ARTICLE

Statistical Analysis of C–H Activation by Oxo Complexes Supports Diverse Thermodynamic Control Over Reactivity†

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Joseph E. Schneider,[a] McKenna K. Goetz, [a] and John S. Anderson* [a]

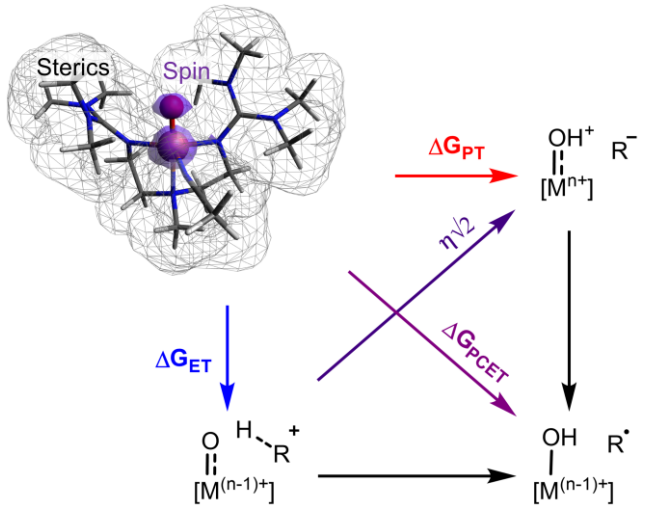
Transition metal oxo species are key intermediates for the activation of strong C–H bonds. As such, there has been interest in understanding which structural or electronic parameters of metal oxo complexes determine their reactivity. Factors such as ground state thermodynamics, spin state, steric environment, oxygen radical character, and asynchronicity have all been cited as key contributors, yet there is no consensus on when each of these parameters is significant or the relative magnitude of their effects. Herein, we present a thorough statistical analysis of parameters that have been proposed to influence transition metal oxo mediated C–H activation. We used density functional theory (DFT) to compute parameters for transition metal oxo complexes and analyzed their ability to explain and predict an extensive data set of experimentally determined reaction barriers. We found that, in general, only thermodynamic parameters play a statistically significant role. Notably, however, there are independent and significant contributions from the oxidation potential and basicity of the oxo complexes which suggest a more complicated thermodynamic picture than what has been shown previously.

Introduction

The activation of C–H bonds through proton-coupled electron transfer (PCET) underpins a wide range of biological and synthetic processes. The applications of this reaction include drug metabolism by cytochrome P450 enzymes as well as synthetic methods for the preparation of fine chemicals.[1–3] In many cases PCET reactions are mediated by transition metal oxo intermediates generated within either protein-based or synthetic ligand scaffolds. In these reactions, both a proton and an electron are transferred to an oxo complex resulting in the net removal of a hydrogen atom from the organic substrate. The generality of this reaction combined with the ubiquity of C–H bonds in synthesis has led to considerable interest in determining what properties govern the PCET reactivity of transition metal oxo species.

A large body of work supports that the free energy of reaction (ΔG_{PCET}) is central to transition metal oxo mediated C–H activation and also offers a great deal of explanatory and predictive power.[4–7] Recently however, additional properties have been cited as important although it is not clear if any have a widespread effect on reactivity. Individual cases support the influence of O-centered spin density,[8] spin state,[9–11] steric environment,[12–14] the free energies of proton and electron transfer (ΔG_{PT} and ΔG_{ET}),[15–20] or the asynchronicity (η) of the reaction,[21–24] but there is a lack of consensus regarding their generality and relative importance (Scheme 1).[8,25–27] Very few studies have explored these parameters outside of a narrow range of complexes,[4,6,10,20,21,28] and none have statistically examined the significance of parameters other than ΔG_{PCET} on the reactivity of a broad set of metal oxo complexes.

We previously found an atypical dependence on ΔG_{PT} in the concerted C–H activation reactivity of a terminal Co^{III} oxo complex which contrasts with the expected rate dependence on ΔG_{PCET}.[15] Given the disparity of this result with the literature, we sought to understand the interplay of characteristics affecting a broad range of transition metal oxo mediated PCET reactions using multivariable linear free energy relationships (LFERs). These models can be used to relate experimentally



Scheme 1 Investigated Parameters of Metal Oxo Species.

[a] *Department of Chemistry, University of Chicago, Chicago, IL 60637, USA. E-mail: jsanderson@uchicago.edu*

† Electronic Supplementary Information (ESI) available: Detailed summary of LFER models; additional methodological details; further analysis of hydrogen bonding, steric effects, and reorganization effects; spreadsheets containing experimental and computational data; optimized geometries and the output of frequency calculations; and python scripts used to perform the regression. See DOI: 10.1039/x0xx00000x

determined data, such as reaction rates, to multiple predictor variables simultaneously. LFER models have recently been used as versatile tools to optimize organic methodology, predict reaction barrier heights, and investigate underlying mechanisms.[29–33]

We have applied this analysis to examine trends in rates of PCET mediated C–H activation for a broad dataset of previously reported metal oxo complexes. This analysis enables a statistical examination of several hypotheses regarding what parameters of metal oxo species determine their PCET reactivity. Unsurprisingly, we observe that ΔG_{PCET} is the most important factor. However, we also observe a significant role for ΔG_{PT} and ΔG_{ET} beyond and independent of their contribution to ΔG_{PCET}. Furthermore, the other parameters investigated do not have broad significance. These results suggest that thermodynamic factors are generally the dominant contributors to transition metal oxo C–H activation reactivity, but also demonstrate that thermodynamic parameters beyond the commonly invoked ΔG_{PCET} are influential.

Results and Discussion

We began our analysis by compiling an extensive data set of second order rate constants (k_2) for the oxidation of 9,10-dihydroanthracene (DHA) by thirty well-characterized metal oxo complexes.[12,13,15,17,34–51] This diverse data set comprises complexes of five different transition metals, tetragonal and trigonal geometries, spin multiplicities from 1 to 5, charges from −1 to +3, and d-electron counts from 0 to 6. We manually divided the data set into a training set of seventeen metal oxo complexes and a test set of thirteen metal oxo complexes such that each set had a diverse mix of species. For our analysis, we calculated the values of parameters that have been hypothesized as important to metal oxo mediated PCET reactivity for each metal oxo complex. The investigated parameters include the steric environment (quantified by percent buried volume, %BV),[52] the spin density on the oxygen atom (via intrinsic bond order (IBO) analysis),[53,54] the energetic cost of accessing a higher spin state if one lies closer to the product spin multiplicity than the reactant ground spin state (spin excitation), the thermodynamic free energies ΔG_{PCET}, ΔG_{PT}, and ΔG_{ET}, and the magnitude of the asynchronicity parameter ($|\eta|$).[21] The detailed approaches used to determine the values for these parameters are provided in the methods section. Importantly, none of these parameters require transition state optimizations which are relatively difficult and less reliable than ground state calculations.[55,56] While this precludes a direct estimation of transition state effects, we expect to indirectly capture some of them. For instance, parameters such as ΔG_{PCET} and $|\eta|$ have been shown to correlate with tunneling.[23,57]

We examined the effect of each of these parameters on experimental reaction barriers by building multivariable free energy models via ordinary least squares regression of the barrier heights against the parameters. Each model consists of a set of coefficients (with variable units such that the product with the respective parameter gives units of kcal/mol) and an intercept (with units kcal/mol). These models were used to generate predicted reaction barriers for each data point, which could be compared with experimental reaction barriers to assess the utility of the model. Because ΔG_{PCET} has strong theoretical and experimental support for affecting reaction barrier heights,[4–6] we analyzed each parameter in combination with ΔG_{PCET} and compared the resulting model to regression against ΔG_{PCET} alone.

We evaluated each regression based on R^2, leave-one-out (LOO) R^2 (sometimes referred to as Q^2), and a statistical F-test.[58–61] R^2 is a goodness of fit measure which quantifies the amount of variation explained by a model. The predictive ability of a model is gauged with LOO R^2, in which each data point is left out and predicted by the remaining data points and the goodness of fit is then reevaluated. Critically, unlike regular R^2 this metric does not necessarily improve with an increase in parameters; overfitted models with too many parameters perform poorly with LOO R^2. For each R^2, a value close to 1 indicates a good fit. Finally, we report the *p*-value from an F-test on each model, which shows the probability the observed correlation arises from statistical noise. The lower this *p*-value is, the more significant a given parameter. Additionally, the calculation of *p*-values considers the number of parameters added to a model, so, as with LOO R^2, an F-test is not biased in favor of adding more parameters.

