Understanding the Role of Geometric and Electronic Structure in Bioinspired Catalyst Design: the Case of Formate Dehydrogenase

Mingjie Liu[1], Azadeh Nazemi[1], Michael G. Taylor[1], Aditya Nandy[1,2], Chenru Duan[1,2], Adam H. Steeves[1], and Heather J. Kulik[1,*]

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139*

[2]*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139*

ABSTRACT: The design of bioinspired synthetic inorganic molecular complexes is challenging, due to a lack of understanding of enzyme action and the degree to which that action can be translated into mimics. Exemplary of this challenge is the reversible conversion of formate into CO_2 by formate dehydrogenase (FDH) enzymes with Mo/W centers in large molybdopterin cofactors. Despite numerous efforts to synthesize Mo/W-containing molecular complexes, none have been demonstrated to reproduce the full reactivity of FDH. Here, we carry out a large-scale, high-throughput screening study on all mononuclear Mo/W complexes currently deposited in Cambridge Structural Database (CSD). Using density functional theory, we systematically investigate the individual effects of metal identity, ligand identity, oxidation state, and coordination number on structural, electronic and catalytic properties. We compare our results on molecular complexes with quantum mechanics/molecular mechanics simulations on a representative FDH enzyme to further elucidate the influence of the enzyme environment. These comparisons reveal that the enzyme environment primarily influences the metal-local geometry, and these metal-local structural variations can improve catalysis. Through a series of computational mutations on molecular complexes, we extend beyond the CSD structures to further identify the limits of varied chalcogen and metal identity. This broad set and comparison reveal relatively little variation of electronic properties of the metal center due to the presence of the enzyme environment or changes in metal-distant ligand chemistry. Instead, these properties are found to be much more sensitive to the identity of the metal and the nature of the bound terminal chalcogen.

1. Introduction

Sustained experimental effort towards the understanding of both biological enzymes and bioinspired synthetic inorganic molecular transition metal catalysts has led to the increased availability of properties and structures of both enzymes and molecular complexes.[1-8] Computational first-principles, quantum mechanical (QM, e.g., with density functional theory or DFT) and multi-scale (i.e., mixed quantum mechanics/molecular mechanics, QM/MM) modeling plays an essential role in revealing properties of short-lived catalytic intermediates and elucidating structure-property relationships. Nevertheless, challenges associated with comparing enzyme properties to their molecular mimics has limited the ability to understand similarities and differences of bioinspired molecular catalysts to enzymes. Due to the dynamic enzyme environment during catalysis, the mechanistically relevant structures in the enzyme cofactor are ill-defined and challenging to identify both by experiment and computation.

The metal-cofactor-dependent enzyme formate dehydrogenase (FDH) catalyzes the reversible conversion of formate into CO_2 on a single Mo/W site coordinated by two bidentate molybdopterin ligands, a terminal oxo or sulfide ligand, and the amino acid side chains of serine (Ser), cysteine (Cys), or selenocysteine (Sec).[9-17] Despite the well-defined coordination environment of this metallocofactor, decades of work has been unable to uncover synthetic inorganic Mo/W molecular complexes with comparable function to FDH enzymes.[9, 18-23] In particular, the mechanism of formate oxidation and CO_2 reduction on FDH enzymes remains elusive, and metal coordinating species and proximal residues in the cofactor remain under debate. Most mechanistic arguments invoke substrate binding to the metal or terminal chalcogen.[24-34] In the most widely accepted mechanism for formate oxidation, the terminal sulfido ligand (M=S) in the FDH enzyme abstracts the hydride in HCO_2^-, to form a terminal

metal-thiol. During this process, the ligands in the primary coordination sphere of the metal remain coordinated[28, 31], although there has been support for dissociation of the coordinating Sec or Cys[25]. Over this catalytic cycle, the oxidation state of the metal is believed to cycle between M(VI), M(V) and M(IV).[35]

Inspired by structural knowledge of the cofactors of FDH enzymes, many synthetic molecular complexes have a mononuclear Mo/W coordinated by two bisdithiolene ligands, a terminal oxo/sulfido/selenido, and a coordinating serinate-, cysteinate-, or selenocysteinate-like ligand.[9, 36-40] Due to the limitations from both synthesis and characterization for each structure by experiment, the exploration on different ligand and metal identities is narrow. For example, some selenolate complexes are difficult to synthesize, and some Mo(VI) complexes with dimethyldithiolene ligands are unstable and are easily reduced to Mo(V), making them challenging to isolate.[9, 41] Within enzymes, Mo- and W-containing subunits are isostructural with same local ligands and residues, but W-containing enzymes oxidize formate faster than Mo-containing counterparts without a fundamental understanding of this observation.[10] Similarly, the selenocysteine-containing enzyme is observed to be faster than cysteine-containing one.[10] The role of the terminal chalcogen also remains unclear, although a terminal metal-sulfido is believed to be essential for the formate oxidation[42, 43], a terminal metal-oxo is sometimes observed[39, 44]. The influence of the terminal chalcogen identity (O or S or Se) towards reactivity remains unclear.[40, 45]

Establishing molecular structure-property relationships will facilitate bioinspired catalyst design. Virtual high throughput screening (VHTS) has unlocked the development of structure-property relationships in inorganic chemistry for catalyst discovery over past few years.[46-49]

VHTS and mechanistic computational modeling are promising paths to provide insights which are otherwise difficult or impossible to obtain from experiments.

We carry out a large-scale screen of synthesized Mo/W complexes from Cambridge Structural Database (CSD).[50] We select the complexes that contain many features of the cofactors in FDH enzymes and generate more configurations by mutating elemental identities beyond those present in CSD complexes. Over this expanded chemical space, we perform density functional theory (DFT) calculations to comprehensively understand structure-property relationships. We draw the first structure-property comparisons between molecular complexes and biological metallocofactors by comparing our DFT calculations on molecular complexes to observations from molecular dynamics (MD) and QM/MM simulations on a representative FDH enzyme structure. We use VHTS to isolate the effects of the metal and ligand identities, metal-local structure, and electronic properties that are all related to formate oxidation catalysis. Through this analysis, we demonstrate that the enzyme environment primarily affects geometric properties, and that the terminal chalcogen moieties primarily influence the local electronic properties.

2. Methods.

2a. Data Set Curation and Computational Details.

A set of mononuclear Mo/W transition metal complexes was curated from the Cambridge Structural Database (CSD)[50]. This procedure employed both the Conquest graphical interface to the CSD as well as the Python application programming interface (API),[50] in all cases applied to the CSD v5.41 from the November 2019 data set. Conquest was used to query for individual molecules with one of either Mo or W bound to 3 atoms of any element in addition to at least one of either S or Se. The initial search produced 3,257 complexes with associated CSD refcodes.

After several filtering steps to remove duplicates and structures with inconsistent oxidation states, we obtained a final set of 404 complexes (Supporting Information Text S1).

We performed gas phase geometry optimizations and single-point energy calculations on the 404 Mo/W complexes using density functional theory (DFT) with a development version of TeraChem[51] v1.9. The B3LYP[52-54] global hybrid functional was employed throughout with the basis set consisted of an LANL2DZ[55] effective core potential for heavy atoms (i.e., Mo, W, Br, I) and the 6-31G* basis for all other atoms. Singlet calculations were carried out in a spin-restricted formalism, whereas all other spin states were performed as unrestricted calculations. Level shifting was employed with a value of 0.25 Ha for both occupied and virtual orbitals. Geometry optimizations were carried out with the translational rotation internal coordinates (TRIC) optimizer[56] using the L-BFGS algorithm in TRIC coordinates. Default convergence thresholds of 4.5×10^{-4} hartree/bohr for the maximum gradient and 1×10^{-6} hartree for the change in self-consistent field (SCF) energy between steps were employed. All calculations are analyzed for consistent connectivity before and after the geometry optimizations using molSimplify[57, 58]. We found that all complexes satisfied connectivity checks[58] and display only minor deviations after optimization. We observed the average of the metal-ligand bond length change after geometry optimization to be 0.05 Å (Supporting Information Figure S1). We only considered the low spin (LS) singlet or doublet states. We determined whether complexes were singlets or doublets by counting electrons (e.g. assuming neutral atoms) and then accounting for the net charge assigned by the user. Thus, systems with an odd number of electrons are assigned as doublets, and those with an even number of electrons are assigned as singlets. From the oxidation states obtained from the CSD and user-defined charges, after counting the expected number of electrons, we identified seven complexes with radical ligands or mislabeled oxidation/charge states and used

their corrected oxidation/charge states in subsequent analysis (Supporting Information Table S1). For all open-shell calculations, no spin contamination was observed, as judged by prior procedures[58-60].

The implicit-aqueous-solvent-corrected energies were used to calculate the vertical ionization potential (IP) and electron affinity (EA). Aqueous solvent effects were modeled using a conductor-like polarizable continuum model[61, 62] (C-PCM), with $\varepsilon = 78.39$ and defaults of 1.2× Bondi's van der Waals radii[63]. Single point energy calculations were performed with C-PCM on gas phase optimized geometries. For complexes with doublet spin states (i.e., all M(V) complexes), we considered only the singlet spin state in the oxidized species and both singlet and triplet spin states in the reduced species (Supporting Information Table S2). Since partial charges can be sensitive to basis set and population scheme[64], we computed partial charges with three schemes: Mulliken[65] and natural bond orbital charge[66], both based on orbital charge partitioning, and Hirshfeld charge[67], based on real space charge distribution. We quantified the Mayer bond valence of the metal as a measure of the degree of covalency[68] for metal-ligand bonds.

