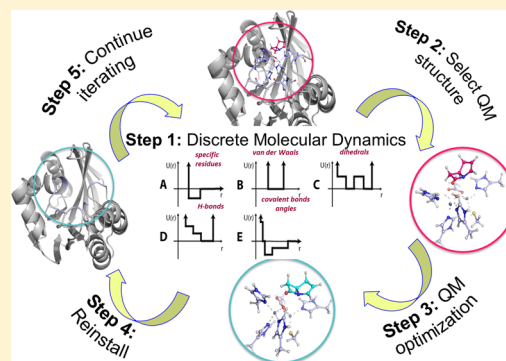


# Computational Treatment of Metalloproteins

Michael R. Nechay,<sup>†,§</sup> Crystal E. Valdez,<sup>†,§</sup> and Anastassia N. Alexandrova<sup>\*,†,‡</sup><sup>†</sup>Department of Chemistry and Biochemistry and <sup>‡</sup>California NanoSystems Institute, University of California, Los Angeles, Los Angeles, California 90095, United States

**ABSTRACT:** Metalloproteins present a considerable challenge for modeling, especially when the starting point is far from thermodynamic equilibrium. Examples include formidable problems such as metalloprotein folding and structure prediction upon metal addition, removal, or even just replacement; metalloenzyme design, where stabilization of a transition state of the catalyzed reaction in the specific binding pocket around the metal needs to be achieved; docking to metal-containing sites and design of metalloenzyme inhibitors. Even more conservative computations, such as elucidations of the mechanisms and energetics of the reaction catalyzed by natural metalloenzymes, are often nontrivial. The reason is the vast span of time and length scales over which these proteins operate, and thus the resultant difficulties in estimating their energies and free energies. It is required to perform extensive sampling, properly treat the electronic structure of the bound metal or metals, and seamlessly merge the required techniques to assess energies and entropies, or their changes, for the entire system. Additionally, the machinery needs to be computationally affordable. Although a great advancement has been made over the years, including some of the seminal works resulting in the 2013 Nobel Prize in chemistry, many aforementioned exciting applications remain far from reach. We review the methodology on the forefront of the field, including several promising methods developed in our lab that bring us closer to the desired modern goals. We further highlight their performance by a few examples of applications.



## INTRODUCTION

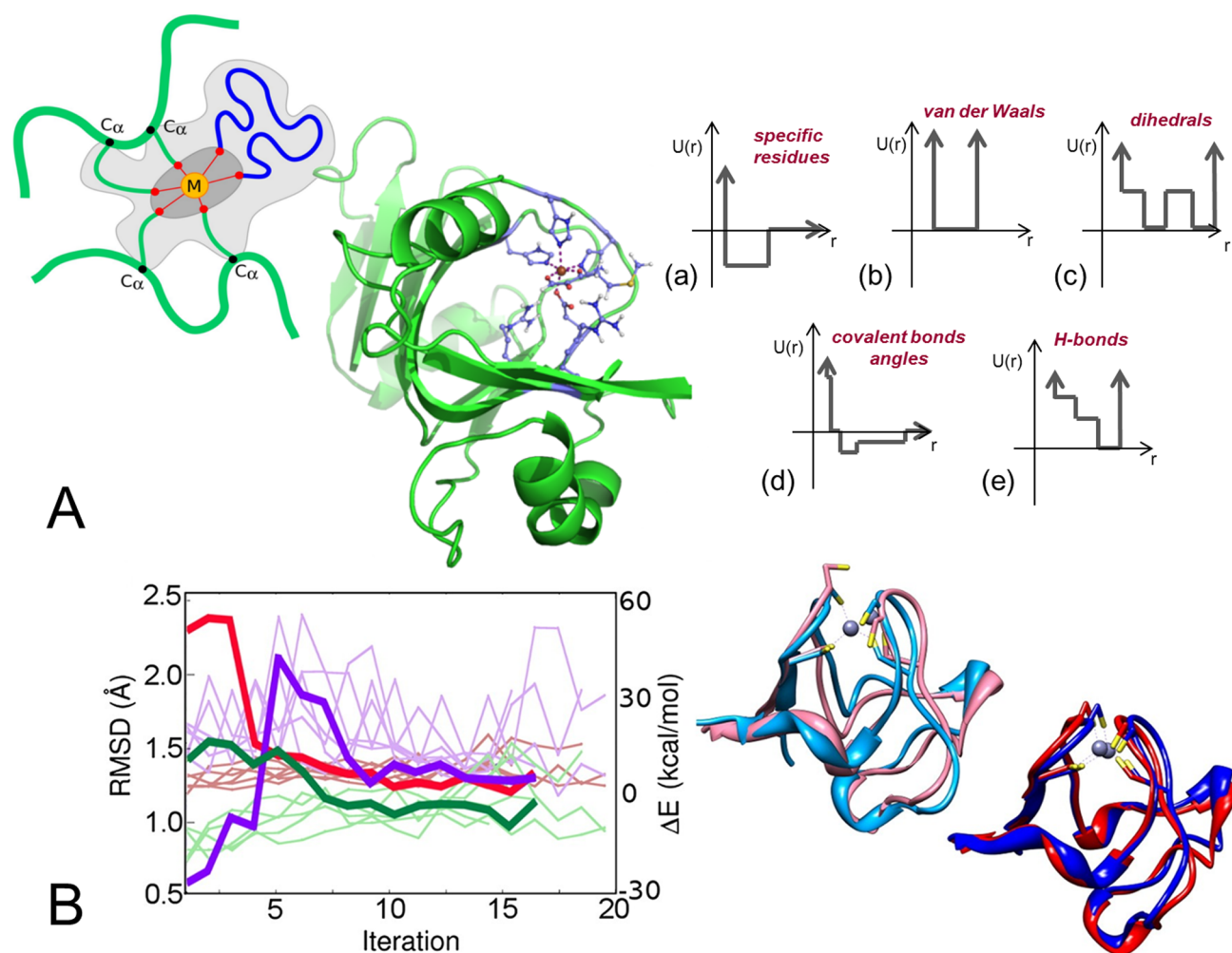
Metalloproteins present many challenges when it comes to computational modeling.<sup>1,2</sup> The strong Coulombic forces of metal cations and charged amino acids induce considerable perturbations to protein tertiary structure. For example, installation or removal of metal cations can yield large conformational changes.<sup>3</sup> In addition, metalloproteins commonly perform electron and proton transfers—a mechanism beyond the most common ways to describe protein dynamics today, i.e., classically. Metalloproteins are large dynamic molecules whose elements operate synergistically and on different time scales during enzymatic catalysis. The challenges in computationally modeling metalloproteins stem from the need to simultaneously describe the complex metal coordination site(s) and efficiently sample the protein backbone. Motions of the protein backbone and smaller metal site(s) are coupled, so the remaining challenge is balancing on a time scale relevant to protein machinery: picoseconds, milliseconds, and beyond. Because of the dynamic nature of proteins, sampling and assessment of free energies or their changes are critical, and it is not a simple task. Many dream applications that stand to benefit from modeling *in silico* include engineering metalloproteins and/or their substrates to enrich catalysis and inhibition. Together with traditional mechanistic studies (especially for proteins that perform electron and proton transfer), these applications rely on an adequate multiscale methodology. Computational treatment of metalloproteins is in a crucial developmental stage, and we believe some of the most

exciting applications are just over the horizon, with more prospects to follow.

Sampling of protein structure typically employs classical force field based methods: molecular dynamics (MD) and Monte Carlo statistical mechanics (MC). Much work has been done in developing more efficient sampling of proteins via MD,<sup>4,5</sup> and MC via the Metropolis–Hastings algorithm.<sup>6,7</sup> Both methods lend themselves well to parallelization, and serve faithfully in areas of computational biophysics such as drug design.<sup>8</sup> However, in modeling metalloproteins, the classical force fields are usually parametrized to handle only a specific metal coordination, which then has no freedom to change in the course of a simulation (by acquiring or losing ligands as in a catalyzed reaction or in response to reduction/oxidations of the metal). Such changes would not be possible within the harmonic oscillator-like description. This limitation is also characteristic of the empirical valence bond (EVB) approach.<sup>2</sup> Merz et al. are currently developing parameters for 24 divalent metal ions; however, they are currently limited to calculating hydration free energies of metal–water clusters within the constraints of a harmonic oscillator-like description of the metal–ligand bonds.<sup>9</sup> On the other hand, fragmentation-based techniques—methods that carefully partition a large system of interest into smaller fragments suitable for *ab initio* QM

Received: January 2, 2015

Revised: February 27, 2015



**Figure 1.** QM/DMD. (A) The unique feature of QM/DMD is the “breathing” QM–DMD boundary: the light gray area on the scheme is managed in both the QM and DMD regions. The alpha carbons of the “breathing” residues are held frozen during the pure QM phase, and the atoms directly coordinated to the metal (red circles) are frozen during the pure DMD phase. The dark gray region is managed exclusively by QM. A real example of the separation into the DMD-only and QM–DMD regions is shown by the green/purple protein (the ARD system). A few step function potentials in DMD are shown. Each is defined with an example in parentheses: (a) Hard-shell interaction potentials (hard-sphere radius, attractive potential well), (b) single-infinite square well (covalent bonds), (c) dihedrals (peptide bonds), (d) discretized van der Waals (solvation nonbonded), and (e) hydrogen-bonding auxiliary distance potential function. (B) The graph shows a representative QM/DMD simulation with converged data such as RMSD (Å) from the X-ray structure (light green lines), the QM energies (pink lines), and the DMD energies (light purple lines). The thick lines illustrate the fast return of a distorted wild type structure of rubredoxin to equilibrium. The structures compare the overlay of the X-ray structure (light blue) and distorted structure (pink) with representative QM/DMD equilibrated structures starting from X-ray (dark blue) and the distorted structure (red). Adapted with permission from ref 42. Copyright 2012 Elsevier.

calculations, with the intent to give the same energy and properties of the complete large system—show promise in expediting the computational time required with massive parallelism for QM calculations.<sup>10</sup> The literature is rich in such techniques, but there are some notable examples for biological systems such as fragment molecular orbital (FMO) methods. A full review on these methods can be found in other works.<sup>11–13</sup> However, the computational power needed for these methods is far from realistic for metalloproteins containing sometimes hundreds of residues.