A summary of our findings is presented in Table 1. In line with previous reports, we find a strong correlation between the experimental reaction barriers and ΔG_{PCET}. This parameter alone explains 70% of the variation in reaction barriers within the training set ($R^2 = 0.70$) and has high predictive ability (LOO $R^2 = 0.60$). Interestingly, most other parameters do not significantly improve the model. While we do observe a small correlation with %BV steric metrics, the magnitude of the effect is too small to be statistically significant. Compared to the ΔG_{PCET} only model, spin-based parameters and $|\eta|$ barely improve R^2 and perform similarly or worse in LOO cross-validation. While it is difficult to rule out the importance of these parameters in individual cases, an F-test indicates they do not have a statistically significant effect across our entire data set.

In contrast, addition of ΔG_{PT} and ΔG_{ET} does significantly improve the fit. For this {ΔG_{PCET}, ΔG_{PT}, ΔG_{ET}} model, R^2 increases from 0.70 to 0.86 and LOO R^2 increases from 0.60 to 0.71, indicating both better explanation of the available data and better predictive ability. An F-test gives $p < 0.01$ which suggests the observed effect is statistically significant. The equation from this fit is $\Delta G^{\ddagger} = 0.31\ \Delta G_{PCET} + 0.07\ \Delta G_{PT} + 0.12\ \Delta G_{ET} - 0.26$ (all coefficients unitless; free energies and intercept in kcal/mol). Typically, ΔG_{PCET} is a negative value while ΔG_{PT} and ΔG_{ET} are positive values. Thus, the positive sign of the ΔG_{PCET} coefficient indicates that a more exergonic reaction will have a lower barrier while increases in ΔG_{PT} and ΔG_{ET} will raise the barrier. The larger coefficient of ΔG_{PCET} indicates the reaction barrier is most sensitive to this free energy. Satisfyingly, the ΔG_{PCET} coefficient agrees with experimental data: for metal oxo complexes that have a demonstrated trend of $\log(k_{obs})$ vs. substrate bond dissociation free energy (BDFE), the average slope of $\Delta G^{\ddagger}$ vs. substrate BDFE is ~0.3 (see Table S1), very similar to the 0.31 observed in our analysis. The intercept of –

Table 1 Statistical Results of Various Models.

Parameter(s) Regressed with ΔG_{PCET}	Training Set on DHA[a] R^2	Training Set on DHA[a] LOO[c] R^2	Training Set on DHA[a] *p*-value[d]	All Data for Multiple Substrates[b] R^2	All Data for Multiple Substrates[b] LOO[e] R^2
ΔG_{PCET} only	0.70	0.60	< 0.001[f]	0.45	0.36
%BV Steric Metrics	0.77	0.64	0.15	0.48	0.28
Oxo Spin Density	0.70	0.55	0.78	0.53	0.37
Spin Excitation	0.71	0.50	0.49	0.50	0.39
$\lvert\eta\rvert$	0.73	0.53	0.22	0.50	0.30
ΔG_{PT}, ΔG_{ET}	0.86	0.71	0.0082 0.023[g] 0.0038[h]	0.64	0.50

[a]A subset of the reactions of 17 metal oxo complexes with DHA. [b]Excluding outlier metal oxo complexes (Ru oxos and oxo complexes of 13-TMC); substrates are DHA, 1,4-cyclohexadiene, xanthene, and fluorene. [c]Leave-One-Out. [d]From an F-test where the null hypothesis is that only ΔG_{PCET} has an effect. [e]Leave-One-Out, slightly modified such that all reactions for a given metal oxo are left out together. [f]From an F-test where the null hypothesis is that ΔG_{PCET} has no effect. [g]From an F-test where the null hypothesis is that ΔG_{PT} has no effect. [h]From an F-test where the null hypothesis is that ΔG_{ET} has no effect.

0.26 contains contributions to the average error not accounted for by the three free energies.

The significance of ΔG_{PT} and ΔG_{ET} is intriguing because the literature discussion of these values has often been framed in terms of how they contribute to ΔG_{PCET} rather than in terms of their intrinsic contribution to reaction barrier heights.[16–19,27] However, ΔG_{PT} and ΔG_{ET} as defined here are the energies to form the intermediates involved in stepwise reactivity – the protonated metal oxo with the deprotonated substrate, or the reduced metal oxo with the oxidized substrate (Scheme 1). Critically, ΔG_{PT} and ΔG_{ET} do not form a full thermodynamic cycle with ΔG_{PCET} and thus are fundamentally distinct. This fact is statistically supported by poor correlations between ΔG_{PCET} and ΔG_{PT} and between ΔG_{PCET} and ΔG_{ET} (–0.12 and 0.31, respectively, see Regression S6). Finally, we find that ΔG_{PT} and ΔG_{ET} have importance independent of a contribution to ΔG_{PCET} as clearly demonstrated by the LOO R^2s and F-tests. All of our analyses therefore suggest that the combination of ΔG_{PT} and ΔG_{ET} is an independent and significant contributor to C–H activation barrier heights.

While the observation of a dependence on ΔG_{PT} and ΔG_{ET} that arises from our linear regressions is principally empirical, it is consistent with prior theoretical models in the literature. The physical underpinning of this dependency on ΔG_{PT} and ΔG_{ET} is likely due to mixing of proton transfer and electron transfer intermediates into the concerted transition state despite these intermediates never being fully realized.[62–64] Within transition state theory, this can be envisioned by a More-O'Ferrall-Jencks plot in which the transition state lies not on a one-dimensional line connecting reactant and product but on a two-dimensional plane containing reactant, product, and intermediates.[65,66] In this case, the intermediates arise from proton transfer and electron transfer, and when either ΔG_{PT} or ΔG_{ET} lowers in energy the transition state character can adopt structural and electronic components of these intermediates resulting in a lower barrier height. While the use of these classical structure-energy relationships to analyse PCET reactions has been questioned recently,[26,27] proton transfer and electron transfer states and their energies also have roles in nonadiabatic rate theories of PCET which treat proton transfer in a quantum mechanical fashion.[67,68] Therefore, the use of ΔG_{PT} and ΔG_{ET} to predict barrier heights of PCET is consistent with prior theoretical foundations.

Assigning a direct role for ΔG_{PT} and ΔG_{ET} is in line with recent computational studies of PCET transition states which invoke off-ΔG_{PCET} diagonal thermodynamic terms from Scheme 1, such as asynchronicity (η), as key contributors to DFT derived reaction barriers.[21–24] Asynchronicity is derived not from the sum of ΔG_{PT} and ΔG_{ET}, but rather their difference. Conversely, we instead find that the sum of ΔG_{PT} and ΔG_{ET} have a more significant effect than $|\eta|$. The reason for this discrepancy is unclear, but a possible explanation is that experimental noise prevents us from observing a comparatively more subtle trend between $|\eta|$ and the experimental reaction barrier heights. Furthermore, the well-controlled nature of the series of complexes previously investigated for asynchronicity may have too little variation in (ΔG_{PT} + ΔG_{ET}) to manifest similarly to the effects we observe here.

Another way in which our data may not be amenable to investigating the effect of $|\eta|$ is the variable reorganization energy of the metal oxo complexes examined here. $|\eta|$ is specifically framed as an adjustment to the Marcus reorganization energy;[21] therefore $|\eta|$'s effect may only be clear when reorganization is properly accounted for. While it is clear that the reorganization energy is important to PCET reactivity, there is no established way to compute it without computationally expensive transition state geometries.[11,67–69] We have made multiple attempts to derive reorganization or deformation parameters using the optimized metal oxo and metal hydroxide geometries and frequencies, but none of these parameters have statistically significant contributions to predicted reaction barriers with or without $|\eta|$ (see ESI). Therefore, a combination of noise in the experimental data and our inability to compute a reliable reorganization parameter could preclude us from observing an effect of $|\eta|$ on the barrier heights. Nonetheless, previous studies as well as this current work offer increasing support that off-ΔG_{PCET} diagonal thermodynamic terms such as ΔG_{PT} and ΔG_{ET} have important effects on reactivity independent of ΔG_{PCET}.

