ACS Catalysis

Open Access
This article is licensed under CC-BY-NC-ND 4.0

pubs.acs.org/acscatalysis
Research Article

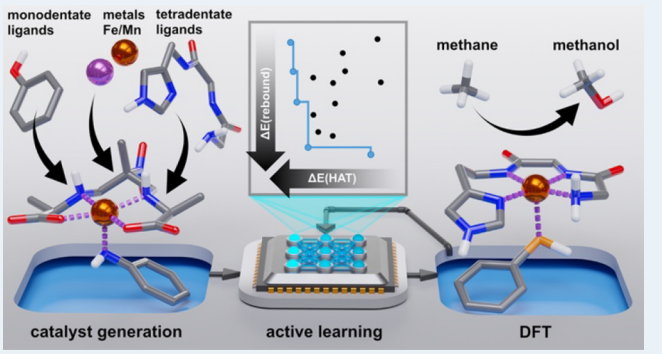
Multiobjective Optimization of Synthetically Accessible Homogeneous Catalysts Reveals Tradeoffs for Methane-to-Methanol Conversion

Husain Adamji, Ilia Kevlishvili, David W. Kastner, Ralf Meyer, Yuriy Román-Leshkov, and Heather J. Kulik*

Cite This: https://doi.org/10.1021/acscatal.6c03365
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The selective conversion of methane to methanol remains a major challenge in catalysis, and no earth-abundant homogeneous catalyst has been demonstrated to carry out this transformation with high turnover. To explore previously synthesized complexes that might be amenable for this transformation, we start from macrocyclic ligands from experimental structures in the Cambridge Structural Database. After expanding this set of ligands with Hammett-tuning functionalization, we construct a design space of 21.9 million Mn and Fe catalysts. We employ active learning using neural network surrogate models with quantified uncertainty to efficiently identify catalysts with optimal methane-to-methanol turnover. We do this by jointly optimizing density functional theory (DFT)-computed hydrogen atom transfer (HAT) and methyl radical rebound energetics. Square planar tetraaza-based macrocycles in combination with low-spin Fe(II) or Mn(II) emerge as a privileged class of transition metal complexes within the Pareto-optimal set, with Mn(II) catalysts exhibiting balanced energetic tradeoffs of the two steps. In contrast, porphyrin catalysts exhibit favorable HAT but unfavorable rebound, and seesaw catalysts, which are not widely employed, are indeed the least favorable in our set. Full reaction coordinate analysis of representative Pareto-optimal catalysts reveals that HAT and rebound thermodynamics no longer govern turnover, which is instead determined by the stability of the metal−oxo intermediate and the rebound kinetic barrier. Nonetheless, the resulting energetic spans for these catalysts remain competitive with known earth-abundant, methane-to-methanol catalysts, underscoring the effectiveness of active learning for discovering previously unexplored, synthetically accessible catalysts and unraveling catalyst design principles for methane-to-methanol reactivity.



KEYWORDS: *methane hydroxylation, active learning, density functional theory, catalyst design, earth-abundant metals*

1. INTRODUCTION

The selective direct conversion of methane to methanol remains a long-standing challenge in catalysis.[1,2] Effective catalyst design requires facile activation of methane while maintaining methanol selectivity and avoiding overoxidation.[3−5] Despite these challenges, heterogeneous methane-to-methanol catalysis[6] has been advanced by the identification of late-transition metal active sites[7−16] and, in some cases, even earth-abundant Fe and Cu sites incorporated within zeolites or on oxide supports that mediate C−H activation via highly reactive metal−oxo intermediates.[17−23] For homogeneous catalysis in contrast, Pt-based systems pioneered by Periana[24] and Shilov[25] are the only well-established examples for converting methane to methanol, and no earth-abundant homogeneous catalysts have been reported. Iron-dependent enzymes such as P450, TauD, and methane monooxygenase (MMO) demonstrate that earth-abundant metal centers can selectively activate strong C−H bonds,[26] inspiring efforts to develop synthetic earth-abundant homogeneous catalysts. However, despite the synthesis of numerous bioinspired Fe-based complexes[26−30] capable of forming the reactive Fe-oxo intermediate, such as Fe-TAML,[31,32] Fe(N4Py),[33] and Fe(bpy),[34] which demonstrate efficient C−H activation of larger aliphatic substrates, methane activation remains elusive.

Received: April 28, 2026
Revised: October 2, 2026
Accepted: October 8, 2026

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

The challenges for the design of novel, earth-abundant transition metal complexes for methane-to-methanol catalysis motivate the exploration of a vast chemical space. Nevertheless, synthesizing and validating each candidate, as well as isolating and characterizing the fleeting metal−oxo intermediates,[35−38] poses an insurmountable challenge. First-principles calculations with density functional theory (DFT) are essential to accelerate this exploration[39,40] by identifying promising active site motifs,[41−44] assessing the feasibility of metal-oxo formation and C−H activation,[45−48] and elucidating the influence of spin state,[49−51] multistate reactivity,[52,53] and metal coordination motifs[54,55] on catalytic performance. However, screening catalysts at the scale of thousands of compounds, even with DFT calculations, can become costly, especially when simulating open-shell transition metal complexes.[56,57] Scaling relations, such as Brønsted−Evans−Polanyi (BEP) relationships[58−61] and linear free-energy relationships (LFERs),[62−65] that draw correlations between catalyst thermodynamics and kinetics[66−73] enable faster estimation of catalytic performance by reducing the number of required DFT calculations. However, these scaling relations break down when applied to diverse chemical spaces,[42] including catalysts with different metals, oxidation or spin states,[42,45,46] varied metal-local structures,[47,60,74] or noncovalent interactions.[55,75] Thus, it is critical to assess the validity of scaling relations within the chemical space under study before relying on them to guide catalyst design.

Nonlinear machine learning (ML) methods provide a promising alternative for navigating extensive catalyst design spaces[76] by directly predicting their reactivity descriptors, such as reaction energetics,[45,77,78] redox and ionization potentials,[79−81] and frontier orbital energies,[82] from their chemical compositions. These models bypass the limitations of linear scaling relations and can be applied to thousands of candidates, dramatically accelerating the evaluation of potential catalysts compared with first-principles calculations. However, developing reliable ML models to predict reactive properties across diverse catalyst chemistries requires broad sampling of chemical space,[83] which in turn demands large, high-quality labeled datasets that are particularly costly to generate for open-shell transition metal complexes.[39,84] Moreover, conventional strategies for generating training data often rely on random sampling, treating all points in a design space as equally informative. When labeling requires expensive DFT calculations, this approach can waste computational resources on redundant or low-information candidates.[85−88]

The inefficiency of random sampling over large design spaces can be addressed with active learning,[89−92] a data-efficient strategy for building accurate models that adaptively selects candidates,[93] balancing exploration of high-uncertainty regions[94,95] with the identification of Pareto-optimal candidates guided by domain-specific objectives. For methane-to-methanol conversion, optimal catalytic performance arises from balancing multiple energetic trade-offs along the reaction coordinate,[96] motivating a multi-objective global Bayesian optimization (MOGBO)[97−100] framework that simultaneously considers competing reaction energetics. Prior work from our group employed this strategy, targeting the joint optimization of hydrogen atom transfer (HAT) and methanol release energetics in the radical rebound mechanism over an extensive design space of mononuclear transition metal complexes containing hypothetically assembled macrocyclic ligands to identify promising earth-abundant methane-to-methanol catalysts.[96] While this prior work successfully identified novel hypothetical catalysts, there remains a need to explore experimentally-grounded catalysts to propose synthetically realizable improvements in overall methane-to-methanol turnover.

In this work, we construct a design space of approximately 21.9 million synthetically accessible five-coordinate Mn and Fe transition metal complexes featuring a tetradentate macrocyclic ligand and a monodentate ligand. We demonstrate that an active learning-driven exploration of this space enables the identification of novel Mn and Fe catalysts that exhibit more balanced reaction energetics along the methane-to-methanol catalytic cycle, resulting in improved turnover. From these lead candidates, we extract key design principles that modulate thermodynamics and kinetics along the reaction coordinate. These include factors such as coordination environment, spin state, Hammett substituents, and reaction conditions. Together, they guide the optimal design of earth-abundant methane-to-methanol catalysts.

2. REACTION MECHANISM

We investigate the reaction energetics for the radical rebound[101] mechanism of methane-to-methanol conversion over mononuclear Mn and Fe transition metal complexes (TMCs) in +2 and +3 oxidation states in the resting state. Starting from the five-coordinate resting state structure (**1**), the catalyst undergoes two-electron oxidation by N_2O, which serves as a stoichiometric oxygen donor with established experimental precedent in alkane oxidation,[102−104] to form a high-valent terminal metal−oxo species (**2**) (Figure 1). If other oxidants (e.g., triplet O_2 or H_2O_2) were instead employed, this would correspond to a rigid shift of reaction energies and would not affect relative energetics (Supporting Information

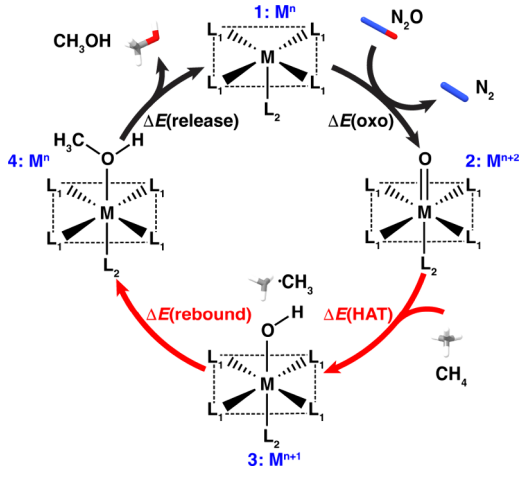


Figure 1. Radical rebound mechanism for the conversion of methane to methanol. Starting from the resting state (1) with the transition metal center (Mn or Fe) at an oxidation state of n = II/III, the cycle proceeds clockwise to form the metal−oxo species (2) from N_2O, the metal−hydroxo species (3) via HAT from methane, and the methanol-bound species (4) through methyl radical rebound. All catalysts contain a tetradentate ligand (L_1) and a monodentate ligand (L_2). The red arrows in the catalytic cycle correspond to the steps we optimize simultaneously in this work. Atom color codes: N in blue, O in red, C in gray, and H in white. In the case of seesaw ligands (not shown), L_2 is not necessarily in an axial position but will be *cis* to the majority of the tetradentate ligand.

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

Table S1). We compute the metal−oxo formation energy, $\Delta E(\text{oxo})$, as:

$$\Delta E(\text{oxo}) = E(\mathbf{2}) + E(N_2) - E(\mathbf{1}) - E(N_2O)$$

The resulting M(IV/V)=O species is highly active and facilitates HAT from methane, yielding a M(III/IV)-OH intermediate (**3**) and a methyl radical (Figure 1). The HAT reaction energy, $\Delta E(\text{HAT})$, is given by:

$$\Delta E(\text{HAT}) = E(\mathbf{3}) + E(\cdot CH_3) - E(\mathbf{2}) - E(CH_4)$$

Following HAT, the methyl radical recombines with the M(III/IV)-OH intermediate in the radical rebound step to form the methanol-bound intermediate (**4**) thereby returning the metal center to its initial M(II/III) oxidation state. The rebound energy, $\Delta E(\text{rebound})$, is given by:

$$\Delta E(\text{rebound}) = E(\mathbf{4}) - E(\mathbf{3}) - E(\cdot CH_3)$$

Finally, methanol dissociates from the catalyst, restoring the original resting state structure (**1**) (Figure 1). The methanol release energy, $\Delta E(\text{release})$, is computed as:

$$\Delta E(\text{release}) = E(\mathbf{1}) + E(CH_3OH) - E(\mathbf{4})$$

Each catalyst is evaluated in up to three spin states, low-spin (LS), intermediate-spin (IS), and high-spin (HS), for a given metal/oxidation state combination. For each selected spin state, we require that it remains valid and conserved across all radical rebound intermediates in the catalytic cycle. Accordingly, Mn(II) and Fe(III) catalysts are studied only in LS doublet and IS quartet states. For similar reasons, Mn(III) catalysts are studied only as LS singlet and IS triplet, while Fe(II) catalysts are studied in LS singlet, IS triplet, and HS quintet states. The metal−oxo and methanol-bound intermediates have identical spin states as the resting state. During the HAT step, we assume ferromagnetic coupling with the transferred hydrogen atom (α-radical transfer) and antiferromagnetic coupling with the resulting $\cdot CH_3$ radical, following prior work.[46] This coupling scheme leads to an increase of one unit in the spin state of the metal−hydroxo intermediate relative to other species, but the total spin of the system remains conserved (Supporting Information, Table S2).

For each catalyst, we perform DFT calculations on all four intermediates and compute all corresponding reaction energies. However, for catalyst optimization, we only simultaneously target the HAT (ΔE(HAT)) and methyl radical rebound (ΔE(rebound)) reaction energies, as detailed next.

3. DESIGN SPACE AND OPTIMIZATION APPROACH

We constructed an expansive catalyst design space of 21.9 million square pyramidal TMCs by pairing experimentally characterized tetradentate macrocyclic ligands with coordinating monodentate ligands (Figure 2). The monodentate ligands correspond to eight ligands selected from prior work[48,96] that comprise a benzene ring bearing a donor substituent from the set of amine, amide, hydroxyl, hydroxide, phosphine, phosphide, thiol, or thiolate groups, coordinating to the metal through nitrogen, phosphorus, oxygen, or sulfur atoms. The benzene ring for the monodentate ligand was chosen to maximize the Hammett-

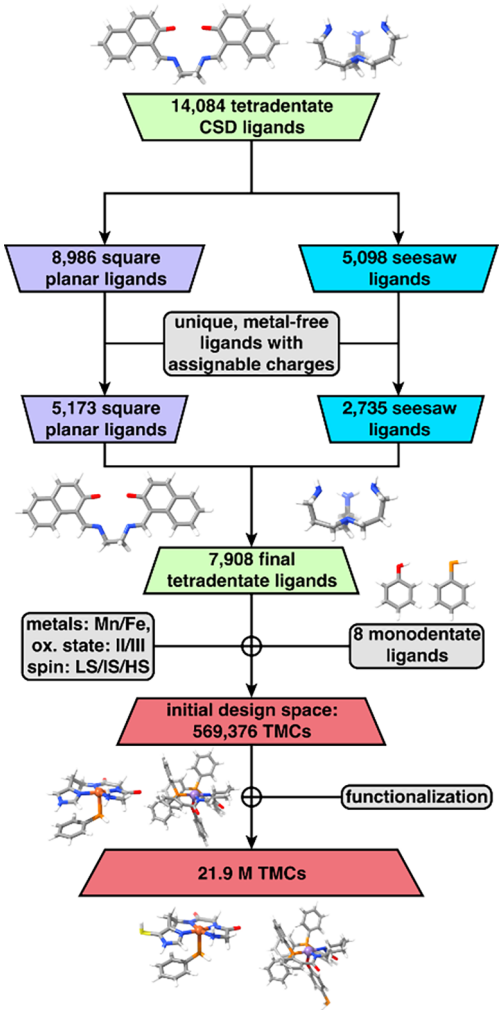


Figure 2. Workflow for constructing the TMC design space used in active learning. Tetradentate ligands are first curated from crystallized TMCs in the CSD and filtered to retain metal-free, charge-assignable ligands that are unique within two metal coordination geometries (seesaw and square planar). These filtered ligands are then combined with monodentate aryl ligands, and when combined with variations in metal identity, oxidation state, and spin state, they generate an initial design space that is subsequently expanded through functionalization to yield a complete design space of 21.9 million TMCs.

tuning effects by substituents in the *para* position as discussed later. The tetradentate macrocyclic ligands were extracted from the Cambridge Structural Database (CSD)[105] by querying all mononuclear TMCs with well-resolved crystal structures and retaining those containing tetradentate ligands. For each resulting complex, we removed all atoms except the metal center and its tetradentate ligand and computed the corresponding Weisfeiler−Lehman graph hash[106] as implemented in NetworkX,[107] which serves as a unique connectivity-based identifier. This enabled the retention of unique metal−ligand combinations (Supporting Information, Table S3). We assigned ligand geometry based on the relative positions of the two remaining coordination sites in an idealized octahedral geometry. When the two sites were *cis* to each other, the ligand was assigned seesaw geometry, while if the two sites were *trans* to each other, the ligand was assigned a square planar geometry. This procedure yielded 14,084 unique tetradentate macrocyclic

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

ligands (5,098 seesaw, 8,986 square planar, Figure 2 and Supporting Information, Table S3).