Hybrid methodologies which bridge the quantum mechanical description of the metal–ligand interactions with classical sampling of protein structure are very promising. There is a rich history behind these methods. The quantum mechanical/molecular mechanical (QM/MM) approach was first introduced by Levitt and Warshel<sup>14</sup> to calculate the energies of intermediate states in enzymatic reactions. QM/MM methods

have undergone remarkable advancement, and now it is possible to study reaction pathways in large systems, such as solvated enzymes.<sup>15,16</sup> Car–Parrinello MD (CPMD)<sup>17,18</sup> techniques account for electronic and nuclear motion in an approximate manner and have substantial success. Presently, CPMD can describe the time evolution of molecular systems with  $\sim 100$  heavy atoms for  $\sim 100$  ps without relying on a force field. Nonetheless, hybrid molecular mechanics/CPMD (MM/CPMD) schemes,<sup>17,19</sup> that do employ a force field for part of the simulation, have been successfully applied to several different metal-containing systems. A few examples include determination of the catalytic role of  $\text{Zn}^{2+}$  and  $\text{Mg}^{2+}$  in metalloenzymes,<sup>20–22</sup> ligand–DNA interactions with transition-metal-dependent anticancer drugs,<sup>23,24</sup> and properties of electron transport proteins.<sup>25</sup> Fragmentation-based methods have been recently extended into the QM/MM formalism, in particular for simple polymers,<sup>26</sup> biopolymers,<sup>27,28</sup> and

zeolites.<sup>28</sup> Approaches that contain integrated QM with MM, such as ONIOM,<sup>29</sup> compute entire system properties at a lower level of theory and add higher level components to chosen areas. For large metal-containing proteins, these methods are not developed enough yet.

A primary issue with even the most efficient QM/MM methods is still speed. Conformational sampling is expensive, restricting the use of QM/MM methods in docking drug candidates, protein structure design, and specific metal-dependent functions. Toward reducing the computational cost, semiempirical methods such as PM7 and PDDG have taken the place of QM in QM/MM in pioneering work by Gao and Truhlar,<sup>30</sup> and in continued expansion.<sup>31–33</sup> However, the reliability of semiempirical methods is known to be more limited. For sensitive quantities, such as reduction potentials of metals in proteins, one needs to advance beyond semiempirics into a quality quantum mechanical description.<sup>34–36</sup> A powerful approach, specific for the description for electron transfer in metalloproteins, was designed by Voth et al.,<sup>37</sup> which skips the QM part all together. Voth instead coupled a coarse-grained version of the classical mechanical description of the protein to quantum dynamics describing electron transfer. This simulation yields electron transfer between FeS clusters [FeFe]-hydrogenase. Despite its indisputable beauty, the method cannot describe the FeS clusters themselves and address such questions as what electronic states are involved in accepting and releasing electrons, and how the cluster-coordination environment impacts those results. In contrast to pure electron transfer, which employs cooperation from experimental optical spectroscopy methods, computation is necessary in studying proton coupled electron transfer (PCET) due to experimental limitations in capturing proton translocation. PCET is commonly considered a suitable reaction mechanism for proteins and metalloproteins such as cytochrome oxidase. Many theoretical studies on PCET have been done by Hammes-Schiffer et al.,<sup>38,39</sup> including capturing reaction mechanisms and reproducing reaction rates in the metalloenzyme soybean lipoxygenase, via vibronically nonadiabatic formulation of PCET reactions in solution and proteins.<sup>40</sup>

In this Feature Article, we discuss selected methodologies on the forefront of this field and how sampling, entropy, and description of the metal centers are currently addressed. In particular, we believe that our QM/DMD method (DMD standing for discrete molecular dynamics<sup>41</sup>) is among the most efficient for sampling of metalloprotein structure.<sup>42</sup> Further, we will present specialized techniques for applications such as metalloprotein design, which demands a robust sampling strategy. We illustrate our discussion with computational studies that showcase the power of our methods, and a look forward to where we envision they could be successfully applied.

## METHODS FOR METALLOPROTEIN TREATMENT

**QM/DMD.** QM/DMD is a metalloprotein sampling engine.<sup>42</sup> It operates through an iterative scheme between QM and DMD machineries (Figure 1). DMD is a flavor of MD that approximates the continuous interaction potentials in classical MD with square-well potentials (Figure 1A), course-graining the potentials and overall reducing the number of calculations needed. Due to these *discretized* potentials, DMD is driven by collision events rather than physical forces as in traditional MM and MD. Therefore, the user saves a tremendous amount of time with DMD by solving ballistic

equations of motions rather than Newtonian equations of motions (Figure 1A). Complete details can be found in earlier works.<sup>43</sup> QM/DMD operates in the following scheme: the simulation begins with a DMD simulation of the entire protein keeping the metal and atoms directly bound to the metal frozen, and a few other constraints possibly being included (Figure 1A). This saves one from the need to parametrize the classical force field for the metal–ligand interactions. Following DMD, a structure is selected from the trajectory representative of the ensemble. A larger, chemically meaningful QM region is extracted from the protein for optimization at the QM level. Most of our work employs density functional theory (DFT) for the QM management area due to the size and transition metal species under consideration. However, any *ab initio* QM formalism, whether it be density of wave function based methods, can be chosen for the QM region. This region can be something normally used for a QM mechanistic study on an enzyme using a cluster model (Figure 1A). During the relaxation, the structural changes in the protein predicted by DMD can influence the metal coordination. The relaxed part is then reinstalled back into the protein, the QM–DMD boundary shrinks back to going right around the metal center or centers, and the small region inside the boundary is again fixed. The simulation continues with DMD. During the DMD stage, the updated geometry of the QM region has a chance to influence the rest of the protein. The simulation proceeds in the analogous iterative manner to convergence (illustrated in Figure 1B). The described “breathing” QM–DMD boundary is a simple solution for the communication between the two simulation machineries. Then, last but not the least, DMD is highly suitable for being incorporated into the hybrid method; DMD does not calculate forces by calculating energy gradients, and operates on discontinuous potentials, and as a result is insensitive facing a discontinuity of the potential at the QM–DMD boundary. To the best of our knowledge, QM/DMD provides record metalloprotein sampling speeds for simulations done on CPUs.

QM/DMD has the ability to recapitulate native protein structure from native and distorted ones,<sup>42,44</sup> and provide finer structural details of the active site at the level needed for, for example, subsequent mechanistic studies.<sup>45</sup> Large-scale motions of protein parts are also captured efficiently.<sup>46</sup> For the equilibrated QM/DMD structures, one can apply most rigorous QM methods to obtain such sensitive properties as changes in the reduction potentials of the metal or barriers of catalyzed reactions.<sup>42</sup> Naturally, since sampling is done on the entire protein while metal coordination is purely in the QM management, events such as ligand attachment<sup>44</sup> or detachment<sup>47</sup> are easily captured. This is often important in mechanistic studies. We are particularly excited about the demonstrated ability of QM/DMD to predict how the structure of a metalloprotein would react to the changes in protein sequence or the nature of the bound metal,<sup>42,44,45,48</sup> removal of one of the metal cations,<sup>47</sup> or binding a metal to a protein that originally did not contain a metal. These capabilities make QM/DMD a good platform for metalloprotein design, as is currently being tested (manuscript in preparation). Undoubtedly, metalloprotein design is a very interesting and exciting goal pursued in the field.<sup>49–54</sup> It is also known that, in the design of specific, buried binding pockets, as opposed to surface- or interface-exposed binding sites, sampling the protein backbone dramatically expands the design repertoire.<sup>55,56</sup> Hence, the idea is to utilize QM/DMD sampling in design,

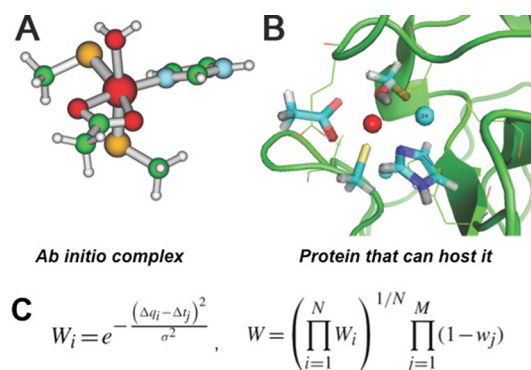
enabling a true design of buried active sites containing transition metals, perhaps for the first time.

### Approach to Metalloenzyme Design: General Outline.

The design process we aim for is based on the “inside-out” protocol developed by Baker et al.<sup>57</sup> The original idea was to design an arrangement of amino acids around the rate-limiting transition state of the catalyzed reaction such that it would be stabilized selectively from the reactants and then incorporate this design into a pocket of an existing protein scaffold. For metalloenzymes, this process acquires an additional dimension: the design of the electronic structure of the metal center or centers. Indeed, the electronic structure of the metal defines catalysis, and it is what makes metalloenzymes such great catalysts, often accomplishing the most dramatic reactions in a single step. Design here means tuning Lewis acidity, reduction potential, or order and population of d-atomic orbitals (d-AOs) on the metal, determined by its ligand environment. Thus, *ab initio* calculations of the catalytic transition metal complex are at the root of the computational design process. It is then needed to find proteins that can arrange for such metal coordination, which may be nontrivial, since only a few (natural) amino acids and the backbone N and C=O groups can be the ligands. After that, the stabilization of the transition state of interest should be done using continuous QM treatment of the metal for its optimal description, and sampling of the protein backbone. Hence, we put together a set of tools enabling every step in this process.

**Erebus.** *Erebus* is one of such tools we found necessary to develop to assist in the design process. To increase the chance of a designed protein to fold into a correct structure, it is safest to use an existing robust protein scaffold that already contains a metal. If desired, it generally can be arranged for the metal to be replaced with another one *in vitro*, or *de novo* synthesis of the protein can be done with the desired metal being supplied. Mutations in the active site are typically kept to a minimum in the design, again to preserve the native fold. *Erebus*<sup>58</sup> is a data mining tool for searching through the Protein Data Bank (PDB) for substructures in proteins. For the design application, *Erebus* can look for arrangements of the metal, its ligands, and possibly other amino acids near the metal, that are close to what is desired. The idea is that one can use *ab initio* calculations to construct a potentially catalytic metal–amino acid residues complex and then look for this arrangement in the full set of available protein structures. For example, the calculated complex shown in Figure 2A was found in the natural protein in Figure 2B. The *Erebus* search is based on a subgraph isomorphism algorithm with a user-defined allowed structural uncertainty, as illustrated in Figure 2C. It browses through the entire PDB in an efficient manner (on the order of under 1 h for a single search). Depending on the allowed uncertainty, many matches can be uncovered.