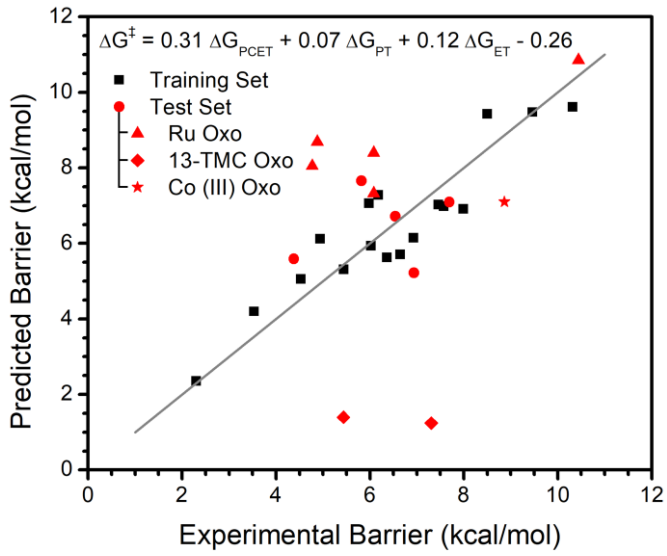


Fig. 1 Regression analysis of the experimental reaction barrier vs. ΔG_{PCET}, ΔG_{PT}, and ΔG_{ET} for various metal oxo complexes reacting with DHA. Specific cases discussed in the main text are given unique symbols. The grey line marks where predicted equals experimental. Predicted values are computed using the given formula.

The all-thermodynamic model we find here provides insights and possible alternative explanations for previously reported trends in PCET reactivity. In one study,[12] steric and spin state effects were invoked to explain the comparatively high reactivity of the S = 2 complex $[Fe^{IV}(O)(TMG_2dien)(CH_3CN)]^{2+}$. A higher rate of C–H activation as compared to S = 2 $[Fe^{IV}(O)(TMG_3tren)]^{2+}$ was ascribed to reduced steric hinderance in the TMG_2dien complex,[13] and the higher rate of C–H activation as compared to the S = 1 complexes $[Fe^{IV}(O)(N4Py)]^{2+}$ and $[Fe^{IV}(O)(TMC)(CH_3CN)]^{2+}$ was ascribed to the S = 2 spin state in the TMG_2dien complex.[36,44] However, it was noted that the even faster reactivity of $[Fe^{IV}(O)(Me_3NTB,CH_3CN)]^{2+}$, which is S = 1 and has a similar %BV profile to $[Fe^{IV}(O)(TMG_2dien)(CH_3CN)]^{2+}$,[34] is not easily explained by either hypothesis. Our analysis suggests that the thermodynamic properties of these complexes may provide an alternative explanation in these comparisons (see Table S9). The Me_3NTB complex has by far the most exergonic reaction with DHA (ΔG_{PCET} = –16 kcal/mol), followed by the TMG_2dien complex (ΔG_{PCET} = –9 kcal/mol), followed by the complexes of TMG_3tren, TMC, and N4Py (ΔG_{PCET} = –7, –6, and –6 kcal/mol, respectively). Thus, thermodynamic parameters would predict the Me_3NTB complex to have the lowest reaction barrier and fastest rate of reaction, with the TMG_2dien complex being the next most reactive, and the remaining complexes the least reactive as is observed experimentally.

In another study, it was observed that the rates of PCET reactions performed by $[Fe^{IV}(O)(TMC)(X)]^{n+}$ decrease with more strongly donating axial ligands X.[36] Variation in ΔG_{PCET} was ruled out as a cause of this trend, as it was calculated to be similar for all complexes investigated. It was suggested that the accessibility of a high-spin state may explain this variation in the rates, as the energy of the quintet excited state decreased with stronger X ligands. However, our calculations indicate that while stronger axial donors increase ΔG_{ET}, ΔG_{PT} decreases more substantially (see Table S9). In our model, these changes result in a net decrease in the reaction barrier, suggesting that despite a similar ΔG_{PCET}, the reactivity trend could be explained by thermodynamic effects. These analyses do not rule out that spin state or steric effects may be important in the previous studies, but suggest that thermodynamics may also play an important role.

The fit of the training data to {ΔG_{PCET}, ΔG_{PT}, ΔG_{ET}} and this model's performance on the test set is depicted graphically in Figure 1. It is clear that the reaction barriers for most metal oxo complexes in the test set are well predicted. Nonetheless, several metal oxo complexes (given unique symbols in Figure 1) deserve further discussion.

The {ΔG_{PCET}, ΔG_{PT}, ΔG_{ET}} model behaves the most poorly in predicting reaction barriers for the Fe^{IV} oxo and Co^{IV} oxo complexes of the ligand 13-TMC.[49,70] Essentially no barrier is predicted for these reactions, which is not observed experimentally. This is due to a large negative calculated ΔG_{PCET} in both cases; in fact, these complexes are outliers even in the ΔG_{PCET} only fit (see Regression S1). The cause of this discrepancy is not entirely clear. However, it appears to be systemic to the particular ligand scaffold rather than the identity of the metal center, which suggests this discrepancy could arise from ambiguity in the primary coordination sphere of these complexes. No structural characterization is reported for the Fe^{IV} complex, and while a short Co–O bond is identified by EXAFS for the Co^{IV} complex, it is difficult to conclusively determine the primary coordination sphere. Any discrepancy in coordination sphere would render our calculated parameters incorrect, potentially explaining their inability to predict the experimental reaction barriers.

The reaction barrier is overestimated for all Ru oxo complexes, and for three of them by more than two kcal/mol. As Ru is the only second row transition metal in our data set, we suspect this overestimation is due to a consistent difference between first and second row transition metals rather than Ru examples not following the same trends. For instance, it is possible that the Ru oxo complexes have relatively low structural reorganization energy or that relativistic effects influence the coefficients. It may also simply be a change in the systemic DFT error upon going to the second row. Regardless, regression of barriers from the kinetics of an individual Ru oxo complex reacting with several different substrates reveals there is a trend with ΔG_{PCET}, ΔG_{PT}, and ΔG_{ET} with similar coefficients to those obtained from the more general model with multiple different oxo complexes (see ESI). This supports that the same trends in free energies are at play in the Ru oxo complexes.

Interestingly, the {ΔG_{PCET}, ΔG_{PT}, ΔG_{ET}} model only moderately underestimates the reaction barrier (by ~2 kcal/mol) for a terminal Co^{III} oxo complex which has unusual trends in its reactivity with various substrates.[15] Unlike most metal oxo complexes, the reactivity of this complex does not have a clear trend with ΔG_{PCET}; its kinetics are instead dominated by ΔG_{PT}. Therefore, its adherence to trends in {ΔG_{PCET}, ΔG_{PT}, ΔG_{ET}} as seen for the broad set of metal oxo complexes deserves further investigation. We regressed the experimental reaction barriers for the reactivity of this complex

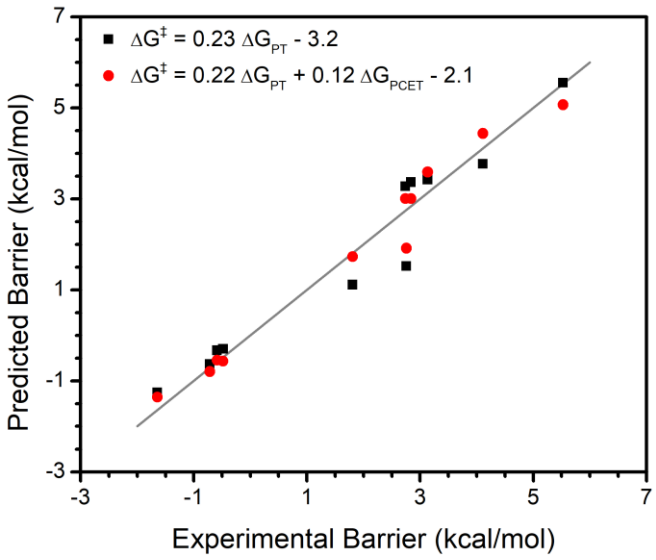


Fig. 2 Regression analysis for a Co^{III} oxo. The $\{\Delta G_{PT}\}$ fit is shown in black with predicted values computed using the indicated formula; $R^2 = 0.94$ and LOO $R^2 = 0.93$. The $\{\Delta G_{PT}, \Delta G_{PCET}\}$ fit is shown in red with predicted values computed using the indicated formula; $R^2 = 0.97$ and LOO $R^2 = 0.95$. The negative barriers are due to overestimation of the entropy of association. The grey line marks where the predicted barrier matches the experimentally determined barrier.

with several substrates against only ΔG_{PT} as well as against $\{\Delta G_{PT}, \Delta G_{PCET}\}$ (Figure 2). We find that the inclusion of ΔG_{PCET} significantly improves the model, increasing R^2 from 0.94 to 0.97 and LOO R^2 from 0.93 to 0.95 and having an F-test *p*-value of 0.02 (see ESI). However, the relative weighting of the contribution from ΔG_{PCET} is quite different than for the broader set of complexes.