We used the hydrogen atom bound to the terminal chalcogen as a screening tool to understand the catalytic activity for formate oxidation to CO_2 (Supporting Information Text S2). To quantify the energetics of structural reorganization, we performed constrained geometry optimizations on hydrogen atom bound structures by fixing the coordinates of all atoms in the transition metal complex except the terminal chalcogen and the adsorbed hydrogen atom, which were allowed to relax. The differences in E_b(H) with and without constraints (ΔE_b) quantify the energetics of structural reorganization.

2b. Protein Preparation and QM Cluster Optimization.

An X-ray crystal structure of an oxidized form of the formate dehydrogenase (FDH) enzyme was obtained from the protein data bank (PDB ID: 1FDO[42]). The active site of 1FDO contains selenocysteine, Mo, two molybdopterin guanine dinucleotide (MGD) cofactors, in addition to a Fe_4S_4 cluster. To prepare the protein, the Fe_4S_4 cluster was removed from the crystal structure and a sulfido ligand was added manually as the sixth coordinating ligand. The coordinating selenocysteine was converted to a cysteine. The protonation states of the apoenzyme residues were assigned using the H++ webserver[69-72] assuming a pH of 7.0 with all other defaults applied. Force-field parameters for the core active site were obtained from a combination of the MCPB.py[73] utility in AMBER18[74] and the parameters available in literature[75]. Standard protein residues were simulated with AMBER ff14SB[76]. The protein was solvated with 10 Å of TIP3P[77] water buffer in a periodic rectangular prism box and neutralized with Na^+ counterions. Finally, the initial AMBER topology and coordinate files were generated for the 77,778-atom system, which are provided in the Supporting Information. These coordinates were used as a starting point to carry out 100 ns of MD simulation (Supporting Information Text S3).

To do further analysis on the active site of enzyme, a 376-atom QM region was selected systematically using charge shift analysis method[78, 79] (Supporting Information Figure S2). The QM cluster optimizations were carried out using a range-separated exchange-correlation functional ωPBEh[80] ($\omega = 0.2$ bohr^{-1}) along with the composite LACVP* basis set (i.e., LANL2DZ effective core potential[81] on Mo and W heavy atoms along with 6-31G*[82] for the remaining atoms) in using the C-PCM[61, 62] model (with $\varepsilon = 4$), while all backbone atoms (i.e., C_α, N, C, O in protein residues) were constrained. The level of theory was chosen based on prior studies that show degraded performance of global hybrids (e.g., B3LYP) with increasing QM region size[78, 83].

2c. Overview of the CSD data set and workflow for virtual high-throughput screening.

Analysis of the synthetic Mo/W complex data set, including preferred coordination numbers, oxidation states, and ligand denticity, provides insight into the chemistry of experimentally-studied Mo/W-containing compounds. Out of 242 Mo and 162 W complexes, the majority (Mo: 76%, W: 74%) are 5- or 6-coordinated. In addition, complexes in the set contain primarily high valent (e.g., IV, V, and VI) oxidation states (Mo: 92%, W: 87%). Most complexes (74%, 300 out of 404) have a maximum ligand denticity of two (Figure 1 and Supporting Information Tables S3–S4). These observations suggest the studied inorganic compounds should share similarities with the cofactors in Mo/W-containing enzymes[12].

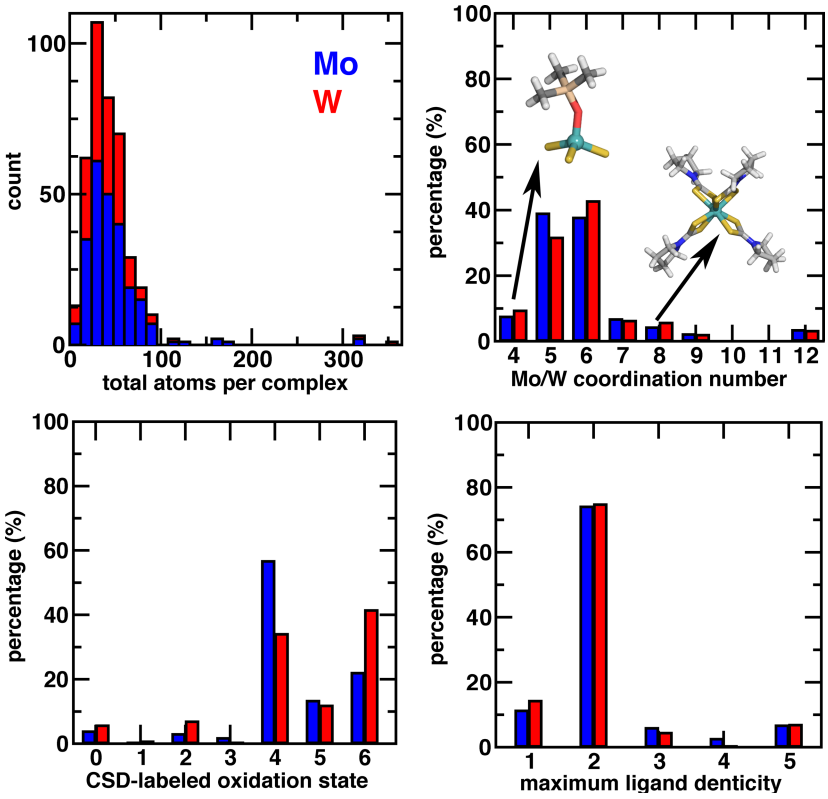

Figure 1. (top, left) Stacked histogram of the total atoms per complex for Mo (blue) and W (red). Statistics of Mo and W based on the Mo/W coordination number (top, right), CSD-labeled oxidation state (bottom, left), and maximum ligand denticity (bottom, right). The percentage was calculated for Mo (242 complexes) and W (162 complexes) separately. The inset structures are CSD refcode: DARGEI (left) and CSD refcode: YAWKEM (right). Structures are shown as sticks with carbon in gray, nitrogen in blue, hydrogen in white, oxygen in red, sulfur in yellow, silicon in light orange, and molybdenum in cyan.

Dithiolene ligands are the most common class of bidentate sulfur ligands that occur in FDH and other Mo/W containing enzymes[35]. We find that only 26% of the 404 Mo/W compounds lack bidentate sulfur ligands (Supporting Information Table S5). We expect this prevalence is due to the good stability of Mo/W molecular complexes formed with bidentate sulfur ligands since analysis of the source papers suggests that only a small number are directly intending to mimic the catalysis of FDH or related enzymes. In FDH enzymes, the mononuclear metal center is coordinated by two dithiolene ligands and at least one monodentate terminal chalcogen (i.e., oxo, sulfido, or selenido) ligand. Several recent studies[27, 28, 84] have demonstrated this monodentate terminal chalcogen (tc) ligand to be the binding site for a hydrogen atom during hydrogen atom transfer in putative FDH catalytic mechanisms[27].

After isolating complexes that contain two dithiolene ligands and at least one terminal chalcogen as the most representative mimics of FDHs, 88 complexes remain. Next, we mutated any terminal chalcogen sites in this set with O, S and Se. Prior to this step, only 7 complexes contained sulfido and only 4 selenido ligands, with the majority of compounds containing terminal oxo moieties. After performing the mutations, we obtained a dataset of 348 complexes that has increased balance between different terminal chalcogens in comparison to the original CSD-derived set (Supporting Information Table S6). On this set, we use the molSimplify-computed molecular graph determinant[85] to perform a connectivity uniqueness check after the chalcogen mutation to avoid repeat simulations. We remove 93 duplicates through this check, producing 255 complexes with unique connectivity, which increases to 276 if also considering CSD-assigned metal oxidation states (Supporting Information Tables S7–S8). The new dataset with 276 complexes derived from 88 bisdithiolene Mo/W complexes are used for geometric and electronic properties analysis (see Sec. 3).

3. Results and Discussion.

3a. Structural Properties of Bisdithiolene Complexes.

Over our curated set of experimental compounds, we observe the dithiolene ligands adopt a range of orientations that could be expected to influence their reactivity. Since the dithiolene ligands and the metal center are known to form a range of geometries[9], this motivates identifying a continuous descriptor of the metal-ligand primary coordination sphere. There are a number of metrics that have been used to described the angles of sulfur ligands with metal centers, including the largest angle between the metal and two sulfur atoms in different dithiolenes[86], explicitly listing all the angles forms by the metal and coordinated sulfur atoms[36, 38, 87], the perpendicular displacement of the metal atom from the S4 least-squares plane[9, 88, 89], and the dihedral angle of four sulfur atoms in bisdithiolene[90]. We selected two general descriptors closely related to those that have been previously proposed but can describe the variation we observe over our larger set of complexes. To describe the geometry of the metal and four sulfur coordinating atoms in bisdithiolene, we define two key angles: the twisting and folding angles (Figure 2). We define the twisting angle, θ_t, as the angle between the sulfur atoms of the two bisdithiolene ligands, capturing the extent to which the coordinating sulfur atoms lie in the same plane. We define the folding angle, θ_f, as 180°-θ, where the θ is the angle formed by the metal at the vertex with the non-bonded S-S midpoint for each bidentate ligand. The folding angle measures the extent to which the metal is puckered out of the plane of the sulfur coordinating atoms. For conjugated (i.e., planar) ligands, the folding angle also captures the extent to which ligands are oriented in the same plane. Both the twisting and folding angles are between 0-90° in our definition. It has been argued that the folding effect is the primary factor that linked with the

functionality of the enzymes, where the folding imposes interaction between the sulfur π orbital and the empty *d* orbital of the metal to lower the energy[91, 92].