From this set of tetradentate macrocycles, we first removed any ligands identical to hypothetical macrocycles constructed in prior work,[96] as determined by comparison of Weisfeiler−Lehman graph hashes computed only on the ligand (Supporting Information, Table S3). Next, we eliminated ligands containing metal atoms, as well as duplicate entries within each geometry class as determined by their graph hashes. Finally, we excluded ligands for which a net molecular charge could not be reliably assigned by the iterative charge scheme utilized in our recently developed dataset.[108] These filtering steps yielded a final set of 7,908 unique tetradentate macrocyclic ligands of which 2,735 were seesaw and 5,173 were square planar (Figure 2 and Supporting Information, Table S3). Combining these macrocyclic ligands with the eight aromatic neutral and anionic monodentate ligands and nine distinct metal/oxidation/spin-state combinations (see Section 2), we obtained an initial design space comprising 569,376 square pyramidal TMCs (Figure 2). This design space was further expanded by introducing ten Hammett-type substituents on either the tetradentate or the monodentate ligand of a given TMC. Where applicable, we functionalized all equivalent carbon atoms in each aromatic ring, as well as across identical aromatic rings within the same ligand. For tetradentate ligands, functionalization was limited to those containing a heuristic cutoff of eight or fewer aromatic rings (Supporting Information, Figure S1). This resulted in a final space of 21.9 million TMCs (Figure 2).

To identify high-performing candidates within this large design space, we target complexes with relatively flat methane-to-methanol reaction energy landscapes that promote high catalytic turnover. We employ MOGBO[97−100] to optimize key methane-to-methanol thermodynamic parameters towards this goal. We use the expected hypervolume improvement (EHVI) acquisition function[100,109] to prioritize simultaneous improvements to the Pareto front across multiple objectives. Our optimization focuses exclusively on thermodynamic parameters, due to their compatibility with high-throughput screening. We aim to minimize the HAT energy, ΔE(HAT). Lowering this energy helps identify catalysts with greater potential to activate the strong C−H bonds in methane. We also minimize the absolute methyl radical rebound energy, $|\Delta E(\text{rebound})|$, to favor thermoneutral rebound. This avoids creating a thermodynamic sink at the methanol-bound intermediate and to ensure methanol release remains favorable. Jointly optimizing ΔE(HAT) and $|\Delta E(\text{rebound})|$ appears to capture the dominant thermodynamic tradeoffs across the overall landscape, as evidenced by the Pareto front for other objective spaces (Supporting Information, Figure S2).

To guide candidate selection, we require surrogate models that output Gaussian predictive distributions for both target properties in the unexplored design space, enabling the evaluation of the 2D-EHVI acquisition function. We use two independent artificial neural networks (ANNs) as our surrogate models to predict the mean values of the two targets using ligand-centered revised autocorrelations (RACs) together with six additional whole-complex descriptors, based on the superior performance of these models over probabilistic alternatives such as Gaussian process regression (GPR) models on the initial dataset (see Section 6b and Supporting Information, Figure S3). To estimate standard errors from our deterministic ANNs, we apply the Laplace approximation,[110−115] modeling the weight posterior as a Gaussian around the learned weights, with uncertainty derived from the loss curvature (Supporting Information, Text S1 and Figure S4). Predictive uncertainties are then obtained by propagating uncertainty on the weights to the outputs[116] followed by a post-hoc empirical calibration to globally correct the uncertainty estimates (Supporting Information, Texts S2 and S3).[117] This framework provides efficient uncertainty estimates without the cost of obtaining an ensemble and yields more robust error bounds relative to optimization-dependent alternatives, such as latent space distance methods, where the parameter estimation for converting latent distances to uncertainty is sensitive to initialization (Supporting Information, Text S3 and Figure S5).[96,118]

We initiate the active learning workflow by training the ΔE(HAT) and ΔE(rebound) surrogate models on an initialization set from prior work (see Section 4a).[96] At each active learning iteration, predictive means and Laplace-approximated uncertainties for ΔE(HAT) and ΔE(rebound) are used to evaluate the 2D-EHVI across the design space, excluding candidates with existing DFT-computed targets (Supporting Information, Text S4). EHVI integrates over the full predictive distribution of both objectives, explicitly incorporating both the predictive means and the well-calibrated surrogate uncertainties rather than ranking by predicted means alone, thereby balancing exploitation of promising candidates against exploration of high-uncertainty regions. From this, the top 10,000 candidates with the highest 2D-EHVI values are identified, and *k*-medoids clustering is applied to select 200 diverse candidates for subsequent DFT evaluation and surrogate model retraining in the next generation.[96] We performed four active learning generations (generations 1−4) on the initial design space (569,376 TMCs) and three subsequent generations (generations 5−7) on the functionalized design space (21.9M TMCs, Supporting Information, Table S4). In generation 2, a more stringent filtering step based on Tanimoto similarity[119,120] was applied to further restrict the candidate pool in the initial design space (Supporting Information, Text S5 and Figures S6 and S7).

4. RESULTS AND DISCUSSION

4a. Initial Design Space Exploration Outcomes

We first compare the initial catalyst design space explored in this work to that from prior active learning studies,[96] focusing on the chemical diversity of the tetradentate ligands. In prior work, macrocyclic ligands were constructed hypothetically from common ligand fragments and were restricted to square planar geometries. In this work, we instead obtain macrocyclic ligands observed in experimentally crystallized TMCs in the CSD,[105] encompassing both square planar and seesaw coordination geometries. Comparison of the metal-local chemical structures by evaluating the metal-agnostic tetradentate ligand substructures at a depth of two from the metal center reveals that the five most common metal-coordinating environments in each dataset have limited overlap, with the corresponding substructure frequencies differing markedly between the two spaces (Figure 3). The current design space is dominated by five-membered metallocyclic motifs and N_4 metal coordination environments, which comprise 32.1% of the tetradentate ligands from the CSD, whereas these motifs are comparatively infrequent in prior work (3.9%, Figure 3). In contrast,

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

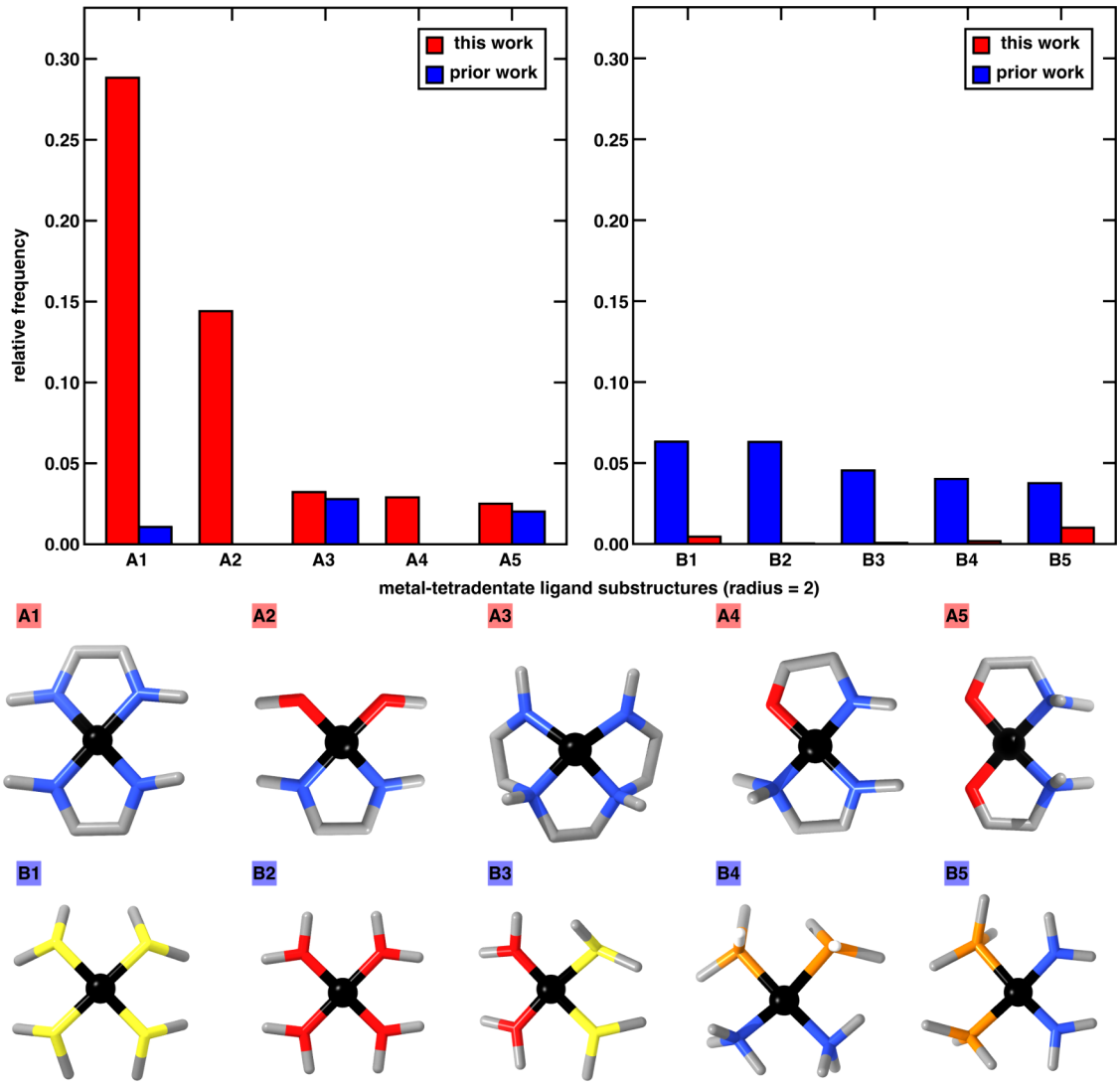


Figure 3. Relative frequency associated with the five most prevalent metal−tetradentate substructures at graph depth 2 from the metal center for (top-left) tetradentate ligands curated from the CSD (A1−A5, red) shown alongside the corresponding substructure relative frequencies for ligands from prior work (blue) and (top-right) tetradentate ligands from prior work (B1−B5, blue) shown alongside the corresponding substructure relative frequencies for CSD-derived ligands (red). (bottom) Insets illustrating the five most prevalent metal−tetradentate substructures from each dataset (A1−A5 and B1−B5). Atom color codes: dummy metal center in black, C in gray, N in blue, O in red, P in orange, S in yellow, and H in white.

macrocyclic ligands in the prior work were generated from exhaustive enumeration regardless of donor atom and thus frequently contained donor sets such as, S_4 (6.3%), O_4 (6.3%), or N_2P_2 (7.8%), which are rare among CSD-derived macrocyclic ligands in this work (0.5%, 0.03%, and 1.2%, respectively, Figure 3). Additionally, tetradentate ligands from our work exhibit a broader range of net molecular charges, heavy atom counts, and aromatic ring counts compared to prior work[96] (Supporting Information, Figures S1 and S8). These differences highlight the distinctions between hypothetical catalysts and synthetically accessible complexes, and motivate a new multi-objective search over catalysts comprised of ligands that have been previously synthesized.

Despite differences in tetradentate ligand chemistry, the two design spaces share key features that make prior reaction energetics a reasonable basis for initializing our active learning campaign. Both spaces contain square pyramidal mononuclear TMCs with a tetradentate macrocyclic ligand and a monodentate ligand, and the two sets share the same metal identities, monodentate ligands, and oxidation/spin states. We therefore began training of the surrogate ΔE(HAT) and ΔE(rebound) ANN models using data from prior work.[96] Specifically, unfunctionalized Mn and Fe TMCs from the first four generations of the prior active learning campaign were retained if they passed all geometric stability criteria, yielding a curated set of 640 TMCs that were not part of our constructed TMC design space and were used solely to initialize surrogate model training (initialization set; see Section 6a).

Starting from a design space of 569,376 unfunctionalized TMCs, we initiated the active learning workflow by applying the ANN models trained on our initialization set to predict ΔE(HAT) and ΔE(rebound) values. We score all points using 2D-EHVI and, from the top 10,000 candidates, we select 200 promising candidates for subsequent DFT evaluation from *k*-medoids sampling. A high 2D-EHVI score is obtained for highly uncertain and promising points that are likely to go

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

beyond the Pareto front, and the 2D-EHVI score is lower when the model becomes more certain and/or if points are predicted to be less likely to pass the Pareto front. Active learning cycles enable improvements in surrogate model performance in a data-efficient manner while simultaneously identifying new Pareto-optimal methane-to-methanol catalysts. To assess how model accuracy evolves across active learning generations, we track lookahead errors,[94,96] which measure the ability of the models trained on one generation to predict target reaction energies for the next. Over four generations, we observe a six-fold reduction in lookahead errors for ΔE(HAT) (from 34 kcal/mol to 6 kcal/mol) and a two-fold reduction in lookahead errors for ΔE(rebound (from 18 kcal/mol to 8 kcal/mol, Figure 4). Uncertainties for the lookahead predictions also generally decrease alongside the lookahead errors, although the uncertainty calibration is poor in generation 1. At generation 1, the ΔE(HAT) model is overconfident while the ΔE(rebound) model is conservative relative to the observed lookahead errors (Supporting Information, Figure S9). Agreement between predicted uncertainties and observed lookahead errors improves in subsequent generations. However, because these lookahead batches differ in composition, we also evaluated uncertainty estimates on a fixed held-out test set. In our active learning campaign, lookahead errors plateau beyond generation 3 (ca. 6 kcal/mol for ΔE(HAT) and ca. 9 kcal/mol for ΔE(rebound)), with only minor fluctuations within 1 kcal/mol in subsequent generations (Figure 4). At the same time as errors improve, 2D-EHVI values decrease markedly with successive generations, suggesting termination of the active learning workflow after four cycles due to diminishing returns in both model accuracy and Pareto front expansion (Figure 4 and Supporting Information, Figures S10 and S11). Interestingly, the models trained on the initialization set exhibit higher lookahead errors on generation 1 than on subsequent generations. This behavior arises because generation 1 introduces previously unseen seesaw macrocyclic ligands in the TMCs (i.e., absent from our initialization set curated from prior work), pushing the model into less familiar chemical space (Supporting Information, Table S5). By generation 2, more new TMCs with square-planar macrocyclic ligands are introduced, balancing errors across geometries and yielding lower overall lookahead error.

We next evaluated how the Pareto front evolved with each active learning generation. In evaluating the ΔE(HAT)−|ΔE(rebound)| Pareto front, we exclude the initialization set because several catalysts in the initialization set exhibit unrealistically favorable |ΔE(rebound)| values (Supporting Information, Figure S12). Examining the Pareto front evolution across generations 1 through 4, we find that successive active learning cycles expand beyond the fronts of earlier generations, highlighting the

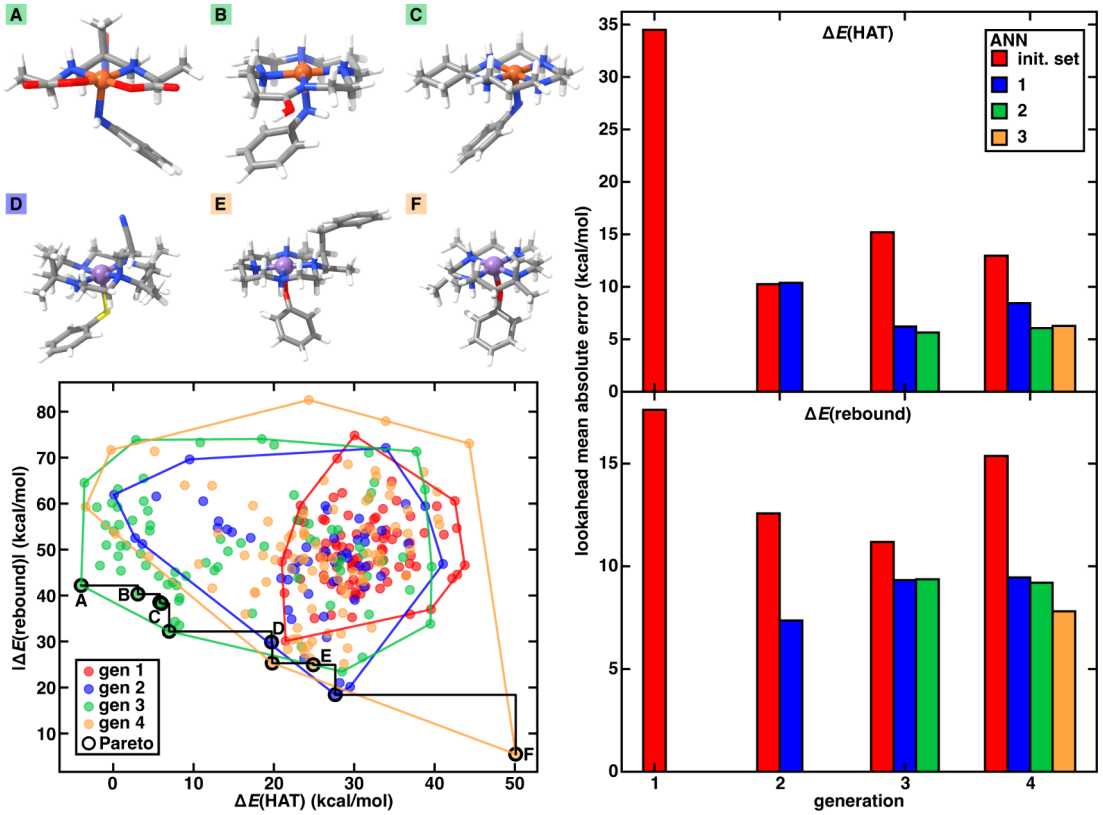


Figure 4. (bottom-left) |ΔE(rebound)| vs ΔE(HAT) scatter plot (kcal/mol) showing TMCs from generation 1 (red), generation 2 (blue), generation 3 (green), and generation 4 (orange) of the active learning workflow. Data points from each generation are enclosed by their convex hulls and Pareto-optimal complexes identified across generations 1−4 are outlined in black and connected by a black stair-step line. (top-left) Insets of representative Pareto-optimal complexes identified across generations 1−4. Atom color codes: Fe in brown, Mn in purple, C in gray, N in blue, O in red, S in yellow, and H in white. (right) Lookahead mean absolute errors (kcal/mol) for (top) ΔE(HAT) and (bottom) ΔE(rebound) from the two independent surrogate ANN models across generations 1−4. As indicated in the inset legend, each bar is colored by the generation at which each ANN was trained, with errors evaluated on all subsequent generations.