In the broader set we observe that ΔG_{PCET} has a larger effect on the reaction barriers than either ΔG_{PT} or ΔG_{ET}, which is reflected in the larger coefficient for the ΔG_{PCET} term than for the ΔG_{PT} and ΔG_{ET} terms in the fit equation (Figure 1). In contrast, ΔG_{PT} has a greater effect than ΔG_{PCET} on the reaction barriers for the Co^{III} oxo complex, again reflected in the magnitude of their coefficients (Figure 2). Furthermore, the addition of ΔG_{ET} significantly improves the model for the broader set of metal oxo complexes (Table 1) but is insignificant for the series of substrates reacting with the Co^{III} oxo complex (*p*-value > 0.05, see Regressions S42 and S43). Overall, this Co^{III} oxo complex is not so dissimilar from the broader set of metal oxo complexes in that the same thermodynamic free energies explain the reactivity of both. However, this individual case demonstrates a different weighting of parameters than that observed in the broad set.

Our analysis of the Co^{III} reactivity rests on the assumption that the coefficients of the model do not change appreciably from substrate to substrate. To test this assumption, we extended our analysis of the larger set of metal oxo complexes to include reactivity with 1,4-cyclohexadiene (CHD), fluorene, and xanthene in addition to DHA. We refit the model with reported data for the reactions between each substrate and all metal oxo complexes (excluding the previously discussed Ru and 13-TMC oxo complexes). As with our regressions for DHA alone, the inclusion of ΔG_{PT} and ΔG_{ET} notably improves the fit (Table 1, Figure 3). Other parameters offer comparably little improvement to the fit and do not perform well by LOO cross validation. The equation for this model is $\Delta G^{\ddagger} = 0.23\ \Delta G_{PCET} + 0.04\ \Delta G_{PT} + 0.10\ \Delta G_{ET} + 2.10$, which is satisfyingly similar to the equation of the fit to DHA data alone, supporting the assumption that the coefficients of the model are not appreciably affected by the identity of the substrate.

Conclusions

Overall, this thorough analysis of the reported C–H activation reactivity of transition metal oxo complexes demonstrates that ΔG_{PCET}, ΔG_{PT}, and ΔG_{ET} have a statistically significant correlation with the reaction barrier. Interestingly, no other parameters examined here, including steric environment and spin-based parameters, provide a significant improvement to a ΔG_{PCET} only model. This is in contrast to previous literature reports which implicate such factors in explaining metal oxo mediated PCET. The $\{\Delta G_{PCET}, \Delta G_{PT}, \Delta G_{ET}\}$ model predicts all but five of the reaction barrier heights for reactivity with DHA within 2 kcal/mol, and predicts most of these barrier heights within 1 kcal/mol. We find that the asynchronous reactivity of a Co^{III} oxo complex is also well-predicted, although a fit to just this complex alone reveals changes in the relative importance of ΔG_{PCET}, ΔG_{PT}, and ΔG_{ET}. The structural and electronic bases for such variation, and whether concomitant changes in selectivity can be leveraged, are exciting avenues for future research.

While the relative importance of these thermodynamic parameters can vary between specific cases, this study on a broad set of metal oxo complexes suggests that thermodynamic parameters provide the most general contribution to reaction barriers. Furthermore, while a strong dependence on ΔG_{PCET} is observed, as is expected based on literature precedent, significant and independent contributions from ΔG_{PT} and ΔG_{ET} are observed. This conclusion adds to the growing body of

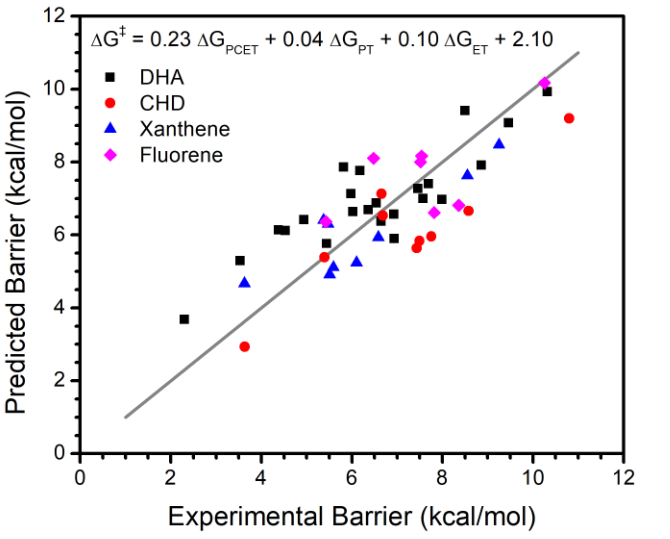


Fig. 3 Regression analysis of the experimental reaction barrier to PCET mediated C–H activation vs. ΔG_{PCET}, ΔG_{PT}, and ΔG_{ET} for all non-outlier metal oxo complexes reacting with DHA, CHD, xanthene, and fluorene. The grey line marks where predicted equals experimental. Predicted values are computed using the given formula.

literature supporting the importance of thermodynamic parameters beyond ΔG_{PCET}.

Methods

Tabulation of Experimental Kinetics

In this study we used thirty reported k_2 values of metal oxo species reacting with DHA.[12,13,15,17,34–51] For each of these oxo complexes, we tabulated various descriptors (metal, valency, d-count, coordination number, etc.), experimental parameters (M–O bond length and vibrational frequency, BDFE, pK_a, etc.), and the reported kinetics for reactions with various substrates (k_2, and, if reported, the statistical correction to this k_2, experimental $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$, and the KIE) along with the conditions these data were reported in (temperature and solvent). A full tabulation is found in the SI data folder. We excluded a few metal oxo species from our analysis despite having reported kinetic data for reactivity with DHA. The reasons for these exclusions were varied: several did not have a well-defined primary coordination sphere,[71–74] we were unable to calculate the reduced form of Mn^{III} oxo complexes without deprotonation of the hydrogen-bonding ligands,[17,20] one V^{V} oxo has too much experimental uncertainty in its k_2 value,[75] corrolazine complexes were too large to calculate their vibrational frequencies using our methods,[16,76] vibrational frequencies did not converge for the reduced form of two Ru^{VI} dioxo complexes,[50] we did not include third row complexes or complexes with ligand radicals,[37,77–79] and in one case saturation was reported at higher concentrations of DHA.[80] We also found several reports of metal oxo mediated C–H activation of substrates besides DHA[14,39,70,81–94] and useful reviews.[95,96]

All rate constants utilized here were reported as k_2 values with the exception of several rate constants used in the Co^{III} oxo reactivity analysis.[15] In this case, for substrates which did not have a reported k_2, the pseudo-first order rate constant k_{obs} at 0.0125 M of substrate was divided by 0.0125 M to obtain an approximate k_2. We used all substrates with reported kinetic data in this analysis except for 1,1,3,3-tetraphenylpropene. This substrate reacts unusually slowly, which we believe to be due to large steric hindrance of the reacting C–H bond. The remaining substrates were sterically similar enough that there is no steric effect on their kinetics (see Regression S41).

