box of the electric field calculation was centered on the Fe−carbene bond. Vargas et al. used the method similarly to obtain representative structures for QM/MM spin density calculations and redox potential across heme families.[385] For the heme families, the field was centered around the Fe. For both studies, the box sizes are parametrized to provide distinguishable electric field clusters, and oriented by the porphyrin nitrogen to have a stable electric field box across the molecular dynamics simulations.

Dimensionality reduction methods, when used in tandem with methods to obtain global features of the vector field, can

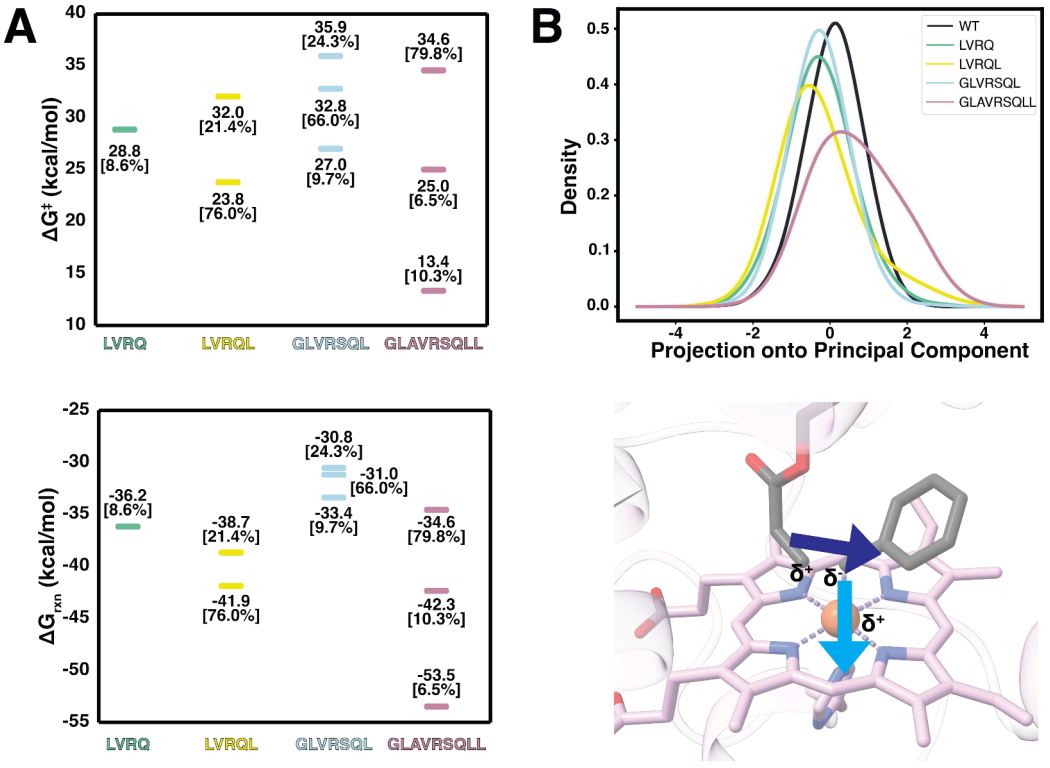


Figure 19. (A) Improved reactivity observed from both reaction stabilization (bottom) and decrease in barrier over directed evolution path. QM/MM calculations were done on frames obtained from electric field clustering. (B) PCA reveals an electric field present prominently in the final variant (top) that corresponds to stabilization of the transition state (bottom).

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

provide chemical intuition to the computed LEFs in proteins. Principal component analysis (PCA) is a particularly promising dimensionality reduction algorithm to understand active site electric fields as it can decompose input features into a manageable set of dimensions.[391] In both studies of evolved protoglobin and heme protein families,[168,385] chemically relevant components can be observed by applying PCA on measured fields across dynamics and even for crystal structures. For directly evolved protoglobin, the PCA component that most differed between evolved variants aligned with carbene−diazirine bond formation and breaking of Fe−carbon bond in the transition state (Figure 19B). In applying machine learning to hemes, the principal component breakdown of the vector field provided a lower-dimensional input to predict function entirely from the 3D field.[385] Principal components from both crystal structures and molecular dynamic simulation 3D fields were shown to accurately assign heme reactivities (Figure 20A). Feature importance analysis showed that the most

A

Model + Input	F1 Score	Accuracy
XGBoost + PCA 3D	0.84	0.84
Balanced RF + PCA 3D	0.75	0.82
XGBoost + Single Point	0.42	0.44

B

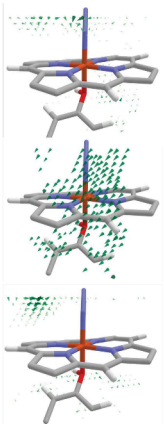


Figure 20. (A) Accuracy and F1 Scores of XGBoost and Random Forest (RF) models on PCA components from point and 3D electric fields. (B) Feature importance yields three principal components that align to Fe−oxyl bond or compression on the plan of oxygen binding. Adapted from ref 385

important components for determining activity in this study partially aligned with the Fe−oxyl bond or showed a compressive field on the plane of oxygen binding (Figure 20B), the former being consistent with prior findings in Bim et al.[130]

3. SUMMARY

Research on protein electrostatics in enzymology is rapidly growing. While experimentally it is not easy to study this effect in isolation from other factors that govern enzymatic catalysis, theoretical methods and models can aid this research. In fact, electrostatics as a driving force in enzymatic catalysis was first noticed in theory, in pioneering works by Warshel. Nowadays, theoretical tools for the analysis and characterization of intramolecular fields in proteins grow in number and sophistication. We review these developments. The methods are complementary, and each has unique strengths and shortcomings.

The EVB theory can rightfully be seen as pioneering in this field, for it was the first tool used to detect fields in proteins. It does so via decomposing the electronic structure of the reactants into the resonance forms, assuming that the ionic form will be affected the most by the field aligned with charge relocation. Through this feature of the model, fields can be detected, and reaction mechanisms can be simulated, as influenced by the fields. Importantly, the method incorporates the notion of dynamics in a sense of averaging over the protein motion upon traversing reaction free energy barrier. The assumption in the strategy is that the mechanism of reaction remains unchanged, and only the height of the barrier may be affected by the intramolecular field—a notion that was later challenged.

Another model-based field detection is done via polarizable FFs. In this case, the polarization of the electronic structure, due to the environment, is reflected in the FF parameters. While traditionally intended to increase the accuracy of simulations (compared to unpolarizable versions), polarizable FFs began to serve also as a measuring device for the field direction and strength.

Vibrational Stark effect was used experimentally to determine the strength of the electric fields experienced by Stark probes placed in enzyme active sites. These experiments, for the first time, confirmed that the fields attainable in proteins can be of an incredible strength, on the order of 100 MV/cm. These studies, coupled with molecular dynamics simulations, also view the fields as an averaged property over the protein dynamics trajectory, and link fields and reactivity. A shortcoming of the method is that it probes the field only at the location of the probe, e.g. a carbonyl attached to an inhibitor. The reason this can be problematic is 2-fold: first, the inhibitor itself alters the field, differently form the native substrate, and second, the more global portrait of the field in the entirety of the active site was later shown to be much more descriptive of the catalytic role of the field. It is unquestionable, however, that experimental observations enabled by the Stark spectroscopy propelled the research field forward, and led to its incredible subsequent growth.

Indeed, field in enzyme active sites are strongly heterogeneous, because they are created by strongly heterogeneous environments. The analysis of full electric fields in large volumes of the active site enabled new insights. These fields are incredibly complex and information-rich. Through AI approaches, such as PCA and clustering algorithms, these fields can be decomposed, analyzed, compared to fields in related proteins, or to fields in the same protein across a dynamics trajectory. The distribution of streamlines method, developed for the analysis of such global fields, revealed that field heterogeneity is meaningful, and the fields contain PCs that align with the direction opposite to electron flow in the reaction mechanisms of any complexity. Dynamics coupled to field analysis revealed that distinct field geometries can be visited by the protein, and each may facilitate a somewhat specific mechanism and barrier. This challenged the notion of averaging over all protein structures in the dynamics, and offered a perspective on the protein dynamics from the point of view of the dynamics of the field that it creates. Finally, the field plotted on a grid has an appearance of an image recognition problem. It was indeed demonstrated (albeit so far on a single prototypical example) that fields can be used as signatures of protein activity, i.e. for protein function recognition, through ML.

The thus revealed complexity of heterogeneous and dynamic electric fields is contextualized and explained in the clearest way by the QTAIM analysis. QTAIM describes the geometry of the full quantum mechanical electronic density in the active site. Since the number of electrons is fixed, and the density

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

(and molecular orbitals) is delocalized, pulling or pushing on any part of it with an electric field unavoidably causes changes in all other parts of the active site. Hence, it is obvious that fields need to be viewed globally, and their effect can be detected in a global way through changes in the global electron density. Electron density in the reactant state reports simultaneously on the global effect of the field experienced by the system, and on the magnitude of the reaction barrier that the system is about to cross. Thus, the density is a rigorous and chemically meaningful proxy linking the fields and reactivity in a chemically intuitive way.

Beyond fundamental analysis of electric fields in enzymes with the purpose of understanding how enzyme operate, the reviewed tools and insights open doors to new aspects of enzyme design. Fields being evidently prominent players in catalysis, have to be incorporated in enzyme design protocols, and that indeed begins to happen. Including the fields in design necessitates the focus on areas of the protein beyond the active site, e.g. the second and further coordination spheres where charged amino acids can be strategically placed. Such an amendment to the theo-zyme based design strategies might let computational enzyme design overcome the current limitations, and bring the performance of artificial enzymes closer to natural. Geometry of heterogeneous electric field in the active site as one of the design targets should be envisioned.

AUTHOR INFORMATION

Corresponding Authors

Mark E. Eberhart − *Chemistry Department, Colorado School of Mines, Golden, Colorado 80401, United States;* orcid.org/0000-0002-7061-3908; Email: meberhar@mines.edu

Anastassia N. Alexandrova − *Department of Chemistry, and Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, United States;* orcid.org/0000-0002-3003-1911; Email: ana@chem.ucla.edu

Authors

Pujan Ajmera − *Department of Chemistry, and Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, United States;* orcid.org/0000-0001-9509-6807

Daniel Bím − *Department of Physical Chemistry, University of Chemistry and Technology, Prague 166 28, Czech Republic;* orcid.org/0000-0003-3100-4293

Shobhit S. Chaturvedi − *Department of Chemistry, and Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, United States;* orcid.org/0000-0003-4977-5600

Santiago Vargas − *Department of Chemistry, and Biochemistry, University of California, Los Angeles, Los Angeles, California 90095, United States*

Timothy R. Wilson − *Chemistry Department, Colorado School of Mines, Golden, Colorado 80401, United States;* orcid.org/0000-0002-4320-9387

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.chemrev.4c00471

Author Contributions

[§]P.A., D.B., S.S.C., S.V., and T.R.W. are equally contributing authors. CRediT: **Mark E. Eberhart** conceptualization, funding acquisition, methodology, project administration, resources, software, supervision, validation, writing - original draft, writing - review & editing; **Anastassia N. Alexandrova** conceptualization, funding acquisition, investigation, methodology, project administration, resources, supervision, writing - original draft, writing - review & editing; **Pujan Ajmera** data curation, formal analysis, investigation, methodology, software, validation, visualization, writing - original draft; **Daniel Bím** data curation, formal analysis, investigation, validation, visualization, writing - original draft; **Shobhit S. Chaturvedi** data curation, formal analysis, investigation, validation, visualization, writing - original draft; **Santiago Vargas** conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, software, visualization, writing - original draft; **Timothy R. Wilson** data curation, formal analysis, investigation, methodology, software, validation, visualization, writing - original draft.

Notes

The authors declare no competing financial interest.

Biographies

Professor Mark Eberhart received a dual B.S. in Chemistry and Applied Mathematics, followed by an M.S. in Physical Biochemistry, from University of Colorado, and a Ph.D. in Materials Science and Engineering from Massachusetts Institute of Technology (MIT) under the mentorship of Professor Keith Johnson. He completed Postdoctoral research at MIT and worked as a staff scientist at Los Alamos National Laboratory before becoming a Full Professor at Colorado School of Mines in 1998. His research career has focused on understanding chemical and material properties from first-principles, specifically by discovering and leveraging electron charge density-based structure property relationships.

Anastassia Alexandrova is a Charles W. Clifford Jr. Professor in Chemistry and Biochemistry, and Professor of Materials Science and Engineering in UCLA. She obtained a B.S./M.S. Diploma with highest honors, from Saratov University, Russia, her Ph.D. in Theoretical Physical Chemistry from Utah State University, and was then a Postdoctoral Associate and an American Cancer Society Postdoctoral Fellow at Yale University. Anastassia joined the faculty of UCLA and CNSI in 2010. The focus of her laboratory is theory and computation for design and multiscale modeling of functional materials: dynamic catalytic interfaces, artificial metalloenzymes, molecular qubits and their assemblies, and quantum materials.

Pujan Ajmera earned his bachelor's degree in Engineering Physics at the University of Michigan, Ann Arbor. He is currently a Ph.D. student in Theoretical Chemistry at University of California, Los Angeles in the Alexandrova Group, where he studies how electric fields and their dynamics tune the electronic structure of metalloenzyme active sites.