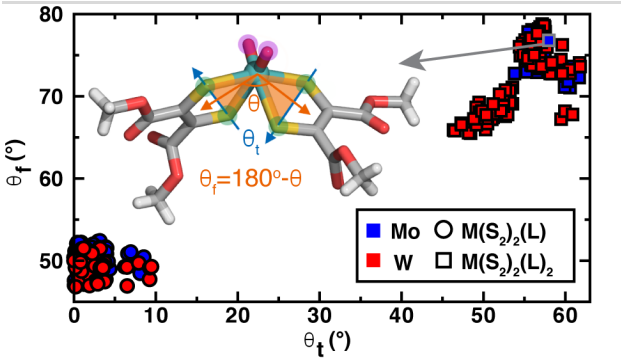

Figure 2. The twisting and folding angles of Mo/W complexes with bisdithiolene ligands. Mo complexes are shown in blue and W complexes are shown in red. Five-coordinated compounds are shown with circles, and six coordinated compounds are shown with squares. An example Mo compound demonstrates how the twisting angle (θ_t) and folding angle (θ_f) are defined, with carbon atoms shown in gray, sulfur in yellow, hydrogen in white, oxygen in red, and molybdenum in cyan.

By using twisting and folding angles to quantify the bisdithiolene ligand orientation, we intend to determine if any catalytic or electronic properties of the complexes correlate with these two angles. We measured the angles for 276 complexes and grouped each complex by metal center (i.e., Mo or W) and metal coordination numbers (i.e., 5 or 6, Figure 2). Overall, the twisting and folding angles appear related to metal coordination numbers, with smaller angles (θ_t: 0-14°, θ_f: 46-53°) for five-coordinated complexes relative to the six-coordinated (θ_t: 46-62°, θ_f: 65-79°). Oxidation state differences, on the other hand, do not distinguish these angles, i.e., all M(VI) and M(V) are five-coordinated complexes, and so their folding and twisting angles are comparable.

To understand if there is a structural difference between molecular complexes and geometries adopted by the cofactors in enzymes, we measured the folding and twisting angles of the crystal structures for the 15 FDH enzymes (i.e., 13 M(VI) and two M(IV)) reported in the

PDB[93] (Supporting Information Figure S3). Comparing the M(VI) protein cofactors with equivalent molecular complexes, the protein cofactors have both smaller twisting angles (21-36° vs 46-62°) and folding angles (56-73° vs 65-79°) than the molecular complexes. For the two Mo(IV) cofactors, the twisting angles (8-9°) are within the range spanned by five-coordinate molecular complexes (0-14°), whereas the cofactor folding angles (39° and 57°) reside on either side of the range (46-53°) spanned by M(IV) complexes.

To investigate how the enzyme environment might influence the structure of the cofactor, we carried out MD simulations on one FDH enzyme (PDB: 1FDO[42]), which includes a Mo(VI)-containing cofactor with twisting and folding angles (27° and 57°, respectively) representative of FDH enzymes. During MD simulations, the twisting and folding angles remain consistent with the crystal structure (i.e., 29 ± 2° and 63 ± 2°, respectively). We extracted a structure from the end of the MD trajectory and carried out QM analysis on the cofactor both in the enzyme and upon extraction from the enzyme. The QM enzyme environment model includes a large number of systematically selected residues (Supporting Information Figures S2 and S4). Comparisons between angles obtained from the QM cluster (θ_t: 38°,θ_f: 64°) and those from optimization of the isolated cofactor (θ_t: 49°, θ_f: 68°) reveals that the residues surrounding the cofactor in the enzyme have the effect of decreasing both angles. When the cofactor is extracted from the enzyme, it adopts angles much more comparable to that observed in molecular complexes.

The terminal chalcogen is the reactive species during formate oxidation. Thus, in addition to characterizing the angles of the metal-dithiolene ligand, we also analyzed the bond lengths between the terminal chalcogen and the metal, d(M-tc). We observed that bond lengths are independent of the metal coordination number and local dithiolene geometry (Supporting Information Figure S5). We also observed the average bond length of the metal and the

dithiolene sulfur atoms, d(M-sd), are shorter for five-coordinated complexes than six-coordinated complexes (i.e., 2.4-2.5 Å vs 2.45-2.62 Å, Supporting Information Figure S5).

Considering the difference in covalent radii among the species studied, we next computed the relative metal-ligand bond length where the d(M-tc) is normalized by the sum of the relevant metal (i.e., Mo or W) and chalcogen (i.e., O, S, or Se) covalent radii[85]. We also divided the complexes into groups, as determined by the metal types (i.e., Mo or W) and the number of terminal chalcogen atoms (i.e., one: 1tc or two: 2tc, Figure 3). We observe period dependence in the metal-chalcogen bond, with the oxo ligand forming the shortest bond with the metal, followed by the sulfido- and selenido- ligands, and shorter bonds for W over Mo. The 2tc complexes generally form longer relative metal-ligand bonds compared to the 1tc complexes.

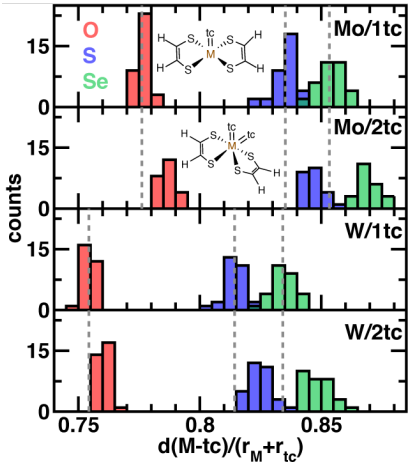

Figure 3. Histograms of relative metal-terminal-chalcogen (tc) bond lengths for Mo and W complexes with one terminal chalcogen (1tc) and two terminal chalcogen (2tc) atoms. The tc with O, S, and Se are shown as translucent bars in red, blue and green respectively. Vertical dotted lines indicate the average values of Mo(W)/1tc for O, S and Se.

Overall, in the bisdithiolene Mo/W complexes, the metal-local bisdithiolene ligand geometries (i.e., angles and bond distances) are closely related to the metal coordination numbers: five-coordinated complexes tend to be planar and compact with smaller folding and twisting angles and shorter metal-dithiolene bonds, while six-coordinated complexes tend to be three-

dimensional and elongated with larger folding and twisting angles and longer metal-dithiolene bonds.

3b. Electronic Properties of Bisdithiolene Complexes.

The differences in chemical bonding between the metal (i.e., Mo or W) and terminal chalcogen atoms (i.e., oxo, sulfido or selenido) can be expected to lead to differences in electronic properties of the model complexes in our set. To uncover trends in electronic properties, we first computed the metal partial charge and bond valence (i.e., degree of covalency between the metal and coordinating atoms) for 276 complexes. We also computed these properties for the cofactor in a representative Mo-containing FDH enzyme[42]. This analysis allows us to understand the influence of the metal and terminal chalcogen on those electronic properties. As in the analysis of relative metal-ligand bond lengths, we divided the complexes based on the number of terminal chalcogen moieties, atom types, and metal identities. Within the 2tc set, the chalcogen atoms can be identical or a combination of two chalcogen elements, and they are always six-coordinate. Both five- and six-coordinate complexes are present in the 1tc set, where the higher-coordinate species have as the sixth ligand a halogen (e.g., Cl) or a monodentate ligand (i.e., O/S/Se-R).

Across our 276 complexes, the metal Mulliken[65] and natural bond orbital (NBO)[66] partial charges of these complexes vary widely and are, at times, negatively valued (Supporting Information Figures S6–S7). Given the high-valent nature of the metal center, large negative values of both charge schemes are surprising and could be sensitive to the choice of basis set and population scheme[64]. We thus also compared to the metal partial charge obtained with real space partitioning (i.e., Hirshfeld charges[67] which instead are always positive and in a much narrower range (from 0.0 to 0.7 e, Supporting Information Figure S8). Despite differences in quantitative

values, the qualitative trends for all three partial charge schemes follow expectations based on differences in terminal chalcogen electronegativities (i.e., O>S>Se), with terminal-oxo complexes having more positive metal partial charges than terminal-sulfido or terminal-selenido complexes. When two terminal chalcogen atom types are present, the effect on partial charge is additive, with values in between the limits of the complexes with two identical chalcogen atoms. We also observe the expected metal formal oxidation state dependence in Hirshfeld partial charge analysis, where metals with higher oxidation states have higher positive partial charges (Supporting Information Figure S8). In addition to these expected trends, the metal identity has a distinct influence on the partial charge, with a lower partial charge on W in most complexes in comparison to Mo (Supporting Information Figure S8). Overall, we observe that the partial charge of the metal center is most strongly determined by the terminal chalcogen atoms that coordinate it but also by the metal identity and oxidation state.