Determination of Experimental Barrier Heights

Before determining barrier heights from experimental k_2 values, we first multiplied each k_2 by any reported stochiometric and statistical adjustments so as to start from consistently unadjusted k_2 rate constants (experimental k_2 rate constants are often reported with statistical corrections to facilitate comparisons between substrates, either for the stoichiometry of the substrate's reactivity or for the number of benzylic C–H bonds). We assume that where no adjustment is noted in a paper, none has been made. The barriers of PCET reactivity were then determined from the unadjusted experimental k_2 values by solving the Eyring equation[97] and subtracting approximate expressions for the free energy of metal oxo-substrate association:

$$\Delta G^{\ddagger}_{PCET} = RT \ln\left(\frac{k_2 h}{n_{CH} n_O RT}\right) - RT\left[\ln\left(\left(\frac{2\pi\,\mu\,RT}{h^2}\right)^{\frac{3}{2}}\frac{1}{C^{\circ}}\right)+\frac{5}{2}\right] \quad (1)$$

where h is Planck's constant, n_{CH} is 2 for DHA and CHD and 1 for fluorene and xanthene, n_O is the number of oxo ligands in the metal oxo complex, RT is the thermal energy, μ is the reduced mass of the metal oxo and the substrate, and C° is 1 M. Our n_{CH} adjusts for the typical stoichiometry of each substrate's oxidation (DHA and CHD tend to lose two H-atoms, fluorene and xanthene one H-atom); we do not adjust for the number of reactive H-atoms, as ring puckering of the substrates means that not all reactive C–H bonds are equivalent. For instance, DHA has four benzylic C–H bonds, with two lying equatorial to the central ring and two lying axial. It is unlikely that the equatorial and axial positions are equally reactive, and it is entirely possible that reactivity predominantly occurs at only one of the positions. It is therefore not necessarily true that DHA is four times as reactive as an otherwise similar substrate with only one benzylic C–H bond.

The second and third terms in Equation 1 are an approximation for the free energy of association of the metal oxo and the substrate.[97] This adjustment allows us to compare kinetic data collected at different temperatures. As C–H bonds are poor hydrogen bond donors, we assume that the cost of association is purely entropic (or at least that enthalpic components vary minimally between different metal oxo complexes and substrates) and further assume this entropy cost is solely the loss of translational entropy. This neglects the loss of rotational entropy and the gain of low frequency metal oxo-substrate vibrational modes, but these effects will partially cancel. Regression with ΔG_{PCET} and RT does not fit DHA reaction barrier heights significantly better than a fit to ΔG_{PCET} alone (see Regression S10), indicating that this adjustment satisfactorily accounts for the temperature dependence of the reaction barrier.

We do not take into account hydrogen bonding between the metal oxo complexes and protic solvents as we were unable to derive a suitably accurate correction. However, in the SI we demonstrate that our best attempt to do so does not change the main conclusions herein (see Table S11).[28,98–100]

Calculation of Parameters

For each of these oxo complexes we calculated the values of parameters that have been proposed to influence metal oxo mediated PCET reactivity. Specifically, the parameters investigated were %BV steric metrics, spin density on the oxygen atom, available spin states, thermodynamic free energies ΔG_{PCET}, ΔG_{PT}, and ΔG_{ET}, and the magnitude of the asynchronicity parameter ($|\eta|$). Geometry optimization and frequency calculations were performed in ORCA using the def2 basis sets of Weigend and Ahlrichs and the O3LYP functional.[101–106] For the wider set of metal oxo complexes, all transition metals were given the def2-TZVPP basis set, all metal-bonded atoms and the transferring hydrogen atom the def2-TZVP basis

set, and the remaining atoms the def2-SV(P) basis set. Substrates were calculated entirely with the def2-TZVP basis set. For calculations regarding the Co^{III} oxo and various substrates, Co, N, O, the carbene carbons of the ligand, and the carbon undergoing C–H activation was given def2-TZVPP; all other atoms were given def2-SVP. For both sets of calculations solvent effects were included as a polarizable continuum (CPCM with the dielectric constant of acetonitrile for the broader set of metal oxo complexes; COSMO with the dielectric constant of THF for the Co^{III} oxo with substrates). This is primarily to mitigate the effect of self-interaction error;[107] we assume that the solvent dielectric has little effect on the rate of PCET, as reported solvent effects on similar reactions are typically limited to hydrogen bonding.[28,100,108] The def2 ECP was used for ruthenium.[109] The resolution of identity approximation was used for coulomb integrals and the chain of spheres approximation for exchange integrals (with def2/J as the auxiliary basis). No change was made to ORCA's default grid settings. Free energies were derived from the electronic energies and vibrational calculations using the quasi-harmonic oscillator formulation of Grimme and coworkers.[110]

Unfortunately, several of our optimized structures have small imaginary frequencies (See ESI), which we believe is due to numerical noise of CPCM solvation. Occasionally these frequencies lie below −100 cm^{-1} but in each of these cases the mode is isolated to a soft dihedral motion, e.g. methyl rotation on an acetonitrile ligand. We used the absolute value of these frequencies when calculating the thermodynamic enthalpy and entropy values, believing that to be a better approximation for these modes than either nonexistence or a frequency of 0 cm^{-1}. We were unable to reoptimize these structures to remove the imaginary frequencies.

In many cases, the correct ground state multiplicity of a species was not immediately clear. In such cases we confirmed our initial assignment by running ten geometry optimization cycles on alternate spin states and confirming these alternate assignments were several kcal/mol uphill of the assigned spin state. In a few cases were the energy was within 5 kcal/mol and the optimization was not close to convergence, we fully optimized the alternate spin state. Whenever two spin states had nearly the same energy, we chose the higher spin state as the ground state due to the typically higher entropy of high spin states.

To quantify the steric environment around each metal oxo center or substrate reactive C–H bond, we calculated percent buried volume (%BV) steric metrics using the online SambVCA web application.[52] We centered the calculation on the oxygen atom (for oxos) or the transferring hydrogen (for substrates), defined the negative z-axis as going through the metal center (for oxos) or the reacting carbon center (for substrates), and defined the xz plane as containing another atom bonded to the metal or carbon (the first such atom in the .xyz file). We had the center oxygen or hydrogen atom deleted from the calculation, included hydrogen atoms in the calculation, and left all other settings to their default value (using Bondi radii scaled by 1.17, a sphere radius of 3.5 Angstroms, and a mesh setting of 0.10 Angstroms). The application returns a total percent buried volume, as well as that for individual quadrants of the sphere. For metal oxo complexes, we used the total percent buried volume (%BV Tot) and the standard deviation of these four quadrants (%BV Dev) in our regressions in order to capture both overall steric bulk and how evenly distributed this bulk is around the metal oxo moiety. For substrates, we solely used %BV Tot. See the ESI for a further discussion of steric parameters and their effect on reaction barrier heights.

To evaluate the effect of spin and spin state on reactivity, we used two parameters that have been discussed in the literature: spin density on the oxo ligand and the energy to excite to a higher spin state.[8,10] Atomic spin populations were determined via IBO analysis using the freely available IBO Viewer software.[53,54] We recorded the spin density on the metal and on oxygen for each metal oxo complex as well as how much spin both atoms gain upon PCET reduction; we also tabulated similar values for the IBO charges. In the regression analysis we solely used the spin density on the oxo ligand. The "Spin Excitation Energy" is the vertical energy from the ground spin state of the initial oxo complex to the lowest lying excited spin state that is within one spin multiplicity of the resulting metal hydroxide ground spin state. If the ground spin state is already one spin multiplicity greater or lower than the product hydroxide spin state, then the spin excitation energy is taken to be zero. For instance, in the case of a triplet Fe^{IV} oxo reacting to give a sextet Fe^{III} hydroxide the spin excitation energy is the energy of the quintet Fe^{IV} oxo relative to the triplet Fe^{IV} oxo at the ground state optimized geometry. This is the scenario for most Fe^{IV} oxos in the data set. But in the case of the two non-heme Fe^{IV} quintet oxos,[12,13] the spin excitation energy is zero because the ground spin state is already within one spin multiplicity of the sextet hydroxide product. Essentially, the spin excitation energy is the energy needed to reach a spin surface on which reduction to the metal hydroxide's ground spin state is spin allowed. While this simple metric ignores the nuances of two state reactivity theory (such as the spin inversion probability) it is relatively simple to compute and has precedent as a quantitative measure of PCET reactivity.[10,36]