Professor Daniel Bím received his B.S./M.S. with honors in Organic Chemistry from the University of Chemistry and Technology, Czech Republic, and his Ph.D. in Theoretical Chemistry from Charles University, Czech Republic. He completed Postdoctoral research at the University of California under the mentorship of Professor Alexandrova, was a Marie Sklodowska-Curie Postdoctoral Fellow at The California Institute of Technology under Professor Ryan G. Hadt, and worked as a MSCA Research Fellow at the Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences (IOCB CAS CR) before joining the IOCB CAS CR faculty as an Assistant Professor in 2024. Daniel's research focuses on theoretical and computational chemistry of biological and enzymatic systems, and on the role of electric fields in catalysis and redox chemistry.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

Dr. Shobhit S. Chaturvedi is a postdoctoral researcher in the Department of Chemistry and Biochemistry at the University of California, Los Angeles, working under the mentorship of Professor Alexandrova. Building on his Ph.D. work at Michigan Technological University under Dr. Christo Z. Christov, where he explored enzyme reaction mechanisms through Molecular Dynamics and hybrid Quantum Mechanics/Molecular Mechanics simulations, Shobhit now advances his research by integrating these techniques with innovative method developments and Machine Learning to investigate the role of protein scaffolds in enzyme catalysis. His ultimate goal is to harness these insights to design enzymes with enhanced catalytic properties and novel functionalities.

Dr. Santiago Vargas graduated from Harvard College in 2019 where he studied Chemistry and Physics. He received his Ph.D. in Chemistry in 2024 from the University of California, Los Angeles. Since graduating, he has conducted research at Lawrence Berkeley National Laboratory as a Darleane C. Hoffman Postdoctoral Fellow. His research is focused on developing interatomic potentials for heavy metal chemistries and foundational models for chemical descriptors.

Dr. Timothy Wilson received his B.S. with honors from Colorado School of Mines in Chemical Engineering, where he began developing the software (Bondalyzer) used for the QTAIM-adjacent method of bond bundle and gradient bundle analysis with Professor Eberhart. He remained at Mines for his Ph.D. in Chemistry, and there continued as a Postdoctoral fellow and now as Research faculty to develop Bondalyzer and apply it to a variety of chemical, biological, and material science problems.

ACKNOWLEDGMENTS

This work was supported by the NSF-CHE grant 2203366 grant to A.N.A. and M.E.E. P.A. was supported in part by the NSF NRT-QISE: Accelerating Interdisciplinary Frontiers in Quantum Sciences and Technologies grant 2125924. S.V. acknowledges the support of the Department of Energy Computational Science Graduate Fellowship under grant DE-SC0021110. M.E.E. and T.W. acknowledge the support of the State of Colorado AIA 2021 grant.

REFERENCES

(1) Kusalik, P. G.; Gracheva, M. E.; Shaik, S.; English, N. J. New Perspectives on Molecular Simulation of Chemistry and Physics in External Electric Fields. *Phys. Chem. Chem. Phys.* **2022**, *24*, 28660−28661.
(2) Vaissier Welborn, V. Structural Dynamics and Computational Design of Synthetic Enzymes. *Chem. Catalysis* **2022**, *2*, 19−28.
(3) Siddiqui, S. A.; Stuyver, T.; Shaik, S.; Dubey, K. D. Designed Local Electric Fields-Promising Tools for Enzyme Engineering. *JACS Au* **2023**, *3*, 3259−3269.
(4) Ruiz-Pernía, J. J.; Świderek, K.; Bertran, J.; Moliner, V.; Tuñón, I. Electrostatics as a Guiding Principle in Understanding and Designing Enzymes. *J. Chem. Theory Comput.* **2024**, *20*, 1783−1795.
(5) Griffiths, D. J. *Introduction to Electrodynamics*, 4th ed.; Pearson: Boston, 2013.
(6) Leontyev, I. V.; Stuchebrukhov, A. A. Electronic Continuum Model for Molecular Dynamics Simulations. *J. Chem. Phys.* **2009**, *130*, 085102.
(7) Bridge, N. J.; Buckingham, A. D.; Linnett, J. W. The Polarization of Laser Light Scattered by Gases. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **1966**, *295*, 334−349.
(8) Lewis, M.; Wu, Z.; Glaser, R. Polarizabilities of Carbon Dioxide and Carbodiimide. Assessment of Theoretical Model Dependencies on Dipole Polarizabilities and Dipole Polarizability Anisotropies. *J. Phys. Chem. A* **2000**, *104*, 11355−11361.
(9) Khaniya, U.; Mao, J.; Wei, R. J.; Gunner, M. R. Characterizing Protein Protonation Microstates Using Monte Carlo Sampling. *J. Phys. Chem. B* **2022**, *126*, 2476−2485.
(10) Warshel, A. Electrostatic Basis of Structure-Function Correlation in Proteins. *Acc. Chem. Res.* **1981**, *14*, 284−290.
(11) Warshel, A. Electrostatic Origin of the Catalytic Power of Enzymes and the Role of Preorganized Active Sites. *J. Biol. Chem.* **1998**, *273*, 27035−27038.
(12) 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.
(13) Roca, M.; Andrés, J.; Moliner, V.; Tuñón, I.; Bertrán, J. On the Nature of the Transition State in Catechol O-Methyltransferase. A Complementary Study Based on Molecular Dynamics and Potential Energy Surface Explorations. *J. Am. Chem. Soc.* **2005**, *127*, 10648−10655.
(14) Kamerlin, S. C.; Sharma, P. K.; Chu, Z. T.; Warshel, A. Ketosteroid Isomerase Provides Further Support for the Idea That Enzymes Work by Electrostatic Preorganization. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 4075−4080.
(15) Fuxreiter, M.; Mones, L. The Role of Reorganization Energy in Rational Enzyme Design. *Curr. Opin. Chem. Biol.* **2014**, *21*, 34−41.
(16) Świderek, K.; Marti, S.; Tuñón, I.; Moliner, V.; Bertran, J. Peptide Bond Formation Mechanism Catalyzed by Ribosome. *J. Am. Chem. Soc.* **2015**, *137*, 12024−12034.
(17) Isaksen, G. V.; Hopmann, K. H.; Åqvist, J.; Brandsdal, B. O. Computer Simulations Reveal Substrate Specificity of Glycosidic Bond Cleavage in Native and Mutant Human Purine Nucleoside Phosphorylase. *Biochemistry* **2016**, *55*, 2153−2162.
(18) Krzemińska, A.; Moliner, V.; Świderek, K. Dynamic and Electrostatic Effects on the Reaction Catalyzed by HIV-1 Protease. *J. Am. Chem. Soc.* **2016**, *138*, 16283−16298.
(19) 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.
(20) Yadav, S.; Shaik, S.; Siddiqui, S. A.; Kalita, S.; Dubey, K. D. Local Electric Fields Dictate Function: The Different Product Selectivities Observed for Fatty Acid Oxidation by Two Deceptively Very Similar P450-Peroxygenases OleT and BSβ. *J. Chem. Inf. Model.* **2022**, *62*, 1025−1035.
(21) Peng, W.; Yan, S.; Zhang, X.; Liao, L.; Zhang, J.; Shaik, S.; Wang, B. How Do Preorganized Electric Fields Function in Catalytic Cycles? The Case of the Enzyme Tyrosine Hydroxylase. *J. Am. Chem. Soc.* **2022**, *144*, 20484−20494.
(22) Kaplan, J.; DeGrado, W. F. De Novo Design of Catalytic Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 11566−11570.
(23) Kiss, G.; Çelebi-Ölçüm, N.; Moretti, R.; Baker, D.; Houk, K. N. Computational Enzyme Design. *Angew. Chem., Int. Ed.* **2013**, *52*, 5700−5725.
(24) Świderek, K.; Tuñón, I.; Moliner, V. Predicting Enzymatic Reactivity: From Theory to Design. *WIREs Computational Molecular Science* **2014**, *4*, 407−421.
(25) Frushicheva, M. P.; Mills, M. J.; Schopf, P.; Singh, M. K.; Prasad, R. B.; Warshel, A. Computer Aided Enzyme Design and Catalytic Concepts. *Curr. Opin. Chem. Biol.* **2014**, *21*, 56−62.
(26) Vaissier, V.; Sharma, S. C.; 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.
(27) Welborn, V. V.; Ruiz Pestana, L.; Head-Gordon, T. Computational Optimization of Electric Fields for Better Catalysis Design. *Nature Catalysis* **2018**, *1*, 649−655.
(28) Bunzel, H. A.; Anderson, J. L. R.; Mulholland, A. J. Designing Better Enzymes: Insights from Directed Evolution. *Curr. Opin. Struct. Biol.* **2021**, *67*, 212−218.
(29) Xie, W. J.; Asadi, M.; Warshel, A. Enhancing Computational Enzyme Design by a Maximum Entropy Strategy. *Proceedings of the National Academy of Sciences U.S.A.* **2022**, *119*, No. e2122355119.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