We also analyzed the metal bond valence as a measure of the degree of covalency between the metal and its ligands both with varying metal identity and changes in formal oxidation state. Intuitively, higher formal oxidation states correspond to increased metal bond valence and correspondingly stronger metal-ligand bonds (Figure 4 and Supporting Information Figure S9). The terminal oxo ligands have reduced metal bond valence relative to sulfido and selenido groups (Figure 4). Similar to the Hirshfeld partial charge, we observe the metal bond valence to be influenced by the metal identity. For the same formal oxidation state, W complexes exhibit larger metal bond valence than Mo complexes (Figure 4 and Supporting Information Figure S9). The additional, non-terminal-chalcogen ligands also influence the bond valence, suggesting the possibility of tuning covalency in transition metal complex design. When this additional ligand is S/Se-R or Cl, the metal bond valence is higher than when an O-R ligand is

present for M(VI) complexes (Figure 4). Unlike the geometric features, the metal partial charge and bond valence have less dependence on the ligand architecture over our screened set, but the ligand and metal types play a more significant role in tuning these electronic properties that can be expected to be related to catalytic and other experimentally-observable, electronic properties.

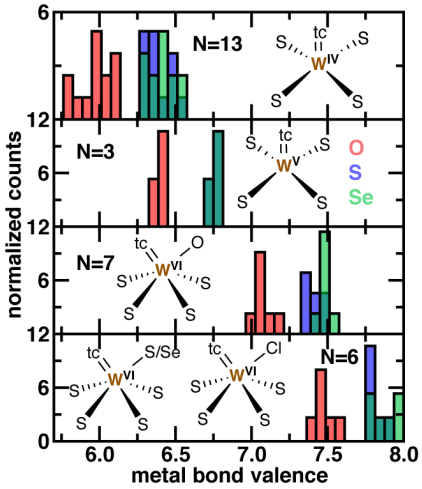

Figure 4. Normalized histograms of the metal bond valence for 87 W-complexes with one terminal chalcogen atom (1tc). The O, S, and Se terminal chalcogen atoms are shown as translucent bars in red, blue and green respectively. The W oxidation states are shown in the inset structures. The total number of complexes with O, S, or Se used to compute each histogram is indicated in each panel.

For the cofactor of the Mo-containing FDH enzyme[42], we also computed the metal Hirshfeld[67] partial charge and Mayer bond valence[68] both in the presence of key enzyme residues and in isolation (see Computational Details and Supporting Information Figures S2 and S4). To compare properties of the enzyme with trends from the complexes, we also mutated the terminal sulfido to both a terminal oxo and terminal selenido in combination with both Mo and W ions. Similar to what we observed in molecular complexes, the metal partial charge on the enzyme cofactor is largest with a terminal oxo followed by sulfido and selenido groups and lower (i.e., less positive) partial charges are observed for W. This good correspondence between the enzyme and the molecular complex electronic structure is perhaps expected given that consistent

qualitative trends are observed whether we study the isolated cofactor or in the presence of the enzyme environment (Supporting Information Table S11 and Figures S2 and S4). Instead focusing on what is most influenced by the enzyme environment within a fixed metal and terminal chalcogen, we note that positive metal partial charges are increased in the presence of the enzyme (Supporting Information Table S11). Thus, while qualitative ordering by metal or terminal chalcogen is consistent across the molecular complexes and enzymes, the enzyme environment plays a role in making the metal more electron poor.

For the closely-related metal bond valence property, we unsurprisingly observe similar qualitative ordering by metal ion or terminal chalcogen for both the enzyme as we did for molecular complexes (Supporting Information Table S12). As with molecular complexes, the bond valence is smallest in the case of a metal-oxo, and it is generally larger for W than Mo enzyme models. Comparing within each type of terminal ligand and metal, we isolate the effect that the enzyme environment has on metal bond valence. In all cases, the metal center bond valence increases in the presence of the enzyme environment, typically by around a comparable 0.2 bond order increment (Supporting Information Table S12). Thus, in contrast to structural differences between the cofactor in the enzyme and in isolation, molecular complexes and the cofactor bear strikingly similar electronic properties at the metal center.

3c. Effect of Complex Structure on H Atom Binding.

Beyond structural and electronic properties of Mo/W complexes, binding energies of catalytic intermediates provide the most direct probe of potential activity of these complexes for formate oxidation to CO_2. Specifically, we focus on H atom binding energetics to the terminal chalcogen atoms as a rapid screening approach and also evaluate resulting structural changes. This allows us to investigate the coupled influence of the metal, catalyst structure, and terminal

chalcogen identity on H atom binding. Since the chosen reference is atomic hydrogen, the computed binding strength should be larger in magnitude than a mechanism involving proton abstraction from formic acid, but we expect trends to be preserved across complexes. First comparing the different terminal chalcogen atoms over this set, we again observe similarities between terminal sulfido and selenido moieties and larger differences of these two to the oxo moiety in terms of favored orientation for the coordinating H atom and, to a lesser extent, the energetics (Supporting Information Table S13). Structurally, the M-tc-H (tc = O, S, or Se) angles are largest for O and smaller for S and Se. Despite structural differences, H atom binding energies, E_b(H), are more comparable among O, S, and Se terminal chalcogens (see Sec. 2 and Supporting Information Table S13).

While largely insensitive to terminal chalcogen identity, we find that E_b(H) is strongly influenced by metal and oxidation state. Over our entire set, the E_b(H) spans a large range from -3.0 to -0.7 eV, where more negative numbers correspond to more favorable binding. The lower oxidation state complexes bind hydrogen less favorably relative to the higher oxidation state complexes, reproducing trends for H atom binding on terminal Fe-oxo compounds[94, 95]. Moreover, Mo complexes typically binds hydrogen more strongly than W in equivalent oxidation states (Figure 5 and Supporting Information Table S14). The metal- and oxidation state-dependence we observe for E_b(H) aligns well with the oxidation-state-dependence that was observed in a recent experimental study of a Mo-containing formate dehydrogenase[26].

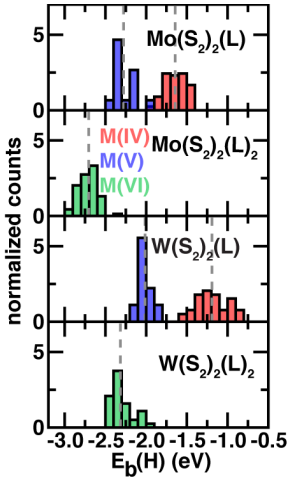

Figure 5. Normalized histograms of the hydrogen atom binding energies for 330 complexes, categorized by metal identity and ligand numbers, shown in the upper right corner of each panel. The metal oxidation states IV, V and VI are shown as translucent bars in red, blue and green, respectively. The number of Mo or W complexes with oxidation states IV/V/VI are used to compute each histogram. Vertical dashed lines indicate the average value of the E_b(H) for each group.

To assess structural changes concomitant with H atom binding, we measure the folding and twisting angles of the relaxed intermediate obtained upon H atom binding. We find that all H atom bound intermediates are more planar (i.e., as judged through reduced folding angles), while twisting angles alternately increase or decrease. Importantly, these structural changes upon H atom adsorption are dependent on metal oxidation state (Figure 6). Independent of the metal or chalcogen atom type, we observe that the M(VI)-tc and M(IV)-tc complexes exhibit larger folding and twisting angle changes than the M(V)-tc complexes upon H atom binding. The smaller change in folding and twisting angles for M(V)-tc complexes is balanced by a more significant reduction in the bond lengths of metal-dithiolene bonds for this oxidation state (Supporting Information Figure S10). We estimate the energetics of this structural reorganization upon H atom binding by computing the difference in E_b(H), ΔE_b, with respect to a structure in which the catalyst complex is constrained (see Sec. 2). The typical energy penalty for structural reorganization is less than 0.85 eV with the average about 0.25 eV, both of which are

significantly smaller than the overall range of energetics of binding (Supporting Information Figure S11).

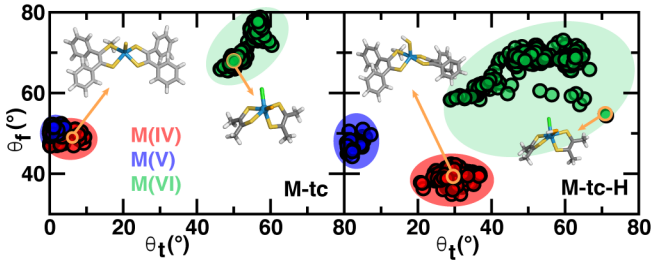

Figure 6. The twisting vs folding angles of 330 complexes before (M-tc) and after hydrogen atom adsorption on terminal chalcogen atoms (M-tc-H). Metal oxidation states are shown in translucent dots in red (M(IV)), blue (M(V)), and green (M(VI)), respectively. Two representative molecules are shown inset (W(IV) CSD refcode: CUMYOY with O mutated by S, W(VI) CSD refcode: IVOJOS with O mutated by Se).