For each metal oxo-substrate combination assessed here, we tabulated the free energies of proton coupled electron transfer (ΔG_{PCET}, Equation 2), proton transfer (ΔG_{PT}, Equation 3), electron transfer (ΔG_{ET}, Equation 4), and the asynchronicity as defined by Srnec and coworkers (η, Equation 5):[21]

$$\Delta G_{PCET} = G_{M-OH} + G_{C\cdot} - G_{M=O} - G_{C-H} \quad (2)$$

$$\Delta G_{PT} = G_{M-OH^+} + G_{C:^-} - G_{M=O} - G_{C-H} \quad (3)$$

$$\Delta G_{ET} = G_{M-O^-} + G_{C-H^+} - G_{M=O} - G_{C-H} \quad (4)$$

$$\eta = {G_{M-OH^+} + G_{C:^-} - G_{M-O^-} - G_{C-H^+}}/{\sqrt{2}} \quad (5)$$

where $G_{M=O}$ is the calculated free energy of the oxo species, G_{C-H} is the calculated free energy of the substrate, and all other free energies are defined analogously. We also tabulated the absolute value of the asynchronicity ($|\eta|$), the average of ΔG_{PT} and ΔG_{ET} ($\Delta G_{CT\ Average}$), and the analogous electronic energies (same notation, with G replaced with E).

Statistical Analysis

All statistical analysis was performed in Python using the Numpy, Scipy, Pandas, Sklearn, and Matplotlib packages.[111–115] A script ran a prescribed set of regression models and reported statistics on each model. All regressions were performed with ordinary least squares. Prior to fitting any regression, we separated the data into a test set and a training set of metal oxo species. While we show both test and training sets for each regression in the ESI, we initially did not plot the test set or calculate statistics with it. We solely used the training set in the earlier stages of our analysis, where we determined which parameters improved a fit to ΔG_{PCET} only and which did not. We then examined if the good fits to the training set extrapolated well to the test set. We had to make a few changes to the initial division of the training and test sets, however. Initially, the Co^{IV} oxo was included in the training set,[49] but that was interfering with the fit to ΔG_{PCET}. We moved it to the test set, and to provide insight into this poor fit we calculated the Fe^{IV} oxo complex of 13-TMC and added it to the test set.[70] We also saw, visually, that the inclusion of a Ru^{IV} oxo in the training set was interfering with the addition of other parameters and moved all Ru oxo complexes to the test set.[38,43,45,50,51]

The simplest metrics reported from these models are the mean square error (MSE) and the goodness of fit R^2.[58–61] These both give an indication of how well a model fits the available data but are prone to overfitting; more complicated models can only improve these metrics, regardless of whether or not the model is actually better.

We also evaluated each model with cross validation (CV) metrics, which can become worse upon overfitting. In K-fold cross validation, the training data is further subdivided into K subsets, and each subset is predicted by the K–1 remaining subsets.[59,61] When K is the number of data points, i.e. each data point being predicted by the rest of the data points, this is known as leave-one-out (LOO) cross validation. These predicted data points can be used to calculate the MSE and R^2 as above. The MSE from LOO cross validation is an approximately unbiased estimate of the expected error of a test set; however, it has high variability from training set to training set because each prediction uses nearly every point in a given training set. By repeatedly subdividing into larger groups and averaging the resultant K-fold MSEs, one obtains a pessimistic but less variable estimate of the expected test error. As we see similar trends for both LOO and 5-fold CV, we only report LOO R^2 in the main text but show all metrics in the ESI.

Another way to determine the significance of the model is to use a statistical F-test.[58,60] This allows one to compare an unrestricted model with a more restricted one (fewer parameters used as regressors, or no parameters regressed, or restrictions placed on the relationship between coefficients, etc.). In the language of hypothesis testing, the null hypothesis is that the unrestricted model offers no improvement on the restricted model and the alternate hypothesis is that there is an improvement. When both models are fit to the data, the unrestricted model will have less total squared error than the restricted model. Assuming said error of each data point is normally distributed (or that there is enough data such that the error is approximately normally distributed), that the average error is zero, and that the model is properly formulated, it is possible to determine the probability that this reduction in total squared error is spurious. This probability is known as the *p*-value. The test relies on a well-defined number of degrees of freedom in both the restricted and unrestricted model to draw out what the statistical distribution of total squared error ought to be.

For regressions on multiple substrates at once, the unequal weighting of different metal oxo complexes (depending on how many substrates are reported for them) renders these statistical metrics unreliable.[61] We ameliorate this issue for LOO cross validation by leaving out all reaction barriers for a given metal oxo complex together rather than one at a time. That is, we leave one metal oxo complex out and predict its reaction barrier heights based on all other metal oxos' reaction barrier heights rather than leave one reaction barrier height out and predict this barrier based off all other barriers. We accordingly only report LOO CV metrics for this set of regressions.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work was funded by the National Institutes of Health (R35 GM133470), the National Science Foundation (1654144), and the University of Chicago. J.E.S. thanks the Department of Defense for a National Defense Science and Engineering Graduate Fellowship (00003765) and J.S.A. thanks the Sloan Foundation for a Research Fellowship (FG-2019-11497). Computations were generously supported by the UChicago Research Computing Center. We also acknowledge useful discussions with Dr. Daniel Bím.

Notes and references

1 M. Milan, M. Salamone, M. Costas and M. Bietti, *Acc. Chem. Res.*, 2018, **51**, 1984–1995.
2 H. B. Gray and J. R. Winkler, *Acc. Chem. Res.*, 2018, **51**, 1850–1857.
3 G. Yin, *Coord. Chem. Rev.*, 2010, **254**, 1826–1842.
4 W. Lai, C. Li, H. Chen and S. Shaik, *Angew. Chem. Int. Ed.*, 2012, **51**, 5556–5578.
5 J. M. Mayer, *J. Phys. Chem. Lett.*, 2011, **2**, 1481–1489.
6 J. M. Mayer, *Acc. Chem. Res.*, 2011, **44**, 36–46.
7 D. Dhar, G. M. Yee, A. D. Spaeth, D. W. Boyce, H. Zhang, B. Dereli, C. J. Cramer and W. B. Tolman, *J. Am. Chem. Soc.*, 2016, **138**, 356–368.
8 N. Dietl, M. Schlangen and H. Schwarz, *Angew. Chem. Int. Ed.*, 2012, **51**, 5544–5555.