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

effectiveness of the workflow (Figure 4). Near-optimal ΔE(HAT) energetics are still accessible within this experimentally derived ligand space (Supporting Information, Figure S12).

Consistent with the spin/oxidation states favored in prior work,[96] the final Pareto front over the unfunctionalized CSD design space comprises five Fe(II) and five Mn(II) LS catalysts. The Pareto-optimal Mn catalysts exhibit balanced ΔE(HAT)−|ΔE(rebound)| tradeoffs relative to the Pareto-optimal Fe catalysts, which primarily occupy HAT-favored regions of the front. All catalysts share a common structural motif involving a square-planar tetradentate ligand on the equatorial plane paired with a negatively charged monodentate ligand (Figure 4 and Supporting Information, Figure S13 and Table S6). While the LS state is not the resting ground spin state for these complexes, it is readily accessible due to relatively low spin-splitting energies (Supporting Information, Table S6). The equatorial plane of the Pareto-optimal TMCs predominantly features N_4 coordination, reflecting the composition of the tetradentate ligand library curated from the CSD, while the axial positions are consistently occupied by N- or O-donor ligands (Figure 4 and Supporting Information, Figure S13). Together, these ligands establish a weak-to-moderate overall ligand field in the Pareto-optimal catalysts, consistent with them having IS and HS ground states in the resting state (Supporting Information, Table S6). Notably, 14-membered tetraaza macrocycles, including functionalized derivatives bearing methyl or hydroxyl groups, emerge as a privileged class in the Pareto-optimal catalysts in both neutral and anionic forms (Supporting Information, Figure S13 and Table S7). These findings demonstrate that the active learning workflow recovers established design principles observed in non-heme enzymes and synthetic C−H activation catalysts,[26] while simultaneously identifying novel optimal catalysts for the conversion of methane to methanol. A review of manuscripts corresponding to the CSD TMCs containing the tetradentate ligands in the Pareto-optimal catalysts confirms that most of these scaffolds have not been previously explored for catalysis. Rather, they were primarily structurally characterized, and in some cases, evaluated for their photochemical properties (Supporting Information, Table S7).[121−133]

Given their prevalence in nature, we also assessed the performance of porphyrin-moiety TMCs within our evaluated design space, which have been widely explored for C−H activation.[134−136] Our dataset includes Fe and Mn complexes bearing meso-substituted porphyrins (Supporting Information, Figure S14). Within this subset, both Fe and Mn complexes exhibit ΔE(HAT) values comparable to their respective Pareto-optimal catalysts, indicating good propensity for C−H activation. However, these same catalysts have |ΔE(rebound)| values that are consistently higher by ca. 20 kcal/mol on average with respect to the Pareto front (Supporting Information, Figure S14).

Given that the catalysts on the Pareto front we identified were not necessarily in their resting ground spin state, we extended the Pareto front analysis by constructing a spin-consistent Pareto front. To obtain this modified Pareto front, we required that the spin state of the computed reaction energetics be consistent with the TMC's ground spin state in the resting state. Imposing this constraint narrows the dataset to 120 of the 315 TMCs evaluated in generations 1−4, the majority of which adopt the IS state in the resting state (99 IS state, 21 LS state, Figure 5). Under these conditions, the spin-consistent Pareto front predominantly consists of LS Fe(II) catalysts with only one LS Mn(II) catalyst (Figure 5 and Supporting Information, Figure S15). Although N-donor ligands remain prevalent, the spin-consistent Pareto-optimal TMCs exhibit a higher frequency of stronger-field P- and S-based donors across both equatorial and axial positions, consistent with stabilization of LS ground states in the resting state (Figure 5 and Supporting Information, Figure S15). Relative to the original Pareto front, the spin-consistent Pareto front generally exhibits ca. 10 kcal/mol higher |ΔE(rebound)| values on average, while ΔE(HAT) values remain comparable. Several IS Mn(III) complexes lie just above the spin-consistent front, displaying only modest (ca. 5 kcal/mol on average) increases in |ΔE(rebound)| and featuring weak-to-moderate N/O donor environments in line with their IS resting ground states (Supporting Information, Figure S16). Interestingly, these Mn(III) TMCs adopt a seesaw tetradentate geometry, suggesting that this coordination motif, while absent from both

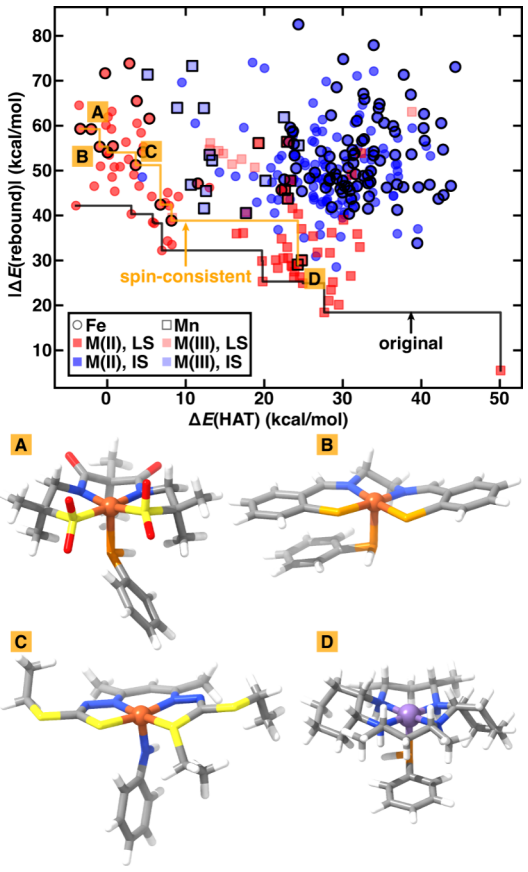


Figure 5. (top) |ΔE(rebound)| vs ΔE(HAT) scatter plot (kcal/mol) showing all TMCs from generations 1−4 grouped by metal (Fe in circles, Mn in squares), colored by spin state (LS in red and IS in blue), and shaded by oxidation state (M(II) solid and M(III) translucent). Complexes whose resting state spin states are identical to the spin state at which the reaction energetics were evaluated are outlined in black. Complexes on the Pareto front and spin-consistent Pareto front are connected by a black and an orange stair-step line, respectively. (bottom) Insets of representative spin-consistent Pareto-optimal complexes identified across generations 1−4. Atom color codes: Fe in brown, Mn in purple, C in gray, N in blue, O in red, P in orange, S in yellow, and H in white.

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

Pareto fronts, can still support competitive reaction energetics for Mn catalysts.

While HS Fe(II) catalysts are often pursued in synthetic C−H activation catalysis due to their similarity to enzymatic active sites,[26,137−139] our active learning workflow consistently favored LS Fe(II) catalysts. In the initialization set, HS Fe(II) TMCs exhibited significantly less favorable target reaction energetics than LS Fe(II), particularly for ΔE(HAT) (ca. 25 kcal/mol less favorable on average) with $|\Delta E$(rebound)$|$ disfavored by ca. 11 kcal/mol (Supporting Information, Figure S17). Consequently, HS Fe(II) catalysts were not sampled from our unfunctionalized catalyst design space. Our workflow initially prioritized IS Fe(II) TMCs, before converging towards LS and IS Fe and Mn TMCs at both oxidation states in later generations (Supporting Information, Table S8). Although HS Fe(II) species do not appear in generations 1−4, comparisons between LS and IS TMCs reveal that LS species consistently occupy more favorable regions of the ΔE(HAT)−$|\Delta E$(rebound)$|$ landscape across both metals and oxidation states (Figure 5). The least favorable region of the ΔE(HAT)−$|\Delta E$(rebound)$|$ landscape explored during our optimization predominantly consists of IS Fe(II) catalysts, which have ΔE(HAT) values that are substantially higher than the LS counterparts (by ca. 27 kcal/mol on average, Figure 5). These catalysts predominantly consist of seesaw tetradentate ligands with a relatively higher prevalence of S-donor atoms appearing on both the equatorial and axial positions, indicating that this combination constitutes a sub-optimal design (Supporting Information, Figure S18). Collectively, these trends highlight LS catalysts as readily accessible yet underexamined, and energetically favorable compared to traditional higher-spin designs.

To assess how the Pareto-optimal catalysts perform beyond the ΔE(HAT)−$|\Delta E$(rebound)$|$ optimization, we examined trends in other key methane-to-methanol reaction energetics. Metal−oxo formation energies are consistently favorable for all Mn and Fe Pareto-optimal TMCs, with Mn catalysts displaying significantly more favorable ΔE(oxo) values than Fe (by ca. 40 kcal/mol on average), a trend that holds across the full dataset and aligns with prior work (Supporting Information, Table S6 and Figure S19).[42,46] In contrast, Pareto-optimal TMC release energies remain moderate (ca. 23 kcal/mol on average) for both metals, despite the weak correlation between ΔE(release) and ΔE(rebound) (Supporting Information, Figure S20). These results indicate that optimizing for ΔE(HAT) and $|\Delta E$(rebound)$|$ does not compromise thermodynamic performance in other key steps of the radical rebound catalytic cycle.

4b. Influence of TMC Functionalization on Reaction Energetics

A common synthetic strategy to modulate catalyst reactivity is through the introduction of functional groups on the ligands. Toward that end, we investigated whether incorporating Hammett-type substituents can further optimize target reaction energetics and uncover new Pareto-optimal catalysts. We begin with a targeted control study on two representative Mn and Fe TMCs drawn from generations 1−4 to assess the sensitivity of ΔE(HAT) and $|\Delta E$(rebound)$|$ to ligand functionalization (Supporting Information, Figures S21 and S22). For each complex, we independently functionalize the tetradentate and monodentate ligands with ten Hammett-type substituents spanning electron-donating (e.g., amino and methyl) to electron-withdrawing groups (e.g., halogens, aldehyde, and cyano) (Figure 6). Across these test systems, functionalization yielded improvements of at most 3 kcal/mol in ΔE(HAT) and 6 kal/mol in $|\Delta E$(rebound)$|$ (Supporting Information, Figures S21 and S22). Although secondary electronic effects obscure direct correlations between substituent properties and reactivity, general trends are nevertheless apparent wherein electron-withdrawing groups favorably shift ΔE(HAT) values for both the Mn and Fe TMCs, while electron-donating groups shift $|\Delta E$(rebound)$|$ favorably, particularly in the Mn TMCs (Supporting Information, Figures S21 and S22). Applying these maximal improvements in ΔE(HAT) and $|\Delta E$(rebound)$|$ observed in these functionalized test systems to the generations 1−4 dataset would bring several dominated (i.e., not on the Pareto front) TMCs within reach of the Pareto front,

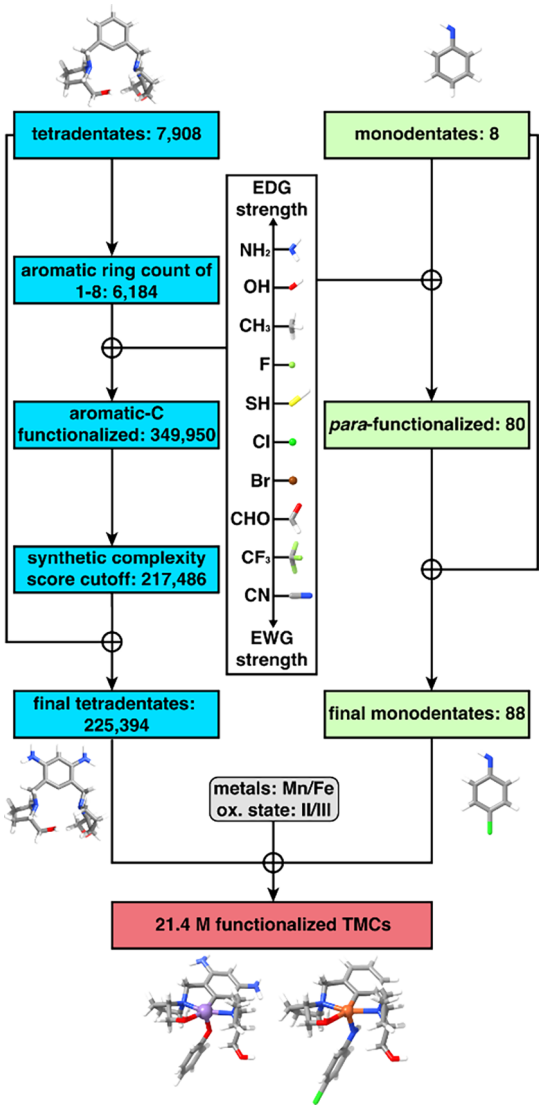


Figure 6. Workflow for constructing the functionalized TMC design space, illustrating the installation of Hammett-type functional groups ranging from electron-donating to electron-withdrawing substituents on the CSD-derived tetradentate ligands and monodentate ligands, followed by their combinatorial assembly with metal identity, oxidation state, and spin state choices.

Downloaded from pubs.acs.org/acscacs/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

corresponding to an approximately four-fold increase in potentially accessible optimal catalyst designs (Supporting Information, Figure S23).

Motivated by these findings, we systematically functionalized the initial TMC design space with the same set of Hammett substituents. The final set of tetradentate ligands used to construct the initial design space were functionalized with each functional group at symmetry-equivalent aromatic carbons and filtered for synthesizability using a heuristic synthetic complexity score (SCScore) cutoff,[140] yielding 217,486 functionalized variants (Figure 6 and Supporting Information, Table S9 and Figure S24). Similarly, the monodentate ligands from the initial design space were functionalized at the *para* position with each functional group, resulting in 80 functionalized monodentate ligands, but these were not screened for their SCScore (Figure 6). To keep the effects of tetradentate and monodentate functionalization separable, functionalized tetradentate ligands were only paired with unfunctionalized monodentate ligands, and vice versa, and combined with nine metal/oxidation-/spin-state configurations, producing 21,352,752 functionalized TMCs (Figure 6). Including the original design space, the final design space comprised 21,922,128 TMCs.

Another consideration is that our functionalization workflow may introduce functional groups at positions distant from the metal center, such that the ligand-centered RACs up to depth 4 in the original feature set may not fully capture the influence of these functional groups (see Section 6b). To address this, we augment the feature set with full-complex RACs to better represent functional group effects (Supporting Information, Table S10).[79,141] Upon updating the feature set, the ΔE(HAT) and ΔE(rebound) ANNs were retrained with hyperparameter optimization on the initialization set and generations 1−4 using this augmented feature set (Supporting Information, Table S11). While the retention of invariant ligand-centered RACs in previous generations did not affect model performance, they were removed here to streamline the expanded feature set and retain only informative descriptors. The updated surrogate models for ΔE(HAT) and ΔE(rebound) were then applied in three additional generations of active learning across the expanded design space, including the new functionalized TMCs as well as TMCs from the initial unfunctionalized design space without DFT evaluations. Due to the nearly 40-fold increase in design space, the top-2D-EHVI selection was increased to 100,000 candidates from previously 10,000, while continuing to sample 200 cluster centers via *k*-medoids from these top candidates for DFT evaluation. To check whether the final surrogate models possess the physical resolution to capture differences in energetics arising from functionalization, we evaluate their predictions on representative Fe and Mn TMCs from generation 3, along with their systematically functionalized variants, which are unseen by the final models. To test whether these differences are quantitatively meaningful, we computed DFT ground truths for all functionalized variants of the representative TMCs and compared the functionalization-induced shifts, $\Delta\Delta E$ (Supporting Information, Figure S25). Evaluation with DFT confirms that functionalization perturbs both energetics by only a few kcal/mol. The predicted and DFT-computed mean $\Delta\Delta E$ agree to within 0.5−2.0 kcal/mol, and individual, per-compound errors on $\Delta\Delta E$ of 1.8−3.6 kcal/mol are smaller than the models' absolute held-out test set errors (8.0 and 5.4 kcal/mol for ΔE(HAT) and ΔE(rebound), respectively) because systematic errors largely cancel in the difference of energies of functionalized derivatives. Consistent with the DFT ground truth, the models also approximately recover the functionalization trends discussed previously for ΔE(HAT) at the extremes, with electron-withdrawing groups (e.g., $-CF_3$) having more favorable ΔE(HAT) than electron-donating groups (e.g., $-NH_2$) for Mn, where the DFT separation is the largest. Nevertheless, the final models cannot reliably rank individual functional groups for which DFT energetics differ by less than the uncertainty in the DFT calculation (e.g., from differences in geometry optimization outcome). This is particularly evident for ΔE(rebound), where the effects of functionalization are predicted to lie close to each other (Supporting Information, Figure S25). In future work, these DFT ground truths could be used to retrospectively repeat Bayesian optimization within the functionalized variants of each representative Fe and Mn TMC, assessing over many independent trials whether model-guided selection consistently outperforms random selection for the same number of candidates evaluated. However, the current limitation on model accuracy does not hinder the broad exploration of Bayesian optimization. As described previously, the campaign requires only that functionalization induces small shifts in energetics, which we confirm here by DFT, rather than that the surrogates rank these shifts a priori. The reported Pareto front comprises DFT-verified complexes throughout.