(30) 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. *Chemical Science* **2023**, *14*, 10997−11011.
(31) Vennelakanti, V.; Nazemi, A.; Mehmood, R.; Steeves, A. H.; Kulik, H. J. Harder, Better, Faster, Stronger: Large-scale QM and QM/MM for Predictive Modeling in Enzymes and Proteins. *Curr. Opin. Struct. Biol.* **2022**, *72*, 9−17.
(32) 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*, 200−204.
(33) Fafarman, A. T.; Sigala, P. A.; Herschlag, D.; Boxer, S. G. Decomposition of Vibrational Shifts of Nitriles into Electrostatic and Hydrogen-Bonding Effects. *J. Am. Chem. Soc.* **2010**, *132*, 12811−12813.
(34) Jha, S. K.; Ji, M.; Gaffney, K. J.; Boxer, S. G. Direct Measurement of the Protein Response to an Electrostatic Perturbation That Mimics the Catalytic Cycle in Ketosteroid Isomerase. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 16612−16617.
(35) Bagchi, S.; Fried, S. D.; Boxer, S. G. A Solvatochromic Model Calibrates Nitriles' Vibrational Frequencies to Electrostatic Fields. *J. Am. Chem. Soc.* **2012**, *134*, 10373−10376.
(36) Fafarman, A. T.; Sigala, P. A.; Schwans, J. P.; Fenn, T. D.; Herschlag, D.; Boxer, S. G. Quantitative, Directional Measurement of Electric Field Heterogeneity in the Active Site of Ketosteroid Isomerase. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, E299−E308.
(37) Stafford, A. J.; Walker, D. M.; Webb, L. J. Electrostatic Effects of Mutations of Ras Glutamine 61 Measured Using Vibrational Spectroscopy of a Thiocyanate Probe. *Biochemistry* **2012**, *51*, 2757−2767.
(38) Wu, Y.; Boxer, S. G. A Critical Test of the Electrostatic Contribution to Catalysis with Noncanonical Amino Acids in Ketosteroid Isomerase. *J. Am. Chem. Soc.* **2016**, *138*, 11890−11895.
(39) Adesina, A. S.; Świderek, K.; Luk, L. Y. P.; Moliner, V.; Allemann, R. K. Electric Field Measurements Reveal the Pivotal Role of Cofactor−Substrate Interaction in Dihydrofolate Reductase Catalysis. *ACS Catal.* **2020**, *10*, 7907−7914.
(40) Kraskov, A.; von Sass, J.; Nguyen, A. D.; Hoang, T. O.; Buhrke, D.; Katz, S.; Michael, N.; Kozuch, J.; Zebger, I.; Siebert, F.; et al. Local Electric Field Changes during the Photoconversion of the Bathy Phytochrome Agp2. *Biochemistry* **2021**, *60*, 2967−2977.
(41) Ji, Z.; Kozuch, J.; Mathews, I. I.; Diercks, C. S.; Shamsudin, Y.; Schulz, M. A.; Boxer, S. G. Protein Electric Fields Enable Faster and Longer-Lasting Covalent Inhibition of β-Lactamases. *J. Am. Chem. Soc.* **2022**, *144*, 20947−20954.
(42) Warshel, A.; Dryga, A. Simulating Electrostatic Energies in Proteins: Perspectives and Some Recent Studies of pKas, Redox, and Other Crucial Functional Properties. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 3469−3484.
(43) Yang, A.-S.; Gunner, M. R.; Sampogna, R.; Sharp, K.; Honig, B. On the Calculation of pKas in Proteins. *Proteins: Struct., Funct., Bioinf.* **1993**, *15*, 252−265.
(44) Del Buono, G. S.; Figueirido, F. E.; Levy, R. M. Intrinsic pKas of Ionizable Residues in Proteins: An Explicit Solvent Calculation for Lysozyme. *Proteins: Struct., Funct., Bioinf.* **1994**, *20*, 85−97.
(45) Sham, Y. Y.; Chu, Z. T.; Warshel, A. Consistent Calculations of pKas of Ionizable Residues in Proteins: Semi-Microscopic and Microscopic Approaches. *J. Phys. Chem. B* **1997**, *101*, 4458−4472.
(46) Mehler, E. L.; Guarnieri, F. A Self-Consistent, Microenvironment Modulated Screened Coulomb Potential Approximation to Calculate pH-Dependent Electrostatic Effects in Proteins. *Biophys. J.* **1999**, *77*, 3−22.
(47) Fitch, C. A.; Karp, D. A.; Lee, K. K.; Stites, W. E.; Lattman, E. E.; García-Moreno, E. B. Experimental pKa Values of Buried Residues: Analysis with Continuum Methods and Role of Water Penetration. *Biophys. J.* **2002**, *82*, 3289−3304.
(48) Kato, M.; Warshel, A. Using a Charging Coordinate in Studies of Ionization Induced Partial Unfolding. *J. Phys. Chem. B* **2006**, *110*, 11566−11570.
(49) Fitzkee, N. C.; García-Moreno, E. B. Electrostatic Effects in Unfolded Staphylococcal Nuclease. *Protein Sci.* **2008**, *17*, 216−227.
(50) Ishikita, H. Origin of the pKa Shift of the Catalytic Lysine in Acetoacetate Decarboxylase. *FEBS Lett.* **2010**, *584*, 3464−3468.
(51) Adam, S.; Bondar, A.-N. Mechanism by Which Water and Protein Electrostatic Interactions Control Proton Transfer at the Active Site of Channelrhodopsin. *PLoS One* **2018**, *13*, No. e0201298.
(52) Lindman, S.; Bauer, M. C.; Lund, M.; Diehl, C.; Mulder, F. A.; Akke, M.; Linse, S. pKa Values for the Unfolded State under Native Conditions Explain the pH-Dependent Stability of PGB1. *Biophys. J.* **2010**, *99*, 3365−3373.
(53) Coskun, D. Free Energy Methods in Drug Discovery: Current State and Future Directions. *ACS Symposium Series; American Chemical Society* **2021**, *1397*, 143−159. Chapter 6
(54) Kwon, I.; Jo, G.; Shin, K.-S. A Deep Neural Network Based on ResNet for Predicting Solutions of Poisson−Boltzmann Equation. *Electronics* **2021**, *10*, 2627.
(55) Chen, J.; Geng, W.; Wei, G.-W. MLIMC: Machine Learning-Based Implicit-Solvent Monte Carlo. *Chinese Journal of Chemical Physics* **2021**, *34*, 683−694.
(56) Rogers, N. K.; Moore, G. R.; Sternberg, M. J. Electrostatic Interactions in Globular Proteins: Calculation of the pH Dependence of the Redox Potential of Cytochrome C551. *J. Mol. Biol.* **1985**, *182*, 613−616.
(57) Mouesca, J.-M.; Chen, J. L.; Noodleman, L.; Bashford, D.; Case, D. A. Density Functional/Poisson-Boltzmann Calculations of Redox Potentials for Iron-Sulfur Clusters. *J. Am. Chem. Soc.* **1994**, *116*, 11898−11914.
(58) Swartz, P.; Beck, B.; Ichiye, T. Structural Origins of Redox Potentials in Fe-S Proteins: Electrostatic Potentials of Crystal Structures. *Biophys. J.* **1996**, *71*, 2958−2969.
(59) Simonson, T.; Archontis, G.; Karplus, M. A Poisson-boltzmann Study of Charge Insertion in an Enzyme Active Site: The Effect of Dielectric Relaxation. *J. Phys. Chem. B* **1999**, *103*, 6142−6156.
(60) Gane, P.; Freedman, R.; Warwicker, J. A Molecular Model for the Redox Potential Difference between Thioredoxin and DsbA, Based on Electrostatics Calculations. *J. Mol. Biol.* **1995**, *249*, 376−387.
(61) Mao, J.; Hauser, K.; Gunner, M. R. How Cytochromes with Different Folds Control Heme Redox Potentials. *Biochemistry* **2003**, *42*, 9829−9840.
(62) Chen, C. G.; Nardi, A. N.; Amadei, A.; D'Abramo, M. Theoretical Modeling of Redox Potentials of Biomolecules. *Molecules* **2022**, *27*, 1077.
(63) Kuznetsov, A. M.; Zueva, E. M.; Masliy, A. N.; Krishtalik, L. I. Redox Potential of the Rieske Iron−Sulfur Protein. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2010**, *1797*, 347−359.
(64) Kanda, T.; Ishikita, H. Redox Potentials of Iron−Sulfur Clusters in Type I Photosynthetic Reaction Centers. *J. Phys. Chem. B* **2023**, *127*, 4998−5004.
(65) Gamiz-Hernandez, A. P.; Kieseritzky, G.; Ishikita, H.; Knapp, E. W. Rubredoxin Function: Redox Behavior from Electrostatics. *J. Chem. Theory Comput.* **2011**, *7*, 742−752.
(66) Gaughan, S. J. H.; Hirst, J. D.; Croft, A. K.; Jäger, C. M. Effect of Oriented Electric Fields on Biologically Relevant Iron−Sulfur Clusters: Tuning Redox Reactivity for Catalysis. *J. Chem. Inf. Model.* **2022**, *62*, 591−601.
(67) Popović, D. M.; Zarić, S. D.; Rabenstein, B.; Knapp, E.-W. Artificial Cytochrome b: Computer Modeling and Evaluation of Redox Potentials. *J. Am. Chem. Soc.* **2001**, *123*, 6040−6053.
(68) Popović, D. M.; Zmirić, A.; Zarić, S. D.; Knapp, E.-W. Energetics of Radical Transfer in DNA Photolyase. *J. Am. Chem. Soc.* **2002**, *124*, 3775−3782.
(69) Ishikita, H.; Knapp, E.-W. Redox Potential of Quinones in Both Electron Transfer Branches of Photosystem I. *J. Biol. Chem.* **2003**, *278*, 52002−52011.
(70) Gámiz-Hernández, A. P.; Kieseritzky, G.; Galstyan, A. S.; Demir-Kavuk, O.; Knapp, E.-W. Understanding Properties of Cofactors in Proteins: Redox Potentials of Synthetic Cytochromes b. *ChemPhysChem* **2010**, *11*, 1196−1206.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

(71) Matsui, T.; Song, J.-W. A Density Functional Theory-Based Scheme to Compute the Redox Potential of a Transition Metal Complex: Applications to Heme Compound. *Molecules* **2019**, *24*, 819.
(72) Szefczyk, B.; Mulholland, A. J.; Ranaghan, K. E.; Sokalski, W. A. 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.
(73) Chojnacka, M.; Feliks, M.; Beker, W.; Sokalski, W. A. Predicting Substituent Effects on Activation Energy Changes by Static Catalytic Fields. *J. Mol. Model.* **2018**, *24*, 28.
(74) Tantillo, D. J.; Jiangang, C.; Houk, K. N. Theozymes and Compuzymes: Theoretical Models for Biological Catalysis. *Curr. Opin. Chem. Biol.* **1998**, *2*, 743−750.
(75) Lassila, J. K.; Privett, H. K.; Allen, B. D.; Mayo, S. L. Combinatorial Methods for Small-Molecule Placement in Computational Enzyme Design. *Proceedings of the National Academy of Sciences U.S.A.* **2006**, *103*, 16710−16715.
(76) Beker, W.; Sokalski, W. A. Bottom-up Nonempirical Approach to Reducing Search Space in Enzyme Design Guided by Catalytic Fields. *J. Chem. Theory Comput.* **2020**, *16*, 3420−3429.
(77) 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.
(78) Bolon, D. N.; Mayo, S. L. Enzyme-like Proteins by Computational Design. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 14274−14279.
(79) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; et al. De Novo Computational Design of Retro-Aldol Enzymes. *Science* **2008**, *319*, 1387−1391.
(80) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; et al. Kemp Elimination Catalysts by Computational Enzyme Design. *Nature* **2008**, *453*, 190−195.
(81) Privett, H. K.; Kiss, G.; Lee, T. M.; Blomberg, R.; Chica, R. A.; Thomas, L. M.; Hilvert, D.; Houk, K. N.; Mayo, S. L. Iterative Approach to Computational Enzyme Design. *Proceedings of the National Academy of Sciences U.S.A.* **2012**, *109*, 3790−3795.
(82) Casey, M. L.; Kemp, D. S.; Paul, K. G.; Cox, D. D. Physical Organic Chemistry of Benzisoxazoles. I. Mechanism of the Base-Catalyzed Decomposition of Benzisoxazoles. *Journal of Organic Chemistry* **1973**, *38*, 2294−2301.
(83) Kemp, D. S.; Casey, M. L. Physical Organic Chemistry of Benzisoxazoles. II. Linearity of the Broensted Free Energy Relation for the Base-Catalyzed Decomposition of Benzisoxazoles. *J. Am. Chem. Soc.* **1973**, *95*, 6670−6680.
(84) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; et al. Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **2010**, *329*, 309−313.
(85) Layfield, J. P.; Hammes-Schiffer, S. Calculation of Vibrational Shifts of Nitrile Probes in the Active Site of Ketosteroid Isomerase upon Ligand Binding. *J. Am. Chem. Soc.* **2013**, *135*, 717−725.
(86) Fried, S. D.; Wang, L.-P.; Boxer, S. G.; Ren, P.; Pande, V. S. Calculations of the Electric Fields in Liquid Solutions. *J. Phys. Chem. B* **2013**, *117*, 16236−16248.
(87) Liu, C. T.; Layfield, J. P.; Stewart, R. J.; French, J. B.; Hanoian, P.; Asbury, J. B.; Hammes-Schiffer, S.; Benkovic, S. J. Probing the Electrostatics of Active Site Microenvironments along the Catalytic Cycle for Escherichia Coli Dihydrofolate Reductase. *J. Am. Chem. Soc.* **2014**, *136*, 10349−10360.
(88) Wang, X.; He, X. An Ab Initio QM/MM Study of the Electrostatic Contribution to Catalysis in the Active Site of Ketosteroid Isomerase. *Molecules (Basel, Switzerland)* **2018**, *23*, 2410.
(89) Bradshaw, R. T.; Dziedzic, J.; Skylaris, C.-K.; Essex, J. W. The Role of Electrostatics in Enzymes: Do Biomolecular Force Fields Reflect Protein Electric Fields? *J. Chem. Inf. Model.* **2020**, *60*, 3131−3144.
(90) Prah, A.; Frančišković, E.; Mavri, J.; Stare, J. Electrostatics as the Driving Force behind the Catalytic Function of the Monoamine Oxidase a Enzyme Confirmed by Quantum Computations. *ACS Catal.* **2019**, *9*, 1231−1240.
(91) Acosta-Silva, C.; Bertran, J.; Branchadell, V.; Oliva, A. Kemp Elimination Reaction Catalyzed by Electric Fields. *Chemphyschem: a European journal of chemical physics and physical chemistry* **2020**, *21*, 295−306.
(92) Huang, P.-S.; Boyken, S. E.; Baker, D. The Coming of Age of de Novo Protein Design. *Nature* **2016**, *537*, 320−327.
(93) Lovelock, S. L.; Crawshaw, R.; Basler, S.; Levy, C.; Baker, D.; Hilvert, D.; Green, A. P. The Road to Fully Programmable Protein Catalysis. *Nature* **2022**, *606*, 49−58.
(94) Yon, J.; Perahia, D.; Ghélis, C. Conformational Dynamics and Enzyme Activity. *Biochimie* **1998**, *80*, 33−42.
(95) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D. M.; Wolf-Watz, M.; Bosco, D. A.; Skalicky, J. J.; Kay, L. E.; Kern, D. Intrinsic Dynamics of an Enzyme Underlies Catalysis. *Nature* **2005**, *438*, 117−121.
(96) Nashine, V. C.; Hammes-Schiffer, S.; Benkovic, S. J. Coupled Motions in Enzyme Catalysis. *Curr. Opin. Chem. Biol.* **2010**, *14*, 644−651.
(97) Bhowmick, A.; Sharma, S. C.; Honma, H.; Head-Gordon, T. The Role of Side Chain Entropy and Mutual Information for Improving the de Novo Design of Kemp Eliminases KE07 and KE70. *Phys. Chem. Chem. Phys.* **2016**, *18*, 19386−19396.
(98) Zoi, I.; Antoniou, D.; Schwartz, S. D. Electric Fields and Fast Protein Dynamics in Enzymes. *J. Phys. Chem. Lett.* **2017**, *8*, 6165−6170.
(99) Osuna, S. The Challenge of Predicting Distal Active Site Mutations in Computational Enzyme Design. *WIREs Computational Molecular Science* **2021**, *11*, No. e1502.
(100) Yang, Z.; Hajlasz, N.; Steeves, A. H.; Kulik, H. J. Quantifying the Long-range Coupling of Electronic Properties in Proteins with Ab Initio Molecular Dynamics. *Chemistry−Methods* **2021**, *1*, 362−373.
(101) Corbella, M.; Pinto, G. P.; Kamerlin, S. C. L. Loop Dynamics and the Evolution of Enzyme Activity. *Nature Reviews Chemistry* **2023**, *7*, 536−547.
(102) Blomberg, R.; Kries, H.; Pinkas, D. M.; Mittl, P. R. E.; Grütter, M. G.; Privett, H. K.; Mayo, S. L.; Hilvert, D. Precision Is Essential for Efficient Catalysis in an Evolved Kemp Eliminase. *Nature* **2013**, *503*, 418−421.
(103) Wrabl, J. O.; Gu, J.; Liu, T.; Schrank, T. P.; Whitten, S. T.; Hilser, V. J. The Role of Protein Conformational Fluctuations in Allostery, Function, and Evolution. *Biophys. Chem.* **2011**, *159*, 129−141.
(104) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. The Ensemble Nature of Allostery. *Nature* **2014**, *508*, 331−339.
(105) Petrović, D.; Risso, V. A.; Kamerlin, S. C. L.; Sanchez-Ruiz, J. M. Conformational Dynamics and Enzyme Evolution. *Journal of The Royal Society Interface* **2018**, *15*, 20180330.
(106) Gardner, J. M.; Biler, M.; Risso, V. A.; Sanchez-Ruiz, J. M.; Kamerlin, S. C. L. Manipulating Conformational Dynamics to Repurpose Ancient Proteins for Modern Catalytic Functions. *ACS Catal.* **2020**, *10*, 4863−4870.
(107) Arnold, F. H. Design by Directed Evolution. *Acc. Chem. Res.* **1998**, *31*, 125−131.
(108) Bolon, D. N.; Voigt, C. A.; Mayo, S. L. De Novo Design of Biocatalysts. *Curr. Opin. Chem. Biol.* **2002**, *6*, 125−129.
(109) Khersonsky, O.; Röthlisberger, D.; Dym, O.; Albeck, S.; Jackson, C. J.; Baker, D.; Tawfik, D. S. Evolutionary Optimization of Computationally Designed Enzymes: Kemp Eliminases of the KE07 Series. *J. Mol. Biol.* **2010**, *396*, 1025−1042.
(110) Khersonsky, O.; Röthlisberger, D.; Wollacott, A. M.; Murphy, P.; Dym, O.; Albeck, S.; Kiss, G.; Houk, K.; Baker, D.; Tawfik, D. S. Optimization of the In-Silico-Designed Kemp Eliminase KE70 by