**Multiscale Design Tool.** The last necessary tool in our arsenal should assess the change in protein free energy upon mutagenesis, to simultaneously optimize for the transition state binding and protein stability. These evaluations are also important beyond protein design, for a wide variety of research areas concerning structure/function relationships of proteins. However, assessing the effect of mutagenesis on protein stability is difficult, since it requires addressing the complex nature of many interactions in a secondary structure. With metalloproteins, the additional complexity arises because every bit of repacking in the binding site can have a serious influence on the delicate electronic structure of the metal through both



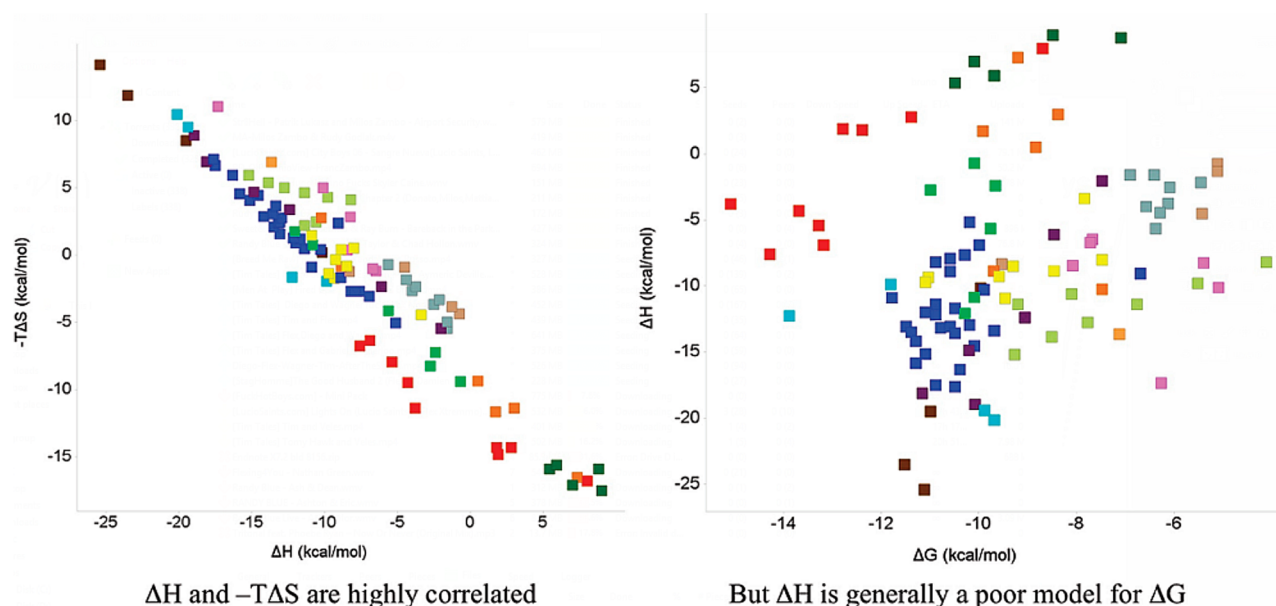
**Figure 2.** Illustration of *Erebus*. (A) A metal–ligand complex calculated using DFT. Hypothetically, this complex is desired for catalysis inside a protein pocket. (B) The active site of di-Zn  $\beta$ -lactamase found by *Erebus* to contain the Asp, two Cys, His residues, and a water molecule in the geometry close to the one desired from part A. Notice that it is already a Zn-binding site. (C)  $W_i$  represents the accuracy of a match to the input scaffold; more specifically,  $W_i$  represents how much the position of every heavy atom  $i$  matches that in the designed structure, subject to user-defined uncertainty  $\sigma$ , and the resultant overall weight of the structure,  $W$ , used by *Erebus* to judge the overall quality of the match.

bonded (shifting ligands) and nonbonded (polarization, electrostatics) interactions. As much as nature fine-tunes the metal to have just the right Lewis acidity, or just the right reduction potential for the reaction it catalyzes,<sup>59</sup> we too must be mindful in a similar way when designing.

Intuition can provide initial guesses for useful mutations to introduce into a protein. However, they only take us so far, where evaluating whether one or two seemingly innocent mutations could cause the entire protein to denature becomes impossible. Additionally, changes beyond a few catalytic residues might be beneficial for protein stability and overall packing of the active site. Typically, human intuition is completely inept at this task, thus, computation is required.

One of the most popular and powerful design tools is classical force-field-based *Rosetta*.<sup>60,61</sup> It samples through the discrete rotameric space of amino acids as well as their chemical nature, in an effort to repack and redesign binding sites for catalysis/stability and has been successful in several cases.<sup>62–65</sup> A competing tool, *Eris*, is a protein stability prediction software that utilizes the slightly coarse grained force field, *Medusa*.<sup>66</sup> *Eris* measures protein stability when a mutation is introduced by calculating the  $\Delta\Delta G$  of the mutation, where  $\Delta\Delta G = \Delta G_{\text{mutant}} - \Delta G_{\text{wild-type}}$ .<sup>55,67</sup> It is beneficial that *Eris* utilizes some form of protein backbone flexibility, as it was shown to be helpful in design,<sup>55,67</sup> and is expected to be of a major importance for metalloproteins. However, none of the existing methods in their current forms capture the structural impact of metals ions on the protein, and in turn allow the protein to impact the electronic structure of the metal.

We have combined our QM/DMD sampling software<sup>42</sup> with *Eris*, creating a recipe for gauging metalloprotein stability induced by mutations called *Eris-QM/DMD*. The *Eris-QM/DMD* method begins with a QM/DMD simulation of the system the user is interested in mutating. After each iteration of QM/DMD, the protein structure is mutated to the desired residue(s) and undergoes an additional sampling step and the  $\Delta\Delta G$  of the mutation is evaluated. This method has demonstrated proof-of-principle native sequence recovery for two structurally different proteins, where each protein is



**Figure 3.** Analysis of the experimental binding thermodynamics for approximately 100 protein–ligand complexes. Reprinted with permission from ref 68. Copyright 2011 American Chemical Society.

mutated in several ways, and *Eris-QM/DMD* is able to recapitulate all native sequences of each protein except for one mutation, which is undergoing experimental validation (manuscript in preparation).

**Entropy Evaluation.** Entropy evaluation is a separate and complicated issue. QM/DMD sampling predicts equilibrium protein structure and energies. Entropic effects are either ignored (for example, in certain mechanistic studies where it is assumed that entropy does not change significantly as the system progresses from reactants to the transition states to products), or are included in an empirical way (as in *Eris* and *Eris-QM/DMD*). However, a quantitative evaluation of configurational entropy is desired, especially for sensitive applications such as drug design, where every fraction of a kcal/mol in the affinity matters greatly. Entropy evaluation presents a fundamental hurdle in the field—protein structures are routinely solved as static crystal structures, and fewer experimental studies are able to extract a structural ensemble. In the case of substrate binding, it is commonly accepted that there is correlation between  $\Delta H$  and  $T\Delta S$ ; however, the relevant free energy in protein and substrate binding is not correlated with either of these individual descriptors (Figure 3).<sup>68</sup>

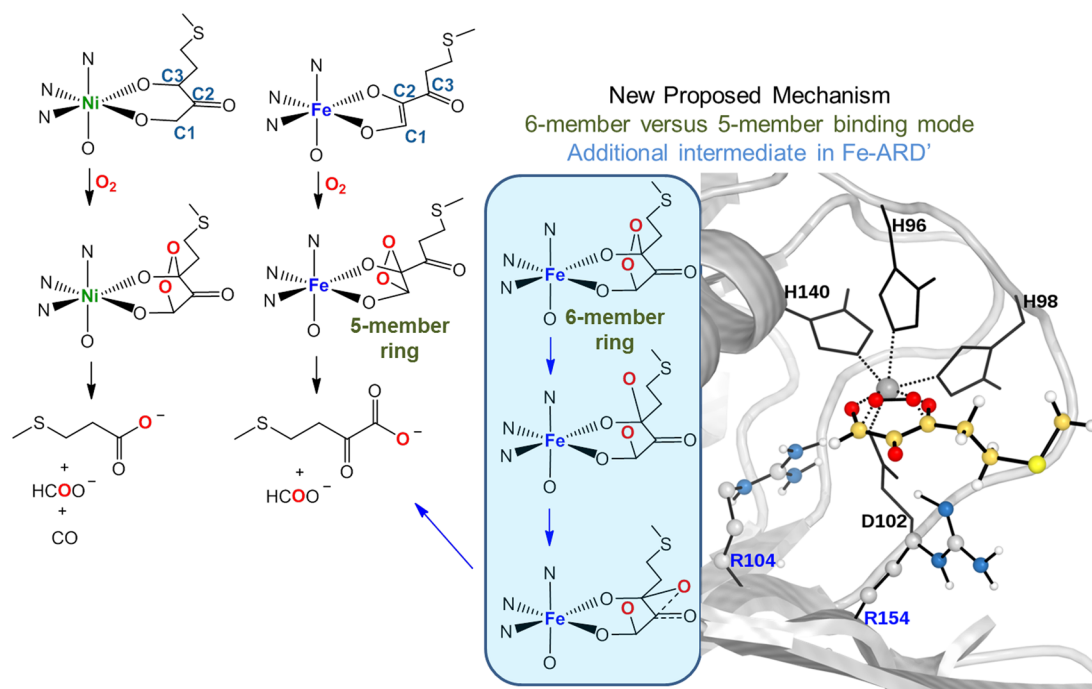
Computational methods allow us to follow precise positions of not only tertiary and secondary structures but also atomic positions, and these motions are coupled with corresponding energies. Provided we have a crystal structure or another good “guess” of a starting point, these motions are accessible. (However, getting to that starting point *in silico* remains an active field in itself.<sup>69</sup>) Normal mode analysis, for example, probes local curvature near a stationary point via the Hessian, and by using quantum mechanical modeling, one can deduce a related entropy component to sum across all modes. Karplus and Kushick laid down the groundwork for quasi-harmonic analysis using internal coordinates<sup>70</sup> which probes the global extent of configurational space accessible to a system at a given temperature. By evaluating a mass-weighted covariance matrix from MD simulation and diagonalizing, the resulting quasi-harmonic frequencies can be used with the QM expression for

entropy of a harmonic oscillator. This approach was later reformulated by Andricioaei and Karplus to work in Cartesian coordinates.<sup>71</sup> Similar methods exist, such as in calculating an “upper” limit to entropy by use of convergence matrices.<sup>72</sup> These methods are common in simulation packages together with MD trajectories; however, they are often more useful for qualitative results than quantitative comparisons of entropy.