Across these additional active learning generations, lookahead errors continued to decrease for both ΔE(HAT) and ΔE(rebound), from ca. 6 kcal/mol to 3 kcal/mol for ΔE(HAT) and ca. 9 kcal/mol to 5 kcal/mol for ΔE(rebound) from generation 4 to 7 with predicted uncertainties broadly corresponding to the observed errors, consistent with progressive improvement of the surrogate models (Supporting Information, Figure S26 and Table S12). Evaluation on an independent held-out test set of complexes drawn from the expanded design space but never selected by 2D-EHVI also showed improved predictive accuracy across successive generations, with the ΔE(HAT) MAE decreasing from ca. 12 kcal/mol to 8 kcal/mol and the ΔE(rebound) MAE decreasing from ca. 18 kcal/mol to 5 kcal/mol (Supporting Information, Figure S27 and Table S13).

Because the held-out set is fixed across generations, it also enables a batch-independent assessment of uncertainty calibration. Normalized-residual analysis on this held-out test set shows that predicted uncertainties for ΔE(rebound) remain well behaved across generations (mean normalized residual $\approx$ 0 and root-mean-square (RMS) normalized residual $\approx$ 1), whereas ΔE(HAT) exhibits a persistent negative mean normalized residual and RMS normalized residual greater than 1 that is stable across generations. The concurrent held-out MAE reductions reflect genuine gains in point-prediction accuracy rather than drifting calibration, suggesting that 2D-EHVI-guided active learning improves surrogate generalization across the design space, despite challenges in uncertainty estimation (Supporting Information, Figure S28 and Table S14).

Ligand functionalization strongly shapes the updated Pareto front obtained after three active learning generations on the expanded design space, with 11 of the 16 Pareto-optimal TMCs containing functionalized ligands, highlighting its effectiveness in enhancing ΔE(HAT)−$|\Delta E$(rebound)$|$

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

energetics (Figure 7). Notably, functionalization adds new points to the Pareto front that improve the most favorable ΔE(HAT) on the initial Pareto front by up to 2 kcal/mol. At the same time, $|\Delta E(\text{rebound})|$ decreases by ca. 5 kcal/mol on average over a comparable ΔE(HAT) range (Supporting Information, Figure S29). Examining the final Pareto-optimal TMCs reveals that Fe(II) complexes predominantly feature electron-withdrawing halogen substituents on ligands occupying the equatorial plane (Figure 7 and Supporting Information, Figure S30 and Table S15). Consistent with our functionalization control study, these functional groups enhance ΔE(HAT), and because Fe complexes already exhibit much more favorable ΔE(HAT) than $|\Delta E(\text{rebound})|$, this functionalization explains their presence on the updated Pareto front. In contrast, Pareto-optimal Mn(II) TMCs feature Hammett substituents on both the equatorial and axial ligands, spanning electron-donating (e.g., amino) to electron-withdrawing groups (e.g., halogens, aldehyde, and cyano, Figure 7 and Supporting Information, Figure S30 and Table S15). With Mn TMCs exhibiting more balanced trade-offs between ΔE(HAT) and $|\Delta E(\text{rebound})|$, a broader variety of substituents can favorably tune either reaction energy, as seen from the functionalization control study, enabling comparatively easier access to Pareto-optimal Mn TMCs.

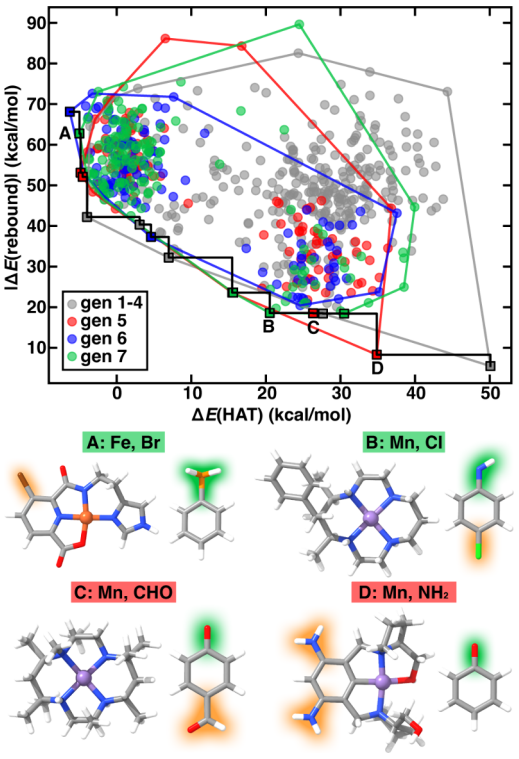


Figure 7. (top) $|\Delta E(\text{rebound})|$ vs ΔE(HAT) scatter plot (kcal/mol) showing TMCs from generation 1−4 (gray), generation 5 (red), generation 6 (blue), and generation 7 (green) of the active learning workflow. Data points from each generation are enclosed by their convex hulls and Pareto-optimal complexes identified across generations 1−7 are depicted as black-outlined squares and connected by a black stair-step line. (bottom) Insets illustrating the metal−tetradentate substructures and monodentate ligands of representative functionalized Pareto-optimal complexes, with the installed functional groups on either ligand highlighted in orange. The green highlight on the monodentate ligands corresponds to the metal-coordinating atoms. Atom color codes: Fe in brown, Mn in purple, C in gray, N in blue, O in red, P in orange, Cl in neon green, Br in dark brown, and H in white.

On the updated Pareto front, functionalized Fe complexes show an increased prevalence of O-coordination on the equatorial plane ligands relative to unfunctionalized Fe complexes on the initial Pareto front, whereas Mn complexes remain predominantly N-coordinated, consistent with the unfunctionalized active learning campaign (Figure 7 and Supporting Information, Figure S30). Although the updated Pareto front again exclusively comprises LS catalysts, four of the LS functionalized TMCs (two Mn and two Fe) also exhibit LS ground states in the resting state (Supporting Information, Table S15). Nevertheless, the small spin splitting energies in these systems indicate marginal LS stabilization, suggesting susceptibility to spin crossover (SCO) to higher-spin states. Analogously, TMCs on the updated front that do not have LS resting ground states may also similarly access the LS state due to small spin splitting energies.

We also revisited the spin-consistent Pareto front after all active learning generations to assess its proximity to the final Pareto front. The final spin-consistent Pareto front is dominated (i.e., nine of eleven) by functionalized Mn(II) and Fe(II) complexes in the LS state, and it has ΔE(HAT) values comparable to the final Pareto front, as seen previously. On the other hand, $|\Delta E(\text{rebound})|$ values are only ca. 5 kcal/mol higher on average, which is half of the upward shift (ca. 10 kcal/mol) observed in the unfunctionalized case (Supporting Information, Figures S31 and S32). This reduced gap between the spin-consistent and final Pareto fronts indicates that functionalization does not hinder access to favorable spin state configurations in the resting state of the catalysts while maintaining optimal target reaction energetics.

4c. Evaluation of Kinetic Trends in Pareto-Optimal Catalysts

We selected six representative catalysts from the final (i.e., not necessarily spin-consistent) Pareto front (**3**, **6**, **8**, **9**, **11**, and **14**) and computed their full methane-to-methanol reaction coordinates with thermodynamic corrections (Supporting Information, Figure S30). The catalysts were chosen to span the $\Delta E(\text{HAT})-|\Delta E(\text{rebound})|$ tradeoff and to have resting state ground states that are either LS or are in low-lying alternative spin states above an LS ground state (Supporting Information, Table S15). The resulting set comprises both Mn(II) (**9**, **11**, and **14**) and Fe(II) (**3**, **6**, and **8**) centers and features tetraaza-based macrocycles on the equatorial plane, except for complex **3**, which contains two O-donor atoms. The axial ligands are negatively charged and feature P- (**3** and **9**), N- (**6**, **8**, and **11**), or O-donor atoms (**14**). The selected Pareto-optimal catalysts include both unfunctionalized systems (**6** and **8**) and complexes functionalized on the equatorial (Br on **3** and F on **9**) and axial planes (Cl on **11** and SH on **14**). Due to thermodynamic accessibility of both the LS and IS states for all selected complexes, we characterize their reaction coordinates on both the LS and IS state surfaces (Figure 8 and Supporting Information, Figure S33). We compute free energy barriers for the metal−oxo formation, HAT, and methyl radical rebound steps (see Section 6a). In general, the resulting reaction coordinates reveal elementary-step thermodynamics broadly consistent with earlier trends, with favorable oxo formation across both metals along with metal-dependent HAT−rebound tradeoffs. They also reveal that HAT and rebound thermodynamics do not influence methane-to-

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

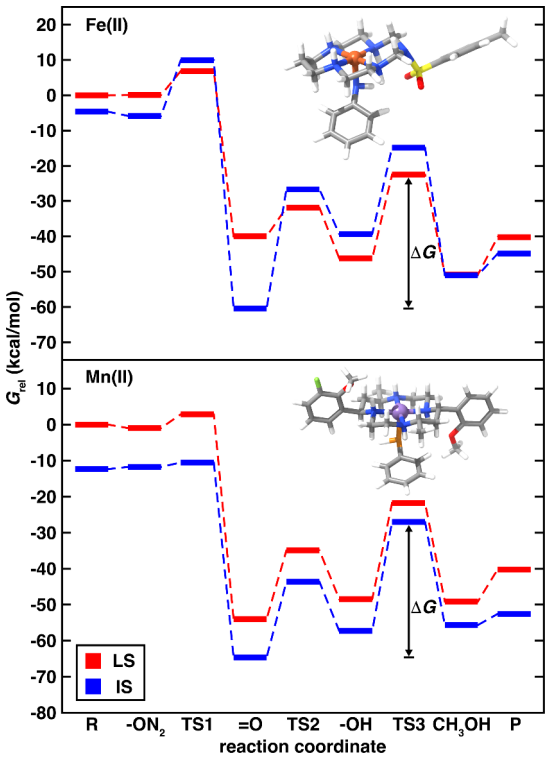


Figure 8. Gibbs free energy landscapes in kcal/mol of representative Pareto-optimal complexes (top) **8** and (bottom) **9** on the LS state surface in red and the IS state surface in blue. All free energies are referenced with respect to the resting state (R) free energy in the LS state. The catalyst structures are shown as insets. The cycle proceeds from the reactants (R) to the products (P) though the N_2O-bound intermediate ($-ON_2$), the metal−oxo formation TS (TS1), the metal−oxo intermediate (=O), the HAT TS (TS2), the metal−hydroxo intermediate (−OH), the radical rebound TS (TS3), and the methanol-bound intermediate (CH_3OH). The energetics spans, ΔG, for each catalyst assuming spin crossover is allowed are shown as annotated insets.

methanol free energy spans in these catalysts, consistent with our optimization strategy. Consequently, turnover is instead governed by steps left outside the optimization objectives: the metal−oxo formation thermodynamics and rebound kinetics, as discussed in detail next.

We first assess the extent to which the BEP relationships hold for the selected Pareto-optimal catalysts. BEPs, which relate activation barriers for elementary steps to readily computed thermodynamic descriptors, are widely used to accelerate catalyst discovery.[47,48,58−61,96] Here, we examine the validity of the BEPs on the electronic surface for the metal−oxo formation, HAT, and methyl radical rebound steps across both metal identities and LS and IS states (Supporting Information, Table S16 and Figure S34). For each step, electronic barriers and reaction energies are referenced to the N_2O-bound, metal−oxo, and metal−hydroxo intermediates, respectively. The BEP for the metal−oxo formation step holds moderately ($R^2 = 0.78$), reflecting structural variability in the oxo formation transition state using N_2O as the oxidant, particularly in the N−N−O angle and N−O and metal−oxo bond lengths (Supporting Information, Table S17). In contrast, we observe a very strong BEP for HAT ($R^2 = 0.94$) with a near-unity slope, consistent with prior work[61,96] and attributable to the highly conserved nature of the HAT transition state geometries (Supporting Information, Table S18). Notably, the rebound step also exhibits a BEP of intermediate strength between these two limits ($R^2 = 0.86$), arising from the relatively high similarity of methyl radical approach and binding motifs during the rebound step (Supporting Information, Table S19). Thus, while the dataset size is limited, catalyst screening for methane-to-methanol across Mn and Fe systems primarily on thermodynamic descriptors, as we have carried out in our optimization, is well motivated across this set.

The free energy landscapes of the selected Pareto-optimal catalysts reveal clear metal-dependent spin−reactivity trends. For the Mn catalysts, both the LS and IS state surfaces remain relevant along the entire reaction coordinate, with spin surface energy differences lying within 5−10 kcal/mol for many intermediates and transition states and none exceeding 16 kcal/mol. In contrast, Fe catalysts largely retain energetically similar LS and IS spin state catalytic surfaces, except during metal−oxo formation where the IS metal−oxo becomes a thermodynamic trap due to its strong stabilization relative to the LS metal−oxo by ca. 27 kcal/mol for **3**, 24 kcal/mol for **6**, and 20 kcal/mol for **8**, respectively (Figure 8 and Supporting Information, Figure S33). The pronounced stabilization of the metal−oxo species in IS Fe catalysts results in unfavorable HAT energetics that are penalized in our optimization, leading to their exclusion from the Pareto front in favor of LS Fe catalysts. Evidenced by the strong HAT BEP, the unfavorable HAT energetics for the IS Fe catalysts translate to high HAT barriers (37 kcal/mol for **3**, 35 kcal/mol for **6**, and 34 kcal/mol for **8**) relative to LS Fe catalysts (barrierless for **3**, 4 kcal/mol for **6**, and 8 kcal/mol for **8**). Notably, the barrierless HAT in the LS state observed for complex **3**, the sole system with an O-coordinating equatorial ligand in our selected Pareto-optimal complexes, is consistent with prior work[96] showing that electron-rich O-donors promote very favorable HAT. In contrast, the Mn catalysts studied have comparable HAT barriers at the LS and IS states (19 vs 21 kcal/mol for **9**, 21 vs 23 kcal/mol for **11**, and 31 vs 21 kcal/mol for **14**, Figure 8 and Supporting Information Figure S33). Beyond HAT, the remaining elementary step barriers display limited sensitivity to spin state for both metals. Metal−oxo formation barriers for the evaluated Mn catalysts are uniformly low across both spin states, spanning 1−12 kcal/mol, with most barriers below 5 kcal/mol (e.g., LS and IS **9**, IS **11**, and LS **14**). The evaluated Fe catalysts show a broader range of 3−31 kcal/mol for oxo formation barriers across both spin states, with most barriers exceeding 15 kcal/mol (e.g., LS and IS **3**, LS **6**, and IS **8**). Rebound barriers differ by no more than 3.5 kcal/mol between LS and IS states for the Mn systems and 2.5 kcal/mol for the Fe systems we evaluated (Figure 8 and Supporting Information, Figure S33).

Notably, rebound barriers in our Pareto-optimal catalysts are uniformly high. For instance, Fe catalyst **8** has a rebound barrier of 24 kcal/mol at both LS and IS while Mn catalyst **9** has rebound barriers of 27 kcal/mol at LS and 30 kcal/mol at IS (Figure 8). Even on the electronic PES (i.e., neglecting entropic contributions), these barriers (11 kcal/mol for Fe catalyst **8** at both LS and IS, 12 kcal/mol at LS and 16 kcal/mol at IS for Mn catalyst **9**) surpass those reported for Pareto-optimal catalysts in prior studies, where rebound was barrierless or at most 4 kcal/mol for LS Fe(II) catalysts and 11 kcal/mol for an IS Mn(II) catalyst (Supporting Information, Table S16).[96] We attribute this to increased steric congestion in the equatorial ligand environments for our Pareto-optimal catalysts, for

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

example from pendant hydrogen atoms that restrict methyl radical approach (Supporting Information, Figure S30). In contrast, prior Pareto-optimal complexes featured less crowded equatorial planes, allowing more facile rebound.[96] Notably, Fe complex **3**, with a relatively less hindered equatorial plane, exhibits electronic rebound barriers comparable to systems studied previously (Supporting Information, Table S16). Inclusion of thermodynamic corrections further raises the rebound barriers on the free energy surface due to significant entropic penalties associated with the constrained approach and binding of the methyl radical, which are considerably larger for rebound (ca. 14 kcal/mol on average) than for HAT (ca. 4 kcal/mol) or oxo formation (ca. 0.5 kcal/mol, Supporting Information, Table S20). We note that the reported rebound free energy barriers were computed assuming infinitely separated reactants under the conventional gas-phase standard state (298 K, 1 bar). Since the methyl radical is a transient intermediate whose effective concentration is expected to be lower under methane-to-methanol catalytic conditions, adopting lower radical concentrations would further increase the reported rebound free energy barriers (Supporting Information, Text S6).