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

Computational Design and Directed Evolution. *J. Mol. Biol.* **2011**, *407*, 391−412.
(111) Khersonsky, O.; Kiss, G.; Röthlisberger, D.; Dym, O.; Albeck, S.; Houk, K. N.; Baker, D.; Tawfik, D. S. Bridging the Gaps in Design Methodologies by Evolutionary Optimization of the Stability and Proficiency of Designed Kemp Eliminase KE59. *Proceedings of the National Academy of Sciences U.S.A.* **2012**, *109*, 10358−10363.
(112) Frushicheva, M. P.; Cao, J.; Chu, Z. T.; Warshel, A. Exploring Challenges in Rational Enzyme Design by Simulating the Catalysis in Artificial Kemp Eliminase. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16869−16874.
(113) Frushicheva, M. P.; Cao, J.; Warshel, A. Challenges and Advances in Validating Enzyme Design Proposals: The Case of Kemp Eliminase Catalysis. *Biochemistry* **2011**, *50*, 3849−3858.
(114) Labas, A.; Szabo, E.; Mones, L.; Fuxreiter, M. Optimization of Reorganization Energy Drives Evolution of the Designed Kemp Eliminase KE07. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2013**, *1834*, 908−917.
(115) Bhowmick, A.; Sharma, S. C.; Head-Gordon, T. The Importance of the Scaffold for de Novo Enzymes: A Case Study with Kemp Eliminase. *J. Am. Chem. Soc.* **2017**, *139*, 5793−5800.
(116) Jindal, G.; Ramachandran, B.; Bora, R. P.; Warshel, A. Exploring the Development of Ground-State Destabilization and Transition-State Stabilization in Two Directed Evolution Paths of Kemp Eliminases. *ACS Catal.* **2017**, *7*, 3301−3305.
(117) Ward, T. R. Artificial Enzymes Made to Order: Combination of Computational Design and Directed Evolution. *Angew. Chem., Int. Ed.* **2008**, *47*, 7802−7803.
(118) Broom, A.; Rakotoharisoa, R. V.; Thompson, M. C.; Zarifi, N.; Nguyen, E.; Mukhametzhanov, N.; Liu, L.; Fraser, J. S.; Chica, R. A. Ensemble-Based Enzyme Design Can Recapitulate the Effects of Laboratory Directed Evolution in Silico. *Nat. Commun.* **2020**, *11*, 4808.
(119) Mariz, B. d. P.; Carvalho, S.; Batalha, I. L.; Pina, A. S. Artificial Enzymes Bringing Together Computational Design and Directed Evolution. *Organic & Biomolecular Chemistry* **2021**, *19*, 1915−1925.
(120) Valdez, C. E.; Morgenstern, A.; Eberhart, M. E.; Alexandrova, A. N. Predictive Methods for Computational Metalloenzyme Redesign − A Test Case with Carboxypeptidase A. *Phys. Chem. Chem. Phys.* **2016**, *18*, 31744−31756.
(121) 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.
(122) Morgenstern, A.; Jaszai, M.; Eberhart, M. E.; Alexandrova, A. N. Quantified Electrostatic Preorganization in Enzymes Using the Geometry of the Electron Charge Density. *Chemical Science* **2017**, *8*, 5010−5018.
(123) 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.
(124) Hennefarth, M. R.; Alexandrova, A. N. Advances in Optimizing Enzyme Electrostatic Preorganization. *Curr. Opin. Struct. Biol.* **2022**, *72*, 1−8.
(125) 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.
(126) Biava, H.; Schreiber, T.; Katz, S.; Völler, J.-S.; Stolarski, M.; Schulz, C.; Michael, N.; Budisa, N.; Kozuch, J.; Utesch, T.; et al. Long-Range Modulations of Electric Fields in Proteins. *J. Phys. Chem. B* **2018**, *122*, 8330−8342.
(127) Wilding, M.; Hong, N.; Spence, M.; Buckle, A. M.; Jackson, C. J. Protein Engineering: The Potential of Remote Mutations. *Biochem. Soc. Trans.* **2019**, *47*, 701−711.
(128) Xu, X.; Yan, S.; Hou, X.; Song, W.; Wang, L.; Wu, T.; Qi, M.; Wu, J.; Rao, Y.; Wang, B.; et al. Local Electric Field Modulated Reactivity of Pseudomonas Aeruginosa Acid Phosphatase for Enhancing Phosphorylation of L-Ascorbic Acid. *ACS Catal.* **2021**, *11*, 13397−13407.
(129) Bím, D.; Alexandrova, A. N. Electrostatic Regulation of Blue Copper Sites. *Chemical Science* **2021**, *12*, 11406−11413.
(130) Bím, D.; Alexandrova, A. N. Local Electric Fields as a Natural Switch of Heme-Iron Protein Reactivity. *ACS Catal.* **2021**, *11*, 6534−6546.
(131) Fried, S. D.; Boxer, S. G. Electric Fields and Enzyme Catalysis. *Annu. Rev. Biochem.* **2017**, *86*, 387−415.
(132) Stark, J. Observation of the Separation of Spectral Lines by an Electric Field. *Nature* **1913**, *92*, 401−401.
(133) Leone, M.; Paoletti, A.; Robotti, N. A Simultaneous Discovery: The Case of Johannes Stark and Antonino Lo Surdo. *Physics in Perspective* **2004**, *6*, 271−294.
(134) Fried, S. D.; Boxer, S. G. Measuring Electric Fields and Noncovalent Interactions Using the Vibrational Stark Effect. *Acc. Chem. Res.* **2015**, *48*, 998−1006.
(135) Wu, Y.; Fried, S. D.; Boxer, S. G. A Preorganized Electric Field Leads to Minimal Geometrical Reorientation in the Catalytic Reaction of Ketosteroid Isomerase. *J. Am. Chem. Soc.* **2020**, *142*, 9993−9998.
(136) Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme Electric Fields Power Catalysis in the Active Site of Ketosteroid Isomerase. *Science* **2014**, *346*, 1510−1514.
(137) Ji, Z.; Boxer, S. G. β-Lactamases Evolve against Antibiotics by Acquiring Large Active-Site Electric Fields. *J. Am. Chem. Soc.* **2022**, *144*, 22289−22294.
(138) First, J. T.; Novelli, E. T.; Webb, L. J. Beyond pKa: Experiments and Simulations of Nitrile Vibrational Probes in Staphylococcal Nuclease Show the Importance of Local Interactions. *J. Phys. Chem. B* **2020**, *124*, 3387−3399.
(139) Błasiak, B.; Ritchie, A. W.; Webb, L. J.; Cho, M. Vibrational Solvatochromism of Nitrile Infrared Probes: Beyond the Vibrational Stark Dipole Approach. *Phys. Chem. Chem. Phys.* **2016**, *18*, 18094−18111.
(140) Kozuch, J.; Schneider, S. H.; Zheng, C.; Ji, Z.; Bradshaw, R. T.; Boxer, S. G. Testing the Limitations of MD-based Local Electric Fields Using the Vibrational Stark Effect in Solution: Penicillin G as a Test Case. *J. Phys. Chem. B* **2021**, *125*, 4415−4427.
(141) Kohler, B. E.; Woehl, J. C. Measuring Internal Electric Fields with Atomic Resolution. *J. Chem. Phys.* **1995**, *102*, 7773−7781.
(142) Bublitz, G. U.; Boxer, S. G. Stark Spectroscopy: Applications in Chemistry, Biology, and Materials Science. *Annu. Rev. Phys. Chem.* **1997**, *48*, 213−242.
(143) Ragain, C. M.; Newberry, R. W.; Ritchie, A. W.; Webb, L. J. Role of Electrostatics in Differential Binding of RalGDS to Rap Mutations E30D and K31E Investigated by Vibrational Spectroscopy of Thiocyanate Probes. *J. Phys. Chem. B* **2012**, *116*, 9326−9336.
(144) Schneider, S. H.; Boxer, S. G. Vibrational Stark Effects of Carbonyl Probes Applied to Reinterpret IR and Raman Data for Enzyme Inhibitors in Terms of Electric Fields at the Active Site. *J. Phys. Chem. B* **2016**, *120*, 9672−9684.
(145) Sandberg, D. J.; Rudnitskaya, A. N.; Gascón, J. A. QM/MM Prediction of the Stark Shift in the Active Site of a Protein. *J. Chem. Theory Comput.* **2012**, *8*, 2817−2823.
(146) Ringer, A. L.; MacKerell, A. D. Calculation of the Vibrational Stark Effect Using a First-Principles Quantum Mechanical/Molecular Mechanical Approach. *J. Phys. Chem. Lett.* **2011**, *2*, 553−556.
(147) Warshel, A.; Weiss, R. M. An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. *J. Am. Chem. Soc.* **1980**, *102*, 6218−6226.
(148) Oanca, G.; van der Ent, F.; Åqvist, J. Efficient Empirical Valence Bond Simulations with GROMACS. *J. Chem. Theory Comput.* **2023**, *19*, 6037−6045.
(149) Marcus, R. A. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. I. *J. Chem. Phys.* **1956**, *24*, 966−978.
(150) Villà, J.; Warshel, A. Energetics and Dynamics of Enzymatic Reactions. *J. Phys. Chem. B* **2001**, *105*, 7887−7907.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