Consequently, free energy is left as the direct target of computation, rather than handling energy and entropy separately. The need for properly handling contributors including roles of disorder, fluctuations, protein dynamics, and multiple pathways in reaction dynamics is crucial and solutions are developing rapidly. Free energy perturbation (FEP) is one technique which can handle direct calculation of free energy changes; however, it is very expensive especially if done in QM/MM settings. Umbrella sampling can be used to include entropic effects in 1-D free energy profiles and 2-D energy landscapes, and Gao, Truhlar, and others use ensemble averaging<sup>73</sup> to include entropic effects in free energy computation.<sup>74–81</sup>

As discussed, DMD efficiently samples configurational space at the atomic level. Bolstered by the efficient and discrete nature of molecular dynamics, Schofield et al. quantitatively measured the free energy and entropy of folding via probabilities of structures defined by their sharp noncovalent bond networks.<sup>82</sup> Basically, the function of the potential in DMD allows for a very straightforward definition of a conformational state: every well is either populated or it is not. Then, entropy evaluation becomes accessible. Schofield showed this on a very simple system (short polypeptide) and using a simpler DMD potential. Extending this idea to larger systems of interest and a more realistic form of DMD is the subject of an ongoing implementation based on QM/DMD, in our lab. Indeed, the sampling of DMD combined with any quantitative method to extract free energy or entropy would provide the missing link between binding free energies and already-robust energy calculations.

With these tools, difficult questions of interest to the metalloenzyme community have been answered and our efforts



**Figure 4.** Old hypothesis on product differentiation of Fe-ARD' and Ni-ARD disproved through our QM/DMD simulations coupled with DFT mechanistic studies and shown to stem purely from the electronic structure of the metal, not the binding mode of the substrate to the metal. However, the residues R104 and R154 stabilize the substrate in the reactive orientation.

to tackle more difficult and multifaceted problems, such as the design of artificial metalloenzymes and inhibitors of them, continue to evolve. We now describe a few successful applications of our methods, and in the end, we will outline what is still missing or under construction, and what future applications are becoming within reach.

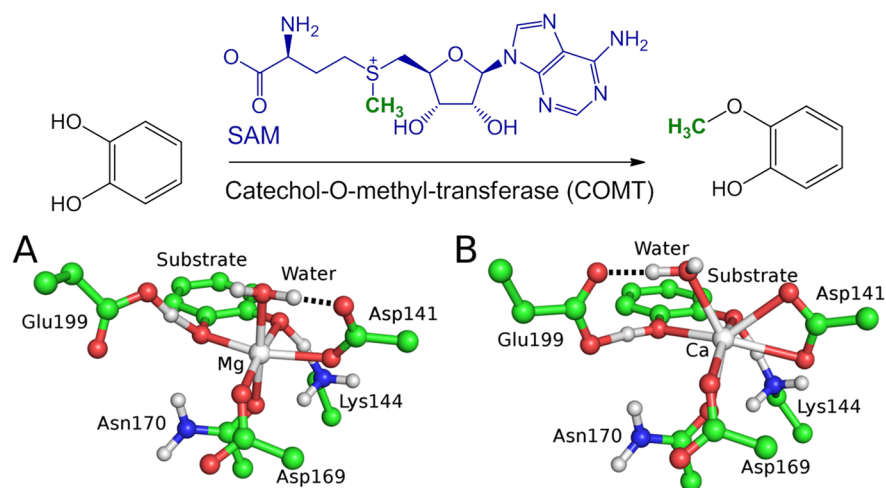
## EXAMPLES OF APPLICATIONS

**Large-Scale Motion Important for Metalloenzyme Catalysis.** The need for adequate sampling for the protein structure is occasionally unnecessary, as the large, structural equilibration does not influence active site chemistry. However, in many instances, backbone motion induces conformational changes around or far away from the active site. These structural motions could close or open channels for substrates to flow through or arrange residues at positions needed for catalysis. One such example of the latter is acireductone dioxygenases (ARDs).

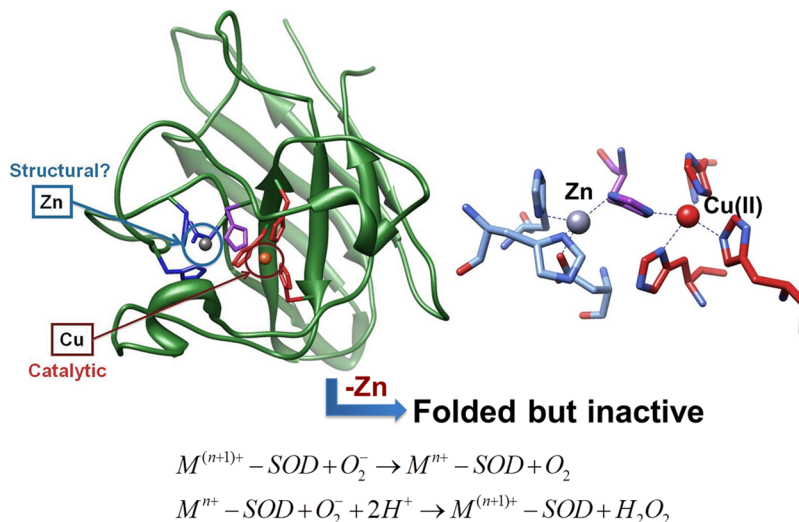
ARD is an enzyme that catalyzes two different oxidation reactions, depending solely on whether  $\text{Fe}^{2+}$  or  $\text{Ni}^{2+}$  is bound to the protein.<sup>83–85</sup> Fe-dependent ARD' recycles methionine in the methionine salvage pathway by oxidizing the substrate, 1,2-dihydroxy-3-keto-5-(methylthio)pentene (acireductone), into two products: the  $\alpha$ -keto acid precursor of methionine and formate. Ni-dependent ARD instead oxidizes acireductone into three products, methylthiopropionate, CO, and formate, and provides a shunt out of the methionine salvage cycle.<sup>2</sup> Interestingly, interconversion between the Fe and Ni forms of ARD is relatively simple as the protein has micromolar affinity for both metals ( $K_d < 0.4 \mu\text{M}$  for Fe and  $K_d < 0.1 \mu\text{M}$  for Ni). The long-standing hypothesis for why Fe and Ni catalyze different oxidation reactions is due exclusively to the coordination mode of the substrate to the metal center (Figure 4). It was proposed that a large conformation change in a nearby protein loop facilitates this different binding together

with the metal replacement.<sup>86</sup> However, it is unclear why two divalent metals of similar radii would bind the substrate in dramatically different ways, and especially also cause large conformational changes in the protein. Indeed, this view has been challenged by recent experiments and our QM/DMD computational studies, and a new mechanism was proposed.<sup>37,69</sup>

Through QM/DMD simulations, we found both Fe and Ni ARD bind the substrate in the same orientation, via O1 and O3 of the substrate (Figure 4). The protein pocket stabilizes this specific orientation by two Arg residues, R104 and R154, which forms a hydrogen bond to the doubly deprotonated substrate bound to the metal in this way. The X-ray structure does not contain a bound substrate or its analogue, and therefore, upon docking the substrate, these hydrogen bonds were not present. DMD sampling during the simulations elucidated the role of these residues and demonstrated the importance of sampling in metalloprotein systems such as this one. However, since substrate orientation is identical in the two systems, the coordination mode to the metal center can no longer be the reason for product differentiation in Fe and Ni. QM mechanistic studies revealed an additional intermediate that forms in the Fe-ARD' mechanistic pathway not observed in Ni-ARD. This extra intermediate is also found in biomimetic complexes mimicking ARD.<sup>87</sup> The ability of Fe but not Ni to stabilize an additional intermediate comes from the redox flexibility of  $\text{Fe}^{2+}$ , allowing for the flow of electrons from the residues to the substrate and the bound  $\text{O}_2$ . Thus, the antibonding  $\pi^*$  orbital in dioxygen gets populated, and  $\text{O}_2$  dissociates. The O atoms easily migrate around, and one accessible epoxy-like transition state produces the mentioned intermediate. The more electron-rich  $\text{Ni}^{2+}$  does not allow for  $\text{O}_2$  dissociation. With ARD, sampling was required to predict the correct substrate binding pose, and dismiss the old



**Figure 5.** Catechol-*O*-methyl-transferase (COMT) catalyzes the methyl transfer from the cofactor SAM to a catechol motif found in neurotransmitters. (A) Native COMT binds  $Mg^{2+}$ , which positions the substrate in the proper orientation toward SAM, for methyl transfer. However, when  $Ca^{2+}$  (B) is bound, the active site distorts, putting reacting parts out of alignment, and thus leaving the enzyme inactive. Reprinted in part with permission from ref 42. Copyright 2012 Elsevier.



**Figure 6.** In the Cu, Zn dependent superoxide dismutase (SOD), Cu plays the role of the catalytic metal; however, Zn, although it plays a purely structural role, plays a vital role in keeping the adequate structure for Cu-mediated dismutation.

mechanistic hypothesis. The protein structure remained consistent upon metal replacement.