To assess the catalytic efficiency of our selected Pareto-optimal catalysts, we applied the energetic span model[142] to the computed methane-to-methanol reaction coordinates. This allows us to identify how the optimization strategy reshapes the turnover-determining states and to provide an indication of the resulting approximate turnover frequencies (TOFs) for the specific Mn and Fe complexes studied. For each catalyst, energetic spans were evaluated on the LS and IS surfaces individually, as well as along pathways permitting SCO between these surfaces (Supporting Information, Table S21). For the evaluated Mn complexes **9**, **11**, and **14**, the turnover-determining transition state (TDTS) is consistently the rebound step, and the turnover-determining intermediate (TDI) is the metal−oxo, reflecting their highly favorable oxo formation energies. These rate-determining states remain unchanged when SCO is allowed. In contrast, LS Fe complexes exhibit variable rate-determining states, with either metal−oxo (**3** and **6**) or rebound transition states (**8**) serving as the TDTS and metal−hydroxo (**8**) or methanol-bound species (**3** and **6**) acting as the TDI. At IS, the stabilized metal−oxo exclusively constitutes the TDI with the rebound transition state as the TDTS for all complexes **3**, **6**, and **8**, a pattern that persists when SCO is considered. While the evaluated Mn catalysts exhibit comparable energetic spans across LS and IS states that are within 5 kcal/mol on average, Fe catalysts **3**, **6**, and **8** show significantly lower spans at LS than IS (by ca. 13 kcal/mol on average), reflecting the bias toward IS metal−oxo stabilization. The strong stabilization of the IS metal−oxo in Fe catalysts constrains their reaction pathway through this state, making energetic spans that account for SCO a more faithful measure of their catalytic turnover. When SCO is considered, the Mn and Fe catalysts we studied have comparable energetic spans, with average values of ca. 33 kcal/mol for the Fe and 38 kcal/mol for the Mn systems (Supporting Information, Table S21).

We note that the energetic spans reported here are systematically higher than those reported in our prior studies on homogeneous[48,96] and heterogeneous[42] catalyst systems by ca. 10−15 kcal/mol, primarily because those prior energetic span analyses did not account for rebound kinetics. By favoring low-magnitude rebound energetics, our optimization also led to Pareto-optimal catalysts with destabilized methanol-bound intermediates, yielding thermoneutral rebound (e.g., ΔG(rebound) values of −12 and −1 kcal/mol for **8** and **9**, respectively on their ground state spin surfaces, Figure 8). This destabilization arises primarily from steric crowding and consequently exposes high rebound barriers that become turnover-determining, consistent with the observed rebound BEP. When rebound barriers are excluded (21−29 kcal/mol) or when entropic penalties are omitted (23−32 kcal/mol), the resulting energetic spans fall within the ranges reported previously. This underscores that explicit treatment of rebound kinetics on the free energy surface is essential for obtaining more reliable insight on catalytic turnover (Supporting Information, Tables S22 and S23). Beyond rebound kinetics, metal−oxo formation, which was not included in the optimization objective due to its anticipated high thermodynamic favorability, also dictates the energetic span as the TDI. The oxo formation step, however, is easily tunable through oxidant choice (i.e., because the thermodynamics of oxo formation includes a contribution from how good an oxo donor the oxidant is and from the catalyst) and does not intrinsically limit improvements in turnover. Notably, the energetic spans obtained here are comparable to or better than those reported for experimentally studied C−H activation catalysts for which full DFT reaction coordinates are available (Supporting Information, Table S24). In contrast, comparisons to catalysts with TOFs derived purely from experiment show that the experimentally measured TOFs are higher than our computed TOF values, which likely reflects the influence of the reaction environment and operating conditions (Supporting Information, Table S24).

Taken together, our analysis suggests that catalytic turnover could be further improved beyond the best-in-class designs by tuning two key levers along the reaction coordinate. First, metal−oxo formation thermodynamics can be tuned by selecting oxidants that maintain downhill oxo formation while avoiding an excessively favorable metal−oxo formation energy, raising the TDI energy and reducing the energetic span. Second, rebound kinetics can be improved through balanced steric tuning of the equatorial ligand environment, which destabilizes the methanol-bound intermediate while preserving an accessible pathway for the methyl radical, enabling near-thermoneutral rebound without prohibitive kinetic barriers. Achieving this likely requires the design of novel macrocycles beyond current synthetically accessible ligands that break the kinetic-thermodynamic scaling at the rebound step. With the EHVI acquisition function being agnostic to the number of objectives, higher-dimensional thermodynamic optimization that also includes oxo formation energetics or constraints based on computed or BEP-predicted kinetic barriers could also aid future optimizations. Solvent effects may further lower rebound barriers by mitigating the entropic penalty of methyl radical capture,[143,144] together helping to lower the TDTS and improve overall energetic spans.

5. CONCLUSIONS

We employed multi-objective global Bayesian optimization to explore a vast catalyst design space and discover novel methane-to-methanol catalysts. The optimization targeted the HAT and methyl radical rebound steps of the radical rebound mechanism. This strategy was designed to promote facile C−H activation alongside thermoneutral rebound, enabling flatter methane-to-methanol reaction energy landscapes. We

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

constructed an initial design space of over half a million synthetically accessible square-pyramidal Mn and Fe mononuclear complexes from tetradentate macrocyclic ligands obtained from the CSD with aromatic monodentate ligands. We navigated this design space using MOGBO with iteratively retrained ANN surrogate models incorporating Laplace-approximated uncertainty, efficiently identifying promising catalysts while progressively improving surrogate model accuracy.

The Pareto front over the initial design space exclusively comprised LS Fe(II) and Mn(II) catalysts, with Mn(II) catalysts exhibiting more balanced ΔE(HAT)−|ΔE(rebound)| tradeoffs. Tetraaza-based macrocycles emerged as a privileged equatorial motif, paired with a negatively charged monodentate aromatic axial ligand. Porphyrin catalysts in the evaluated space had comparable ΔE(HAT) values to the Pareto-optimal catalysts, as expected, but their |ΔE(rebound)| values were considerably higher. Expansion of the design space to 21.9 million catalysts via Hammett-type functionalization, followed by additional active learning generations confirmed that introducing substituents further enhanced catalytic performance for both Mn and Fe catalysts. Future work could test how efficiently active learning selects among closely related functionalized variants compared with random search and develop models that better capture the small energetic differences introduced by functionalization.

Finally, we computed free energy landscapes for representative Mn and Fe Pareto-optimal catalysts on both LS and IS surfaces to assess catalytic turnover. Because of the optimization strategy used to identify these catalysts, HAT and rebound thermodynamics no longer limited turnover, as expected. Instead, turnover was consistently governed by metal−oxo formation thermodynamics and rebound kinetics when spin crossover was considered. Penalizing high-magnitude rebound energetics during optimization favored sterically congested equatorial environments that hindered methyl radical approach, making rebound transition states turnover-determining. Nevertheless, the resulting energetic spans were comparable across the Mn and Fe systems studied and competitive with reported values for other earth-abundant catalysts, indicating that our Pareto-optimal candidates are already on par with the best experimentally characterized catalysts. While the spans may be reduced by tuning oxo formation thermodynamics through oxidant choice, sterically crowded macrocyclic ligands derived from experimentally realized TMCs remain limited in achieving low rebound barriers while maintaining near-thermoneutral rebound. This motivates the design of TMCs with novel macrocyclic ligands that escape this thermodynamic-kinetic scaling at the rebound step to further improve methane-to-methanol turnover.

6. COMPUTATIONAL DETAILS

6a. DFT Calculations

All gas-phase geometry optimizations and single-point energy calculations were performed using DFT with a development version of TeraChem v1.9.[145,146] We employed the B3LYP[147−149] global hybrid functional with empirical D3 dispersion correction[150] using Becke−Johnson damping.[151] We used the LACVP* composite basis set, comprising the LANL2DZ effective core potential[152,153] for atoms heavier than Ar and a 6-31G* basis set[154] for all lighter atoms. Solvent effects, zero-point energies, and entropy corrections were omitted to enable computational scalability, in accordance with prior work.[46,48,96]

Singlet calculations were treated using the spin-restricted formalism, whereas higher-spin states were computed with the unrestricted formalism. To aid self-consistent field (SCF) convergence for open-shell systems, we applied a 0.25 Ha level shift[155] to both majority- and minority-spin virtual orbitals (Supporting Information, Text S7). Geometry optimizations used the translation-rotation internal coordinate (TRIC) optimizer[156] with the BFGS algorithm and default convergence criteria of a maximum energy gradient of 4.5×10^{-4} Ha/bohr and an energy difference threshold of 10^{-6} Ha between steps.

Catalyst structures were generated using a combination of tetradentate ligands extracted from mononuclear TMCs in the CSD[105] version 5.45 update 1 (March 2024), and monodentate ligands from prior work.[48,96] We positioned the reactive moiety on the axial plane, such that square planar ligands exclusively occupied the equatorial plane with the monodentate ligand placed axially, while seesaw ligands had three donors on the equatorial plane and one on the axial plane, with the monodentate ligand placed equatorially (see Section 3). Intermediates in the catalytic cycle were generated sequentially, starting with the metal−oxo species and subsequent intermediates were constructed only if the preceding structure's geometry optimization was successful. Initial metal−oxo geometries were generated using custom scripts in molSimplify,[157,158] starting from the metal-tetradentate ligand substructures extracted from the TMCs (see Section 3), with the metal replaced by Mn or Fe, followed by adding one of eight monodentate ligands and placing an oxo group at a predefined metal−oxo distance of 1.6 Å. Following prior protocols[46,96] metal-hydroxo initial geometries were built by protonating the optimized metal−oxo structure, and methanol-bound initial geometries by methylating the optimized metal-hydroxo species (Supporting Information, Figure S35). Resting state structures were obtained via single-point calculations after removing the methanol ligand from the methanol-bound species.

Job submission was automated using molSimplify, with each run allocated a 24 h wall time and up to three resubmissions. Prior to each resubmission, structures that failed any geometry checks[159] were discarded (Supporting Information Table S25). Open-shell structures were eliminated based on established criteria:[45,141,159] i) if the ⟨$\hat{S}^2$⟩ value deviated from $S(S + 1)$ by >1 or ii) if the Mulliken spin density on the metal and the oxygen atom for the metal−oxo species or just the metal for other intermediates differed from the total spin by >1 μ_B. After optimizing all intermediates, two additional structural checks were applied to ensure geometric stability across the cycle (Supporting Information, Table S25). Most calculation failures arose from non-convergence or failed geometry checks, with further attrition from the structural filters applied at the end of the cycle (Supporting Information, Table S26).

Reaction coordinates for Mn and Fe complexes were computed using ORCA v6.0.0.[160,161] We compute these coordinates through a relaxed potential energy surface (PES) scan along the N−O bond distance for metal−oxo, the O−H bond distance for HAT, and the C−O bond distance for rebound, followed by a partitioned rational-function optimization (P-RFO)[162] on the highest-energy structure from the scan. For HAT and rebound, we ensure that the transition states are associated with an α-radical transfer with antiferromagnetic methyl coupling.[46,48] We employed ORCA rather than TeraChem due to the former code's support of analytical Hessians. All calculations employed the B3LYP functional with D3 dispersion corrections

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

and Becke−Johnson damping. The def2-TZVPP[163] basis set was applied to Mn and Fe atoms while the def2-SVP[163] basis set was used for all other atoms. Geometries of methane-to-methanol intermediates previously optimized in TeraChem were used as starting structures and subsequently reoptimized in ORCA to ensure consistency with the change in basis set and differences in functional implementation between the two software packages. Analytical Hessian calculations were then performed to obtain thermochemical corrections. Transition states were located by first performing PES scans where the putative transition state mode, such as a bond length, was incrementally changed while geometry optimizing all other degrees of freedom. Highest-energy points identified from these scans were used as initial guesses for P-RFO[162] to refine the transition state structures. All transition states were validated via vibrational frequency analysis, confirming the presence of a single imaginary frequency corresponding to the expected transition state mode. Transition states for the HAT and methyl radical rebound steps were performed via broken-symmetry DFT,[164] enforcing an α-radical transfer with an antiferromagnetically coupled methyl radical, conserving the overall spin state at these steps.[46,48]

6b. Featurization

Our models use geometry-free, connectivity-based RACs[79,141] computed on initial metal−oxo intermediates, following their prior success in TMC property prediction,[79,94] including for methane-to-methanol catalyst design.[45,96] RACs are defined as sums of products or differences of heuristic atomic properties (i.e., nuclear charge, electronegativity, topology, identity, and covalent radius) between atoms separated by a bond path of depth d. For initial active learning generations (1−4), we only include ligand-centered RACs up to depth 4, totaling 100 features (Supporting Information, Table S10). Of these, 20 were invariant but retained, as the multi-layer perceptron architecture effectively neglects these uninformative features, as indicated by comparable performance when those features were explicitly removed (Supporting Information, Figure S36 and Table S27). Full-complex RACs were excluded from the initial feature set, and metal-centered RACs were excluded throughout this study. Instead, metal information was captured through 6 additional full-complex features: metal identity (Mn vs Fe, one-hot encoded), oxidation state, spin state, and the charges of the tetradentate and monodentate ligands, yielding 106 features total. For subsequent active learning generations (4−7), the feature set was augmented with full-complex RACs up to depth 4 to better capture distal ligand functionalization effects, and invariant ligand-centered RACs were removed (Supporting Information Table S10).

6c. ML Model Training

We trained two independent ANNs to predict the HAT energy ΔE(HAT), and the methyl radical rebound energy ΔE(rebound), using Keras[165] with TensorFlow as the backend.[166] Input features and target values were Z-score normalized using training set statistics. At each active learning training iteration, data were randomly split into a 70%/10%/20% training/validation/test split. During initial training, optimal hyperparameters for both models were selected via Bayesian optimization in KerasTuner[167] over 200 trials, minimizing the validation set mean squared error (Supporting Information, Table S11). In subsequent generations, optimized hyperparameters defining model architectures and training settings were retained. Models were initialized from previously converged weights and retrained on the updated data with the learning rate unchanged from initial training, since they were already sufficiently low for stable training (Supporting Information, Table S11). Both ANNs were trained with the Adam optimizer for up to 4000 epochs, with dropout, batch normalization, and early stopping (patience = 100) to mitigate overfitting.

GPR models were trained using scikit-learn,[168] with Z-score normalization applied to inputs and targets. We used a radial basis function (RBF) kernel summed with a WhiteKernel component to account for noise in the data. Kernel hyperparameters, including the RBF length scale and noise level, were optimized during training, with a fixed alpha of 0.5 to ensure numerical stability.

■ ASSOCIATED CONTENT

Data Availability Statement

All data presented in the paper is available as supporting information in both a PDF summary as well as a ZIP file on Zenodo (https://doi.org/10.5281/zenodo.21828678). The code is available at: https://github.com/hjkgrp/tmc-m2m-al.git.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.6c03365.