(151) Asadi, M.; Warshel, A. Analyzing the Reaction of Orotidine 5′-Phosphate Decarboxylase as a Way to Examine Some Key Catalytic Proposals. *J. Am. Chem. Soc.* **2023**, *145*, 1334−1341.
(152) Prah, A.; Purg, M.; Stare, J.; Vianello, R.; Mavri, J. How Monoamine Oxidase A Decomposes Serotonin: An Empirical Valence Bond Simulation of the Reactive Step. *J. Phys. Chem. B* **2020**, *124*, 8259−8265.
(153) Lameira, J.; Bora, R. P.; Chu, Z. T.; Warshel, A. Methyltransferases Do Not Work by Compression, Cratic, or Desolvation Effects, but by Electrostatic Preorganization. *Proteins: Struct., Funct., Bioinf.* **2015**, *83*, 318−330.
(154) Sharma, P. K.; Chu, Z. T.; Olsson, M. H. M.; Warshel, A. A New Paradigm for Electrostatic Catalysis of Radical Reactions in Vitamin B $_{12}$ Enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 9661−9666.
(155) Schopf, P.; Mills, M. J. L.; Warshel, A. The Entropic Contributions in Vitamin B_{12} Enzymes Still Reflect the Electrostatic Paradigm. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 4328−4333.
(156) Frushicheva, M. P.; Warshel, A. Towards Quantitative Computer-Aided Studies of Enzymatic Enantioselectivity: The Case of *Candida Antarctica* Lipase A. *ChemBioChem.* **2012**, *13*, 215−223.
(157) B, R. P.; Plotnikov, N. V.; Lameira, J.; Warshel, A. Quantitative Exploration of the Molecular Origin of the Activation of GTPase. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 20509−20514.
(158) Liu, H.; Warshel, A. The Catalytic Effect of Dihydrofolate Reductase and Its Mutants Is Determined by Reorganization Energies. *Biochemistry* **2007**, *46*, 6011−6025.
(159) Adamczyk, A. J.; Cao, J.; Kamerlin, S. C. L.; Warshel, A. Catalysis by Dihydrofolate Reductase and Other Enzymes Arises from Electrostatic Preorganization, Not Conformational Motions. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 14115−14120.
(160) Singh, M. K.; Chu, Z. T.; Warshel, A. Simulating the Catalytic Effect of a Designed Mononuclear Zinc Metalloenzyme That Catalyzes the Hydrolysis of Phosphate Triesters. *J. Phys. Chem. B* **2014**, *118*, 12146−12152.
(161) 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. *Biochemistry* **2007**, *46*, 1466−1476.
(162) Roca, M.; Vardi-Kilshtain, A.; Warshel, A. Toward Accurate Screening in Computer-Aided Enzyme Design. *Biochemistry* **2009**, *48*, 3046−3056.
(163) Fuxreiter, M.; Mones, L. *Theory and Applications of the Empirical Valence Bond Approach*; John Wiley & Sons, Ltd, 2017; Chapter 7, pp 173−198.
(164) Hong, N.-S.; Petrović, D.; Lee, R.; Gryn'ova, G.; Purg, M.; Saunders, J.; Bauer, P.; Carr, P. D.; Lin, C.-Y.; Mabbitt, P. D.; et al. The Evolution of Multiple Active Site Configurations in a Designed Enzyme. *Nat. Commun.* **2018**, *9*, 3900.
(165) Kamerlin, S. C. L.; Warshel, A. Reply to Karplus: Conformational Dynamics Have No Role in the Chemical Step. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, No. e72.
(166) Mincer, J. S.; Schwartz, S. D. Protein Promoting Vibrations in Enzyme CatalysisA Conserved Evolutionary Motif. *J. Proteome Res.* **2003**, *2*, 437−439.
(167) Antoniou, D.; Schwartz, S. D. Role of Protein Motions in Catalysis by Formate Dehydrogenase. *J. Phys. Chem. B* **2020**, *124*, 9483−9489.
(168) 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*, 16670−16680.
(169) Warshel, A.; Kato, M.; Pisliakov, A. V. Polarizable Force Fields: History, Test Cases, and Prospects. *J. Chem. Theory Comput.* **2007**, *3*, 2034−2045.
(170) Ren, P.; Chun, J.; Thomas, D. G.; Schnieders, M. J.; Marucho, M.; Zhang, J.; Baker, N. A. Biomolecular Electrostatics and Solvation: A Computational Perspective. *Q. Rev. Biophys.* **2012**, *45*, 427−491.
(171) Cisneros, G. A.; Karttunen, M.; Ren, P.; Sagui, C. Classical Electrostatics for Biomolecular Simulations. *Chem. Rev.* **2014**, *114*, 779−814.
(172) Vianello, R.; Domene, C.; Mavri, J. The Use of Multiscale Molecular Simulations in Understanding a Relationship between the Structure and Function of Biological Systems of the Brain: The Application to Monoamine Oxidase Enzymes. *Frontiers in Neuroscience* **2016**, *10*, 327.
(173) Zhou, H.-X.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **2018**, *118*, 1691−1741.
(174) Huang, J.; Lopes, P. E. M.; Roux, B.; MacKerell, A. D. Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model. *J. Phys. Chem. Lett.* **2014**, *5*, 3144−3150.
(175) Baker, C. M. Polarizable Force Fields for Molecular Dynamics Simulations of Biomolecules. *WIREs Computational Molecular Science* **2015**, *5*, 241−254.
(176) Jing, Z.; Liu, C.; Cheng, S. Y.; Qi, R.; Walker, B. D.; Piquemal, J.-P.; Ren, P. Polarizable Force Fields for Biomolecular Simulations: Recent Advances and Applications. *Annual Review of Biophysics* **2019**, *48*, 371−394.
(177) Lin, F.-Y.; MacKerell, A. D. In *Biomolecular Simulations: Methods and Protocols*; Bonomi, M., Camilloni, C., Eds.; Methods Mol. Biol.; Springer: New York, NY, 2019; pp 21−54.
(178) Inakollu, V. S.; Geerke, D. P.; Rowley, C. N.; Yu, H. Polarisable Force Fields: What Do They Add in Biomolecular Simulations? *Curr. Opin. Struct. Biol.* **2020**, *61*, 182−190.
(179) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107*, 5933−5947.
(180) Ren, P.; Ponder, J. W. Temperature and Pressure Dependence of the AMOEBA Water Model. *J. Phys. Chem. B* **2004**, *108*, 13427−13437.
(181) Ponder, J. W.; Wu, C.; Ren, P.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; et al. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, *114*, 2549−2564.
(182) Ren, P.; Wu, C.; Ponder, J. W. Polarizable Atomic Multipole-based Molecular Mechanics for Organic Molecules. *J. Chem. Theory Comput.* **2011**, *7*, 3143−3161.
(183) Lemkul, J. A.; Huang, J.; Roux, B.; MacKerell, A. D. An Empirical Polarizable Force Field Based on the Classical Drude Oscillator Model: Development History and Recent Applications. *Chem. Rev.* **2016**, *116*, 4983−5013.
(184) Patel, S.; Brooks, C. L., III CHARMM Fluctuating Charge Force Field for Proteins: I Parameterization and Application to Bulk Organic Liquid Simulations. *J. Comput. Chem.* **2004**, *25*, 1−16.
(185) Chen, J.; Martínez, T. J. QTPIE: Charge Transfer with Polarization Current Equalization. A Fluctuating Charge Model with Correct Asymptotics. *Chem. Phys. Lett.* **2007**, *438*, 315−320.
(186) Chen, J.; Hundertmark, D.; Martínez, T. J. A Unified Theoretical Framework for Fluctuating-Charge Models in Atom-Space and in Bond-Space. *J. Chem. Phys.* **2008**, *129*, 214113.
(187) Huang, J.; Simmonett, A. C.; Pickard, F. C.; MacKerell, A. D.; Brooks, B. R. Mapping the Drude Polarizable Force Field onto a Multipole and Induced Dipole Model. *J. Chem. Phys.* **2017**, *147*, 161702.
(188) Zhao, D.-X.; Liu, C.; Wang, F.-F.; Yu, C.-Y.; Gong, L.-D.; Liu, S.-B.; Yang, Z.-Z. Development of a Polarizable Force Field Using Multiple Fluctuating Charges per Atom. *J. Chem. Theory Comput.* **2010**, *6*, 795−804.
(189) Stern, H. A.; Rittner, F.; Berne, B. J.; Friesner, R. A. Combined Fluctuating Charge and Polarizable Dipole Models: Application to a Five-Site Water Potential Function. *J. Chem. Phys.* **2001**, *115*, 2237−2251.
(190) Patel, S.; Brooks, C. L., III Fluctuating Charge Force Fields: Recent Developments and Applications from Small Molecules to Macromolecular Biological Systems. *Mol. Simul.* **2006**, *32*, 231−249.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

(191) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545−1614.
(192) Jana, M.; MacKerell, A. D. J. CHARMM Drude Polarizable Force Field for Aldopentofuranoses and Methyl-aldopentofuranosides. *J. Phys. Chem. B* **2015**, *119*, 7846−7859.
(193) Patel, D. S.; He, X.; MacKerell, A. D. J. Polarizable Empirical Force Field for Hexopyranose Monosaccharides Based on the Classical Drude Oscillator. *J. Phys. Chem. B* **2015**, *119*, 637−652.
(194) Small, M. C.; Aytenfisu, A. H.; Lin, F.-Y.; He, X.; MacKerell, A. D. Drude Polarizable Force Field for Aliphatic Ketones and Aldehydes, and Their Associated Acyclic Carbohydrates. *Journal of Computer-Aided Molecular Design* **2017**, *31*, 349−363.
(195) Case, D. A.; Aktulga, H. M.; Belfon, K.; Cerutti, D. S.; Cisneros, G. A.; Cruzeiro, V. W. D.; Forouzesh, N.; Giese, T. J.; Götz, A. W.; Gohlke, H.; et al. AmberTools. *J. Chem. Inf. Model.* **2023**, *63*, 6183−6191.
(196) Cieplak, P.; Caldwell, J.; Kollman, P. Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases. *J. Comput. Chem.* **2001**, *22*, 1048−1057.
(197) Wang, Z.-X.; Zhang, W.; Wu, C.; Lei, H.; Cieplak, P.; Duan, Y. Strike a Balance: Optimization of Backbone Torsion Parameters of AMBER Polarizable Force Field for Simulations of Proteins and Peptides. *J. Comput. Chem.* **2006**, *27*, 781−790.
(198) Wang, J.; Cieplak, P.; Li, J.; Cai, Q.; Hsieh, M.-J.; Luo, R.; Duan, Y. Development of Polarizable Models for Molecular Mechanical Calculations. 4. van Der Waals Parametrization. *J. Phys. Chem. B* **2012**, *116*, 7088−7101.
(199) Grossfield, A.; Ren, P.; Ponder, J. W. Ion Solvation Thermodynamics from Simulation with a Polarizable Force Field. *J. Am. Chem. Soc.* **2003**, *125*, 15671−15682.
(200) Gresh, N.; Cisneros, G. A.; Darden, T. A.; Piquemal, J.-P. Polarizable Molecular Mechanics Studies Anisotropic, Polarizable Molecular Mechanics Studies of Inter- and Intramolecular Interactions and Ligand−Macromolecule Complexes. A Bottom-Up Strategy. *J. Chem. Theory Comput.* **2007**, *3*, 1960−1986.
(201) Naseem-Khan, S.; Lagardère, L.; Narth, C.; Cisneros, G. A.; Ren, P.; Gresh, N.; Piquemal, J.-P. Development of the Quantum-Inspired SIBFA Many-Body Polarizable Force Field: Enabling Condensed-Phase Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2022**, *18*, 3607−3621.
(202) Goddard, W. A., III; Zhang, Q.; Uludogan, M.; Strachan, A.; Cagin, T. The ReaxFF Polarizable Reactive Force Fields for Molecular Dynamics Simulation of Ferroelectrics. *AIP Conf. Proc.* **2002**, *626*, 45−55.
(203) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; et al. LAMMPS - a Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Comput. Phys. Commun.* **2022**, *271*, 108171.
(204) Polarizable Models — LAMMPS Documentation. https://docs.lammps.org/Howto_polarizable.html, 2024.
(205) Rappe, A. K.; Goddard, W. A. I. Charge Equilibration for Molecular Dynamics Simulations. *J. Phys. Chem.* **1991**, *95*, 3358−3363.
(206) Rick, S. W.; Stuart, S. J.; Berne, B. J. Dynamical Fluctuating Charge Force Fields: Application to Liquid Water. *J. Chem. Phys.* **1994**, *101*, 6141−6156.
(207) Nakano, A. Parallel Multilevel Preconditioned Conjugate-Gradient Approach to Variable-Charge Molecular Dynamics. *Comput. Phys. Commun.* **1997**, *104*, 59−69.
(208) Mitchell, P. J.; Fincham, D. Shell Model Simulations by Adiabatic Dynamics. *J. Phys.: Condens. Matter* **1993**, *5*, 1031.
(209) Fincham, D.; Mackrodt, W. C.; Mitchell, P. J. MgO at High Temperatures and Pressures: Shell-Model Lattice Dynamics and Molecular Dynamics. *J. Phys.: Condens. Matter* **1994**, *6*, 393.
(210) Lamoureux, G.; Roux, B. Modeling Induced Polarization with Classical Drude Oscillators: Theory and Molecular Dynamics Simulation Algorithm. *J. Chem. Phys.* **2003**, *119*, 3025−3039.
(211) Lemkul, J. A.; MacKerell, A. D. J. Polarizable Force Field for DNA Based on the Classical Drude Oscillator: I. Refinement Using Quantum Mechanical Base Stacking and Conformational Energetics. *J. Chem. Theory Comput.* **2017**, *13*, 2053−2071.
(212) Li, H.; Chowdhary, J.; Huang, L.; He, X.; MacKerell, A. D. J.; Roux, B. Drude Polarizable Force Field for Molecular Dynamics Simulations of Saturated and Unsaturated Zwitterionic Lipids. *J. Chem. Theory Comput.* **2017**, *13*, 4535−4552.
(213) Zhang, C.; Lu, C.; Jing, Z.; Wu, C.; Piquemal, J.-P.; Ponder, J. W.; Ren, P. AMOEBA Polarizable Atomic Multipole Force Field for Nucleic Acids. *J. Chem. Theory Comput.* **2018**, *14*, 2084−2108.
(214) Shi, Y.; Xia, Z.; Zhang, J.; Best, R.; Wu, C.; Ponder, J. W.; Ren, P. Polarizable Atomic Multipole-Based AMOEBA Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 4046−4063.
(215) Gresh, N.; Perahia, D.; de Courcy, B.; Foret, J.; Roux, C.; El-Khoury, L.; Piquemal, J.-P.; Salmon, L. Complexes of a Zn-metalloenzyme Binding Site with Hydroxamate-Containing Ligands. A Case for Detailed Benchmarkings of Polarizable Molecular Mechanics/Dynamics Potentials When the Experimental Binding Structure Is Unknown. *J. Comput. Chem.* **2016**, *37*, 2770−2782.
(216) Hage, K. E.; Piquemal, J.-P.; Hobaika, Z.; Maroun, R. G.; Gresh, N. Could an Anisotropic Molecular Mechanics/Dynamics Potential Account for Sigma Hole Effects in the Complexes of Halogenated Compounds? *J. Comput. Chem.* **2013**, *34*, 1125−1135.
(217) Gresh, N.; Sponer, J. E.; Devereux, M.; Gkionis, K.; de Courcy, B.; Piquemal, J.-P.; Sponer, J. Stacked and H-Bonded Cytosine Dimers. Analysis of the Intermolecular Interaction Energies by Parallel Quantum Chemistry and Polarizable Molecular Mechanics. *J. Phys. Chem. B* **2015**, *119*, 9477−9495.
(218) Chenoweth, K.; van Duin, A. C. T.; Persson, P.; Cheng, M.-J.; Oxgaard, J.; Goddard, W. A. I. Development and Application of a ReaxFF Reactive Force Field for Oxidative Dehydrogenation on Vanadium Oxide Catalysts. *J. Phys. Chem. C* **2008**, *112*, 14645−14654.
(219) Shin, Y. K.; Ashraf, C. M.; van Duin, A. C. T. *In Computational Materials, Chemistry, and Biochemistry: From Bold Initiatives to the Last Mile: In Honor of William A. Goddard's Contributions to Science and Engineering*; Shankar, S., Muller, R., Dunning, T., Chen, G. H., Eds.; Springer International Publishing: Cham, 2021; pp 157−182.
(220) Zheng, M.; Li, X.; Guo, L. Algorithms of GPU-enabled Reactive Force Field (ReaxFF) Molecular Dynamics. *Journal of Molecular Graphics and Modelling* **2013**, *41*, 1−11.
(221) Plimpton, S. J.; Thompson, A. P. Computational Aspects of Many-Body Potentials. *MRS Bull.* **2012**, *37*, 513−521.
(222) Kamat, A. M.; van Duin, A. C. T.; Yakovlev, A. Molecular Dynamics Simulations of Laser-Induced Incandescence of Soot Using an Extended ReaxFF Reactive Force Field. *J. Phys. Chem. A* **2010**, *114*, 12561−12572.
(223) Aktulga, H. M.; Fogarty, J. C.; Pandit, S. A.; Grama, A. Y. Parallel Reactive Molecular Dynamics: Numerical Methods and Algorithmic Techniques. *Parallel Computing* **2012**, *38*, 245−259.
(224) Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; et al. The ReaxFF Reactive Force-Field: Development, Applications and Future Directions. *npj Computational Materials* **2016**, *2*, 15011.
(225) Kumar, A.; Simmonett, A.; MacKerell, A. D. Introduction of Polarized Embedding QMMM Technique in CHARMM with PSI4. *Biophys. J.* **2021**, *120*, 175a−176a.
(226) Lin, F.-Y.; MacKerell, A. D. J. Polarizable Empirical Force Field for Halogen-Containing Compounds Based on the Classical Drude Oscillator. *J. Chem. Theory Comput.* **2018**, *14*, 1083−1098.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