9 S. N. Dhuri, M. S. Seo, Y.-M. Lee, H. Hirao, Y. Wang, W. Nam and S. Shaik, *Angew. Chem. Int. Ed.*, 2008, **47**, 3356–3359.
10 K.-B. Cho, E. J. Kim, M. S. Seo, S. Shaik and W. Nam, *Chem. Eur. J.*, 2012, **18**, 10444–10453.
11 D. Usharani, D. Janardanan, C. Li and S. Shaik, *Acc. Chem. Res.*, 2013, **46**, 471–482.
12 J. England, Y. Guo, K. M. Van Heuvelen, M. A. Cranswick, G. T. Rohde, E. L. Bominaar, E. Münck and L. Que, *J. Am. Chem. Soc.*, 2011, **133**, 11880–11883.
13 J. England, M. Martinho, E. R. Farquhar, J. R. Frisch, E. L. Bominaar, E. Münck and L. Que, *Angew. Chem. Int. Ed.*, 2009, **48**, 3622–3626.
14 J. P. Bigi, W. H. Harman, B. Lassalle-Kaiser, D. M. Robles, T. A. Stich, J. Yano, R. D. Britt and C. J. Chang, *J. Am. Chem. Soc.*, 2012, **134**, 1536–1542.
15 M. K. Goetz and J. S. Anderson, *J. Am. Chem. Soc.*, 2019, **141**, 4051–4062.
16 D. E. Lansky and D. P. Goldberg, *Inorg. Chem.*, 2006, **45**, 5119–5125.
17 T. H. Parsell, M.-Y. Yang and A. S. Borovik, *J. Am. Chem. Soc.*, 2009, **131**, 2762–2763.
18 T. H. Yosca, J. Rittle, C. M. Krest, E. L. Onderko, A. Silakov, J. C. Calixto, R. K. Behan and M. T. Green, *Science*, 2013, **342**, 825–829.
19 R. A. Baglia, C. M. Krest, T. Yang, P. Leeladee and D. P. Goldberg, *Inorg. Chem.*, 2016, **55**, 10800–10809.
20 S. K. Barman, J. R. Jones, C. Sun, E. A. Hill, J. W. Ziller and A. S. Borovik, *J. Am. Chem. Soc.*, 2019, **141**, 11142–11150.
21 D. Bím, M. Maldonado-Domínguez, L. Rulíšek and M. Srnec, *Proc. Natl. Acad. Sci. U.S.A.*, 2018, **115**, E10287–E10294.
22 D. Bím, M. Maldonado-Domínguez, R. Fučík and M. Srnec, *J. Phys. Chem. C*, 2019, **123**, 21422–21428.
23 M. Maldonado-Domínguez, D. Bím, R. Fučík, R. Čurík and M. Srnec, *Phys. Chem. Chem. Phys.*, 2019, **21**, 24912–24918.
24 M. Maldonado-Domínguez and M. Srnec, *J. Am. Chem. Soc.*, 2020, **142**, 3947–3958.
25 C. T. Saouma and J. M. Mayer, *Chem. Sci.*, 2013, **5**, 21–31.
26 C. Costentin and J.-M. Savéant, *Chem. Sci.*, 2020, **11**, 1006–1010.
27 R. Tyburski, T. Liu, S. D. Glover and L. Hammarström, *J. Am. Chem. Soc.*, , DOI:10.1021/jacs.0c09106.
28 J. J. Warren and J. M. Mayer, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 5282–5287.
29 J. P. Reid and M. S. Sigman, *Nature*, 2019, **571**, 343–348.
30 A. G. Maldonado and G. Rothenberg, *Chem. Soc. Rev.*, 2010, **39**, 1891–1902.
31 K. Jorner, T. Brinck, P.-O. Norrby and D. Buttar, , DOI:10.26434/chemrxiv.12758498.v1.
32 C. B. Santiago, J.-Y. Guo and M. S. Sigman, *Chem. Sci.*, 2018, **9**, 2398–2412.
33 B. P. Roberts and A. J. Steel, *J. Chem. Soc., Perkin Trans. 2*, 1994, **0**, 2155–2162.
34 M. Sook Seo, N. Hee Kim, K.-B. Cho, J. Eun So, S. Kyung Park, M. Clémancey, R. Garcia-Serres, J.-M. Latour, S. Shaik and W. Nam, *Chem. Sci.*, 2011, **2**, 1039–1045.
35 S. Kundu, J. V. K. Thompson, L. Q. Shen, M. R. Mills, E. L. Bominaar, A. D. Ryabov and T. J. Collins, *Chem. Eur. J.*, 2015, **21**, 1803–1810.
36 C. V. Sastri, J. Lee, K. Oh, Y. J. Lee, J. Lee, T. A. Jackson, K. Ray, H. Hirao, W. Shin, J. A. Halfen, J. Kim, L. Que, S. Shaik and W. Nam, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 19181–19186.
37 C. Fertinger, N. Hessenauer-Ilicheva, A. Franke and R. van Eldik, *Chem. Eur. J.*, 2009, **15**, 13435–13440.
38 T. Kojima, K. Nakayama, K. Ikemura, T. Ogura and S. Fukuzumi, *J. Am. Chem. Soc.*, 2011, **133**, 11692–11700.
39 Y. J. Jeong, Y. Kang, A.-R. Han, Y.-M. Lee, H. Kotani, S. Fukuzumi and W. Nam, *Angew. Chem. Int. Ed.*, 2008, **47**, 7321–7324.
40 K. A. Gardner, L. L. Kuehnert and J. M. Mayer, *Inorg. Chem.*, 1997, **36**, 2069–2078.
41 C. Arunkumar, Y.-M. Lee, J. Y. Lee, S. Fukuzumi and W. Nam, *Chem. Eur. J.*, 2009, **15**, 11482–11489.
42 K.-B. Cho, H. Kang, J. Woo, Y. J. Park, M. S. Seo, J. Cho and W. Nam, *Inorg. Chem.*, 2014, **53**, 645–652.
43 S. N. Dhuri, Y.-M. Lee, M. S. Seo, J. Cho, D. D. Narulkar, S. Fukuzumi and W. Nam, *Dalton Trans.*, 2015, **44**, 7634–7642.
44 D. Wang, K. Ray, M. J. Collins, E. R. Farquhar, J. R. Frisch, L. Gómez, T. A. Jackson, M. Kerscher, A. Waleska, P. Comba, M. Costas and L. Que, *Chem. Sci.*, 2013, **4**, 282–291.
45 J. R. Bryant and J. M. Mayer, *J. Am. Chem. Soc.*, 2003, **125**, 10351–10361.
46 G. Yin, A. M. Danby, D. Kitko, J. D. Carter, W. M. Scheper and D. H. Busch, *J. Am. Chem. Soc.*, 2008, **130**, 16245–16253.
47 S. Hong, H. So, H. Yoon, K.-B. Cho, Y.-M. Lee, S. Fukuzumi and W. Nam, *Dalton Trans.*, 2013, **42**, 7842–7845.