Shifts in ΔE(oxo) when using other oxidants instead of N_2O; oxidation and spin states of Mn and Fe in different intermediates; accounting of CSD TMC and tetradentate macrocycle eliminations; distribution of aromatic ring counts in final CSD tetradentate ligands; comparison of initialization set Pareto fronts for key reaction energetics; comparison of ANN and GP model performances; quantifying model parameter uncertainty via Laplace approximation; ΔE(HAT) error vs Laplace uncertainty for different Hessian approximations; propagating model parameter uncertainty to predictive uncertainty; details on post-hoc error calibration; ΔE(HAT) error vs uncertainty comparison for different uncertainty metrics; details on 2D-EHVI calculation; dataset summary for training ΔE(HAT) and ΔE(rebound) models; details on active learning generation 2 selection; maximum Tanimoto similarities across design space post-generation 1; t-SNE of initial design space with generation 1 and Tanimoto-filtered subsets; heavy-atom count and net charge distributions on CSD tetradentate ligands; lookahead errors for ΔE(HAT) and ΔE(rebound) across generations 1−4; 2D-EHVI distributions at representative active learning stages; average 2D-EHVI of top-10,000 candidates across generations 1−4; generations 1−4 lookahead errors by tetradentate ligand coordination geometry; pareto front over the initial design space with initialization set included; structures of Pareto-optimal TMCs across generations 1−4; properties of Pareto-optimal TMCs across generations 1−4; catalysis by CSD TMCs with tetradentate ligands from initial Pareto front; representative TMCs with meso-substituted porphyrins (generations 1−4); structures of spin-consistent Pareto-optimal TMCs across generations 1−4; IS Mn(III) TMCs near spin-consistent Pareto front (generations 1−4); $|\Delta E$(rebound)$|$ vs ΔE(HAT) for initialization set showing HS Fe(II) TMCs; electronic and geometric properties of sampled TMCs (generations 1−4); structures of representative sub-optimal TMCs across generations 1−4;

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX

ΔE(oxo) distributions in Mn vs Fe TMCs (generations 1−4); ΔE(release) vs |ΔE(rebound)| (generations 1−4); Hammett substituent effects on target energies in a representative Mn TMC; Hammett substituent effects on target energies in a representative Fe TMC; pareto-accessible TMCs (generations 1−4) via functionalization energy shifts; summary of workflow for functionalizing CSD tetradentate ligands; SCScore distribution of CSD tetradentate ligands; breakdown of RACs used for TMC featurization across all generations; hyperparameters for ΔE(HAT) and ΔE(rebound) surrogate ANNs; DFT vs predicted functionalization effects for example Fe and Mn TMCs; lookahead errors for ΔE(HAT) and ΔE(rebound) across generations 1−7; lookahead metrics for ΔE(HAT) and ΔE(rebound) across generations 1−7; held-out test set errors for all ΔE(HAT) and ΔE(rebound) model iterations; held-out test set metrics for all ΔE(HAT) and ΔE(rebound) model iterations; held-out test set *z*-scores for final ΔE(HAT) and ΔE(rebound) models; Held-out test set *z*-score statistics for all ΔE(HAT) and ΔE(rebound) models; comparison of Pareto fronts on initial vs functionalized design spaces; structures of Pareto-optimal TMCs across generations 1−7; properties of Pareto-optimal TMCs across generations 1−7; spin-consistent Pareto front across generations 1−7; structures of spin-consistent Pareto-optimal TMCs across generations 1−7; free energy landscapes of select Pareto-optimal TMCs; electronic reaction coordinate data for select Pareto-optimal TMCs; BEPs for oxo formation, HAT; and rebound steps; key geometric parameters of oxo formation transition states; key geometric parameters of HAT transition states; key geometric parameters of rebound transition states; entropic penalties for oxo formation, HAT, and rebound steps; details on standard-state correction for rebound free energies; energetic spans for select Pareto-optimal TMCs; energetic spans (no rebound transition states) for select Pareto-optimal TMCs; electronic energetic spans for select Pareto-optimal TMCs; turnover frequencies for experimental C−H activation catalysts; details on job rescue; approach of constructing catalytic cycle intermediates; metrics for geometry checks; DFT job statistics at each active learning generation; comparison of ANN performance with vs without invariant RAC features; comparison of GPR performance with vs without invariant RAC features; optimized structures of all reactive intermediates of TMCs in the final dataset at each active learning generation; optimized intermediates and approximate transition state structures from calculated reaction coordinates; features and energetics of TMCs used to train surrogate models at each generation; calculation failure modes, raw DFT energies, and electronic properties for each TMC intermediate at each generation; features and predicted energetics with associated standard deviations for TMCs selected for DFT in subsequent generations; trained models at each generation; monodentate ligands; initial CSD-derived tetradentate ligands with relevant properties (PDF)

AUTHOR INFORMATION

Corresponding Author

Heather J. Kulik − *Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America; Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America;* orcid.org/0000-0001-9342-0191; Email: hjkulik@mit.edu

Authors

Husain Adamji − *Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America;* orcid.org/0000-0003-3058-680X

Ilia Kevlishvili − *Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America;* orcid.org/0000-0001-8133-4165

David W. Kastner − *Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America; Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America;* orcid.org/0000-0002-7766-4249

Ralf Meyer − *Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America;* orcid.org/0000-0003-2236-0261

Yuriy Román-Leshkov − *Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America; Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States of America;* orcid.org/0000-0002-0025-4233

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acscatal.6c03365

Author Contributions

H.A.: conceptualization, writing−initial draft, writing−review and editing, investigation, formal analysis, and data curation. I.K.: supervision, conceptualization, writing−review and editing, investigation, and formal analysis. D.W.K.: investigation, data curation. R.M.: investigation. Y.R.-L.: supervision, project administration, and funding acquisition. H.J.K.: conceptualization, writing−review and editing, supervision, project administration, and funding acquisition.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was initially supported by the National Science Foundation under grant number CBET-1846426 (to H.A., I.K., and H.J.K.). The authors also acknowledge support for the screening effort by the National Institutes of Health under grant number R35GM152027 (to H.A., D.W.K., and H.J.K.). This work was also supported by the Department of Energy, Office of Basic Energy Sciences under award number DE-SC0016214 (to H.A. and Y.R.-L.). We also acknowledge support from a seed grant from the MIT Energy Initiative (to H.A., R.M., and H.J.K.) and a fellowship from the MIT Energy Initiative, supported by Shell (to H.A.). The authors acknowledge the MIT SuperCloud and Lincoln Laboratory

Downloaded from pubs.acs.org/acscacs/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

Supercomputing Center as well as MIT Office of Research Computing and Data for providing high performance computing resources that have contributed to the research results reported within this paper. The authors thank Adam H. Steeves for providing a critical reading of the manuscript.

REFERENCES

(1) Olah, G. A. Beyond Oil and Gas: The Methanol Economy. *Angew. Chem., Int. Ed.* **2005**, *44*, 2636−2639.
(2) Lunsford, J. H. Catalytic Conversion of Methane to More Useful Chemicals and Fuels: a Challenge for the 21st Century. *Catal. Today* **2000**, *63*, 165−174.
(3) Ravi, M.; Ranocchiari, M.; van Bokhoven, J. A. The Direct Catalytic Oxidation of Methane to Methanol—A Critical Assessment. *Angew. Chem., Int. Ed.* **2017**, *56*, 16464−16483.
(4) Kondratenko, E. V.; Peppel, T.; Seeburg, D.; Kondratenko, V. A.; Kalevaru, N.; Martin, A.; Wohlrab, S. Methane Conversion into Different Hydrocarbons or Oxygenates: Current Status and Future Perspectives in Catalyst Development and Reactor Operation. *Catal. Sci. Technol.* **2017**, *7*, 366−381.
(5) Dinh, K. T.; Sullivan, M. M.; Serna, P.; Meyer, R. J.; Dincă, M.; Román-Leshkov, Y. Viewpoint on the Partial Oxidation of Methane to Methanol Using Cu- and Fe-Exchanged Zeolites. *ACS Catal.* **2018**, *8*, 8306−8313.
(6) Dummer, N. F.; Willock, D. J.; He, Q.; Howard, M. J.; Lewis, R. J.; Qi, G.; Taylor, S. H.; Xu, J.; Bethell, D.; Kiely, C. J.; Hutchings, G. J. Methane Oxidation to Methanol. *Chem. Rev.* **2023**, *123*, 6359−6411.
(7) Jiménez, J. D.; Lustemberg, P. G.; Danielis, M.; Fernández-Villanueva, E.; Hwang, S.; Waluyo, I.; Hunt, A.; Wierzbicki, D.; Zhang, J.; Qi, L.; Trovarelli, A.; Rodriguez, J. A.; Colussi, S.; Ganduglia-Pirovano, M. V.; Senanayake, S. D. From Methane to Methanol: Pd-iC-CeO_2 Catalysts Engineered for High Selectivity via Mechanochemical Synthesis. *J. Am. Chem. Soc.* **2024**, *146*, 25986−25999.
(8) Ab Rahim, M. H.; Forde, M. M.; Jenkins, R. L.; Hammond, C.; He, Q.; Dimitratos, N.; Lopez-Sanchez, J. A.; Carley, A. F.; Taylor, S. H.; Willock, D. J.; Murphy, D. M.; Kiely, C. J.; Hutchings, G. J. Oxidation of Methane to Methanol with Hydrogen Peroxide using Supported Gold-Palladium Alloy Nanoparticles. *Angew. Chem., Int. Ed.* **2013**, *52*.
(9) Carter, J. H.; Lewis, R. J.; Demetriou, N.; Williams, C.; Davies, T. E.; Qin, T.; Dummer, N. F.; Morgan, D. J.; Willock, D. J.; Liu, X.; Taylor, S. H.; Hutchings, G. J. The Selective Oxidation of Methane to Methanol using in situ Generated H_2O_2 over Palladium-Based Bimetallic Catalysts. *Catal. Sci. Technol.* **2023**, *13*, 5848−5858.
(10) He, Y.; Luan, C.; Fang, Y.; Feng, X.; Peng, X.; Yang, G.; Tsubaki, N. Low-Temperature Direct Conversion of Methane to Methanol over Carbon Materials Supported Pd-Au Nanoparticles. *Catal. Today* **2020**, *339*, 48−53.
(11) Huang, W.; Zhang, S.; Tang, Y.; Li, Y.; Nguyen, L.; Li, Y.; Shan, J.; Xiao, D.; Gagne, R.; Frenkel, A. I.; Tao, F. Low-Temperature Transformation of Methane to Methanol on Pd_1O_4 Single Sites Anchored on the Internal Surface of Microporous Silicate. *Angew. Chem., Int. Ed.* **2016**, *55*, 13441−13445.
(12) Shan, J.; Li, M.; Allard, L. F.; Lee, S.; Flytzani-Stephanopoulos, M. Mild Oxidation of Methane to Methanol or Acetic Acid on Supported Isolated Rhodium Catalysts. *Nature* **2017**, *551*, 605−608.
(13) Tang, Y.; Li, Y.; Fung, V.; Jiang, D.-e.; Huang, W.; Zhang, S.; Iwasawa, Y.; Sakata, T.; Nguyen, L.; Zhang, X.; Frenkel, A. I.; Tao, F. Single Rhodium Atoms Anchored in Micropores for Efficient Transformation of Methane under Mild Conditions. *Nature Comm.* **2018**, *9*, 1231.
(14) Kwon, Y.; Kim, T. Y.; Kwon, G.; Yi, J.; Lee, H. Selective Activation of Methane on Single-Atom Catalyst of Rhodium Dispersed on Zirconia for Direct Conversion. *J. Am. Chem. Soc.* **2017**, *139*, 17694−17699.
(15) Qi, G.; Davies, T. E.; Nasrallah, A.; Sainna, M. A.; Howe, A. G. R.; Lewis, R. J.; Quesne, M.; Catlow, C. R. A.; Willock, D. J.; He, Q.; Bethell, D.; Howard, M. J.; Murrer, B. A.; Harrison, B.; Kiely, C. J.; Zhao, X.; Deng, F.; Xu, J.; Hutchings, G. J. Au-ZSM-5 Catalyses the Selective Oxidation of CH_4 to CH_3OH and CH_3COOH using O_2. *Nature Catal.* **2022**, *5*, 45−54.
(16) Agarwal, N.; Freakley, S. J.; McVicker, R. U.; Althahban, S. M.; Dimitratos, N.; He, Q.; Morgan, D. J.; Jenkins, R. L.; Willock, D. J.; Taylor, S. H.; Kiely, C. J.; Hutchings, G. J. Aqueous Au-Pd Colloids Catalyze Selective CH_4 Oxidation to CH_3OH with O_2 under Mild Conditions. *Science* **2017**, *358*, 223−227.
(17) Dubkov, K. A.; Sobolev, V. I.; Talsi, E. P.; Rodkin, M. A.; Watkins, N. H.; Shteinman, A. A.; Panov, G. I. Kinetic Isotope Effects and Mechanism of Biomimetic Oxidation of Methane and Benzene on FeZSM-5 zeolite. *J. Mol. Catal. A Chem.* **1997**, *123*, 155−161.
(18) Bols, M. L.; Hallaert, S. D.; Snyder, B. E. R.; Devos, J.; Plessers, D.; Rhoda, H. M.; Dusselier, M.; Schoonheydt, R. A.; Pierloot, K.; Solomon, E. I.; Sels, B. F. Spectroscopic Identification of the α-Fe/α-O Active Site in Fe-CHA Zeolite for the Low-Temperature Activation of the Methane C−H Bond. *J. Am. Chem. Soc.* **2018**, *140*, 12021−12032.
(19) Snyder, B. E. R.; Vanelderen, P.; Bols, M. L.; Hallaert, S. D.; Böttger, L. H.; Ungur, L.; Pierloot, K.; Schoonheydt, R. A.; Sels, B. F.; Solomon, E. I. The Active Site of Low-Temperature Methane Hydroxylation in Iron-Containing Zeolites. *Nature* **2016**, *536*, 317−321.
(20) Groothaert, M. H.; Smeets, P. J.; Sels, B. F.; Jacobs, P. A.; Schoonheydt, R. A. Selective Oxidation of Methane by the Bis(μ-oxo)-dicopper Core Stabilized on ZSM-5 and Mordenite Zeolites. *J. Am. Chem. Soc.* **2005**, *127*, 1394−1395.
(21) Narsimhan, K.; Iyoki, K.; Dinh, K.; Román-Leshkov, Y. Catalytic Oxidation of Methane into Methanol over Copper-Exchanged Zeolites with Oxygen at Low Temperature. *ACS Cent. Sci.* **2016**, *2*, 424−429.
(22) Dinh, K. T.; Sullivan, M. M.; Narsimhan, K.; Serna, P.; Meyer, R. J.; Dincă, M.; Román-Leshkov, Y. Continuous Partial Oxidation of Methane to Methanol Catalyzed by Diffusion-Paired Copper Dimers in Copper-Exchanged Zeolites. *J. Am. Chem. Soc.* **2019**, *141*, 11641−11650.
(23) Zhang, H.; Zhang, X.; Shi, X.; Liu, X.; Liang, J.; Li, S.; Ma, S.; Xiao, X.; Guo, J.; Zhang, L.; Tan, M. Single-Step Conversion of Methane-Steam to Methanol on Single-Atom Cu1/γ-Al_2O_3 Catalyst Prepared via Electrostatic Anchoring. *Appl. Catal., B* **2025**, *361*, No. 124648.
(24) Periana, R. A.; Taube, D. J.; Gamble, S.; Taube, H.; Satoh, T.; Fujii, H. Platinum Catalysts for the High-Yield Oxidation of Methane to a Methanol Derivative. *Science* **1998**, *280*, 560−564.
(25) Kushch, L.; Lavrushko, V.; Misharin, Y. S.; Moravsky, A.; Shilov, A. Kinetics and Mechanism of Methane Oxidation in Aqueous Solutions of Platinum Complexes. Direct Evidence for a Methylplatinum Intermediate. *Nouv. J. Chim.* **1983**, *7*, 729−733.
(26) Nandy, A.; Adamji, H.; Kastner, D. W.; Vennelakanti, V.; Nazemi, A.; Liu, M.; Kulik, H. J. Using Computational Chemistry To Reveal Nature's Blueprints for Single-Site Catalysis of C−H Activation. *ACS Catal.* **2022**, *12*, 9281−9306.
(27) Dydio, P.; Key, H. M.; Nazarenko, A.; Rha, J. Y.-E.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F. An Artificial Metalloenzyme with the Kinetics of Native Enzymes. *Science* **2016**, *354*, 102−106.
(28) Rohde, J.-U.; In, J.-H.; Lim, M. H.; Brennessel, W. W.; Bukowski, M. R.; Stubna, A.; Münck, E.; Nam, W.; Que, L. Crystallographic and Spectroscopic Characterization of a Nonheme Fe(IV)═O Complex. *Science* **2003**, *299*, 1037−1039.
(29) Cho, J.; Sarangi, R.; Nam, W. Mononuclear Metal−O_2 Complexes Bearing Macrocyclic N-Tetramethylated Cyclam Ligands. *Acc. Chem. Res.* **2012**, *45*, 1321−1330.
(30) Cho, K.; Leeladee, P.; McGown, A. J.; DeBeer, S.; Goldberg, D. P. A High-Valent Iron−Oxo Corrolazine Activates C−H Bonds via Hydrogen-Atom Transfer. *J. Am. Chem. Soc.* **2012**, *134*, 7392−7399.