To explore the effect of enzyme environment on H atom binding, we also modeled H atom binding to a metal-sulfido moiety in a QM cluster model of the FDH enzyme (Supporting Information Figure S12). We find that the H-atom-induced decrease in twisting and folding angles (θ_t: 28°,θ_f: 58° compared to original θ_t: 38°,θ_f: 64°) is comparable to the molecular complexes, even though the original angles were distinct. The M-S-H angle of the enzyme model (100°) is even more comparable to the M-S-H angles in molecular complexes with sulfido ligands (Supporting Information Table S13). These structural comparisons indicate the similarities between enzyme cofactors and molecular complexes for H atom binding.

Our calculations show that the H binding strength is primarily influenced by the metal oxidation states. In Mo/W-containing FDH enzymes, the oxidized species (i.e., M(VI)) is believed to be the catalytically active intermediate that accepts the H (i.e., whether it be a hydrogen atom or hydride) from formate[10, 26]. This is consistent with our observation that M(VI) has the strongest H atom binding energy and is thus a good H atom acceptor. Our calculations indicate that the H atom binding strength weakens in the reduced species (i.e., M(V) or M(IV)),

in which case the catalyst can release H atoms, and the catalytic cycle starts over. Our results also show that the stronger H atom binding energies on Mo compared to W. It indicates more favorable H atom or H-containing species binding to Mo complexes, which would make H atom release less thermodynamically favorable. Experimental observations indicate that W-containing enzymes are faster for formate oxidation than Mo-containing counterparts[10], indicating that H atom release is likely to be a limiting step. Moreover, the debate on whether Sec/Cys dissociation[25, 32, 33, 96] is catalytically relevant can be partially explained based on our results. When the Sec/Cys is not coordinated to the metal, the metal coordination number becomes five. In turn, the metal is either in formal oxidation state +4 or +5, based on the statistics in our dataset. The lower oxidation states weaken H atom binding and assist H atom (or H containing species) release. Therefore, the Sec/Cys dissociation is catalytically relevant during the process.

Secondly, our observations on the local reorganization of the catalyst upon H atom binding are more sensitive to metal oxidation state than the ligand chemistry. The M(V) complex structures are qualitatively unchanged upon H atom binding (i.e., no significant change in the folding and twisting angles upon H atom adsorption). The reduced structural variation of M(V) complexes relative to M(IV) and M(VI) complexes indicate the possibility that M(V)-SH intermediates could be experimentally characterized. In a Mo-dependent FDH enzyme, the Mo(V)-SH group was experimentally observed[31, 97, 98] and in a recent kinetic study[84] Mo(VI)=S(formate) was converted to Mo(V)-SH rapidly. In our study, although we only considered H atom adsorption and not H-atom containing species such as formate, the geometric invariance of M(V)-SH structures supports its importance as a potential abundant intermediate in the catalytic cycle.

3d. Effect of Terminal Chalcogen on Electronic Structure.

Since the terminal chalcogen is the reactive species during formate oxidation, we next mainly discuss the role of the terminal chalcogen. Thus far, we have demonstrated that the terminal chalcogen has a more significant effect on the electronic properties of the metal, as judged by the partial charge and bond valence in comparison to the global structural properties, as judged by twisting or folding angles. To further isolate the terminal chalcogen effects, we compare a subset of complexes that are identical except for their terminal chalcogen atoms. To make this comparison, we revisit the folding and twisting angles and also investigate the HOMO-LUMO gap and implicit-solvent-corrected vertical ionization potential (IP) and electron affinity (EA), which are closely related to the catalytically-relevant redox potential[7] (see Sec. 2).

As was observed over the general set of complexes, we again find that terminal chalcogen identity has little effect on twisting angles (Supporting Information Figure S13). Folding angles, however, show moderate dependence on terminal chalcogen identity, with oxo-containing complexes demonstrating smaller folding angles relative to sulfido and selenido counterparts (Figure 7 and Supporting Information Figure S14). Over this set, we find no differences in folding angles between sulfido and selenido complexes derived from the same parent structure, strengthening our conclusion that sulfido and selenido complexes behave more similarly relative to the same structures with oxo groups. The terminal chalcogen dependence we observe is general for five- and six-coordinate complexes, although the terminal oxo ligand has a greater influence on the folding angles of the more flexible five-coordinate complexes relative to six-coordinate complexes (Supporting Information Figure S14).

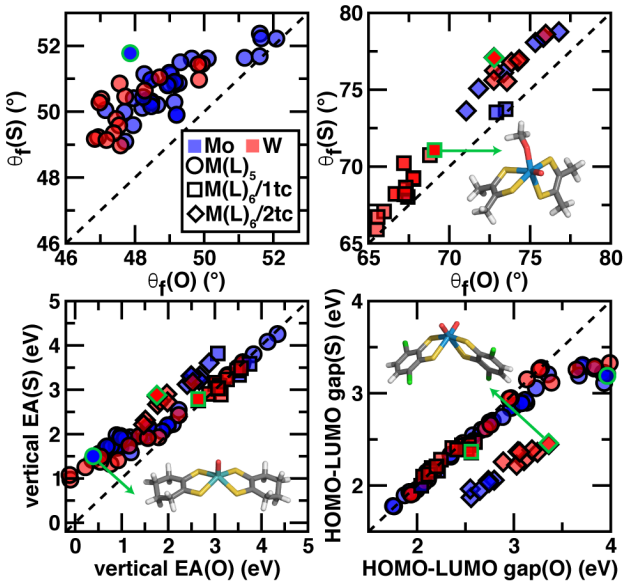

Figure 7. The folding angle, vertical EA and HOMO-LUMO gap of 276 complexes with identical structures apart from the terminal chalcogen atom which is either O or S. Symbols are colored based on their metal type (Mo: blue, W: red) and grouped based on their coordination number and terminal chalcogen (tc) number. (five-coordinated: circle, six-coordinated with 1tc: square, six-coordinated with 2tc: diamond). Three representative renders for each type are CAVZOP (blue circle), DEHQAJ (red square), YIMFIK (red diamond).

Turning to electronic properties relevant for redox potential, we find that the terminal chalcogen identity has no effect on vertical IP while the vertical EA has moderate terminal chalcogen dependence (Supporting Information Figures S15–S16). Complexes with terminal oxo ligands demonstrate reduced vertical EA and a narrower range relative to the same complexes with terminal sulfido or selenido ligands, which contain similar vertical EA values (Figure 7 and Supporting Information Figure S16). Moreover, there are considerable overlaps between Mo and W in both vertical IP and vertical EA, while more W complexes are in the lower range comparing to more Mo complexes are in higher range. It indicates some W complexes would show lower redox potential than Mo complexes which is in line with experimental observations.[9, 99] We observe similar trends for HOMO-LUMO gaps: terminal oxo ligands typically have increased HOMO-LUMO gaps that sample a wider range. Moreover, we observe that six-coordinate complexes with two terminal oxo ligands have larger HOMO-LUMO gaps than the

same complexes with two sulfido or two selenido groups (Figure 7 and Supporting Information Figure S17).

Overall, we observe that the terminal chalcogen identity influences the folding angle, vertical EA, and HOMO-LUMO gap. For these properties, the terminal sulfido and selenido ligands behave similarly, but both are distinct from oxo ligands. Previous studies indicate that terminal sulfido and selenido ligands behave much more similarly than oxo ligands, due to their greater covalency with the metal and lower electronegativity[9, 37, 45, 100]. Our findings based on hundreds of CSD complexes are in line with these observations but expand upon those prior studies. Our results can facilitate further explorations for catalyst design or mechanistic studies, where the terminal chalcogen can govern molecular properties[10].

4. Conclusions.

To understand the structure-property relationships of formate-dehydrogenase-inspired Mo/W complexes, we carried out a large-scale screen of over 400 Mo/W complexes in the CSD. From an 88 bisdithiolene complex subset, we generated a 276 complex dataset by mutating terminal chalcogen sites to contain O, S and Se atoms and performed DFT calculations. Over this set, we observed that the metal-local environment of bisdithiolene Mo/W complexes (e.g., angles accessed by ligands) depends strongly on metal coordination number. In contrast, the electronic properties of the complexes depend less on the metal-local geometry but more on the ligand and metal identities as well as metal formal oxidation states. We compared transition metal complex properties to those observed from MD and QM/MM simulation of a representative FDH enzyme. The protein environment influences the geometry of the cofactor by decreasing both folding and twisting angles around the metal decrease. Trends in electronic properties observed for molecular complexes, however, hold for the cofactor in the enzyme

environment, suggesting that molecular complexes are faithful mimics of the enzyme from an electronic structure perspective.

To rapidly screen and understand catalytic activities in these systems, we also explored H atom binding energies. Here, we observed that H atom binding strength is strongly related to the metal formal oxidation state and is independent of the terminal chalcogen identity. The change in metal-local structure (e.g., the shift in folding and twisting angles) before and after H atom binding on the terminal chalcogen are comparable between molecular complexes and the FDH cofactor in the enzyme environment.