However, sometimes, a simple metal replacement can induce large-scale protein motions and repacking, as showcased by catechol-*O*-methyltransferase (COMT). COMT is a  $Mg^{2+}$ -dependent enzyme involved in the biology of pain.<sup>88,89</sup> COMT catalyzes the transfer of the methyl group from the cofactor *S*-adenosyl-*L*-methionine (SAM) to catechol subsequently, regulating the amount of catecholamine neurotransmitters in the brain and other organs (Figure 5). The metal in COMT binds and positions the catechol substrate in the correct reactive orientation toward SAM for the methylation step.<sup>90–92</sup> Native COMT contains  $Mg^{2+}$ , which however can be replaced with  $Co^{2+}$ ,  $Mn^{2+}$ ,  $Zn^{2+}$ ,  $Cd^{2+}$ ,  $Fe^{2+}$ ,  $Fe^{3+}$ ,  $Ni^{2+}$ , and  $Sn^{2+}$ .<sup>93</sup> Metal replacement leads to varying changes in the activity and structure of the enzyme. Surprisingly, the replacement of  $Mg^{2+}$  with  $Ca^{2+}$  leads to the inhibition of COMT, while replacement of  $Mg^{2+}$  with  $Fe^{2+}$  creates an only slightly weaker catalyst compared to  $Mg^{2+}$ .  $Fe^{3+}$  is also a complete inhibitor. The structure and function of these three metal variants of COMT

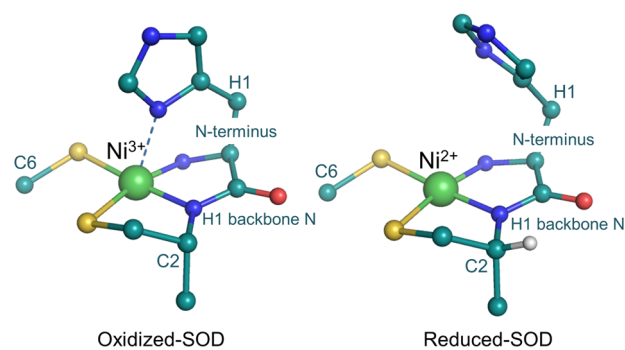
were investigated computationally.<sup>44</sup> Through the course of the simulation, the larger cation,  $Ca^{2+}$ , is shown to coordinate one additional ligand as compared to  $Mg^{2+}$  (Figure 5).  $Ca^{2+}$  also sits deeper into the binding cavity and distorts the reacting parts of catechol and SAM out of alignment for methyl transfer (Figure 5). These structural rearrangements with  $Ca^{2+}$  are predicted to produce disfavored reaction energetics. Thus, inhibition is a purely geometric effect in this case. Importantly, without sampling, and using just a cluster model, this geometric change is not observed: the  $Ca^{2+}$  containing site incorrectly looks identical to the  $Mg^{2+}$  containing site.<sup>44</sup> Predictably, the inhibitory effect of  $Fe^{3+}$  comes solely from the electronic properties of the metal, specifically its high electrophilicity.

COMT also is an illustration of how ligand capture can be important in addressing a mechanistic difference. This would not be done without the QM treatment of the metal, accompanied by adjustment of the portion of the backbone, which are the key features of QM/DMD. There are further, more dramatic, examples of this sort.

### Changes in the Coordination Geometry of the Metal.

Intriguingly, the class of superoxide dismutase (SOD) enzymes exhibits a wide range of redox activities modulated by ligand detachment or attachment. This review will highlight two specific ones, the first one being Cu, Zn dependent superoxide dismutase in humans (SOD1). SODs catalyze the dismutation/disproportionation of superoxide (Figure 6), a dangerous species linked to aging and other oxidation stress processes in organisms. The loss of SOD1 function and subsequent aggregation is known to lead to the neurodegenerative disease Amyotrophic Lateral Sclerosis (ALS) that affects the motor neurons of afflicted patients.<sup>94–98</sup> The metal-dependent nature of SOD is twofold: catalysis occurs at the Cu site, whereas the Zn site is believed to serve an important structural role for the whole protein. Experiments show that monomers without Zn are precursors to the SOD aggregates,<sup>99</sup> while Cu-less SOD1 does not show aggregation and retains its structure.<sup>100</sup> Without the presence of Zn, Cu is not able to catalyze the dismutation reaction. A combined computational and experimental study addressed the effect of Zn removal on the protein structure, electronic structure of the Cu site, and overall catalytic function of SOD.<sup>47</sup> The results show that Zn plays a structural role in SOD1 and directly influences the catalysis, enabling proper coordination and reduction potential of the Cu site. Removal of Zn causes the elimination of the catalytic activity of SOD1 even without protein unfolding and aggregation. It also makes the Cu center prone to deactivation due to immediate reduction to Cu(I) in the resting state. QM/DMD sampling was required to obtain the proper, folded but inactive structure of SOD1. In it, Cu lost one ligand.

**Ni-SODs.** Ni-SODs exist only in *Streptomyces* and cyanobacteria. Unlike all other metals used in SOD, Ni<sup>2+</sup> does not catalyze superoxide dismutation in aqueous solution due to an improper redox potential (a calculated +2.26 V when the optimum reduction potential is 0.36 V).<sup>101</sup> The active site of Ni-SOD is strikingly different from those of the other SODs.<sup>102</sup> The Ni ion, coordinated by a “Ni-hook” motif,<sup>103</sup> has a square-pyramidal coordination geometry when Ni is in its oxidized form (Ni<sup>3+</sup>). It has four equatorial ligands, two thiolates from Cys2 and Cys4, a deprotonated amide of the Cys2 backbone, and the N-terminal group of His1, and one axial imidazole from His1 (Figure 7). Upon reduction, Ni<sup>2+</sup> loses the His1 ligand and becomes square planar, making Ni-SOD the only observed SOD with a coordination number that changes as a function of metal oxidation state.<sup>43,104–107</sup> The flexible coordination geometry activates Ni and confers the proper reduction potential needed to function as a SOD. To



**Figure 7.** Redox reaction in Ni-SOD modulated through the changing coordination environment of the histidine (H1) residue.

address such a structural effect, an extensive sampling within a quantum-classical formalism would be essential.

**Metalloenzyme Design.** Metalloenzyme design is the hallmark challenge, where all difficulties revealed need to be addressed. We make the first attempts to design specific buried binding pockets in metalloproteins for desired catalysis. In all cases, QM/DMD sampling for structure prediction followed by QM mechanistic studies are the core of the effort. In a more conservative approach, the substrate of a known enzyme gets modified, and the binding pocket then gets redesigned slightly to accommodate it. These works involve mostly intuitive mutations done by hand, sampling, and mechanistic investigation, followed by experimental validations. We also use the full machinery described in the section Methods for Metalloprotein Treatment, including *Erebus*, *Eris*-QM/DMD, QM/DMD, and QM, to design artificial enzymes catalyzing non-natural reactions of interest, using nonphysiological metals in the active sites. These efforts are also in the stage of experimental testing of several promising designs. Details of the approach and results will be coming in future publications.

### ■ FUTURE ASPIRATIONS

Perhaps one of the most practical and significant applications of an efficient metalloprotein simulation engine would be computational design of drugs/inhibitors. In this case, it is necessary to predict the binding free energy of a substrate to a protein, or relative free energies of binding of one substrate with respect to another. The field of computational drug design has seen tremendous growth over the past decades.<sup>8</sup> By far, the most widely used approaches hinge on continuously evolving scoring functions,<sup>108</sup> classical MD, Monte Carlo, and sampling via normal-mode analysis, among other non-QM treatments of large molecules. However, with just these classically based methods, metalloenzyme targets remain largely inaccessible. In order to assess the induced fit, i.e., the conformational response of the protein to the bound ligand, the QM or parametrized MM treatment of the metal is paramount. While there are experimental methods that help bridge this gap, for example, fragment-based lead design (FBLD) which has been used to identify the first building blocks of inhibitors to several metalloenzymes,<sup>109</sup> effective computational approaches would advantageously complement metalloenzyme inhibitor design. Insight into preferred ligand poses without a cocrystal and “growing” the best targets within the active site of a metalloenzyme for optimum affinity and protein dynamics is currently not commonly available *in silico*. Leveraging these methods would lessen the need to synthesize ligand candidates every step of the way (as in FBLD).

We are starting to see QM/MM methods used in screening small libraries, including so-called “on-the-fly quantum mechanical/molecular mechanical (QM/MM) docking” which reveals a significantly improved success rate for a zinc-binding protein data set via QM’s comprehensive description of the charge-transfer effects and QM-refinement of binding poses.<sup>110</sup> Indeed, QM offers precise relaxation of ligands into protein active sites and measurement of electronic interactions, from the basic potentials (vdW, electrostatics, H-bonds, ligand strain, etc.) that scoring functions already perform well with to the unparamaterized metals. Even if interactions and poses in the QM active site are well-defined, a method to perform extensive sampling and estimate other key effects including entropy and desolvation is required to help complete the thermodynamic picture. This is no easy task, as we’ve seen for many years drug



design was stuck with rigid models of proteins, showing appreciable effectiveness, however, clearly inadequate for the more interesting cases.<sup>111,112</sup>

While efficient drug design is still an open question with many pending answers, only a few of which have been discussed here, a couple things have become abundantly clear. One barrier to any current approach in drug design is efficient sampling of design space, which quickly becomes expensive when targeting larger systems, and has recently been incorporated into many docking and scoring methods primarily by sampling of the larger protein movements followed by rigid substructure searches. Methods such as DMD show promise in quickly and accurately sampling the design environment, not only of the pure protein but with a potential ligand. When coupled with accurate QM estimation of binding site energies, new opportunities in drug design arise, and are being explored in our lab.

To conclude, the computational treatment of metalloproteins has been accelerating, and in this article, we have introduced many current challenges and developments, especially within our own work. Ongoing improvements to current methodologies and expansion to new methods continue to push the frontier of both structural and mechanistic studies of metalloproteins. This frontier spans exciting directions such as artificial metalloenzymes and metalloenzyme inhibitor design where the efforts in our lab are pushing these boundaries.

## AUTHOR INFORMATION

### Corresponding Author

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

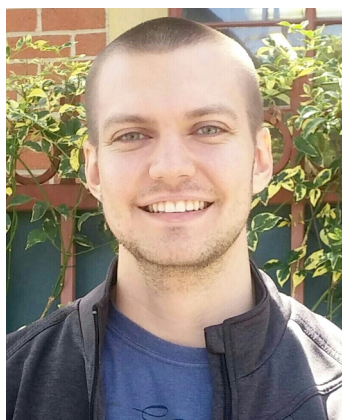
### Author Contributions

<sup>§</sup>M.R.N., C.E.V.: These authors equally contributed to this work.