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

(31) Collins, T. J.; Ryabov, A. D. Targeting of High-Valent Iron-TAML Activators at Hydrocarbons and Beyond. *Chem. Rev.* **2017**, *117*, 9140−9162.
(32) Tang, L. L.; Gunderson, W. A.; Weitz, A. C.; Hendrich, M. P.; Ryabov, A. D.; Collins, T. J. Activation of Dioxygen by a TAML Activator in Reverse Micelles: Characterization of an $Fe^{III}Fe^{IV}$ Dimer and Associated Catalytic Chemistry. *J. Am. Chem. Soc.* **2015**, *137*, 9704−9715.
(33) Lubben, M.; Meetsma, A.; Wilkinson, E. C.; Feringa, B.; Que, L. Jr. Nonheme Iron Centers in Oxygen Activation: Characterization of an Iron (III) Hydroperoxide Intermediate. *Angew. Chem., Int. Ed.* **1995**, *34*, 1512−1514.
(34) Menage, S.; Vincent, J. M.; Lambeaux, C.; Chottard, G.; Grand, A.; Fontecave, M. Alkane Oxidation Catalyzed by mu-Oxo-Bridged Diferric Complexes: a Structure/Reactivity Correlation Study. *Inorg. Chem.* **1993**, *32*, 4766−4773.
(35) Snyder, B. E. R.; Bols, M. L.; Schoonheydt, R. A.; Sels, B. F.; Solomon, E. I. Iron and Copper Active Sites in Zeolites and Their Correlation to Metalloenzymes. *Chem. Rev.* **2018**, *118*, 2718−2768.
(36) England, J.; Martinho, M.; Farquhar, E. R.; Frisch, J. R.; Bominaar, E. L.; Münck, E.; Que, L. Jr. A Synthetic High-Spin Oxoiron(IV) Complex: Generation, Spectroscopic Characterization, and Reactivity. *Angew. Chem., Int. Ed.* **2009**, *48*, 3622−3626.
(37) Kaizer, J.; Klinker, E. J.; Oh, N. Y.; Rohde, J.-U.; Song, W. J.; Stubna, A.; Kim, J.; Münck, E.; Nam, W.; Que, L. Nonheme $Fe^{IV}O$ Complexes That Can Oxidize the C−H Bonds of Cyclohexane at Room Temperature. *J. Am. Chem. Soc.* **2004**, *126*, 472−473.
(38) Biswas, A. N.; Puri, M.; Meier, K. K.; Oloo, W. N.; Rohde, G. T.; Bominaar, E. L.; Münck, E.; Que, L. Jr. Modeling TauD-J: A High-Spin Nonheme Oxoiron(IV) Complex with High Reactivity toward C−H Bonds. *J. Am. Chem. Soc.* **2015**, *137*, 2428−2431.
(39) Foscato, M.; Jensen, V. R. Automated in Silico Design of Homogeneous Catalysts. *ACS Catal.* **2020**, *10*, 2354−2377.
(40) Raugei, S.; DuBois, D. L.; Rousseau, R.; Chen, S.; Ho, M.-H.; Bullock, R. M.; Dupuis, M. Toward Molecular Catalysts by Computer. *Acc. Chem. Res.* **2015**, *48*, 248−255.
(41) Osadchii, D. Y.; Olivos-Suarez, A. I.; Szécsényi, Á.; Li, G.; Nasalevich, M. A.; Dugulan, I. A.; Crespo, P. S.; Hensen, E. J. M.; Veber, S. L.; Fedin, M. V.; Sankar, G.; Pidko, E. A.; Gascon, J. Isolated Fe Sites in Metal Organic Frameworks Catalyze the Direct Conversion of Methane to Methanol. *ACS Catal.* **2018**, *8*, 5542−5548.
(42) Adamji, H.; Nandy, A.; Kevlishvili, I.; Román-Leshkov, Y.; Kulik, H. J. Computational Discovery of Stable Metal−Organic Frameworks for Methane-to-Methanol Catalysis. *J. Am. Chem. Soc.* **2023**, *145*, 14365−14378.
(43) Verma, P.; Vogiatzis, K. D.; Planas, N.; Borycz, J.; Xiao, D. J.; Long, J. R.; Gagliardi, L.; Truhlar, D. G. Mechanism of Oxidation of Ethane to Ethanol at Iron(IV)−Oxo Sites in Magnesium-Diluted Fe_2(dobdc). *J. Am. Chem. Soc.* **2015**, *137*, 5770−5781.
(44) Song, W. J.; Seo, M. S.; DeBeer George, S.; Ohta, T.; Song, R.; Kang, M.-J.; Tosha, T.; Kitagawa, T.; Solomon, E. I.; Nam, W. Synthesis, Characterization, and Reactivities of Manganese(V)−Oxo Porphyrin Complexes. *J. Am. Chem. Soc.* **2007**, *129*, 1268−1277.
(45) Nandy, A.; Zhu, J.; Janet, J. P.; Duan, C.; Getman, R. B.; Kulik, H. J. Machine Learning Accelerates the Discovery of Design Rules and Exceptions in Stable Metal−Oxo Intermediate Formation. *ACS Catal.* **2019**, *9*, 8243−8255.
(46) Nandy, A.; Kulik, H. J. Why Conventional Design Rules for C−H Activation Fail for Open-Shell Transition-Metal Catalysts. *ACS Catal.* **2020**, *10*, 15033−15047.
(47) Gani, T. Z. H.; Kulik, H. J. Understanding and Breaking Scaling Relations in Single-Site Catalysis: Methane to Methanol Conversion by Fe^{IV}=O. *ACS Catal.* **2018**, *8*, 975−986.
(48) Adamji, H.; Kevlishvili, I.; Nandy, A.; Román-Leshkov, Y.; Kulik, H. J. Large-Scale Comparison of Fe and Ru Polyolefin C−H Activation Catalysts. *J. Catal.* **2024**, *431*, No. 115361.
(49) Shaik, S.; Chen, H.; Janardanan, D. Exchange-Enhanced Reactivity in Bond Activation by Metal−Oxo Enzymes and Synthetic Reagents. *Nature Chem.* **2011**, *3*, 19−27.
(50) Janardanan, D.; Wang, Y.; Schyman, P.; Que, L. Jr.; Shaik, S. The Fundamental Role of Exchange-Enhanced Reactivity in C-H Activation by S=2 Oxo Iron(IV) Complexes. *Angew. Chem., Int. Ed.* **2010**, *49*, 3342−3345.
(51) Gupta, R.; Lacy, D. C.; Bominaar, E. L.; Borovik, A. S.; Hendrich, M. P. Electron Paramagnetic Resonance and Mössbauer Spectroscopy and Density Functional Theory Analysis of a High-Spin Fe^{IV}−Oxo Complex. *J. Am. Chem. Soc.* **2012**, *134*, 9775−9784.
(52) Geng, C.; Ye, S.; Neese, F. Analysis of Reaction Channels for Alkane Hydroxylation by Nonheme Iron(IV)−Oxo Complexes. *Angew. Chem., Int. Ed.* **2010**, *49*, 5717−5720.
(53) Kupper, C.; Mondal, B.; Serrano-Plana, J.; Klawitter, I.; Neese, F.; Costas, M.; Ye, S.; Meyer, F. Nonclassical Single-State Reactivity of an Oxo-Iron(IV) Complex Confined to Triplet Pathways. *J. Am. Chem. Soc.* **2017**, *139*, 8939−8949.
(54) Singh, R.; Ganguly, G.; Malinkin, S. O.; Demeshko, S.; Meyer, F.; Nordlander, E.; Paine, T. K. A Mononuclear Nonheme Iron(IV)-Oxo Complex of a Substituted N4Py Ligand: Effect of Ligand Field on Oxygen Atom Transfer and C−H Bond Cleavage Reactivity. *Inorg. Chem.* **2019**, *58*, 1862−1876.
(55) Vitillo, J. G.; Lu, C. C.; Cramer, C. J.; Bhan, A.; Gagliardi, L. Influence of First and Second Coordination Environment on Structural Fe(II) Sites in MIL-101 for C−H Bond Activation in Methane. *ACS Catal.* **2021**, *11*, 579−589.
(56) Vogiatzis, K. D.; Polynski, M. V.; Kirkland, J. K.; Townsend, J.; Hashemi, A.; Liu, C.; Pidko, E. A. Computational Approach to Molecular Catalysis by 3d Transition Metals: Challenges and Opportunities. *Chem. Rev.* **2019**, *119*, 2453−2523.
(57) Harvey, J. N.; Poli, R.; Smith, K. M. Understanding the Reactivity of Transition Metal Complexes Involving Multiple Spin States. *Coord. Chem. Rev.* **2003**, *238-239*, 347−361.
(58) Bligaard, T.; Nørskov, J. K.; Dahl, S.; Matthiesen, J.; Christensen, C. H.; Sehested, J. The Brønsted−Evans−Polanyi Relation and the Volcano Curve in Heterogeneous Catalysis. *J. Catal.* **2004**, *224*, 206−217.
(59) Latimer, A. A.; Kulkarni, A. R.; Aljama, H.; Montoya, J. H.; Yoo, J. S.; Tsai, C.; Abild-Pedersen, F.; Studt, F.; Nørskov, J. K. Understanding Trends in C−H Bond Activation in Heterogeneous Catalysis. *Nature Mater.* **2017**, *16*, 225−229.
(60) Andrikopoulos, P. C.; Michel, C.; Chouzier, S.; Sautet, P. In Silico Screening of Iron-Oxo Catalysts for CH Bond Cleavage. *ACS Catal.* **2015**, *5*, 2490−2499.
(61) Schneider, J. E.; Goetz, M. K.; Anderson, J. S. Statistical Analysis of C−H Activation by Oxo Complexes Supports Diverse Thermodynamic Control over Reactivity. *Chem. Sci.* **2021**, *12*, 4173−4183.
(62) Abild-Pedersen, F.; Greeley, J.; Studt, F.; Rossmeisl, J.; Munter, T. R.; Moses, P. G.; Skúlason, E.; Bligaard, T.; Nørskov, J. K. Scaling Properties of Adsorption Energies for Hydrogen-Containing Molecules on Transition-Metal Surfaces. *Phys. Rev. Lett.* **2007**, *99*, No. 016105.
(63) Greeley, J. Theoretical Heterogeneous Catalysis: Scaling Relationships and Computational Catalyst Design. *Annu. Rev. Chem. Biomol. Eng.* **2016**, *7*, 605−635.
(64) Wodrich, M. D.; Busch, M.; Corminboeuf, C. Accessing and Predicting the Kinetic Profiles of Homogeneous Catalysts from Volcano Plots. *Chem. Sci.* **2016**, *7*, 5723−5735.
(65) Chantarojsiri, T.; Reath, A. H.; Yang, J. Y. Cationic Charges Leading to an Inverse Free-Energy Relationship for N−N Bond Formation by MnVI Nitrides. *Angew. Chem., Int. Ed.* **2018**, *57*, 14037−14042.
(66) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Structure-Activity Relationships That Identify Metal-Organic Framework Catalysts for Methane Activation. *ACS Catal.* **2019**, *9*, 3576−3587.
(67) Latimer, A. A.; Kakekhani, A.; Kulkarni, A. R.; Nørskov, J. K. Direct Methane to Methanol: The Selectivity−Conversion Limit and Design Strategies. *ACS Catal.* **2018**, *8*, 6894−6907.

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

(68) Calle-Vallejo, F.; Loffreda, D.; Koper, M. T. M.; Sautet, P. Introducing Structural Sensitivity into Adsorption-Energy Scaling Relations by Means of Coordination Numbers. *Nat. Chem.* **2015**, *7*, 403−410.
(69) Busch, M.; Wodrich, M. D.; Corminboeuf, C. Linear Scaling Relationships and Volcano Plots in Homogeneous Catalysis: Revisiting the Suzuki Reaction. *Chem. Sci.* **2015**, *6*, 6754−6761.
(70) Barona, M.; Ahn, S.; Morris, W.; Hoover, W.; Notestein, J. M.; Farha, O. K.; Snurr, R. Q. Computational Predictions and Experimental Validation of Alkane Oxidative Dehydrogenation by Fe_2M MOF Nodes. *ACS Catal.* **2020**, *10*, 1460−1469.
(71) Liao, P.; Getman, R. B.; Snurr, R. Q. Optimizing Open Iron Sites in Metal-Organic Frameworks for Ethane Oxidation: A First-Principles Study. *ACS Appl. Mater. Interfaces* **2017**, *9*, 33484−33492.
(72) Cordova, M.; Wodrich, M. D.; Meyer, B.; Sawatlon, B.; Corminboeuf, C. Data-Driven Advancement of Homogeneous Nickel Catalyst Activity for Aryl Ether Cleavage. *ACS Catal.* **2020**, *10*, 7021−7031.
(73) Pellizzeri, S.; Barona, M.; Bernales, V.; Miró, P.; Liao, P.; Gagliardi, L.; Snurr, R. Q.; Getman, R. B. Catalytic Descriptors and Electronic Properties of Single-Site Catalysts for Ethene Dimerization to 1-Butene. *Catal. Today* **2018**, *312*, 149−157.
(74) Khorshidi, A.; Violet, J.; Hashemi, J.; Peterson, A. A. How Strain Can Break the Scaling Relations of Catalysis. *Nat. Catal.* **2018**, *1*, 263−268.
(75) Szécsényi, Á.; Khramenkova, E.; Chernyshov, I. Y.; Li, G.; Gascon, J.; Pidko, E. A. Breaking Linear Scaling Relationships with Secondary Interactions in Confined Space: A Case Study of Methane Oxidation by Fe/ZSM-5 Zeolite. *ACS Catal.* **2019**, *9*, 9276−9284.
(76) Nandy, A.; Duan, C.; Taylor, M. G.; Liu, F.; Steeves, A. H.; Kulik, H. J. Computational Discovery of Transition-metal Complexes: From High-throughput Screening to Machine Learning. *Chem. Rev.* **2021**, *121*, 9927−10000.
(77) Friederich, P.; dos Passos Gomes, G.; De Bin, R.; Aspuru-Guzik, A.; Balcells, D. Machine Learning Dihydrogen Activation in the Chemical Space Surrounding Vaska's Complex. *Chem. Sci.* **2020**, *11*, 4584−4601.
(78) Li, X.; Chiong, R.; Page, A. J. Group and Period-Based Representations for Improved Machine Learning Prediction of Heterogeneous Alloy Catalysts. *J. Phys. Chem. Lett.* **2021**, *12*, 5156−5162.
(79) Janet, J. P.; Kulik, H. J. Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure−Property Relationships. *J. Phys. Chem. A* **2017**, *121*, 8939−8954.
(80) 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*, 10592−10606.
(81) Janet, J. P.; Gani, T. Z. H.; Steeves, A. H.; Ioannidis, E. I.; Kulik, H. J. Leveraging Cheminformatics Strategies for Inorganic Discovery: Application to Redox Potential Design. *Ind. Eng. Chem. Res.* **2017**, *56*, 4898−4910.
(82) Li, Z.; Omidvar, N.; Chin, W. S.; Robb, E.; Morris, A.; Achenie, L.; Xin, H. Machine-Learning Energy Gaps of Porphyrins with Molecular Graph Representations. *J. Phys. Chem. A* **2018**, *122*, 4571−4578.
(83) Coley, C. W. Defining and Exploring Chemical Spaces. *Trends Chem.* **2021**, *3*, 133−145.
(84) Keith, J. A.; Vassilev-Galindo, V.; Cheng, B.; Chmiela, S.; Gastegger, M.; Müller, K.-R.; Tkatchenko, A. Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems. *Chem. Rev.* **2021**, *121*, 9816−9872.
(85) Li, K.; Persaud, D.; Choudhary, K.; DeCost, B.; Greenwood, M.; Hattrick-Simpers, J. Exploiting Redundancy in Large Materials Datasets for Efficient Machine Learning with Less Data. *Nat. Comm.* **2023**, *14*, 7283.
(86) Liu, Y.; Yu, X. Farthest Point Sampling in Property Designated Chemical Feature Space as a General Strategy for Enhancing the Machine Learning Model Performance for Small Scale Chemical Dataset. *arXiv* **2024**.
(87) Hossain, S.; Thiagarajan, P.; Pathrudkar, S.; Taylor, S.; Gangan, A. S.; Banerjee, A. S.; Ghosh, S. Surprisingly High Redundancy in Electronic Structure Data. *arXiv* **2025**.
(88) Li, Q.; Fu, N.; Omee, S. S.; Hu, J. MD-HIT: Machine Learning for Material Property Prediction with Dataset Redundancy Control. *npj Comput. Mater.* **2024**, *10*, 245.
(89) Lookman, T.; Balachandran, P. V.; Xue, D.; Yuan, R. Active Learning in Materials Science with Emphasis on Adaptive Sampling using Uncertainties for Targeted Design. *npj Comput. Mater.* **2019**, *5*, 21.
(90) Häse, F.; Roch, L. M.; Kreisbeck, C.; Aspuru-Guzik, A. Phoenics: A Bayesian Optimizer for Chemistry. *ACS Cent. Sci.* **2018**, *4*, 1134−1145.
(91) Okamoto, Y. Applying Bayesian Approach to Combinatorial Problem in Chemistry. *J. Phys. Chem. A* **2017**, *121*, 3299−3304.
(92) Gubaev, K.; Podryabinkin, E. V.; Shapeev, A. V. Machine Learning of Molecular Properties: Locality and Active Learning. *J. Chem. Phys.* **2018**, *148*, No. 241727.
(93) Smith, J. S.; Nebgen, B.; Lubbers, N.; Isayev, O.; Roitberg, A. E. Less is More: Sampling Chemical Space with Active Learning. *J. Chem. Phys.* **2018**, *148*, No. 241733.
(94) Janet, J. P.; Ramesh, S.; Duan, C.; Kulik, H. J. Accurate Multiobjective Design in a Space of Millions of Transition Metal Complexes with Neural-Network-Driven Efficient Global Optimization. *ACS Cent. Sci.* **2020**, *6*, 513−524.
(95) Duan, C.; Nandy, A.; Terrones, G. G.; Kastner, D. W.; Kulik, H. J. Active Learning Exploration of Transition-Metal Complexes to Discover Method-Insensitive and Synthetically Accessible Chromophores. *JACS Au* **2023**, *3*, 391−401.
(96) Nandy, A.; Duan, C.; Goffinet, C.; Kulik, H. J. New Strategies for Direct Methane-to-Methanol Conversion from Active Learning Exploration of 16 Million Catalysts. *JACS Au* **2022**, *2*, 1200−1213.
(97) Žilinskas, A.; Mockus, J. On One Bayesian Method of Search of the Minimum. *Avtomatica i Vychislitel'naya Teknika* **1972**, *4*, 42−44.
(98) Močkus, J. *Optimization Techniques IFIP Technical Conference Novosibirsk*, **2013**, *6*, pp 400−404.
(99) Mockus, J. The Application of Bayesian Methods for Seeking the Extremum. *Towards Global Optimization* **1998**, *2*, 117−129.
(100) Emmerich, M.; Yang, K.; Deutz, A.; Wang, H.; Fonseca, C. M. *Advances in Stochastic and Deterministic Global Optimization*; Springer International Publishing: Switzerland, **2016**, 229−242.
(101) Groves, J. T.; McClusky, G. A. Aliphatic Hydroxylation via Oxygen Rebound. Oxygen Transfer Catalyzed by Iron. *J. Am. Chem. Soc.* **1976**, *98*, 859−861.
(102) Pannov, G. I.; Sobolev, V. I.; Kharitonov, A. S. The Role of Iron in N_2O Decomposition on ZSM-5 Zeolite and Reactivity of the Surface Oxygen Formed. *J. Mol. Catal.* **1990**, *61*, 85−97.
(103) Xiao, D. J.; Bloch, E. D.; Mason, J. A.; Queen, W. L.; Hudson, M. R.; Planas, N.; Borycz, J.; Dzubak, A. L.; Verma, P.; Lee, K.; Bonino, F.; Crocellà, V.; Yano, J.; Bordiga, S.; Truhlar, D. G.; Gagliardi, L.; Brown, C. M.; Long, J. R. Oxidation of Ethane to Ethanol by N_2O in a Metal−Organic Framework with Coordinatively Unsaturated Iron(II) Sites. *Nat. Chem.* **2014**, *6*, 590−595.
(104) Simons, M. C.; Vitillo, J. G.; Babucci, M.; Hoffman, A. S.; Boubnov, A.; Beauvais, M. L.; Chen, Z.; Cramer, C. J.; Chapman, K. W.; Bare, S. R.; Gates, B. C.; Lu, C. C.; Gagliardi, L.; Bhan, A. Structure, Dynamics, and Reactivity for Light Alkane Oxidation of Fe(II) Sites Situated in the Nodes of a Metal−Organic Framework. *J. Am. Chem. Soc.* **2019**, *141*, 18142−18151.
(105) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr., Sect. B:Struct. Sci., Cryst. Eng. Mater.* **2016**, *72*, 171−179.
(106) Weisfeiler, B.; Leman, A. The Reduction of a Graph to Canonical Form and the Algebra which Appears Therein. *Nauchno-Tekhn. Inform.* **1968**, *2*, 12−16.
(107) Hagberg, A. A.; Swart, P. J.; Schult, D. A. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of*