48 D. F. Leto, R. Ingram, V. W. Day and T. A. Jackson, *ChemComm.*, 2013, **49**, 5378–5380.
49 B. Wang, Y.-M. Lee, W.-Y. Tcho, S. Tussupbayev, S.-T. Kim, Y. Kim, M. S. Seo, K.-B. Cho, Y. Dede, B. C. Keegan, T. Ogura, S. H. Kim, T. Ohta, M.-H. Baik, K. Ray, J. Shearer and W. Nam, *Nat. Commun.*, 2017, **8**, 14839.
50 C.-M. Che, J.-L. Zhang, R. Zhang, J.-S. Huang, T.-S. Lai, W.-M. Tsui, X.-G. Zhou, Z.-Y. Zhou, N. Zhu and C. K. Chang, *Chemistry – A European Journal*, 2005, **11**, 7040–7053.
51 W. W. Y. Lam, S.-M. Yiu, D. T. Y. Yiu, T.-C. Lau, W.-P. Yip and C.-M. Che, *Inorg. Chem.*, 2003, **42**, 8011–8018.
52 L. Falivene, Z. Cao, A. Petta, L. Serra, A. Poater, R. Oliva, V. Scarano and L. Cavallo, *Nat. Chem.*, 2019, **11**, 872–879.
53 G. Knizia, *J. Chem. Theory Comput.*, 2013, **9**, 4834–4843.
54 G. Knizia and J. E. M. N. Klein, *Angew. Chem. Int. Ed.*, 2015, **54**, 5518–5522.
55 F. Jensen, *Introduction to computational chemistry*, John Wiley & Sons, Chichester, 2007.
56 J. N. Harvey, *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 2006, **102**, 203–226.
57 R. P. Bell, *The tunnel effect in chemistry*, Chapman and Hall, London; New York, 1980.
58 D. Freedman, *Statistical models: theory and practice*, Cambridge University Press, Cambridge, 2005.
59 T. Hastie, J. Friedman and R. Tisbshirani, *The Elements of statistical learning: data mining, inference, and prediction*, Springer, New York, 2018.
60 R. Carter Hill, William E. Griffiths and Guay C. Lim, *Principles of econometrics*, Wiley, Hoboken, NJ, 4th edn., 2011.
61 L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. Vanderplas, A. Joly, B. Holt and G. Varoquaux, *arXiv:1309.0238 [cs]*.
62 D. Usharani, D. C. Lacy, A. S. Borovik and S. Shaik, *J. Am. Chem. Soc.*, 2013, **135**, 17090–17104.
63 J. W. Darcy, S. S. Kolmar and J. M. Mayer, *J. Am. Chem. Soc.*, 2019, **141**, 10777–10787.
64 A. Cembran, M. R. Provorse, C. Wang, W. Wu and J. Gao, *J. Chem. Theory Comput.*, 2012, **8**, 4347–4358.
65 E. V. Anslyn and D. A. Dougherty, *Modern Physical Organic Chemistry*, University Science Books, Mill Valley, California, 2006.
66 W. P. Jencks, *Chem. Rev.*, 1985, **85**, 511–527.
67 S. Hammes-Schiffer and A. A. Stuchebrukhov, *Chem. Rev.*, 2010, **110**, 6939–6960.
68 S. Hammes-Schiffer, *Acc. Chem. Res.*, 2001, **34**, 273–281.
69 C. R. Waidmann, X. Zhou, E. A. Tsai, W. Kaminsky, D. A. Hrovat, W. T. Borden and J. M. Mayer, *J. Am. Chem. Soc.*, 2009, **131**, 4729–4743.
70 S. Hong, Y.-M. Lee, K.-B. Cho, K. Sundaravel, J. Cho, M. J. Kim, W. Shin and W. Nam, *J. Am. Chem. Soc.*, 2011, **133**, 11876–11879.
71 A. Company, I. Prat, J. R. Frisch, D. R. Mas-Ballesté, M. Güell, G. Juhász, X. Ribas, D. E. Münck, J. M. Luis, L. Que and M. Costas, *Chem. Eur. J.*, 2011, **17**, 1622–1634.
72 A. T. Fiedler and L. Que, *Inorg. Chem.*, 2009, **48**, 11038–11047.
73 O. Planas, M. Clémancey, J.-M. Latour, A. Company and M. Costas, *ChemComm.*, 2014, **50**, 10887–10890.
74 S. C. Sawant, X. Wu, J. Cho, K.-B. Cho, S. H. Kim, M. S. Seo, Y.-M. Lee, M. Kubo, T. Ogura, S. Shaik and W. Nam, *Angew. Chem. Int. Ed.*, 2010, **49**, 8190–8194.
75 C. R. Waidmann, X. Zhou, E. A. Tsai, W. Kaminsky, D. A. Hrovat, W. T. Borden and J. M. Mayer, *J. Am. Chem. Soc.*, 2009, **131**, 4729–4743.
76 K. A. Prokop, S. P. de Visser and D. P. Goldberg, *Angew. Chem. Int. Ed.*, 2010, **49**, 5091–5095.
77 Y. Liu, S.-M. Ng, W. W. Y. Lam, S.-M. Yiu and T.-C. Lau, *Angew. Chem.*, 2016, **128**, 296–299.
78 K. Cho, P. Leeladee, A. J. McGown, S. DeBeer and D. P. Goldberg, *J. Am. Chem. Soc.*, 2012, **134**, 7392–7399.
79 J. P. T. Zaragoza, M. A. Siegler and D. P. Goldberg, *ChemComm.*, 2016, **52**, 167–170.
80 I. Garcia-Bosch, A. Company, C. W. Cady, S. Styring, W. R. Browne, X. Ribas and M. Costas, *Angew. Chem. Int. Ed.*, 2011, **50**, 5648–5653.
81 X. Wu, M. S. Seo, K. M. Davis, Y.-M. Lee, J. Chen, K.-B. Cho, Y. N. Pushkar and W. Nam, *J. Am. Chem. Soc.*, 2011, **133**, 20088–20091.
82 J. Chen, Y.-M. Lee, K. M. Davis, X. Wu, M. S. Seo, K.-B. Cho, H. Yoon, Y. J. Park, S. Fukuzumi, Y. N. Pushkar and W. Nam, *J. Am. Chem. Soc.*, 2013, **135**, 6388–6391.
83 J. England, J. O. Bigelow, K. M. V. Heuvelen, E. R. Farquhar, M. Martinho, K. K. Meier, J. R. Frisch, E. Münck and L. Que, *Chem. Sci.*, 2014, **5**, 1204–1215.

84 G. K. Cook and J. M. Mayer, *J. Am. Chem. Soc.*, 1995, **117**, 7139–7156.
85 S. N. Dhuri, M. S. Seo, Y.-M. Lee, H. Hirao, Y. Wang, W. Nam and S. Shaik, *Angew. Chem. Int. Ed.*, 2008, **47**, 3356–3359.
86 S. A. Wilson, J. Chen, S. Hong, Y.-M. Lee, M. Clémancey, R. Garcia-Serres, T. Nomura, T. Ogura, J.-M. Latour, B. Hedman, K. O. Hodgson, W. Nam and E. I. Solomon, *J. Am. Chem. Soc.*, 2012, **134**, 11791–11806.
87 W. Nam, S.-E. Park, I. K. Lim, M. H. Lim, J. Hong and J. Kim, *J. Am. Chem. Soc.*, 2003, **125**, 14674–14675.
88 M. Ghosh, K. K. Singh, C. Panda, A. Weitz, M. P. Hendrich, T. J. Collins, B. B. Dhar and S. Sen Gupta, *J. Am. Chem. Soc.*, 2014, **136**, 9524–9527.
89 A. N. Biswas, M. Puri, K. K. Meier, W. N. Oloo, G. T. Rohde, E. L. Bominaar, E. Münck and L. Que, *J. Am. Chem. Soc.*, 2015, **137**, 2428–2431.
90 M. Mitra, H. Nimir, S. Demeshko, S. S. Bhat, S. O. Malinkin, M. Haukka, J. Lloret-Fillol, G. C. Lisensky, F. Meyer, A. A. Shteinman, W. R. Browne, D. A. Hrovat, M. G. Richmond, M. Costas and E. Nordlander, *Inorg. Chem.*, 2015, **54**, 7152–7164.
91 J. Kaizer, E. J. Klinker, N. Y. Oh, J.-U. Rohde, W. J. Song, A. Stubna, J. Kim, E. Münck, W. Nam and L. Que, *J. Am. Chem. Soc.*, 2004, **126**, 472–473.
92 K. Nehru, M. S. Seo, J. Kim and W. Nam, *Inorg. Chem.*, 2007, **46**, 293–298.
93 J. Serrano-Plana, W. N. Oloo, L. Acosta-Rueda, K. K. Meier, B. Verdejo, E. García-España, M. G. Basallote, E. Münck, L. Que, A. Company and M. Costas, *J. Am. Chem. Soc.*, 2015, **137**, 15833–15842.
94 T. Kurahashi, A. Kikuchi, Y. Shiro, M. Hada and H. Fujii, *Inorg. Chem.*, 2010, **49**, 6664–6672.
95 A. Gunay and K. H. Theopold, *Chem. Rev.*, 2010, **110**, 1060–1081.
96 X.-S. Xue, P. Ji, B. Zhou and J.-P. Cheng, *Chem. Rev.*, 2017, **117**, 8622–8648.
97 T. L. Hill, *An Introduction to Statistical Thermodynamics.*, Dover Publications, Newburyport, 2012.
98 M. H. Abraham, P. L. Grellier, D. V. Prior, P. P. Duce, J. J. Morris and P. J. Taylor, *J. Chem. Soc., Perkin Trans. 2*, 1989, 699–711.
99 M. H. Abraham, P. L. Grellier, D. V. Prior, J. J. Morris and P. J. Taylor, *Perkin Trans. 2*, 1990, 521–529.
100 D. W. Snelgrove, J. Lusztyk, J. T. Banks, P. Mulder and K. U. Ingold, *J. Am. Chem. Soc.*, 2001, **123**, 469–477.
101 F. Weigend, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1057–1065.
102 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
103 N. C. HANDY and A. J. COHEN, *Mol. Phys.*, 2001, **99**, 403–412.
104 A. J. COHEN and N. C. HANDY, *Mol. Phys.*, 2001, **99**, 607–615.
105 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
106 F. Neese, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2018, **8**, e1327.
107 A. J. Johansson, M. R. A. Blomberg and P. E. M. Siegbahn, *J. Chem. Phys.*, 2008, **129**, 154301.
108 D. V. Avila, C. E. Brown, K. U. Ingold and J. Lusztyk, *J. Am. Chem. Soc.*, 1993, **115**, 466–470.
109 D. Andrae, U. Häußermann, M. Dolg, H. Stoll and H. Preuß, *Theoret. Chim. Acta*, 1990, **77**, 123–141.
110 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
111 S. van der Walt, S. C. Colbert and G. Varoquaux, *Comput. Sci. Eng.*, 2011, **13**, 22–30.
112 W. McKinney, *Proc. Python Sci. Conf.*, 2010, 56–61.
113 J. D. Hunter, *Comput. Sci. Eng.*, 2007, **9**, 90–95.
114 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and É. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.
115 P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa and P. van Mulbregt, *Nat. Methods*, 2020, **17**, 261–272.