### Notes

The authors declare no competing financial interest.

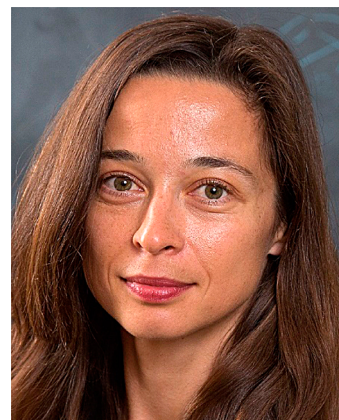
### Biographies



Michael R. Nechay was born and raised in Cuyahoga County, Ohio. He earned his B.S. in Chemical Engineering with Honors in Engineering and minor in Statistics at The Ohio State University. He was first exposed to Computational Chemistry research as an undergraduate under the supervision of Professor Christopher Hadad. Michael R. Nechay has been a Ph.D. student under the mentorship of Professor Anastassia N. Alexandrova since 2012 at UCLA. His research interests include the multiscale description of metalloproteins, development associated with utilizing the large amount of data generated from simulation, and the methodical design of metalloenzymes.



Crystal E. Valdez was born and raised in Los Angeles County, California. She earned her B.S. in Biochemistry at the California Polytechnic State University (*cum laude*) in 2009. She was first exposed to computational chemistry research during a Research Experience for Undergraduates (REU) at the University of Oregon under the supervision of Professor Marina Guenza, where she modeled biomolecules. Since 2010, Crystal has been a Ph.D. student under the mentorship of Professor Anastassia N. Alexandrova at UCLA. Her research interests encompass investigating metal-dependent functionalities of natural metalloproteins and utilizing and developing multiscale tools to design artificial metalloenzymes.



Anastassia N. Alexandrova was born in USSR and completed her undergraduate education in Saratov University, Russia (*summa cum laude*). She obtained a Ph.D. in physical chemistry in 2005, under the supervision of Professor Alexander Boldyrev, at Utah State University. Her graduate research was on chemical physics of small clusters and advancements of fundamental theory of chemical bonding. She did her postdoctoral work at Yale University, first with Professor William Jorgensen focusing on QM/MM statistical mechanical simulations of organic and enzymatic reactions in solution, and enzyme design, and then with Professor John Tully, focusing on nonadiabatic processes related to DNA photodamage. Since 2010, she has been an Assistant Professor in the Department of Chemistry and Biochemistry at the University of California, Los Angeles, and California NanoSystems Institute. Her laboratory focuses on multiscale modeling and design of new materials, ranging from metalloenzymes to heterogeneous catalysts, and novel surfaces and interfaces.

## ACKNOWLEDGMENTS

Financial support of the NSF Graduate Fellowship #2011115747 (to C.E.V.), the DARPA Young Faculty Award N66001-11-1-4138, and Alfred P. Sloan Research Fellowship (to A.N.A.) are gratefully acknowledged.

## REFERENCES

- (1) Chung, L. W.; Li, X.; Morokuma, K. Modeling Enzymatic Reactions in Metalloenzymes and Photobiology by Quantum Mechanics (QM) and Quantum Mechanics/Molecular Mechanics (QM/MM) Calculations. In *Quantum Biochemistry*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2010; pp 85–130.
- (2) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley: New York, 1997.
- (3) Ding, F.; Dokholyan, N. V. Dynamical Roles of Metal Ions and the Disulfide Bond in Cu, Zn Superoxide Dismutase Folding and Aggregation. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 19696–19701.
- (4) Adcock, S. A.; McCammon, J. A. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.* **2006**, *106*, 1589–1615.
- (5) Ben-Nun, M.; Martinez, T. J. Ab Initio Quantum Molecular Dynamics. *Adv. Chem. Phys.* **2002**, *121*, 439–512.
- (6) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
- (7) Hastings, W. K. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika* **1970**, *57*, 97–109.
- (8) Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science* **2004**, *303*, 1813–1818.
- (9) Li, P.; Roberts, B. P.; Chakravorty, D. K.; Merz, K. M. Rational Design of Particle Mesh Ewald Compatible Lennard-Jones Parameters for +2 Metal Cations in Explicit Solvent. *J. Chem. Theory Comput.* **2013**, *9*, 2733–2748.
- (10) Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Systems. *Chem. Rev.* **2012**, *112*, 632–672.
- (11) Fedorov, D. G.; Kitaura, K. Extending the Power of Quantum Chemistry to Large Systems with the Fragment Molecular Orbital Method. *J. Phys. Chem. A* **2007**, *111*, 6904–6914.
- (12) Dmitri, G. F.; Kazuo, K. Introduction. *The Fragment Molecular Orbital Method*; CRC Press: Boca Raton, FL, 2009; pp 1–3.
- (13) Massa, L. Book Review. *Int. J. Quantum Chem.* **2011**, *111*, 3251–3251.
- (14) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (15) Hu, H.; Yang, W. Free Energies of Chemical Reactions in Solution and in Enzymes with Ab Initio Quantum Mechanics/Molecular Mechanics Methods. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.
- (16) Kamerlin, S. C. L.; Vicatos, S.; Dryga, A.; Warshel, A. Coarse-Grained (Multiscale) Simulations in Studies of Biophysical and Chemical Systems. *Annu. Rev. Phys. Chem.* **2011**, *62*, 41–64.
- (17) Car, R.; Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
- (18) Marx, D.; Hutter, J. Ab Initio Molecular Dynamics: Theory and Implementation. *Modern Methods and Algorithms of Quantum Chemistry*; Jülich, 2000; Vol. 1, pp 301–449.
- (19) Laio, A.; VandeVondele, J.; Rothlisberger, U. A Hamiltonian Electrostatic Coupling Scheme for Hybrid Car–Parrinello Molecular Dynamics Simulations. *J. Chem. Phys.* **2002**, *116*, 6941–6947.
- (20) Dal Peraro, M.; Vila, A. J.; Carloni, P.; Klein, M. L. Role of Zinc Content on the Catalytic Efficiency of B1Metallo  $\beta$ -Lactamases. *J. Am. Chem. Soc.* **2007**, *129*, 2808–2816.
- (21) De Vivo, M.; Ensing, B.; Klein, M. L. Computational Study of Phosphatase Activity in Soluble Epoxide Hydrolase: High Efficiency through a Water Bridge Mediated Proton Shuttle. *J. Am. Chem. Soc.* **2005**, *127*, 11226–11227.
- (22) De Vivo, M.; Ensing, B.; Dal Peraro, M.; Gomez, G. A.; Christianson, D. W.; Klein, M. L. Proton Shuttles and Phosphatase Activity in Soluble Epoxide Hydrolase. *J. Am. Chem. Soc.* **2007**, *129*, 387–394.
- (23) Magistrato, A.; Ruggerone, P.; Spiegel, K.; Carloni, P.; Reedijk, J. Binding of Novel Azole-Bridged Dinuclear Platinum(II) Anticancer Drugs to DNA: Insights from Hybrid QM/MM Molecular Dynamics Simulations. *J. Phys. Chem. B* **2006**, *110*, 3604–3613.
- (24) Spiegel, K.; Rothlisberger, U.; Carloni, P. Cisplatin Binding to DNA Oligomers from Hybrid Car-Parrinello/Molecular Dynamics Simulations. *J. Phys. Chem. B* **2004**, *108*, 2699–2707.
- (25) Cascella, M.; Magistrato, A.; Tavernelli, I.; Carloni, P.; Rothlisberger, U. Role of Protein Frame and Solvent for the Redox Properties of Azurin from *Pseudomonas Aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 19641–19646.
- (26) Li, H.; Li, W.; Li, S.; Ma, J. Fragmentation-Based QM/MM Simulations: Length Dependence of Chain Dynamics and Hydrogen Bonding of Polyethylene Oxide and Polyethylene in Aqueous Solutions. *J. Phys. Chem. B* **2008**, *112*, 7061–7070.
- (27) He, X.; Wang, B.; Merz, K. M. Protein Nmr Chemical Shift Calculations Based on the Automated Fragmentation QM/MM Approach. *J. Phys. Chem. B* **2009**, *113*, 10380–10388.
- (28) Guo, W.; Xu, A.; Xu, X. XO: An Extended ONIOM Method for Accurate and Efficient Geometry Optimization of Large Molecules. *Chem. Phys. Lett.* **2010**, *498*, 203–208.
- (29) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. ONIOM: A Multilayered Integrated MO + MM Method for Geometry Optimizations and Single Point Energy Predictions. A Test for Diels–Alder Reactions and Pt(P(T-Bu)<sub>3</sub>)<sub>2</sub> + H<sub>2</sub> Oxidative Addition. *J. Phys. Chem.* **1996**, *100*, 19357–19363.
- (30) Gao, J.; Truhlar, D. G. Quantum Mechanical Methods for Enzyme Kinetics. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505.
- (31) Guimarães, C. R. W.; Repasky, M. P.; Chandrasekhar, J.; Tirado-Rives, J.; Jorgensen, W. L. Contributions of Conformational Compression and Preferential Transition State Stabilization to the Rate Enhancement by Chorismate Mutase. *J. Am. Chem. Soc.* **2003**, *125*, 6892–6899.
- (32) Tubert-Brohman, I.; Acevedo, O.; Jorgensen, W. L. Elucidation of Hydrolysis Mechanisms for Fatty Acid Amide Hydrolase and Its Lys142Ala Variant Via QM/MM Simulations. *J. Am. Chem. Soc.* **2006**, *128*, 16904–16913.
- (33) Alexandrova, A. N.; Jorgensen, W. L. Origin of the Activity Drop with the E50D Variant of Catalytic Antibody 34E4 for Kemp Elimination. *J. Phys. Chem. B* **2008**, *113*, 497–504.
- (34) Heimdal, J.; Kaukonen, M.; Srncic, M.; Rulisek, L.; Ryde, U. Reduction Potentials and Acidity Constants of Mn Superoxide Dismutase Calculated by QM/MM Free-Energy Methods. *Chem-PhysChem* **2011**, *12*, 3337–3347.
- (35) Gámiz-Hernández, A. P.; Galstyan, A. S.; Knapp, E.-W. Understanding Rubredoxin Redox Potentials: Role of H-Bonds on Model Complexes. *J. Chem. Theory Comput.* **2009**, *5*, 2898–2908.
- (36) Gámiz-Hernández, A. P.; Kieseritzky, G.; Ishikita, H.; Knapp, E. W. Rubredoxin Function: Redox Behavior from Electrostatics. *J. Chem. Theory Comput.* **2011**, *7*, 742–752.
- (37) McCullagh, M.; Voth, G. A. Unraveling the Role of the Protein Environment for [FeFe]-Hydrogenase: A New Application of Coarse-Graining. *J. Phys. Chem. B* **2013**, *117*, 4062–4071.
- (38) Hammes-Schiffer, S. Introduction: Proton-Coupled Electron Transfer. *Chem. Rev.* **2010**, *110*, 6937–6938.
- (39) Hammes-Schiffer, S.; Stuchebrukhov, A. A. Theory of Coupled Electron and Proton Transfer Reactions. *Chem. Rev.* **2010**, *110*, 6939–6960.
- (40) Hatcher, E.; Soudackov, A. V.; Hammes-Schiffer, S. Proton-Coupled Electron Transfer in Soybean Lipoyxygenase: Dynamical Behavior and Temperature Dependence of Kinetic Isotope Effects. *J. Am. Chem. Soc.* **2006**, *129*, 187–196.
- (41) Ding, F.; Tsao, D.; Nie, H.; Dokholyan, N. V. Ab Initio Folding of Proteins Using All-Atom Discrete Molecular Dynamics. *Structure* **2008**, *16*, 1010–1018.
- (42) Sparta, M.; Shirvanyants, D.; Ding, F.; Dokholyan, N. V.; Alexandrova, A. N. Hybrid Dynamics Simulation Engine for Metalloproteins. *Biophys. J.* **2012**, *103*, 767–776.

