Understanding Catalytic Enantioselective C–H Bond Oxidation at Nonactivated Methylenes Through Predictive Statistical Modeling Analysis

Arnau Call,[a,*,&] *Andrea Palone,*[a,&] *Jordan P. Liles,*[b] *Natalie P. Romer,*[b] *Jacquelyne A. Read,*[b] *Josep M. Luis,*[a] *Matthew S. Sigman,*[b,*] *Massimo Bietti,*[c,*] *and Miquel Costas*[a,*]

[a] *Institut de Química Computacional i Catàlisi (IQCC) and Departament de Química, Universitat de Girona, Campus Montilivi, Girona E-17071, Catalonia, Spain.*

[b] *Department of Chemistry, University of Utah, Salt Lake City, Utah, 84112, United States.*

[c] *Dipartimento di Scienze e Tecnologie Chimiche, Università "Tor Vergata", Via della Ricerca Scientifica, 1 I-00133 Rome, Italy.*

[&] *Equally contributed to this work*

e-mail: arnau.call@udg.edu, sigman@chem.utah.edu, bietti@uniroma2.it, miquel.costas@udg.edu

Abstract

Enantioselective C(sp^3)–H bond oxidation is a powerful strategy for installing functionality in C(sp^3)–H rich molecules. Site- and enantioselective oxidation of strong C–H bonds in monosubstituted cyclohexanes with hydrogen peroxide catalyzed by aminopyridine manganese catalysts in combination with alkanoic acids has been recently described. Mechanistic uncertainties and nonobvious enantioselectivity trends challenge development of the full potential of this reaction as a powerful synthetic tool. Herein, we apply predictive statistical analysis to identify mechanistically informative correlations that provide valuable reaction understanding and will guide the development and optimization of new enantioselective C–H oxidation reactions.

Introduction

The enantioselective oxidation of C(sp^3)–H bonds converts readily available and simple organic frameworks into high added-value and chemically versatile enantiomerically enriched oxygenated compounds. Notwithstanding, site- and enantioselective C–H oxidation reactions are challenging transformations due to the presence of numerous aliphatic C–H bonds in organic molecules and their intrinsic low reactivity that generally require highly reactive reagents operating via paths seldom compatible with asymmetric catalysis. These challenges render enantioselective C–H oxidation functionally limited to enzymatic systems and few synthetic catalysts that operate in substrates containing activated (benzylic[1-11] or α-to-heteroatom[12-16]) C(sp^3)–H bonds. In this context, the most promising candidates to develop enantioselective C–H oxidation reactions are iron and manganese complexes containing tetradentate ligands due to their ability to oxidize strong C–H bonds in a stereospecific manner.[17-20]

Traditional optimization protocols in stereoselective oxidation reactions often mimic aspects of high-throughput experimentation (HTE) techniques, a resource-intensive and time-consuming method. In this regard, systematic modification of the properties of the species involved in the system can be regarded as an alternative strategy for reaction optimization which—using a simple trial and error approach—has found use in asymmetric epoxidation[21] and C–H oxidation reactions[22] developed with this type of catalysts. While in some instances satisfactory performance has been achieved from these approaches, limited mechanistic information regarding the factors controlling the activity and selectivity has been attained. Identification of these factors has allowed some degree of predictability in the site selectivity of C-H oxidation reactions, an aspect that represents a paradigm shift in the standard logic of organic synthesis.[23, 24] In this context, statistical modeling tools can afford mechanistic interpretability and simultaneously identify better performances from collected datasets. Such an approach has been successfully implemented in a wide range of catalytic asymmetric transformations, accelerating reaction optimization while meticulously designing new improved catalytic performers through predictive modeling.[25, 26] We envision that the use of these *in silico* modeling tools can overcome the limitations of traditional optimization methods so far employed in C–H oxidation catalyst development, and thus may grow the potential of this reaction by unraveling the key parameters impacting yield and selectivity.