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

7th Python in Science Conference Varoquaux, G.; Vaught, T.; Millman, J., Eds.; Pasadena, CA, USA, **2008**, 11−15.
(108) St. Michel, R. G.; Jang, R. J.; Garrison, A. G.; Kevlishvili, I.; Kulik, H. J. The BOS-Lig Dataset: Accurate Ligand Charges from a Consensus Approach for 66,810 Experimentally Synthesized Ligands. *arXiv* **2026**, http://arxiv.org/abs/2604.06043.
(109) Emmerich, M. T.; Deutz, A. H.; Klinkenberg, J. W. Hypervolume-based expected improvement: Monotonicity properties and exact computation. *2011 IEEE Congress of Evolutionary Computation*, **2011**, 2147−2154.
(110) MacKay, D. J. A Practical Bayesian Framework for Backpropagation Networks. *Neural Comput.* **1992**, *4*, 448−472.
(111) Gawlikowski, J.; Tassi, C. R. N.; Ali, M.; Lee, J.; Humt, M.; Feng, J.; Kruspe, A.; Triebel, R.; Jung, P.; Roscher, R.; Shahzad, M.; Yang, W.; Bamler, R.; Zhu, X. X. A Survey of Uncertainty in Deep Neural Networks. *Artif. Intell. Rev.* **2023**, *56*, 1513−1589.
(112) Bigi, F.; Chong, S.; Ceriotti, M.; Grasselli, F. A Prediction Rigidity Formalism for Low-Cost Uncertainties in Trained Neural Networks. *Mach. Learn.: Sci. Technol.* **2024**, *5*, No. 045018.
(113) Chong, S.; Bigi, F.; Grasselli, F.; Loche, P.; Kellner, M.; Ceriotti, M. Prediction Rigidities for Data-Driven Chemistry. *Faraday Discussi.* **2025**, *256*, 322−344.
(114) Martens, J.; Grosse, R. Optimizing Neural Networks with Kronecker-factored Approximate Curvature. *Proceedings of the 32nd International Conference on Machine Learning* **2015**, 2408−2417.
(115) Lee, J.; Humt, M.; Feng, J.; Triebel, R. Estimating Model Uncertainty of Neural Networks in Sparse Information Form. *Proceedings of the 37th International Conference on Machine Learning* **2020**, 5702−5713.
(116) Kristiadi, A.; Hein, M.; Hennig, P. Learnable uncertainty under Laplace approximations. *Proceedings of the Thirty-Seventh Conference on Uncertainty in Artificial Intelligence*, **2021**, 344−353.
(117) Kellner, M.; Ceriotti, M. Uncertainty Quantification by Direct Propagation of Shallow Ensembles. *Mach. Learn.: Sci. Technol.* **2024**, *5*, No. 035006.
(118) Janet, J. P.; Duan, C.; Yang, T.; Nandy, A.; Kulik, H. J. A Quantitative Uncertainty Metric Controls Error in Neural Network-Driven Chemical Discovery. *Chem. Sci.* **2019**, *10*, 7913−7922.
(119) Baldi, P.; Nasr, R. When is Chemical Similarity Significant? The Statistical Distribution of Chemical Similarity Scores and its Extreme Values. *J. Chem. inf. Model.* **2010**, *50*, 1205−1222.
(120) Bajusz, D.; Rácz, A.; Héberger, K. Why is Tanimoto Index an Appropriate Choice for Fingerprint-Based Similarity Calculations? *J. Cheminform.* **2015**, *7*, 1−13.
(121) Comba, P.; Hambley, T. W.; Lawrance, G. A.; Martin, L. L.; Renold, P.; Váarnagy, K. Template Syntheses of Chiral Tetradentate Ligands Derived from L-Amino Acids. Structural and Spectroscopic Characterization of the Free Ligands and of their Copper(II) Complexes. *J. Chem. Soc., Dalton Trans.* **1991**, , 277−283.
(122) Siegfried, L.; Neuburger, M.; Zehnder, M.; Kaden, T. A. Structural Evidence for Carbonyl-Oxygen Protonation in the Copper(II) Complex of 1,4,8,11-Tetraaz-5-oxocyclotetradecane. *J. Chem. Soc., Chem. Commun.* **1994**, 951−952.
(123) Gajda, T.; Henry, B.; Aubry, A.; Delpuech, J.-J. Proton and Metal Ion Interactions with Glycylglycylhistamine, a Serum Albumin Mimicking Pseudopeptide. *Inorg. Chem.* **1996**, *35*, 586−593.
(124) Hashimoto, M.; Sakata, K.; Tanoue, Y. Crystal Structures and Spectral Properties of (2,3-*cis*- and 2,3-*trans*-Cyclohexano-1,4,8,11-Tetraazacyclotetradecane)Nickel(II) Complexes. *Polyhedron* **2001**, *20*, 409−413.
(125) Kobiro, K.; Nakayama, A.; Hiro, T.; Suwa, M.; Tobe, Y. Synthesis and Molecular Structures of Nickel(II) Alkyl-Substituted Cyclam Complexes. *Inorg. Chem.* **1992**, *31*, 676−685.
(126) Abba, F.; De Santis, G.; Fabbrizzi, L.; Licchelli, M.; Manotti Lanfredi, A. M.; Pallavicini, P.; Poggi, A.; Ugozzoli, F. Nickel(II) Complexes of Azacyclams: Oxidation and Reduction Behavior and Catalytic Effects in the Electroreduction of Carbon Dioxide. *Inorg. Chem.* **1994**, *33*, 1366−1375.
(127) Aneetha, H.; Lai, Y.-H.; Lin, S.-C.; Panneerselvam, K.; Lu, T.-H.; Chung, C.-S. Copper(II) Complexes of Tetraaza Macrocycles Bearing Pendant Arms: Syntheses, Structures and Properties. *J. Chem. Soc., Dalton Trans.* **1999**, *28*, 2885−2892.
(128) Hay, R. W.; Clifford, T.; Klein, J.; Lightfoot, P. Bibrachial Ligands from Sterically Protected Tetra-Azamacrocycles. Synthesis of Bis(N-2-Cyanoethyl)Tet-*a* and the Crystal Structure of its Copper(II) Complex. *Polyhedron* **1996**, *15*, 2315−2319.
(129) Dumas, S.; Lastra, E.; Hegedus, L. S. Synthesis and Metal Complexation Reactions of Bis-Dioxocyclams from Photochemical Reaction of Bis-Chromium Alkoxycarbene Complexes with Imidazolines. *J. Am. Chem. Soc.* **1995**, *117*, 3368−3379.
(130) DeRosa, F.; Bu, X.; Pohaku, K.; Ford, P. C. Synthesis and Luminescence Properties of Cr(III) Complexes with Cyclam-Type Ligands Having Pendant Chromophores, trans-$[Cr(L)Cl_2]Cl^1$. *Inorg. Chem.* **2005**, *44*, 4166−4174.
(131) DeRosa, F.; Bu, X.; Ford, P. C. Chromium(III) Complexes for Photochemical Nitric Oxide Generation from Coordinated Nitrite: Synthesis and Photochemistry of Macrocyclic Complexes with Pendant Chromophores, *trans*-$[Cr(L)(ONO)_2]BF_4$. *Inorg. Chem.* **2005**, *44*, 4157−4165.
(132) Wagner, H. K.; Ansmann, N.; Gentner, T.; Wadepohl, H.; Ballmann, J. The Multifaceted Palladium Chemistry of 2,2′-Diphosphinotolanes. *Organometallics* **2021**, *40*, 804−812.
(133) Curtis, N. F.; Wikaira, J.; Butcher, R. (3,5,7,7,12,14,14-Heptamethyl-1,4,8,11-Tetraazacyclotetradec-4-11-Diene)Nickel(II) Revisited and some (2,5,5,7,12,12,14-Heptamethyl-1,4,8,11-Tetraazacyclotetradecane)Nickel(II) Compounds. *Inorg. Chim. Acta* **2015**, *436*, 103−117.
(134) Groves, J. T.; Nemo, T. E.; Myers, R. S. Hydroxylation and Epoxidation Catalyzed by Iron-Porphine Complexes. Oxygen Transfer from Iodosylbenzene. *J. Am. Chem. Soc.* **1979**, *101*, 1032−1033.
(135) Nam, W.; Goh, Y. M.; Lee, Y. J.; Lim, M. H.; Kim, C. Biomimetic Alkane Hydroxylations by an Iron(III) Porphyrin Complex with H_2O_2 and by a High-Valent Iron(IV) Oxo Porphyrin Cation Radical Complex. *Inorg. Chem.* **1999**, *38*, 3238−3240.
(136) Nam, W. High-Valent Iron(IV)−Oxo Complexes of Heme and Non-Heme Ligands in Oxygenation Reactions. *Acc. Chem. Res.* **2007**, *40*, 522−531.
(137) Galonić Fujimori, D.; Barr, E. W.; Matthews, M. L.; Koch, G. M.; Yonce, J. R.; Walsh, C. T.; Bollinger, J. M. Jr.; Krebs, C.; Riggs-Gelasco, P. J. Spectroscopic Evidence for a High-Spin Br-Fe(IV)-Oxo Intermediate in the α-Ketoglutarate-Dependent Halogenase CytC3 from Streptomyces. *J. Am. Chem. Soc.* **2007**, *129*, 13408−13409.
(138) Krebs, C.; Galonić Fujimori, D.; Walsh, C. T.; Bollinger, J. M. Jr. Non-Heme Fe(IV)−Oxo Intermediates. *Acc. Chem. Res.* **2007**, *40*, 484−492.
(139) Wong, S. D.; Srnec, M.; Matthews, M. L.; Liu, L. V.; Kwak, Y.; Park, K.; Bell, C. B. III; Alp, E. E.; Zhao, J.; Yoda, Y.; Kitao, S.; Seto, M.; Krebs, C.; Bollinger, J. M.; Solomon, E. I. Elucidation of the Fe(IV)=O Intermediate in the Catalytic Cycle of the Halogenase SyrB2. *Nature* **2013**, *499*, 320−323.
(140) Coley, C. W.; Rogers, L.; Green, W. H.; Jensen, K. F. SCScore: Synthetic Complexity Learned from a Reaction Corpus. *J. Chem. Inf. Model.* **2018**, *58*, 252−261.
(141) 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*, 13973−13986.
(142) Kozuch, S.; Shaik, S. How to Conceptualize Catalytic Cycles? The Energetic Span Model. *Acc. Chem. Res.* **2011**, *44*, 101−110.
(143) Qiu, Z.; Neumann, C. N. Influence of Cage Effects in Directing the Outcome of C−X Bond Forming Reactions. *ACS Org. Inorg. Au* **2024**, *4*, 1−25.
(144) Snyder, B. E. R.; Bols, M. L.; Rhoda, H. M.; Plessers, D.; Schoonheydt, R. A.; Sels, B. F.; Solomon, E. I. Cage Effects Control the Mechanism of Methane Hydroxylation in Zeolites. *Science* **2021**, *373*, 327−331.

Downloaded from pubs.acs.org/acscatal/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

(145) Seritan, S.; Bannwarth, C.; Fales, B. S.; Hohenstein, E. G.; Isborn, C. M.; Kokkila-Schumacher, S. I. L.; Li, X.; Liu, F.; Luehr, N.; Snyder, J. W. Jr.; Song, C.; Titov, A. V.; Ufimtsev, I. S.; Wang, L.-P.; Martínez, T. J. TeraChem: A Graphical Processing Unit-Accelerated Electronic Structure Package for Large-Scale Ab Initio Molecular Dynamics. *WIREs Comput. Mol. Sci.* **2021**, *11*, No. e1494.
(146) Petachem. *TeraChem*, http://www.petachem.com. (accessed July 12, 2023).
(147) Becke, A. D. Density-Functional Thermochemistry III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(148) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785−789.
(149) 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. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(150) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, No. 154104.
(151) Becke, A. D.; Johnson, E. R. A Density-Functional Model of the Dispersion Interaction. *J. Chem. Phys.* **2005**, *123*, No. 154101.
(152) Wadt, W. R.; Hay, P. J. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for Main Group Elements Na to Bi. *J. Chem. Phys.* **1985**, *82*, 284−298.
(153) Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations: Potentials for the Transition Metal Atoms Sc to Hg. *J. Chem. Phys.* **1985**, *82*, 270−283.
(154) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. 6-31G* Basis Set for Atoms K Through Zn. *J. Chem. Phys.* **1998**, *109*, 1223−1229.
(155) Saunders, V. R.; Hillier, I. H. A "Level−Shifting" Method for Converging Closed Shell Hartree−Fock Wave Functions. *Int. J. Quantum Chem.* **1973**, *7*, 699−705.
(156) Wang, L.-P.; Song, C. Geometry Optimization Made Simple with Translation and Rotation Coordinates. *J. Chem. Phys.* **2016**, *144*, No. 214108.
(157) Ioannidis, E. I.; Gani, T. Z.; Kulik, H. J. molSimplify: A Toolkit for Automating Discovery in Inorganic Chemistry. *J. Comput. Chem.* **2016**, *37*, 2106−2117.
(158) Kulik Group. *molSimplify & molSimplify Automatic Design*. https://github.com/hjkgrp/molsimplify. (accessed July 25, 2022).
(159) Duan, C.; Janet, J. P.; Liu, F.; Nandy, A.; Kulik, H. J. Learning from Failure: Predicting Electronic Structure Calculation Outcomes with Machine Learning Models. *J. Chem. Theory Comput.* **2019**, *15*, 2331−2345.
(160) Neese, F. Software Update: The ORCA Program System—Version 6.0. *WIREs Comput. Mol. Sci.* **2025**, *15*, No. e70019.
(161) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA Quantum Chemistry Program Package. *J. Chem. Phys.* **2020**, *152*, No. 224108.
(162) Baker, J. An Algorithm for the Location of Transition States. *J. Comput. Chem.* **1986**, *7*, 385−395.
(163) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.
(164) Neese, F. Definition of Corresponding Orbitals and the Diradical Character in Broken Symmetry DFT Calculations on Spin Coupled Systems. *J. Phys. Chem. Solids* **2004**, *65*, 781−785.
(165) Chollet, F. K. *The Keras Blog*, https://keras.io. (accessed July 6, 2025).
(166) Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G. S.; Davis, A.; Dean, J.; Devin, M. Tensorflow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv* **2016**.
(167) O'Malley, T.; Bursztein, E.; Long, J.; Chollet, F.; Jin, H.; Invernizzi, L. *KerasTuner*. **2019**. https://github.com/keras-team/keras-tuner. (accessed July 6, 2025).
(168) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.

Downloaded from pubs.acs.org/acscacs/article-pdf/doi/10.1021/acscatal.6c03365/6890716/cs-2026-03365s.pdf by guest on 26 October 2026

https://doi.org/10.1021/acscatal.6c03365
ACS Catal. XXXX, XXX, XXX−XXX