(227) Lin, F.-Y.; Huang, J.; Pandey, P.; Rupakheti, C.; Li, J.; Roux, B.; MacKerell, A. D. J. Further Optimization and Validation of the Classical Drude Polarizable Protein Force Field. *J. Chem. Theory Comput.* **2020**, *16*, 3221−3239.
(228) Wang, L.-P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117*, 9956−9972.
(229) Wang, Q.; Rackers, J. A.; He, C.; Qi, R.; Narth, C.; Lagardere, L.; Gresh, N.; Ponder, J. W.; Piquemal, J.-P.; Ren, P. General Model for Treating Short-Range Electrostatic Penetration in a Molecular Mechanics Force Field. *J. Chem. Theory Comput.* **2015**, *11*, 2609−2618.
(230) Rackers, J. A.; Wang, Q.; Liu, C.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. J. An Optimized Charge Penetration Model for Use with the AMOEBA Force Field. *Phys. Chem. Chem. Phys.* **2017**, *19*, 276−291.
(231) Qi, R.; Wang, Q.; Ren, P. General van Der Waals Potential for Common Organic Molecules. *Bioorg. Med. Chem.* **2016**, *24*, 4911−4919.
(232) Polêto, M. D.; Lemkul, J. A. Integration of Experimental Data and Use of Automated Fitting Methods in Developing Protein Force Fields. *Communications Chemistry* **2022**, *5*, 1−10.
(233) Kirsh, J. M.; Weaver, J. B.; Boxer, S. G.; Kozuch, J. Critical Evaluation of Polarizable and Nonpolarizable Force Fields for Proteins Using Experimentally Derived Nitrile Electric Fields. *J. Am. Chem. Soc.* **2024**, *146*, 6983−6991.
(234) Amin, K. S.; Hu, X.; Salahub, D. R.; Baldauf, C.; Lim, C.; Noskov, S. Benchmarking Polarizable and Non-Polarizable Force Fields for Ca^{2+}−Peptides against a Comprehensive QM Dataset. *J. Chem. Phys.* **2020**, *153*, 144102.
(235) Sakharov, D. V.; Lim, C. Zn Protein Simulations Including Charge Transfer and Local Polarization Effects. *J. Am. Chem. Soc.* **2005**, *127*, 4921−4929.
(236) Sakharov, D. V.; Lim, C. Force fields including charge transfer and local polarization effects: Application to proteins containing multi/heavy metal ions. *J. Comput. Chem.* **2009**, *30*, 191−202.
(237) Zhang, X.; Wang, Z.; Li, Z.; Shaik, S.; Wang, B. [4Fe−4S]-Mediated Proton-Coupled Electron Transfer Enables the Efficient Degradation of Chloroalkenes by Reductive Dehalogenases. *ACS Catal.* **2023**, *13*, 1173−1185.
(238) Yan, S.; Ji, X.; Peng, W.; Wang, B. Evaluating the Transition State Stabilization/Destabilization Effects of the Electric Fields from Scaffold Residues by a QM/MM Approach. *J. Phys. Chem. B* **2023**, *127*, 4245−4253.
(239) Song, C.; Wang, L.-P. A Polarizable QM/MM Model That Combines the State-Averaged CASSCF and AMOEBA Force Field for Photoreactions in Proteins. *J. Chem. Theory Comput.* **2024**, *20*, 6632−6651.
(240) Ahmadi, S.; Barrios Herrera, L.; Chehelamirani, M.; Hostaš, J.; Jalife, S.; Salahub, D. R. Multiscale Modeling of Enzymes: QM-cluster, QM/MM, and QM/MM/MD: A Tutorial Review. *Int. J. Quantum Chem.* **2018**, *118*, No. e25558.
(241) de la Lande, A.; Alvarez-Ibarra, A.; Hasnaoui, K.; Cailliez, F.; Wu, X.; Mineva, T.; Cuny, J.; Calaminici, P.; López-Sosa, L.; Geudtner, G.; et al. Molecular Simulations with In-deMon2k QM/MM, a Tutorial-Review. *Molecules* **2019**, *24*, 1653.
(242) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.* **2016**, *116*, 7898−7936.
(243) Klein, F.; Soñora, M.; Helene Santos, L.; Nazareno Frigini, E.; Ballesteros-Casallas, A.; Rodrigo Machado, M.; Pantano, S. The SIRAH Force Field: A Suite for Simulations of Complex Biological Systems at the Coarse-Grained and Multiscale Levels. *J. Struct. Biol.* **2023**, *215*, 107985.
(244) Souza, P. C. T.; Alessandri, R.; Barnoud, J.; Thallmair, S.; Faustino, I.; Grünewald, F.; Patmanidis, I.; Abdizadeh, H.; Bruininks, B. M. H.; Wassenaar, T. A.; et al. Martini 3: A General Purpose Force Field for Coarse-Grained Molecular Dynamics. *Nat. Methods* **2021**, *18*, 382−388.
(245) Liwo, A.; Czaplewski, C.; Sieradzan, A. K.; Lipska, A. G.; Samsonov, S. A.; Murarka, R. K. Theory and Practice of Coarse-Grained Molecular Dynamics of Biologically Important Systems. *Biomolecules* **2021**, *11*, 1347.
(246) Yin, Y.; Sieradzan, A. K.; Liwo, A.; He, Y.; Scheraga, H. A. Physics-Based Potentials for Coarse-Grained Modeling of Protein−DNA Interactions. *J. Chem. Theory Comput.* **2015**, *11*, 1792−1808.
(247) Kolinski, A. Protein Modeling and Structure Prediction with a Reduced Representation. *Acta Biochimica Polonica* **2019**, *51*, 349−371.
(248) Kar, P.; Gopal, S. M.; Cheng, Y.-M.; Predeus, A.; Feig, M. PRIMO: A Transferable Coarse-Grained Force Field for Proteins. *J. Chem. Theory Comput.* **2013**, *9*, 3769−3788.
(249) Cheon, M.; Chang, I.; Hall, C. K. Extending the PRIME Model for Protein Aggregation to All 20 Amino Acids. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 2950−2960.
(250) Derreumaux, P. From Polypeptide Sequences to Structures Using Monte Carlo Simulations and an Optimized Potential. *J. Chem. Phys.* **1999**, *111*, 2301−2310.
(251) Latham, A. P.; Zhang, B. Unifying Coarse-grained Force Fields for Folded and Disordered Proteins. *Curr. Opin. Struct. Biol.* **2022**, *72*, 63−70.
(252) Arnarez, C.; Uusitalo, J. J.; Masman, M. F.; Ingólfsson, H. I.; de Jong, D. H.; Melo, M. N.; Periole, X.; de Vries, A. H.; Marrink, S. J. Dry Martini, a Coarse-Grained Force Field for Lipid Membrane Simulations with Implicit Solvent. *J. Chem. Theory Comput.* **2015**, *11*, 260−275.
(253) Bryer, A. J.; Rey, J. S.; Perilla, J. R. Performance Efficient Macromolecular Mechanics via Sub-Nanometer Shape Based Coarse Graining. *Nat. Commun.* **2023**, *14*, 2014.
(254) Caceres-Delpiano, J.; Wang, L.-P.; Essex, J. W. The Automated Optimisation of a Coarse-Grained Force Field Using Free Energy Data. *Phys. Chem. Chem. Phys.* **2021**, *23*, 24842−24851.
(255) Kanada, R.; Tokuhisa, A.; Nagasaka, Y.; Okuno, S.; Amemiya, K.; Chiba, S.; Bekker, G.-J.; Kamiya, N.; Kato, K.; Okuno, Y. Enhanced Coarse-Grained Molecular Dynamics Simulation with a Smoothed Hybrid Potential Using a Neural Network Model. *J. Chem. Theory Comput.* **2024**, *20*, 7−17.
(256) Navarro, C.; Majewski, M.; De Fabritiis, G. Top-Down Machine Learning of Coarse-Grained Protein Force Fields. *J. Chem. Theory Comput.* **2023**, *19*, 7518−7526.
(257) Carrer, M.; Cezar, H. M.; Bore, S. L.; Ledum, M.; Cascella, M. Learning Force Field Parameters from Differentiable Particle-Field Molecular Dynamics. *J. Chem. Inf. Model.* **2024**, *64*, 5510−5520.
(258) Singh, N.; Li, W. Recent Advances in Coarse-Grained Models for Biomolecules and Their Applications. *International Journal of Molecular Sciences* **2019**, *20*, 3774.
(259) Srivastava, A.; Tiwari, S. P.; Miyashita, O.; Tama, F. Integrative/Hybrid Modeling Approaches for Studying Biomolecules. *J. Mol. Biol.* **2020**, *432*, 2846−2860.
(260) Roel-Touris, J.; Bonvin, A. M. J. J. Coarse-Grained (Hybrid) Integrative Modeling of Biomolecular Interactions. *Computational and Structural Biotechnology Journal* **2020**, *18*, 1182−1190.
(261) Behler, J. Four Generations of High-Dimensional Neural Network Potentials. *Chem. Rev.* **2021**, *121*, 10037−10072.
(262) Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*, 146401.
(263) Ghasemi, S. A.; Hofstetter, A.; Saha, S.; Goedecker, S. Interatomic Potentials for Ionic Systems with Density Functional Accuracy Based on Charge Densities Obtained by a Neural Network. *Phys. Rev. B* **2015**, *92*, 045131.
(264) Xie, X.; Persson, K. A.; Small, D. W. Incorporating Electronic Information into Machine Learning Potential Energy Surfaces via Approaching the Ground-State Electronic Energy as a Function of Atom-Based Electronic Populations. *J. Chem. Theory Comput.* **2020**, *16*, 4256−4270.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