- (43) Dokholyan, N. V.; Buldyrev, S. V.; Stanley, H. E.; Shakhnovich, E. I. Discrete Molecular Dynamics Studies of the Folding of a Protein-Like Model. *Folding Des.* **1998**, *3*, 577–587.
- (44) Sparta, M.; Alexandrova, A. N. How Metal Substitution Affects the Enzymatic Activity of Catechol-O-Methyltransferase. *PLoS One* **2012**, *7*, e47172.
- (45) Valdez, C. E.; Alexandrova, A. N. Why Urease Is a Di-Nickel Enzyme Whereas the CcrA  $\beta$ -Lactamase Is a Di-Zinc Enzyme. *J. Phys. Chem. B* **2012**, *116*, 10649–10656.
- (46) Valdez, C. E.; Sparta, M.; Alexandrova, A. N. The Role of the Flexible L43-S54 Protein Loop in the CcrA Metallo- $\beta$ -Lactamase in Binding Structurally Dissimilar  $\beta$ -Lactam Antibiotics. *J. Chem. Theory Comput.* **2012**, *9*, 730–737.
- (47) Nedd, S.; Redler, R. L.; Proctor, E. A.; Dokholyan, N. V.; Alexandrova, A. N. Cu,Zn-Superoxide Dismutase without Zn Is Folded but Catalytically Inactive. *J. Mol. Biol.* **2014**, *426*, 4112–4124.
- (48) Sparta, M.; Valdez, C. E.; Alexandrova, A. N. Metal-Dependent Activity of Fe and Ni Acireductone Dioxygenases: How Two Electrons Reroute the Catalytic Pathway. *J. Mol. Biol.* **2013**, *425*, 3007–3018.
- (49) Calhoun, J. R.; Natri, F.; Maglio, O.; Pavone, V.; Lombardi, A.; DeGrado, W. F. Artificial Diiron Proteins: From Structure to Function. *Biopolymers* **2005**, *80*, 264–278.
- (50) Bell, C. B.; Calhoun, J. R.; Bobyr, E.; Wei, P.-p.; Hedman, B.; Hodgson, K. O.; DeGrado, W. F.; Solomon, E. I. Spectroscopic Definition of the Biferrous and Biferic Sites in De Novo Designed Four-Helix Bundle DFsc Peptides: Implications for O<sub>2</sub> Reactivity of Binuclear Non-Heme Iron Enzymes. *Biochemistry* **2008**, *48*, 59–73.
- (51) Kaplan, J.; DeGrado, W. F. De Novo Design of Catalytic Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 11566–11570.
- (52) Fry, H. C.; Lehmann, A.; Sinks, L. E.; Asselberghs, I.; Tronin, A.; Krishnan, V.; Blasie, J. K.; Clays, K.; DeGrado, W. F.; Saven, J. G.; Therien, M. J. Computational De Novo Design and Characterization of a Protein That Selectively Binds a Highly Hyperpolarizable Abiological Chromophore. *J. Am. Chem. Soc.* **2013**, *135*, 13914–13926.
- (53) Der, B. S.; Jha, R. K.; Lewis, S. M.; Thompson, P. M.; Guntas, G.; Kuhlman, B. Combined Computational Design of a Zinc-Binding Site and a Protein–Protein Interaction: One Open Zinc Coordination Site Was Not a Robust Hotspot for De Novo Ubiquitin Binding. *Proteins: Struct., Funct., Bioinf.* **2013**, *81*, 1245–1255.
- (54) Pabo, C. O.; Peisach, E.; Grant, R. A. Design and Selection of Novel Cys2His2 Zinc Finger Proteins. *Annu. Rev. Biochem.* **2001**, *70*, 313–340.
- (55) Yin, S.; Ding, F.; Dokholyan, N. V. Modeling Backbone Flexibility Improves Protein Stability Estimation. *Structure* **2007**, *15*, 1567–1576.
- (56) Liao, R.-Z.; Thiel, W. On the Effect of Varying Constraints in the Quantum Mechanics Only Modeling of Enzymatic Reactions: The Case of Acetylene Hydratase. *J. Phys. Chem. B* **2013**, *117*, 3954–3961.
- (57) Han, K. E.; Bystrhoff, C.; Baker, D. Three-Dimensional Structures and Contexts Associated with Recurrent Amino Acid Sequence Patterns. *Protein Sci.* **1997**, *6*, 1587–1590.
- (58) Shirvanyants, D.; Alexandrova, A. N.; Dokholyan, N. V. Rigid Substructure Search. *Bioinformatics* **2011**, *27*, 1327–1329.
- (59) Valdez, C. E.; Smith, Q. A.; Nechay, M. R.; Alexandrova, A. N. Mysteries of Metals in Metalloenzymes. *Acc. Chem. Res.* **2014**, *47*, 3110–3117.
- (60) Kuhlman, B.; O'Neill, J. W.; Kim, D. E.; Zhang, K. Y. J.; Baker, D. Accurate Computer-Based Design of a New Backbone Conformation in the Second Turn of Protein L. *J. Mol. Biol.* **2002**, *315*, 471–477.
- (61) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. Design of a Novel Globular Protein Fold with Atomic-Level Accuracy. *Science* **2003**, *302*, 1364–1368.
- (62) DiMaio, F.; Terwilliger, T. C.; Read, R. J.; Wlodawer, A.; Oberdorfer, G.; Wagner, U.; Valkov, E.; Alon, A.; Fass, D.; Axelrod, H. L.; Das, D.; Vorobiev, S. M.; Iwai, H.; Pokkuluri, P. R.; Baker, D. Improved Molecular Replacement by Density- and Energy-Guided Protein Structure Optimization. *Nature* **2011**, *473*, 540–543.
- (63) Huang, P.-S.; Oberdorfer, G.; Xu, C.; Pei, X. Y.; Nannenga, B. L.; Rogers, J. M.; DiMaio, F.; Gonen, T.; Luisi, B.; Baker, D. High Thermodynamic Stability of Parametrically Designed Helical Bundles. *Science* **2014**, *346*, 481–485.
- (64) Dahiyat, B. I.; Mayo, S. L. De Novo Protein Design: Fully Automated Sequence Selection. *Science* **1997**, *278*, 82–87.
- (65) Kuhlman, B.; Baker, D. Native Protein Sequences Are Close to Optimal for Their Structures. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10383–10388.
- (66) Ding, F.; Dokholyan, N. V. Emergence of Protein Fold Families through Rational Design. *PLoS Comput. Biol.* **2006**, *2*, e85.
- (67) Yin, S.; Ding, F.; Dokholyan, N. V. Eris: An Automated Estimator of Protein Stability. *Nat. Methods* **2007**, *4*, 466–467.
- (68) Reynolds, C. H.; Holloway, M. K. Thermodynamics of Ligand Binding and Efficiency. *ACS Med. Chem. Lett.* **2011**, *2*, 433–437.
- (69) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science* **2012**, *338*, 1042–1046.
- (70) Karplus, M.; Kushick, J. N. Method for Estimating the Configurational Entropy of Macromolecules. *Macromolecules* **1981**, *14*, 325–332.
- (71) Andricioaei, I.; Karplus, M. On the Calculation of Entropy from Covariance Matrices of the Atomic Fluctuations. *J. Chem. Phys.* **2001**, *115*, 6289–6292.
- (72) Schlitter, J. Estimation of Absolute and Relative Entropies of Macromolecules Using the Covariance Matrix. *Chem. Phys. Lett.* **1993**, *215*, 617–621.
- (73) Masgrau, L.; Truhlar, D. G. The Importance of Ensemble Averaging in Enzyme Kinetics. *Acc. Chem. Res.* **2015**, *48*, 431–438.
- (74) Poulsen, T. D.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. Free Energy Surface, Reaction Paths, and Kinetic Isotope Effect of Short-Chain Acyl-Coa Dehydrogenase. *J. Phys. Chem. B* **2003**, *107*, 9567–9578.
- (75) Pang, J.; Pu, J.; Gao, J.; Truhlar, D. G.; Allemann, R. K. Hydride Transfer Reaction Catalyzed by Hyperthermophilic Dihydrofolate Reductase Is Dominated by Quantum Mechanical Tunneling and Is Promoted by Both Inter- and Intramonomeric Correlated Motions. *J. Am. Chem. Soc.* **2006**, *128*, 8015–8023.
- (76) Hu, H.; Lu, Z.; Parks, J. M.; Burger, S. K.; Yang, W. Quantum Mechanics/Molecular Mechanics Minimum Free-Energy Path for Accurate Reaction Energetics in Solution and Enzymes: Sequential Sampling and Optimization on the Potential of Mean Force Surface. *J. Chem. Phys.* **2008**, *128*, 034105.
- (77) Ruiz-Pernía, J. J.; Garcia-Viloca, M.; Bhattacharyya, S.; Gao, J.; Truhlar, D. G.; Tuñón, I. Critical Role of Substrate Conformational Change in the Proton Transfer Process Catalyzed by 4-Oxalocrotonate Tautomerase. *J. Am. Chem. Soc.* **2009**, *131*, 2687–2698.
- (78) Senn, H. M.; Kästner, J.; Breidung, J.; Thiel, W. Finite-Temperature Effects in Enzymatic Reactions: Insights from QM/MM Free-Energy Simulations. *Can. J. Chem.* **2009**, *87*, 1322–1337.
- (79) Kanaan, N.; Ferrer, S.; Martí, S.; Garcia-Viloca, M.; Kohen, A.; Moliner, V. Temperature Dependence of the Kinetic Isotope Effects in Thymidylate Synthase. A Theoretical Study. *J. Am. Chem. Soc.* **2011**, *133*, 6692–6702.
- (80) Riniker, S.; Christ, C. D.; Hansen, H. S.; Hünenberger, P. H.; Oostenbrink, C.; Steiner, D.; van Gunsteren, W. F. Calculation of Relative Free Energies for Ligand-Protein Binding, Solvation, and Conformational Transitions Using the Gromos Software. *J. Phys. Chem. B* **2011**, *115*, 13570–13577.
- (81) Huang, W.; Lin, Z.; van Gunsteren, W. F. Use of Enveloping Distribution Sampling to Evaluate Important Characteristics of Biomolecular Force Fields. *J. Phys. Chem. B* **2014**, *118*, 6424–6430.
- (82) Movahed, H. B.; van Zon, R.; Schofield, J. Free Energy Landscape of Protein-Like Chains with Discontinuous Potentials. *J. Chem. Phys.* **2012**, *136*, 245103.
- (83) Martin-Diaconescu, V.; Maroney, M. J. *Comprehensive Inorganic Chemistry II*; Elsevier: Oxford, U.K., 2013.
- (84) Dai, Y.; Wensink, P. C.; Abeles, R. H. One Protein, Two Enzymes. *J. Biol. Chem.* **1999**, *274*, 1193–1195.