While natural FDH enzymes can have either Mo or W centers with a range of terminal ligands (i.e., O or S or Se, Sec or Cys), isolating the role that metal or ligand identity plays has been challenging. The present data set has allowed a means of isolating the individual effects of metal center identity, terminal chalcogens, and coordinating ligands. Since over this set we observe that the metal-local structure of the metal-bisdithiolene is primarily contingent upon metal coordination number, we expect that a labile Sec or Cys ligand could be important in the catalytic cycle of FDH enzymes. Although different terminal chalcogen atoms have distinct electronic structure (i.e., metal partial charge and bond valence), the identity of the terminal chalcogen is not critical for determining the H atom binding strength. Our results suggest that all terminal chalcogen moieties (e.g., M=O, M=S, M=Se) can be active if the reaction is governed by this adsorption/desorption. Overall, the enzyme environment influences structures more than local electronic properties. This suggests that distinct reactivity between enzymes and molecular complexes may be due to differences in geometric structure rather than differences in electronic properties. Overall, a plausible mechanism for the enzyme involving our observations on structural rearrangements suggests that formate oxidation occurs when formate binds to the

active site M(VI)=tc, and Sec/Cys is displaced to reduce the metal center to M(V), weakening the bound H atom interaction for its release.

Our analysis based on VHTS of Mo/W bisdithiolene complexes from the CSD enables the extraction of structure-property relationships that may not be directly accessible from comparison of available FDH enzyme or molecular omplex structures. This observation that the enzyme environment mostly affects geometries but the electronic properties of the metal are primarily related to the terminal chalcogen and metal identities could be used to design mimics with more constrained geometries or labile ligands that capture the unique features of the enzyme.

ASSOCIATED CONTENT

AUTHOR INFORMATION

Corresponding Author

*email: hjkulik@mit.edu phone: 617-253-4584

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors acknowledge primary funding from the ExxonMobil Research and Engineering Company. H.J.K. holds a Career Award at the Scientific Interface from the Burroughs Wellcome Fund, an AAAS Marion Milligan Mason Award, and an Alfred P. Sloan fellowship in chemistry, which supported this work. A. Nandy was partially supported by a National Science Foundation Graduate Research Fellowship under Grant #1122374.

Supporting Information Available. The workflow for screening Mo/W complexes from the CSD; statistics of CSD-derived Mo/W complexes; complexes with radical ligands or mislabeled charge/oxidation states; changes in metal-ligand bond lengths after geometry optimization; workflow for selecting bisdithiolene ligands and terminal chalcogen mutations; spin state assignments for oxidized and reduced states for vertical IP/EA; comparison of geometric and electronic properties between complexes and FDH cofactors; normalized histograms of metal Mulliken, NBO, and Hirshfeld charges and bond valence; details of MD simulations for the FDH enzyme; the QM cluster generated from FDH enzyme; Hirshfeld metal partial charge and metal bond valence for the FDH cofactor; workflow for H atom binding on chalcogen and H atom binding energies; statistics of complexes after H atom adsorption; geometric properties and H atom binding energies after H atom adsorption on terminal chalcogens; histogram of H atom binding energy differences between constrained and unconstrained calculations; parity plot of geometric and electronic properties for terminal chalcogen mutations with O/S/Se. (PDF)

Optimized structures and electronic energies for all molecules presented in this work; initial AMBER topology and coordinate files for the FDH enzyme. (ZIP)

References

1. Ye, R.; Hurlburt, T. J.; Sabyrov, K.; Alayoglu, S.; Somorjai, G. A., Molecular catalysis science: Perspective on unifying the fields of catalysis. *P Natl Acad Sci USA* **2016,** *113* (19), 5159-5166.
2. Gamba, I., Biomimetic Approach to CO2 Reduction. *Bioinorg Chem Appl* **2018,** *2018*.
3. Marchetti, L.; Levine, M., Biomimetic Catalysis. *Acs Catal* **2011,** *1* (9), 1090-1118.
4. Simmons, T. R.; Berggren, G.; Bacchi, M.; Fontecave, M.; Artero, V., Mimicking hydrogenases: From biomimetics to artificial enzymes. *Coordin Chem Rev* **2014,** *270*, 127-150.
5. Appel, A. M.; Bercaw, J. E.; Bocarsly, A. B.; Dobbek, H.; DuBois, D. L.; Dupuis, M.; Ferry, J. G.; Fujita, E.; Hille, R.; Kenis, P. J. A.; Kerfeld, C. A.; Morris, R. H.; Peden, C. H. F.; Portis, A. R.; Ragsdale, S. W.; Rauchfuss, T. B.; Reek, J. N. H.; Seefeldt, L. C.; Thauer,

R. K.; Waldrop, G. L., Frontiers, Opportunities, and Challenges in Biochemical and Chemical Catalysis of CO2 Fixation. *Chem Rev* **2013,** *113* (8), 6621-6658.
6. Cook, S. A.; Hill, E. A.; Borovik, A. S., Lessons from Nature: A Bio-Inspired Approach to Molecular Design. *Biochemistry-Us* **2015,** *54* (27), 4167-4180.
7. Holm, R. H.; Kennepohl, P.; Solomon, E. I., Structural and functional aspects of metal sites in biology. *Chem Rev* **1996,** *96* (7), 2239-2314.
8. Laureanti, J. A.; Ginovska, B.; Buchko, G. W.; Schenter, G. K.; Hebert, M.; Zadvornyy, O. A.; Peters, J. W.; Shaw, W. J., A Positive Charge in the Outer Coordination Sphere of an Artificial Enzyme Increases CO2 Hydrogenation. *Organometallics* **2020,** *39* (9), 1532-1544.
9. Groysman, S.; Holm, R. H., Synthesis and structures of bis(dithiolene)tungsten(IV,VI) thiolate and selenolate complexes: Approaches to the active sites of molybdenum and tungsten formate dehydrogenases. *Inorg Chem* **2007,** *46* (10), 4090-4102.
10. Niks, D.; Hille, R., Molybdenum- and tungsten-containing formate dehydrogenases and formylmethanofuran dehydrogenases: Structure, mechanism, and cofactor insertion. *Protein Sci* **2019,** *28* (1), 111-122.
11. Bassegoda, A.; Madden, C.; Wakerley, D. W.; Reisner, E.; Hirst, J., Reversible Interconversion of CO2 and Formate by a Molybdenum-Containing Formate Dehydrogenase (vol 136, pg 15473, 2014). *J Am Chem Soc* **2015,** *137* (13), 4592-4592.
12. Hille, R., The mononuclear molybdenum enzymes. *Chem Rev* **1996,** *96* (7), 2757-2816.
13. Vorholt, J. A.; Thauer, R. K., Molybdenum and tungsten enzymes in C1 metabolism. *Met Ions Biol Syst* **2002,** *39*, 571-619.
14. Jormakka, M.; Byrne, B.; Iwata, S., Formate dehydrogenase - a versatile enzyme in changing environments. *Curr Opin Struc Biol* **2003,** *13* (4), 418-423.
15. Moura, J. J. G.; Brondino, C. D.; Trincao, J.; Romao, M. J., Mo and W bis-MGD enzymes: nitrate reductases and formate dehydrogenases. *J Biol Inorg Chem* **2004,** *9* (7), 791-799.
16. Dobbek, H., Structural aspects of mononuclear Mo/W-enzymes. *Coordin Chem Rev* **2011,** *255* (9-10), 1104-1116.
17. Maia, L. B.; Moura, I.; Moura, J. J. G., Molybdenum and tungsten-containing formate dehydrogenases: Aiming to inspire a catalyst for carbon dioxide utilization. *Inorg Chim Acta* **2017,** *455*, 350-363.
18. Seo, J.; Shearer, J.; Williard, P. G.; Kim, E., Reactivity of a biomimetic W(IV) bis-dithiolene complex with CO2 leading to formate production and structural rearrangement. *Dalton T* **2019,** *48* (47), 17441-17444.
19. Lim, B. S.; Donahue, J. P.; Holm, R. H., Synthesis and structures of bis(dithiolene)molybdenum complexes related to the active sites of the DMSO reductase enzyme family. *Inorg Chem* **2000,** *39* (2), 263-273.
20. Fogeron, T.; Retailleau, P.; Chamoreau, L. M.; Li, Y.; Fontecave, M., Pyranopterin Related Dithiolene Molybdenum Complexes as Homogeneous Catalysts for CO2 Photoreduction. *Angew Chem Int Edit* **2018,** *57* (52), 17033-17037.
21. Sung, K. M.; Holm, R. H., Synthesis and structures of bis(dithiolene)-tungsten(IV) complexes related to the active sites of tungstoenzymes. *Inorg Chem* **2000,** *39* (6), 1275-1281.
22. Shiekh, B. A.; Kaur, D.; Kumar, S., Bio-mimetic self-assembled computationally designed catalysts of Mo and W for hydrogenation of CO2/dehydrogenation of HCOOH inspired by the active site of formate dehydrogenase. *Phys Chem Chem Phys* **2019,** *21* (38), 21370-21380.