(265) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, Th.; Suhai, S.; Seifert, G. Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties. *Phys. Rev. B* **1998**, *58*, 7260−7268.
(266) Kipf, T. N.; Welling, M. *Semi-Supervised Classification with Graph Convolutional Networks. 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24−26, 2017, Conference Track Proceedings*. 2017.
(267) Kozinsky, B.; Musaelian, A.; Johansson, A.; Batzner, S. Scaling the Leading Accuracy of Deep Equivariant Models to Biomolecular Simulations of Realistic Size. Association for Computing Machinery. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. New York, NY, USA, 2023; pp 1−12.
(268) Unke, O. T.; Stöhr, M.; Ganscha, S.; Unterthiner, T.; Maennel, H.; Kashubin, S.; Ahlin, D.; Gastegger, M.; Medrano Sandonas, L.; Berryman, J. T.; et al. Biomolecular Dynamics with Machine-Learned Quantum-Mechanical Force Fields Trained on Diverse Chemical Fragments. 10.1126/Sciadv.Adn4397. *Science Advances* **2024**, *10*, No. eadn4397.
(269) Christensen, A. S.; Faber, F. A.; von Lilienfeld, O. A. Operators in Quantum Machine Learning: Response Properties in Chemical Space. *J. Chem. Phys.* **2019**, *150*, 064105.
(270) Gastegger, M.; Schütt, K. T.; Müller, K.-R. Machine Learning of Solvent Effects on Molecular Spectra and Reactions. *Chemical Science* **2021**, *12*, 11473−11483.
(271) Gao, A.; Remsing, R. C. Self-Consistent Determination of Long-Range Electrostatics in Neural Network Potentials. *Nat. Commun.* **2022**, *13*, 1572.
(272) Zhang, Y.; Jiang, B. Universal Machine Learning for the Response of Atomistic Systems to External Fields. *Nat. Commun.* **2023**, *14*, 6424.
(273) Zinovjev, K. Electrostatic Embedding of Machine Learning Potentials. *J. Chem. Theory Comput.* **2023**, *19*, 1888−1897.
(274) Giese, T. J.; Zeng, J.; Ekesan, Ş.; York, D. M. Combined QM/MM, Machine Learning Path Integral Approach to Compute Free Energy Profiles and Kinetic Isotope Effects in RNA Cleavage Reactions. *J. Chem. Theory Comput.* **2022**, *18*, 4304−4317.
(275) Jaffrelot Inizan, T.; Plé, T.; Adjoua, O.; Ren, P.; Gökcan, H.; Isayev, O.; Lagardère, L.; Piquemal, J.-P. Scalable Hybrid Deep Neural Networks/Polarizable Potentials Biomolecular Simulations Including Long-Range Effects. *Chemical Science* **2023**, *14*, 5438−5452.
(276) Heal, J. W.; Bartlett, G. J.; Wood, C. W.; Thomson, A. R.; Woolfson, D. N. Applying Graph Theory to Protein Structures: An Atlas of Coiled Coils. *Bioinformatics* **2018**, *34*, 3316−3323.
(277) Kantelis, K. F.; Asteriou, V.; Papadimitriou-Tsantarliotou, A.; Petrou, A.; Angelis, L.; Nicopolitidis, P.; Papadimitriou, G.; Vizirianakis, I. S. Graph Theory-Based Simulation Tools for Protein Structure Networks. *Simulation Modelling Practice and Theory* **2022**, *121*, 102640.
(278) Gligorijević, V.; Renfrew, P. D.; Kosciolek, T.; Leman, J. K.; Berenberg, D.; Vatanen, T.; Chandler, C.; Taylor, B. C.; Fisk, I. M.; Vlamakis, H.; et al. Structure-Based Protein Function Prediction Using Graph Convolutional Networks. *Nat. Commun.* **2021**, *12*, 3168.
(279) Zhou, Z.; Hu, G. Applications of Graph Theory in Studying Protein Structure, Dynamics, and Interactions. *J. Math. Chem.* **2024**, *62*, 2562.
(280) Meng, X.-Y.; Zhang, H.-X.; Mezei, M.; Cui, M. Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current computer-aided drug design* **2011**, *7*, 146−157.
(281) Tian, J.; Woodard, J. C.; Whitney, A.; Shakhnovich, E. I. Thermal Stabilization of Dihydrofolate Reductase Using Monte Carlo Unfolding Simulations and Its Functional Consequences. *PLOS Computational Biology* **2015**, *11*, No. e1004207.
(282) Kuhlman, B.; Bradley, P. Advances in Protein Structure Prediction and Design. *Nature reviews. Molecular cell biology* **2019**, *20*, 681−697.
(283) Zhou, J.; Huang, M. Navigating the Landscape of Enzyme Design: From Molecular Simulations to Machine Learning. *Chem. Soc. Rev.* **2024**, *53*, 8202−8239.
(284) Gilabert, J. F.; Lecina, D.; Estrada, J.; Guallar, V. *Biomolecular Simulations in Structure-Based Drug Discovery*; John Wiley & Sons, Ltd, 2018; Chapter 5, pp 87−103.
(285) Opuu, V.; Simonson, T. Enzyme Redesign and Genetic Code Expansion. *Protein Engineering, Design and Selection* **2023**, *36*, gzad017.
(286) Cabeza de Vaca, I.; Qian, Y.; Vilseck, J. Z.; Tirado-Rives, J.; Jorgensen, W. L. Enhanced Monte Carlo Methods for Modeling Proteins Including Computation of Absolute Free Energies of Binding. *J. Chem. Theory Comput.* **2018**, *14*, 3279−3288.
(287) Heilmann, N.; Wolf, M.; Kozlowska, M.; Sedghamiz, E.; Setzler, J.; Brieg, M.; Wenzel, W. Sampling of the Conformational Landscape of Small Proteins with Monte Carlo Methods. *Sci. Rep.* **2020**, *10*, 18211.
(288) Pan, X.; Kortemme, T. Recent Advances in de Novo Protein Design: Principles, Methods, and Applications. *J. Biol. Chem.* **2021**, *296*, 100558.
(289) Pesce, F.; Bremer, A.; Tesei, G.; Hopkins, J. B.; Grace, C. R.; Mittag, T.; Lindorff-Larsen, K. Design of Intrinsically Disordered Protein Variants with Diverse Structural Properties. *Science Advances* **2024**, *10*, No. eadm9926.
(290) Ross, G. A.; Bruce Macdonald, H. E.; Cave-Ayland, C.; Cabedo Martinez, A. I.; Essex, J. W. Replica-Exchange and Standard State Binding Free Energies with Grand Canonical Monte Carlo. *J. Chem. Theory Comput.* **2017**, *13*, 6373−6381.
(291) Bodnarchuk, M. S.; Packer, M. J.; Haywood, A. Utilizing Grand Canonical Monte Carlo Methods in Drug Discovery. *ACS Med. Chem. Lett.* **2020**, *11*, 77−82.
(292) Gasic, A. G.; Sarkar, A.; Cheung, M. S. Understanding Protein-Complex Assembly through Grand Canonical Maximum Entropy Modeling. *Physical Review Research* **2021**, *3*, 033220.
(293) Ge, Y.; Melling, O. J.; Dong, W.; Essex, J. W.; Mobley, D. L. Enhancing Sampling of Water Rehydration upon Ligand Binding Using Variants of Grand Canonical Monte Carlo. *Journal of Computer-Aided Molecular Design* **2022**, *36*, 767−779.
(294) Ge, Y.; Wych, D. C.; Samways, M. L.; Wall, M. E.; Essex, J. W.; Mobley, D. L. Enhancing Sampling of Water Rehydration on Ligand Binding: A Comparison of Techniques. *J. Chem. Theory Comput.* **2022**, *18*, 1359−1381.
(295) Melling, O. J.; Samways, M. L.; Ge, Y.; Mobley, D. L.; Essex, J. W. Enhanced Grand Canonical Sampling of Occluded Water Sites Using Nonequilibrium Candidate Monte Carlo. *J. Chem. Theory Comput.* **2023**, *19*, 1050−1062.
(296) Samways, M. L.; Bruce Macdonald, H. E.; Taylor, R. D.; Essex, J. W. Water Networks in Complexes between Proteins and FDA-Approved Drugs. *J. Chem. Inf. Model.* **2023**, *63*, 387−396.
(297) Zhang, W.; Bell, E. W.; Yin, M.; Zhang, Y. EDock: Blind Protein−Ligand Docking by Replica-Exchange Monte Carlo Simulation. *Journal of Cheminformatics* **2020**, *12*, 37.
(298) Siebenmorgen, T.; Engelhard, M.; Zacharias, M. Prediction of Protein−Protein Complexes Using Replica Exchange with Repulsive Scaling. *J. Comput. Chem.* **2020**, *41*, 1436−1447.
(299) Mortuza, S. M.; Zheng, W.; Zhang, C.; Li, Y.; Pearce, R.; Zhang, Y. Improving Fragment-Based Ab Initio Protein Structure Assembly Using Low-Accuracy Contact-Map Predictions. *Nat. Commun.* **2021**, *12*, 5011.
(300) Harmalkar, A.; Mahajan, S. P.; Gray, J. J. Induced Fit with Replica Exchange Improves Protein Complex Structure Prediction. *PLOS Computational Biology* **2022**, *18*, No. e1010124.
(301) Tang, Y.; Moretti, R.; Meiler, J. Recent Advances in Automated Structure-Based De Novo Drug Design. *J. Chem. Inf. Model.* **2024**, *64*, 1794−1805.
(302) Onufriev, A. V.; Case, D. A. Generalized Born Implicit Solvent Models for Biomolecules. *Annual Review of Biophysics* **2019**, *48*, 275−296.
(303) Dong, L.; Qu, X.; Zhao, Y.; Wang, B. Prediction of Binding Free Energy of Protein−Ligand Complexes with a Hybrid Molecular

Mechanics/Generalized Born Surface Area and Machine Learning Method. *ACS Omega* **2021**, *6*, 32938−32947.
(304) Nyambo, K.; Tapfuma, K. I.; Adu-Amankwaah, F.; Julius, L.; Baatjies, L.; Niang, I. S.; Smith, L.; Govender, K. K.; Ngxande, M.; Watson, D. J.; et al. Molecular Docking, Molecular Dynamics Simulations and Binding Free Energy Studies of Interactions between Mycobacterium Tuberculosis Pks13, PknG and Bioactive Constituents of Extremophilic Bacteria. *Sci. Rep.* **2024**, *14*, 6794.
(305) Virtanen, S. I.; Niinivehmas, S. P.; Pentikäinen, O. T. Case-Specific Performance of MM-PBSA, MM-GBSA, and SIE in Virtual Screening. *Journal of Molecular Graphics and Modelling* **2015**, *62*, 303−318.
(306) Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. *Expert Opinion on Drug Discovery* **2015**, *10*, 449−461.
(307) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119*, 9478−9508.
(308) Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilizaion of Ab Initio Molecular Potentials for the Prevision of Solvent Effects. *Chem. Phys.* **1981**, *55*, 117−129.
(309) Liu, F.; Luehr, N.; Kulik, H. J.; Martínez, T. J. Quantum Chemistry for Solvated Molecules on Graphical Processing Units Using Polarizable Continuum Models. *J. Chem. Theory Comput.* **2015**, *11*, 3131−3144.
(310) Wang, X.; Li, Y.; Gao, Y.; Yang, Z.; Lu, C.; Zhu, T. A Quantum Mechanical Computational Method for Modeling Electrostatic and Solvation Effects of Protein. *Sci. Rep.* **2018**, *8*, 5475.
(311) Michael, E.; Saint-Jalme, R.; Mignon, D.; Simonson, T. Computational Protein Design Repurposed to Explore Enzyme Vitality and Help Predict Antibiotic Resistance. *Frontiers in Molecular Biosciences* **2023**, *9*, 905588.
(312) Ennist, N. M.; Wang, S.; Kennedy, M. A.; Curti, M.; Sutherland, G. A.; Vasilev, C.; Redler, R. L.; Maffeis, V.; Shareef, S.; Sica, A. V.; et al. De Novo Design of Proteins Housing Excitonically Coupled Chlorophyll Special Pairs. *Nat. Chem. Biol.* **2024**, *20*, 906−915.
(313) He, X.; Zhu, T.; Wang, X.; Liu, J.; Zhang, J. Z. H. Fragment Quantum Mechanical Calculation of Proteins and Its Applications. *Acc. Chem. Res.* **2014**, *47*, 2748−2757.
(314) Raghavachari, K.; Saha, A. Accurate Composite and Fragment-Based Quantum Chemical Models for Large Molecules. *Chem. Rev.* **2015**, *115*, 5643−5677.
(315) Jin, X.; Glover, W. J.; He, X. Fragment Quantum Mechanical Method for Excited States of Proteins: Development and Application to the Green Fluorescent Protein. *J. Chem. Theory Comput.* **2020**, *16*, 5174−5188.
(316) Vornweg, J. R.; Wolter, M.; Jacob, C. R. A Simple and Consistent Quantum-Chemical Fragmentation Scheme for Proteins That Includes Two-Body Contributions. *J. Comput. Chem.* **2023**, *44*, 1634−1644.
(317) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. The Effective Fragment Potential Method: A QM-Based MM Approach to Modeling Environmental Effects in Chemistry. *J. Phys. Chem. A* **2001**, *105*, 293−307.
(318) Gurunathan, P. K.; Acharya, A.; Ghosh, D.; Kosenkov, D.; Kaliman, I.; Shao, Y.; Krylov, A. I.; Slipchenko, L. V. Extension of the Effective Fragment Potential Method to Macromolecules. *J. Phys. Chem. B* **2016**, *120*, 6562−6574.
(319) Slipchenko, L. V.; Gurunathan, P. K. *Fragmentation*; John Wiley & Sons, Ltd, 2017; Chapter 6, pp 183−208.
(320) Viquez Rojas, C. I.; Slipchenko, L. V. Exchange Repulsion in Quantum Mechanical/Effective Fragment Potential Excitation Energies: Beyond Polarizable Embedding. *J. Chem. Theory Comput.* **2020**, *16*, 6408−6417.
(321) Pranami, G.; Slipchenko, L.; Lamm, M. H.; Gordon, M. S. In *Multi-Scale Quantum Models for Biocatalysis: Modern Techniques and Applications*; York, D. M., Lee, T.-S., Eds.; Springer Netherlands: Dordrecht, 2009; pp 197−218.
(322) Adamovic, I.; Freitag, M. A.; Gordon, M. S. Density Functional Theory Based Effective Fragment Potential Method. *J. Chem. Phys.* **2003**, *118*, 6725−6732.
(323) Netzloff, H. M.; Gordon, M. S. Fast Fragments: The Development of a Parallel Effective Fragment Potential Method. *J. Comput. Chem.* **2004**, *25*, 1926−1936.
(324) Li, H.; Gordon, M. S.; Jensen, J. H. Charge Transfer Interaction in the Effective Fragment Potential Method. *J. Chem. Phys.* **2006**, *124*, 214108.
(325) Slipchenko, L. V.; Gordon, M. S. Damping Functions in the Effective Fragment Potential Method. *Mol. Phys.* **2009**, *107*, 999−1016.
(326) Bertoni, C.; Slipchenko, L. V.; Misquitta, A. J.; Gordon, M. S. Multipole Moments in the Effective Fragment Potential Method. *J. Phys. Chem. A* **2017**, *121*, 2056−2067.
(327) Adamovic, I.; Gordon, M. S. Dynamic Polarizability, Dispersion Coefficient C6 and Dispersion Energy in the Effective Fragment Potential Method. *Mol. Phys.* **2005**, *103*, 379−387.
(328) Xu, P.; Zahariev, F.; Gordon, M. S. The R−7 Dispersion Interaction in the General Effective Fragment Potential Method. *J. Chem. Theory Comput.* **2014**, *10*, 1576−1587.
(329) Xu, P.; Leonard, S. L.; O'Brien, W.; Gordon, M. S. R−8 Dispersion Interaction: Derivation and Application to the Effective Fragment Potential Method. *J. Phys. Chem. A* **2024**, *128*, 292−327.
(330) Tazhigulov, R. N.; Gurunathan, P. K.; Kim, Y.; Slipchenko, L. V.; Bravaya, K. B. Polarizable Embedding for Simulating Redox Potentials of Biomolecules. *Phys. Chem. Chem. Phys.* **2019**, *21*, 11642−11650.
(331) Slipchenko, L. V. Detangling Solvatochromic Effects by the Effective Fragment Potential Method. *J. Phys. Chem. A* **2024**, *128*, 656−669.
(332) Sattasathuchana, T.; Xu, P.; Bertoni, C.; Kim, Y. L.; Leang, S. S.; Pham, B. Q.; Gordon, M. S. The Effective Fragment Molecular Orbital Method: Achieving High Scalability and Accuracy for Large Systems. *J. Chem. Theory Comput.* **2024**, *20*, 2445−2461.
(333) Debye, P.; Hückel, E. Zur Theorie Der Electrolyte. *Phyikalishce Zeitschrift.* **1923**, *9*, 185−206.
(334) Gray, C. G.; Stiles, P. J. Nonlinear Electrostatics: The Poisson−Boltzmann Equation. *European Journal of Physics* **2018**, *39*, 053002.
(335) Blossey, R. *The Poisson-Boltzmann Equation: An Introduction*; SpringerBriefs in Physics; Springer International Publishing: Cham, 2023.
(336) Ullmann, G. M.; Knapp, E.-W. Electrostatic Models for Computing Protonation and Redox Equilibria in Proteins. *Eur. Biophys. J.* **1999**, *28*, 533−551.
(337) Meyer, T.; Knapp, E.-W. pKa Values in Proteins Determined by Electrostatics Applied to Molecular Dynamics Trajectories. *J. Chem. Theory Comput.* **2015**, *11*, 2827−2840.
(338) Aleksandrov, A.; Roux, B.; MacKerell, A. D. J. pKa Calculations with the Polarizable Drude Force Field and Poisson−Boltzmann Solvation Model. *J. Chem. Theory Comput.* **2020**, *16*, 4655−4668.
(339) Briskot, T.; Hillebrandt, N.; Kluters, S.; Wang, G.; Studts, J.; Hahn, T.; Huuk, T.; Hubbuch, J. Modeling the Gibbs−Donnan Effect during Ultrafiltration and Diafiltration Processes Using the Poisson−Boltzmann Theory in Combination with a Basic Stern Model. *J. Membr. Sci.* **2022**, *648*, 120333.
(340) Gama, M. d. S.; Barreto, A. G.; Tavares, F. W. The Binding Interaction of Protein on a Charged Surface Using Poisson−Boltzmann Equation: Lysozyme Adsorption onto SBA-15. *Adsorption* **2021**, *27*, 1137−1148.
(341) Alexov, E.; Mehler, E. L.; Baker, N.; Baptista, A. M.; Huang, Y.; Milletti, F.; Erik Nielsen, J.; Farrell, D.; Carstensen, T.; Olsson, M. H. M.; et al. Progress in the Prediction of pKa Values in Proteins. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 3260−3275.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