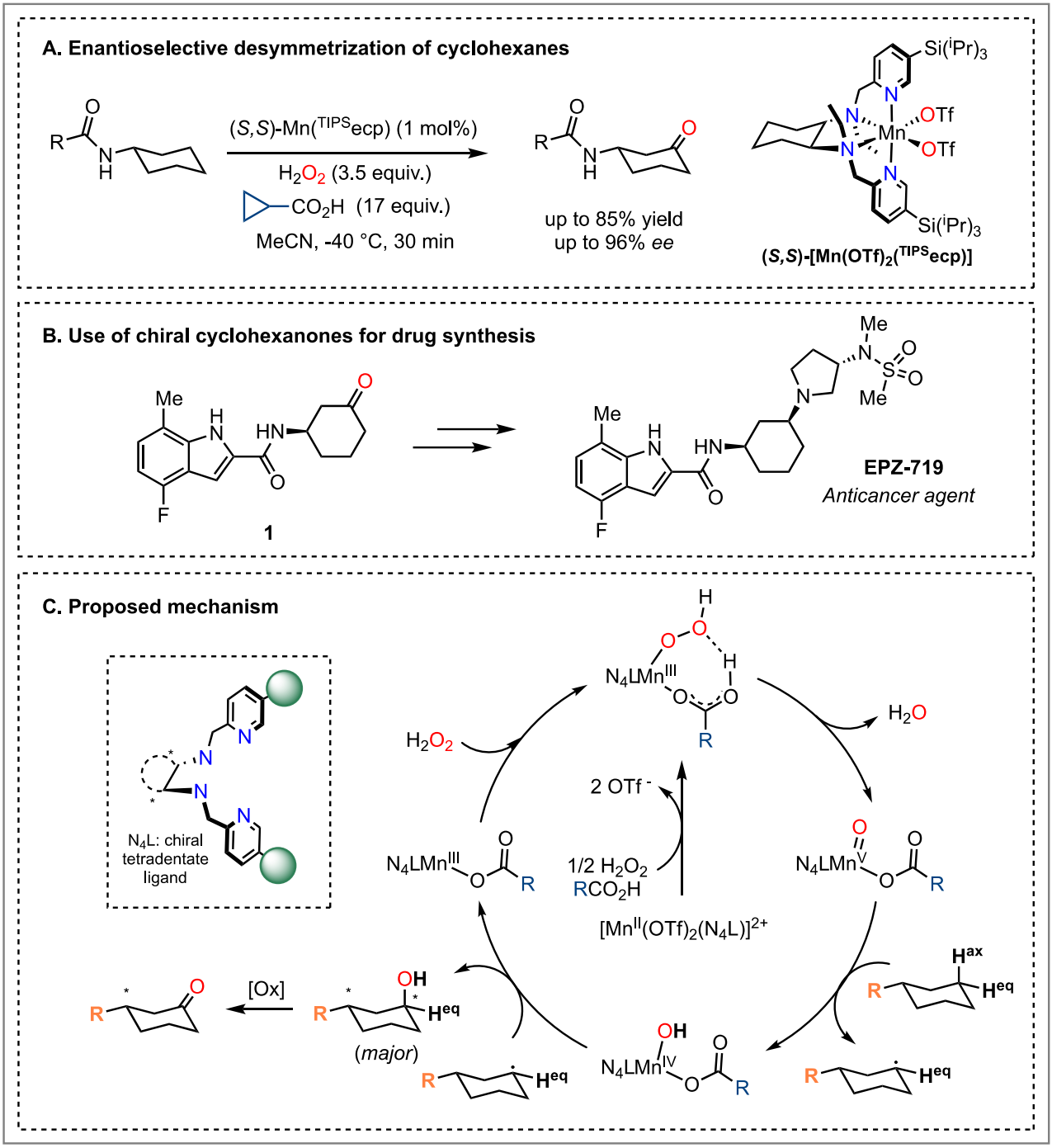

Figure 1. A) Model reaction studied in this work. B) Representative example demonstrating the importance of chiral cyclohexanones for the synthesis of a pharmaceutically approved drug. C) Proposed mechanism for the desymmetrization of substituted cyclohexanes via enantioselective oxidation of nonactivated methylenic units catalyzed by Mn complexes.