- (85) Pochapsky, T. C.; Ju, T.; Dang, M.; Beaulieu, R.; Pagani, G. M.; OuYang, B. Nickel in Acireductone Dioxygenase. *Nickel and Its Surprising Impact in Nature*; John Wiley & Sons, Ltd: Hoboken, NJ, 2007; pp 473–500.
- (86) Ju, T.; Goldsmith, R. B.; Chai, S. C.; Maroney, M. J.; Pochapsky, S. S.; Pochapsky, T. C. One Protein, Two Enzymes Revisited: A Structural Entropy Switch Interconverts the Two Isoforms of Acireductone Dioxygenase. *J. Mol. Biol.* **2006**, *363*, 823–834.
- (87) Allpress, C. J.; Grubel, K.; Szajna-Fuller, E.; Arif, A. M.; Berreau, L. M. Regioselective Aliphatic Carbon–Carbon Bond Cleavage by a Model System of Relevance to Iron-Containing Acireductone Dioxygenase. *J. Am. Chem. Soc.* **2012**, *135*, 659–668.
- (88) Männistö, P. T.; Kaakkola, S. Catechol-O-Methyltransferase (COMT): Biochemistry, Molecular Biology, Pharmacology, and Clinical Efficacy of the New Selective COMT Inhibitors. *Pharmacol. Rev.* **1999**, *51*, 593–628.
- (89) Guldberg, H. C.; Marsden, C. A. Catechol-O-Methyl Transferase: Pharmacological Aspects and Physiological Role. *Pharmacol. Rev.* **1975**, *27*, 135–206.
- (90) Lotta, T.; Vidgren, J.; Tilgmann, C.; Ulmanen, I.; Melen, K.; Julkunen, I.; Taskinen, J. Kinetics of Human Soluble and Membrane-Bound Catechol O-Methyltransferase: A Revised Mechanism and Description of the Thermolabile Variant of the Enzyme. *Biochemistry* **1995**, *34*, 4202–4210.
- (91) Vidgren, J.; Svensson, L. A.; Liljas, A. Crystal Structure of Catechol O-Methyltransferase. *Nature* **1994**, *368*, 354–358.
- (92) Rutherford, K.; Le Trong, I.; Stenkamp, R. E.; Parson, W. W. Crystal Structures of Human 108V and 108M Catechol O-Methyltransferase. *J. Mol. Biol.* **2008**, *380*, 120–130.
- (93) Axelrod, J.; Tomchick, R. Enzymatic O-Methylation of Epinephrine and Other Catechols. *J. Biol. Chem.* **1958**, *233*, 702–705.
- (94) Khare, S. D.; Dokholyan, N. V. Common Dynamical Signatures of Familial Amyotrophic Lateral Sclerosis-Associated Structurally Diverse Cu, Zn Superoxide Dismutase Mutants. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 3147–3152.
- (95) Pardo, C. A.; Xu, Z.; Borchelt, D. R.; Price, D. L.; Sisodia, S. S.; Cleveland, D. W. Superoxide Dismutase Is an Abundant Component in Cell Bodies, Dendrites, and Axons of Motor Neurons and in a Subset of Other Neurons. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 954–958.
- (96) Valentine, J. S.; Hart, P. J. Misfolded CuZnSOD and Amyotrophic Lateral Sclerosis. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3617–3622.
- (97) Wilcox, K. C.; Zhou, L.; Jordon, J. K.; Huang, Y.; Yu, Y.; Redler, R. L.; Chen, X.; Caplow, M.; Dokholyan, N. V. Modifications of Superoxide Dismutase (SOD1) in Human Erythrocytes: A Possible Role in Amyotrophic Lateral Sclerosis. *J. Biol. Chem.* **2009**, *284*, 13940–13947.
- (98) Sheng, Y.; Abreu, I. A.; Cabelli, D. E.; Maroney, M. J.; Miller, A.-F.; Teixeira, M.; Valentine, J. S. Superoxide Dismutases and Superoxide Reductases. *Chem. Rev.* **2014**, *114*, 3854–3918.
- (99) Lyons, T. J.; Liu, H.; Goto, J. J.; Nersissian, A.; Roe, J. A.; Graden, J. A.; Café, C.; Ellerby, L. M.; Bredesen, D. E.; Gralla, E. B.; Valentine, J. S. Mutations in Copper-Zinc Superoxide Dismutase That Cause Amyotrophic Lateral Sclerosis Alter the Zinc Binding Site and the Redox Behavior of the Protein. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12240–12244.
- (100) Strange, R. W.; Antonyuk, S.; Hough, M. A.; Doucette, P. A.; Rodriguez, J. A.; Hart, P. J.; Hayward, L. J.; Valentine, J. S.; Hasnain, S. S. The Structure of Holo and Metal-Deficient Wild-Type Human Cu, Zn Superoxide Dismutase and Its Relevance to Familial Amyotrophic Lateral Sclerosis. *J. Mol. Biol.* **2003**, *328*, 877–891.
- (101) Meunier, B. *Biomimetic Oxidations Catalyzed by Transition Metal Complexes*; Imperial College Press: London, 2000.
- (102) Fiedler, A. T.; Bryngelson, P. A.; Maroney, M. J.; Brunold, T. C. Spectroscopic and Computational Studies of Ni Superoxide Dismutase: Electronic Structure Contributions to Enzymatic Function. *J. Am. Chem. Soc.* **2005**, *127*, 5449–5462.
- (103) Barondeau, D. P.; Kassmann, C. J.; Bruns, C. K.; Tainer, J. A.; Getzoff, E. D. Nickel Superoxide Dismutase Structure and Mechanism. *Biochemistry* **2004**, *43*, 8038–8047.
- (104) Dokholyan, N. V. Studies of Folding and Misfolding Using Simplified Models. *Curr. Opin. Struct. Biol.* **2006**, *16*, 79–85.
- (105) Ding, F.; Guo, W.; Dokholyan, N. V.; Shakhnovich, E. I.; Shea, J. E. Reconstruction of the Src-Sh3 Protein Domain Transition State Ensemble Using Multiscale Molecular Dynamics Simulations. *J. Mol. Biol.* **2005**, *350*, 1035–1050.
- (106) Ding, F.; Tsao, D.; Nie, H.; Dokholyan, N. V. Ab Initio Folding of Proteins with All-Atom Discrete Molecular Dynamics. *Structure* **2008**, *16*, 1010–1018.
- (107) Wuerges, J.; Lee, J.-W.; Yim, Y.-I.; Yim, H.-S.; Kang, S.-O.; Carugo, K. D. Crystal Structure of Nickel-Containing Superoxide Dismutase Reveals Another Type of Active Site. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8569–8574.
- (108) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nat. Rev. Drug Discovery* **2004**, *3*, 935–949.
- (109) Jacobsen, J. A.; Fullagar, J. L.; Miller, M. T.; Cohen, S. M. Identifying Chelators for Metalloprotein Inhibitors Using a Fragment-Based Approach. *J. Med. Chem.* **2010**, *54*, 591–602.
- (110) Chaskar, P.; Zoete, V.; Röhrig, U. F. Toward on-the-Fly Quantum Mechanical/Molecular Mechanical (QM/MM) Docking: Development and Benchmark of a Scoring Function. *J. Chem. Inf. Model.* **2014**, *54*, 3137–3152.
- (111) Alvarez-Garcia, D.; Barril, X. Relationship between Protein Flexibility and Binding: Lessons for Structure-Based Drug Design. *J. Chem. Theory Comput.* **2014**, *10*, 2608–2614.
- (112) Feixas, F.; Lindert, S.; Sinko, W.; McCammon, J. A. Exploring the Role of Receptor Flexibility in Structure-Based Drug Discovery. *Biophys. Chem.* **2014**, *186*, 31–45.