(342) Coskun, D.; Chen, W.; Clark, A. J.; Lu, C.; Harder, E. D.; Wang, L.; Friesner, R. A.; Miller, E. B. Reliable and Accurate Prediction of Single-Residue pKa Values through Free Energy Perturbation Calculations. *J. Chem. Theory Comput.* **2022**, *18*, 7193−7204.
(343) Onufriev, A. V.; Alexov, E. Protonation and pK Changes in Protein−Ligand Binding. *Q. Rev. Biophys.* **2013**, *46*, 181−209.
(344) Kim, M. O.; McCammon, J. A. Computation of pH-dependent Binding Free Energies. *Biopolymers* **2016**, *105*, 43−49.
(345) Wang, L.; Li, L.; Alexov, E. pKa Predictions for Proteins, RNAs, and DNAs with the Gaussian Dielectric Function Using DelPhi pKa. *Proteins: Struct., Funct., Bioinf.* **2015**, *83*, 2186−2197.
(346) Gokcan, H.; Isayev, O. Prediction of Protein pKa with Representation Learning. *Chemical Science* **2022**, *13*, 2462−2474.
(347) Reis, P. B. P. S.; Vila-Viçosa, D.; Rocchia, W.; Machuqueiro, M. PypKa: A Flexible Python Module for Poisson−Boltzmann-Based pKa Calculations. *J. Chem. Inf. Model.* **2020**, *60*, 4442−4448.
(348) Fogolari, F.; Brigo, A.; Molinari, H. The Poisson−Boltzmann Equation for Biomolecular Electrostatics: A Tool for Structural Biology. *Journal of Molecular Recognition* **2002**, *15*, 377−392.
(349) Grochowski, P.; Trylska, J. Continuum Molecular Electrostatics, Salt Effects, and Counterion Binding—A Review of the Poisson−Boltzmann Theory and Its Modifications. *Biopolymers* **2008**, *89*, 93−113.
(350) Wang, C.; Greene, D.; Xiao, L.; Qi, R.; Luo, R. Recent Developments and Applications of the MMPBSA Method. *Frontiers in Molecular Biosciences* **2018**, *4*, 87.
(351) Su, M.; Wang, Y. A Brief Review of Continuous Models for Ionic Solutions: The Poisson−Boltzmann and Related Theories. *Commun. Theor. Phys.* **2020**, *72*, 067601.
(352) Bader, R. F. W.; Preston, H. J. T. The Kinetic Energy of Molecular Charge Distributions and Molecular Stability. *Int. J. Quantum Chem.* **1969**, *3*, 327−347.
(353) Anderson, J. S. M.; Ayers, P. W.; Hernandez, J. I. R. How Ambiguous Is the Local Kinetic Energy? *J. Phys. Chem. A* **2010**, *114*, 8884−8895.
(354) 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.
(355) Eberhart, M. The Metallic Bond: Elastic Properties. *Acta Mater.* **1996**, *44*, 2495−2504.
(356) Eberhart, M. E.; Jones, T. E. Cauchy Pressure and the Generalized Bonding Model for Nonmagnetic BCC Transition Metals. *Phys. Rev. B* **2012**, *86*, 134106.
(357) 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.
(358) Eberhart, M. E.; Wilson, T. R.; Johnston, N. W.; Alexandrova, A. N. Geometry of Charge Density as a Reporter on the Role of the Protein Scaffold in Enzymatic Catalysis: Electrostatic Preorganization and Beyond. *J. Chem. Theory Comput.* **2023**, *19*, 694−704.
(359) Abramov, Yu. A. On the Possibility of Kinetic Energy Density Evaluation from the Experimental Electron-Density Distribution. *Acta Crystallographica Section A Foundations of Crystallography* **1997**, *53*, 264−272.
(360) Yang, H.; Boulet, P.; Record, M.-C. A Rapid Method for Analyzing the Chemical Bond from Energy Densities Calculations at the Bond Critical Point. *Computational and Theoretical Chemistry* **2020**, *1178*, 112784.
(361) Arnold, W. D.; Oldfield, E. The Chemical Nature of Hydrogen Bonding in Proteins via NMR: *J*-Couplings, Chemical Shifts, and AIM Theory. *J. Am. Chem. Soc.* **2000**, *122*, 12835−12841.
(362) Jenkins, S.; Morrison, I. The Chemical Character of the Intermolecular Bonds of Seven Phases of Ice as Revealed by Ab Initio Calculation of Electron Densities. *Chem. Phys. Lett.* **2000**, *317*, 97−102.
(363) Espinosa, E.; Alkorta, I.; Elguero, J.; Molins, E. From Weak to Strong Interactions: A Comprehensive Analysis of the Topological and Energetic Properties of the Electron Density Distribution Involving X−H··· F−Y Systems. *J. Chem. Phys.* **2002**, *117*, 5529−5542.
(364) Grabowski, S. J.; Sokalski, W. A.; Dyguda, E.; Leszczyński, J. Quantitative Classification of Covalent and Noncovalent H-Bonds. *J. Phys. Chem. B* **2006**, *110*, 6444−6446.
(365) Pakiari, A.; Eskandari, K. The Chemical Nature of Very Strong Hydrogen Bonds in Some Categories of Compounds. *Journal of Molecular Structure: THEOCHEM* **2006**, *759*, 51−60.
(366) Duarte, D. J. R.; Angelina, E. L.; Peruchena, N. M. Physical Meaning of the QTAIM Topological Parameters in Hydrogen Bonding. *J. Mol. Model.* **2014**, *20*, 2510.
(367) Fradera, X.; Austen, M. A.; Bader, R. F. W. The Lewis Model and Beyond. *J. Phys. Chem. A* **1999**, *103*, 304−314.
(368) Tecplot, Inc. *Tecplot 360 2013R1* (Accessed 2024-10-07). https://www.tecplot.com, 2013.
(369) Wilson, T. R.; Eberhart, M. E. Bondalyzer (Accessed 2024-10-07). https://github.com/moltheorygroup/BondalyzerTecplotAddon, 2024.
(370) Eberhart, M. E.; Jones, T. E. The Two Faces of Chemistry: Can They Be Reconciled? *Foundations of Chemistry* **2013**, *15*, 277−285.
(371) Wilson, T. R.; Eberhart, M. Quantum Theory of Atoms in Molecules in Condensed Charge Density Space. *Can. J. Chem.* **2019**, *97*, 757−762.
(372) Morgenstern, A.; Wilson, T.; Miorelli, J.; Jones, T.; Eberhart, M. In Search of an Intrinsic Chemical Bond. *Computational and Theoretical Chemistry* **2015**, *1053*, 31−37.
(373) Morgenstern, A.; Eberhart, M. Bond Dissociation Energies from the Topology of the Charge Density Using Gradient Bundle Analysis. *Phys. Scr.* **2016**, *91*, 023012.
(374) Gribben, J.; Wilson, T. R.; Eberhart, M. E. Unicorns, Rhinoceroses and Chemical Bonds. *Molecules (Basel, Switzerland)* **2023**, *28*, 1746.
(375) Hammond, G. S. A Correlation of Reaction Rates. *J. Am. Chem. Soc.* **1955**, *77*, 334−338.
(376) Nechay, M. R.; Gallup, N. M.; Morgenstern, A.; Smith, Q. A.; Eberhart, M. E.; Alexandrova, A. N. Histone Deacetylase 8: Characterization of Physiological Divalent Metal Catalysis. *J. Phys. Chem. B* **2016**, *120*, 5884−5895.
(377) Yang, Z.; Liu, F.; Steeves, A. H.; Kulik, H. J. Quantum Mechanical Description of Electrostatics Provides a Unified Picture of Catalytic Action across Methyltransferases. *J. Phys. Chem. Lett.* **2019**, *10*, 3779−3787.
(378) 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.
(379) 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.
(380) Freindorf, M.; Tao, Y.; Kraka, E. A Closer Look at the Isomerization of 5-Androstene-3,17-Dione to 4-Androstene-3,17-Dione in Ketosteroid Isomerase. *Journal of Computational Biophysics and Chemistry* **2022**, *21*, 313−333.
(381) Wilson, T. R.; Morgenstern, A.; Alexandrova, A. N.; Eberhart, M. Bond Bundle Analysis of Ketosteroid Isomerase. *J. Phys. Chem. B* **2022**, *126*, 9443−9456.
(382) Wilson, T. R.; Eberhart, M. E. *Advances in Quantum Chemical Topology Beyond QTAIM*; Elsevier, 2023; pp 407−430.
(383) 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.
(384) Hennefarth, M. R.; Alexandrova, A. N. Direct Look at the Electric Field in Ketosteroid Isomerase and Its Variants. *ACS Catal.* **2020**, *10*, 9915−9924.
(385) Vargas, S.; Chaturvedi, S. S.; Alexandrova, A. N. Machine-Learning Prediction of Protein Function from the Portrait of Its

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813

Intramolecular Electric Field. *J. Am. Chem. Soc.* **2024**, DOI: 10.1021/jacs.4c09549.
(386) Helman, J.; Hesselink, L. Surface Representations of Two- and Three-Dimensional Fluid Flow Topology. In *Proceedings of the First IEEE Conference on Visualization: Visualization '90*; IEEE: San Francisco, CA, USA, 1990; pp 6−13.
(387) Scheuermann, G.; Kruger, H.; Menzel, M.; Rockwood, A. Visualizing Nonlinear Vector Field Topology. *IEEE Transactions on Visualization and Computer Graphics* **1998**, *4*, 109−116.
(388) Ebling, J.; Scheuermann, G. *Clifford Convolution and Pattern Matching on Vector Fields*; IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control; IEEE: Seattle, WA, USA, 2003; pp 193−200.
(389) Polthier, K.; Preuß, E. In *Visualization and Mathematics III*; Farin, G., Hege, H.-C., Hoffman, D., Johnson, C. R., Polthier, K., Hege, H.-C., Polthier, K., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2003; pp 113−134.
(390) Dinh, H. Q.; Xu, L. 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.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2008; Vol. *5342*; pp 187−196.
(391) Bishop, C. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*; Springer-Verlag: Berlin, Heidelberg, 2006.
(392) Vargas, S.; Hennefarth, M. R.; Liu, Z.; Alexandrova, A. N. Machine Learning to Predict Reaction Barriers from the Reactant State Electron Density. *J. Chem. Theor. Comput.* **2021**, *17*, 6203−6213.

https://doi.org/10.1021/acs.chemrev.4c00471
Chem. Rev. 2025, 125, 3772−3813