23. Gan, L.; Jennings, D.; Laureanti, J.; Jones, A. K., Biomimetic Complexes for Production of Dihydrogen and Reduction of CO2. *Homo- and Heterobimetallic Complexes in Catalysis: Cooperative Catalysis* **2016,** *59*, 233-272.
24. Tiberti, M.; Papaleo, E.; Russo, N.; De Gioia, L.; Zampella, G., Evidence for the Formation of a Mo-H Intermediate in the Catalytic Cycle of Formate Dehydrogenase. *Inorg Chem* **2012,** *51* (15), 8331-8339.
25. Mota, C. S.; Rivas, M. G.; Brondino, C. D.; Moura, I.; Moura, J. J. G.; Gonzalez, P. J.; Cerqueira, N. M. F. S. A., The mechanism of formate oxidation by metal-dependent formate dehydrogenases. *J Biol Inorg Chem* **2011,** *16* (8), 1255-1268.
26. Robinson, W. E.; Bassegoda, A.; Reisner, E.; Hirst, J., Oxidation-State-Dependent Binding Properties of the Active Site in a Mo-Containing Formate Dehydrogenase. *J Am Chem Soc* **2017,** *139* (29), 9927-9936.
27. Dong, G.; Ryde, U., Reaction mechanism of formate dehydrogenase studied by computational methods. *J Biol Inorg Chem* **2018,** *23* (8), 1243-1254.
28. Maia, L. B.; Fonseca, L.; Moura, I.; Moura, J. J. G., Reduction of Carbon Dioxide by a Molybdenum-Containing Formate Dehydrogenase: A Kinetic and Mechanistic Study. *J Am Chem Soc* **2016,** *138* (28), 8834-8846.
29. Schrapers, P.; Hartmann, T.; Kositzki, R.; Dau, H.; Reschke, S.; Schulzke, C.; Leimkuhler, S.; Haumann, M., 'Sulfido and Cysteine Ligation Changes at the Molybdenum Cofactor during Substrate Conversion by Formate Dehydrogenase (FDH) from Rhodobacter capsulatus. *Inorg Chem* **2015,** *54* (7), 3260-3271.
30. Oliveira, A. R.; Mota, C.; Mourato, C.; Domingos, R. M.; Santos, M. F. A.; Gesto, D.; Guigliarelli, B.; Santos-Silva, T.; Romao, M. J.; Pereira, I. A. C., Toward the Mechanistic Understanding of Enzymatic CO2 Reduction. *Acs Catal* **2020,** *10* (6), 3844-3856.
31. Niks, D.; Duvvuru, J.; Escalona, M.; Hille, R., Spectroscopic and Kinetic Properties of the Molybdenum-containing, NAD(+) - dependent Formate Dehydrogenase from Ralstonia eutropha. *J Biol Chem* **2016,** *291* (3), 1162-1174.
32. Raaijmakers, H. C. A.; Romao, M. J., Formate-reduced E-coli formate dehydrogenase H: the reinterpretation of the crystal structure suggests a new reaction mechanism. *J Biol Inorg Chem* **2006,** *11* (7), 849-854.
33. Hartmann, T.; Schrapers, P.; Utesch, T.; Nimtz, M.; Rippers, Y.; Dau, H.; Mroginski, M. A.; Haumann, M.; Leimkuhler, S., The Molybdenum Active Site of Formate Dehydrogenase Is Capable of Catalyzing C-H Bond Cleavage and Oxygen Atom Transfer Reactions. *Biochemistry-Us* **2016,** *55* (16), 2381-2389.
34. Yang, J. Y.; Kerr, T. A.; Wang, X. S.; Barlow, J. M., Reducing CO2 to HCO2– at Mild Potentials: Lessons from Formate Dehydrogenase. *J Am Chem Soc* **2020,** *142* (46), 19438-19445.
35. Hille, R.; Hall, J.; Basu, P., The Mononuclear Molybdenum Enzymes. *Chem Rev* **2014,** *114* (7), 3963-4038.
36. Jiang, J. F.; Holm, R. H., An expanded set of functional groups in bis(dithiolene)tungsten(IV,VI) complexes related to the active sites of tungstoenzymes, including W-IV-SR and W-VI-O(SR). *Inorg Chem* **2004,** *43* (4), 1302-1310.
37. Basu, P.; Nemykin, V. N.; Sengar, R. S., Syntheses, spectroscopy, and redox chemistry of encapsulated Oxo-Mo(V) centers: Implications for pyranopterin-containing molybdoenzymes. *Inorg Chem* **2003,** *42* (23), 7489-7501.

38. Nagarajan, K.; Joshi, H. K.; Chaudhury, P. K.; Pal, K.; Cooney, J. J. A.; Enemark, J. H.; Sarkar, S., Structural and functional analogue of the active site of polysulfide reductase from Wolinella succinogenes. *Inorg Chem* **2004,** *43* (15), 4532-4533.
39. Donahue, J. P.; Goldsmith, C. R.; Nadiminti, U.; Holm, R. H., Synthesis, structures, and reactivity of bis(dithiolene)molybdenum(IV,VI) complexes related to the active sites of molybdoenzymes. *J Am Chem Soc* **1998,** *120* (49), 12869-12881.
40. Musgrave, K. B.; Lim, B. S.; Sung, K. M.; Holm, R. H.; Hedman, B.; Hodgson, K. O., X-ray spectroscopy of enzyme active site analogues and related molecules: Bis(dithiolene)molybdenum(IV) and -tungsten(IV,VI) complexes with variant terminal ligands. *Inorg Chem* **2000,** *39* (23), 5238-5247.
41. Arnoux, P.; Ruppelt, C.; Oudouhou, F.; Lavergne, J.; Siponen, M. I.; Toci, R.; Mendel, R. R.; Bittner, F.; Pignol, D.; Magalon, A.; Walburger, A., Sulphur shuttling across a chaperone during molybdenum cofactor maturation. *Nat Commun* **2015,** *6*.
42. Boyington, J. C.; Gladyshev, V. N.; Khangulov, S. V.; Stadtman, T. C.; Sun, P. D., Crystal structure of formate dehydrogenase H: Catalysis involving Mo, molybdopterin, selenocysteine, and an Fe4S4 cluster. *Science* **1997,** *275* (5304), 1305-1308.
43. Khangulov, S. V.; Gladyshev, V. N.; Dismukes, G. C.; Stadtman, T. C., Selenium-containing formate dehydrogenase H from Escherichia coli: A molybdopterin enzyme that catalyzes formate oxidation without oxygen transfer. *Biochemistry-Us* **1998,** *37* (10), 3518-3528.
44. Garton, S. D.; Hilton, J.; Oku, H.; Crouse, B. R.; Rajagopalan, K. V.; Johnson, M. K., Active site structures and catalytic mechanism of Rhodobacter sphaeroides dimethyl sulfoxide reductase as revealed by resonance Raman spectroscopy. *J Am Chem Soc* **1997,** *119* (52), 12906-12916.
45. McNaughton, R. L.; Lim, B. S.; Knottenbelt, S. Z.; Holm, R. H.; Kirk, M. L., Spectroscopic and electronic structure studies of symmetrized models for reduced members of the dimethylsulfoxide reductase enzyme family. *J Am Chem Soc* **2008,** *130* (14), 4628-4636.
46. Li, Z.; Wang, S. W.; Chin, W. S.; Achenie, L. E.; Xin, H. L., High-throughput screening of bimetallic catalysts enabled by machine learning. *J Mater Chem A* **2017,** *5* (46), 24131-24138.
47. Potyrailo, R.; Rajan, K.; Stoewe, K.; Takeuchi, I.; Chisholm, B.; Lam, H., Combinatorial and High-Throughput Screening of Materials Libraries: Review of State of the Art. *Acs Comb Sci* **2011,** *13* (6), 579-633.
48. Bricker, M. L.; Sachtler, J. W. A.; Gillespie, R. D.; McGonegal, C. P.; Vega, H.; Bem, D. S.; Holmgren, J. S., Strategies and applications of combinatorial methods and high throughput screening to the discovery of non-noble metal catalyst. *Appl Surf Sci* **2004,** *223* (1-3), 109-117.
49. Zhang, X.; Zhang, Z. H.; Wu, D. H.; Zhang, X.; Zhao, X. D.; Zhou, Z., Computational Screening of 2D Materials and Rational Design of Heterojunctions for Water Splitting Photocatalysts. *Small Methods* **2018,** *2* (5).
50. Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C., The Cambridge Structural Database. *Acta Crystallogr B* **2016,** *72*, 171-179.
51. Ufimtsev, I. S.; Martinez, T. J., Quantum chemistry on graphical processing units. 3. Analytical energy gradients, geometry optimization, and first principles molecular dynamics. *J Chem Theory Comput* **2009,** *5* (10), 2619-2628.
52. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **1988,** *37*, 785--789.

53. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. *Journal of Chemical Physics* **1993,** *98* (7), 5648-5652.
54. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J., Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *The Journal of Physical Chemistry* **1994,** *98* (45), 11623-11627.
55. Hay, P. J.; Wadt, W. R., Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *The Journal of Chemical Physics* **1985,** *82* (1), 270-283.
56. Wang, L. P.; Song, C. C., Geometry optimization made simple with translation and rotation coordinates. *Journal of Chemical Physics* **2016,** *144* (21).
57. Ioannidis, E. I.; Gani, T. Z. H.; Kulik, H. J., molSimplify: A toolkit for automating discovery in inorganic chemistry. *J Comput Chem* **2016,** *37* (22), 2106-2117.
58. Nandy, A.; Duan, C.; Janet, J. P.; Gugler, S.; Kulik, H. J., Strategies and Software for Machine Learning Accelerated Discovery in Transition Metal Chemistry. *Ind Eng Chem Res* **2018,** *57* (42), 13973-13986.
59. Janet, J. P.; Liu, F.; Nandy, A.; Duan, C.; Yang, T.; Lin, S.; Kulik, H. J., Designing in the face of uncertainty: Exploiting electronic structure and machine learning models for discovery in inorganic chemistry. *Inorg Chem* **2019,** *58* (16), 10592-10606.
60. Janet, J. P.; Kulik, H. J., Predicting electronic structure properties of transition metal complexes with neural networks. *Chem Sci* **2017,** *8* (7), 5137-5152.
61. Klamt, A.; Schuurmann, G., Cosmo - a New Approach to Dielectric Screening in Solvents with Explicit Expressions for the Screening Energy and Its Gradient. *J Chem Soc Perk T 2* **1993,** (5), 799-805.
62. 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* (7), 3131-3144.
63. Bondi, A., Van Der Waals Volumes + Radii. *J Phys Chem-Us* **1964,** *68* (3), 441-+.
64. Fonseca Guerra, C.; Handgraaf, J. W.; Baerends, E. J.; Bickelhaupt, F. M., Voronoi deformation density (VDD) charges: Assessment of the Mulliken, Bader, Hirshfeld, Weinhold, and VDD methods for charge analysis. *Journal of computational chemistry* **2004,** *25* (2), 189-210.
65. Mulliken, R. S., Electronic population analysis on LCAO–MO molecular wave functions. I. *The Journal of Chemical Physics* **1955,** *23* (10), 1833-1840.
66. Weinhold, F.; Landis, C. R., Natural bond orbitals and extensions of localized bonding concepts. *Chemistry Education Research and Practice* **2001,** *2* (2), 91-104.
67. Hirshfeld, F. L., Bonded-atom fragments for describing molecular charge densities. *Theoretica chimica acta* **1977,** *44* (2), 129-138.
68. Mayer, I., Bond order and valence indices: A personal account. *J Comput Chem* **2007,** *28* (1), 204-221.
69. Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V., H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res* **2012,** *40* (Web Server issue), W537-41.
70. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A., H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res* **2005,** *33* (Web Server issue), W368-71.

71. Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A., A simple clustering algorithm can be accurate enough for use in calculations of pKs in macromolecules. *Proteins* **2006,** *63* (4), 928-38.
72. Labahn, J.; Granzin, J.; Schluckebier, G.; Robinson, D. P.; Jack, W. E.; Schildkraut, I.; Saenger, W., Three-dimensional structure of the adenine-specific DNA methyltransferase M.Taq I in complex with the cofactor S-adenosylmethionine. *Proc Natl Acad Sci U S A* **1994,** *91* (23), 10957-61.
73. Li, P.; Merz, K. M., MCPB.py: A Python Based Metal Center Parameter Builder. *J Chem Inf Model* **2016,** *56* (4), 599-604.
74. Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, I., T.E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER 2018*, University of California, San Francisco.
75. Ferreira, P.; Cerqueira, N. M. F. S.; Brás, N. F.; Fernandes, P. A.; Ramos, M. J., Parametrization of Molybdenum Cofactors for the AMBER Force Field. *J Chem Theory Comput* **2018,** *14* (5), 2538-2548.
76. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **2015,** *11* (8), 3696-713.
77. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R.; W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983,** *79*, 926-935.
78. Karelina, M.; Kulik, H. J., Systematic quantum mechanical Region determination in QM/MM simulation. *J Chem Theory Comput* **2017,** *13* (2), 563-576.
79. Kulik, H. J.; Zhang, J.; Klinman, J. P.; Martinez, T. J., How large should the QM region be in QM/MM calculations? The case of catechol O-methyltransferase. *The Journal of Physical Chemistry B* **2016,** *120* (44), 11381-11394 %@ 1520-6106.
80. Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M., A long-range-corrected density functional that performs well for both ground-state properties and time-dependent density functional theory excitation energies, including charge-transfer excited states. *J Chem Phys* **2009,** *130* (5), 054−112.
81. Hay, P. J.; Wadt, W. R., Ab initio effective core potetentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *J. Chem. Phys.* **1985,** *82*, 270−283.
82. Hariharan, P. C.; Pople, J. A., The influence of polarization functions on molecular-orbital hydrogenation energie. *Theor. Chim. Acta.* **1973,** *28*, 213−222.
83. Mehmood, R.; Kulik, H. J., Both configuration and QM region size matter: zinc stability in QM/MM models of DNA methyltransferase. *J Chem Theory Comput* **2020,** *16* (5), 3121-3134.
84. Robinson, W. E.; Bassegoda, A.; Blaza, J. N.; Reisner, E.; Hirst, J., Understanding How the Rate of C-H Bond Cleavage Affects Formate Oxidation Catalysis by a Mo-Dependent Formate Dehydrogenase. *J Am Chem Soc* **2020,** *142* (28), 12226-12236.

85. Taylor, M. G.; Yang, T.; Lin, S.; Nandy, A.; Janet, J. P.; Duan, C.; Kulik, H. J., Seeing Is Believing: Experimental Spin States from Machine Learning Model Structure Predictions. *The Journal of Physical Chemistry A* **2020,** *124* (16), 3286-3299.
86. Webster, C. E.; Hall, M. B., The theoretical transition state structure of a model complex bears a striking resemblance to the active site structure of DMSO reductase. *J Am Chem Soc* **2001,** *123* (24), 5820-5821.
87. Das, S. K.; Biswas, D.; Maiti, R.; Sarkar, S., Modeling the tungsten sites of inactive and active forms of hyperthermophilic Pyrococcus furiosus aldehyde ferredoxin oxidoreductase. *J Am Chem Soc* **1996,** *118* (6), 1387-1397.
88. Moula, G.; Bose, M.; Datta, H.; Sarkar, S., Photoluminescent Mo(IV) and W(IV) bis-dithiolene complexes with bidentate phosphonodithioato ligand derived from Lawesson's reagent. *Polyhedron* **2013,** *52*, 900-908.
89. Lim, B. S.; Holm, R. H., Bis(dithiolene)molybdenum analogues relevant to the DMSO reductase enzyme family: Synthesis, structures, and oxygen atom transfer reactions and kinetics. *J Am Chem Soc* **2001,** *123* (9), 1920-1930.
90. Habib, U.; Hoffman, M., Effect of molybdenum and tungsten on the reduction of nitrate in nitrate reductase, a DFT study. *Chem Cent J* **2017,** *11*.
91. Joshi, H. K.; Cooney, J. J. A.; Inscore, F. E.; Gruhn, N. E.; Lichtenberger, D. L.; Enemark, J. H., Investigation of metal-dithiolate fold angle effects: Implications for molybdenum and tungsten enzymes. *P Natl Acad Sci USA* **2003,** *100* (7), 3719-3724.
92. Drew, S. C.; Hanson, G. R., Determination of the Metal-Dithiolate Fold Angle in Mononuclear Molybdenum(V) Centers by EPR Spectroscopy. *Inorg Chem* **2009,** *48* (5), 2224-2232.
93. Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Meyer, E. F.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M., Protein Data Bank - Computer-Based Archival File for Macromolecular Structures. *J Mol Biol* **1977,** *112* (3), 535-542.
94. Dantignana, V.; Serrano-Plana, J.; Draksharapu, A.; Magallón, C.; Banerjee, S.; Fan, R.; Gamba, I.; Guo, Y.; Que, L.; Costas, M.; Company, A., Spectroscopic and Reactivity Comparisons between Nonheme Oxoiron(IV) and Oxoiron(V) Species Bearing the Same Ancillary Ligand. *Journal of the American Chemical Society* **2019,** *141* (38), 15078-15091.
95. Geng, C.; Ye, S.; Neese, F., Does a higher metal oxidation state necessarily imply higher reactivity toward H-atom transfer? A computational study of C–H bond oxidation by high-valent iron-oxo and -nitrido complexes. *Dalton Transactions* **2014,** *43* (16), 6079-6086.
96. Cerqueira, N. M. F. S. A.; Fernandes, P. A.; Gonzalez, P. J.; Moura, J. J. G.; Ramos, M. J., The Sulfur Shift: An Activation Mechanism for Periplasmic Nitrate Reductase and Formate Dehydrogenase. *Inorg Chem* **2013,** *52* (19), 10766-10772.
97. Gladyshev, V. N.; Khangulov, S. V.; Axley, M. J.; Stadtman, T. C., Coordination of Selenium to Molybdenum in Formate Dehydrogenase-H from Escherichia-Coli. *P Natl Acad Sci USA* **1994,** *91* (16), 7708-7711.
98. Rivas, M. G.; Gonzalez, P. J.; Brondino, C. D.; Moura, J. J. G.; Moura, I., EPR characterization of the molybdenum(V) forms of formate dehydrogenase from Desulfovibrio desulfuricans ATCC 27774 upon formate reduction. *J Inorg Biochem* **2007,** *101* (11-12), 1617-1622.
99. Tucci, G. C.; Donahue, J. P.; Holm, R. H., Comparative kinetics of oxo transfer to substrate mediated by bis(dithiolene)dioxomolybdenum and -tungsten complexes. *Inorg Chem* **1998,** *37* (7), 1602-1608.

100. van Stipdonk, M. J.; Basu, P.; Dille, S. A.; Gibson, J. K.; Berden, G.; Oomens, J., Infrared Multiple Photon Dissociation Spectroscopy of a Gas-Phase Oxo-Molybdenum Complex with 1,2-Dithiolene Ligands. *J Phys Chem A* **2014,** *118* (29), 5407-5418.