In this context, we recently reported the site- and enantioselective oxidation of methylenic C–H bonds in monosubstituted cyclohexanes with H_2O_2 catalyzed by chiral manganese complexes.[20] One interesting facet of the reaction is the furnishing of the corresponding ketoamides derived from the oxidation at C3 of the cyclohexane ring, producing a desymmetrized product with high enantioselectivity (Figure 1A). Besides being a rare case of site-selective and enantioselective oxidation of strong C–H bonds, the resulting enantioenriched ketoamides have several diversification

handles and are important building blocks for the synthesis of bioactive compounds. For example, product **1** is a valuable intermediate en route to the synthesis of the anticancer agent EPZ-719, a potent SET domain-containing protein 2 (SETD2) inhibitor (Figure 1B).[27]

Enantioselectivities and product yields were found to be quite sensitive to apparently modest changes in catalyst structure, substrate, and carboxylic acid co-catalyst, with the best results in terms of enantioselectivity being obtained with *N*-cyclohexylalkanamides, a sterically encumbered manganese catalyst ([Mn(OTf)$_2$(^{TIPS}ecp)]) (ecp = *N,N′*-diethyl *N,N′*-bis(2- pyridylmethyl)-1,2-*trans*-diamino cyclohexane) and cyclopropanecarboxylic acid. The proposed reaction mechanism for this class of catalysts entails heterolytic cleavage of H_2O_2 in a $[Mn^{III}(OOH)(OCOR)(N_4L)]^{2+}$ intermediate (where N_4L stands for a tetradentate ligand such as ^{TIPS}ecp) facilitated by the carboxylic acid, to generate an electrophilic high-valent manganese oxo carboxylato species, $[Mn^{V}O(OCOR)(N_4L)]^{2+}$, that has been shown to execute C_3–H hydroxylation via a hydrogen atom transfer (HAT)/rebound mechanism as exhibited by heme and non-heme iron-dependent hydroxylases.[28-31] Overoxidation of the first formed alcohol moiety to the corresponding ketone eliminates the stereocenter at C3, but the product remains chiral because of the initial desymmetrization that uncovers a remote chiral center at C1 (Figure 1C).

Herein we implement a data-driven workflow to understand the basic principles governing this reaction class. Leveraging the structural modularity of the species involved in the model reaction, we expanded the initially reported experimental data set and built parameter libraries for the carboxylic acids, substrates, and catalysts.[20] The new dataset was designed based on common chemical modifications involved in conventional empiricism-driven optimization routes in asymmetric C–H oxidation reactions.[28, 32, 33] Using fundamental principles of physical organic chemistry in combination with modern data science tools, we identified the parameters controlling the product yield and enantioselectivity with demonstrated predictive ability. The statistical models provide useful mechanistic information and significantly contribute to reaction comprehension. We envision that guidelines derived from this work can be generalized and successively employed to develop new asymmetric C–H oxidation reactions.

Statistical model of the carboxylic acids

Given the well-established importance of carboxylic acids in governing reactivity and imparting selectivity in these asymmetric oxidation reactions,[28] we first sought to understand which structural features of these co-ligands underpin performance. Therefore, an extended array of carboxylic acids differing in electronic and structural properties were tested by taking the previously investigated oxidation of *N*-cyclohexylpivalamide (**S1**) catalyzed by [Mn(OTf)$_2$(^{TIPS}mcp)] (mcp =

N,N′-dimethyl *N,N′*-bis(2- pyridylmethyl)-1,2-*trans*-diamino cyclohexane) as the model reaction.[20] As depicted in Figure 2, the yield of *N*-(3-oxocyclohexyl)pivalamide (**P1**) oxidation product is markedly affected by acid choice, resulting in yields ranging from 0 to 90%. Enantioselectivities also span a relatively wide range (33–86% *ee*), but are prominently clustered towards moderately high values (78–86% *ee*).

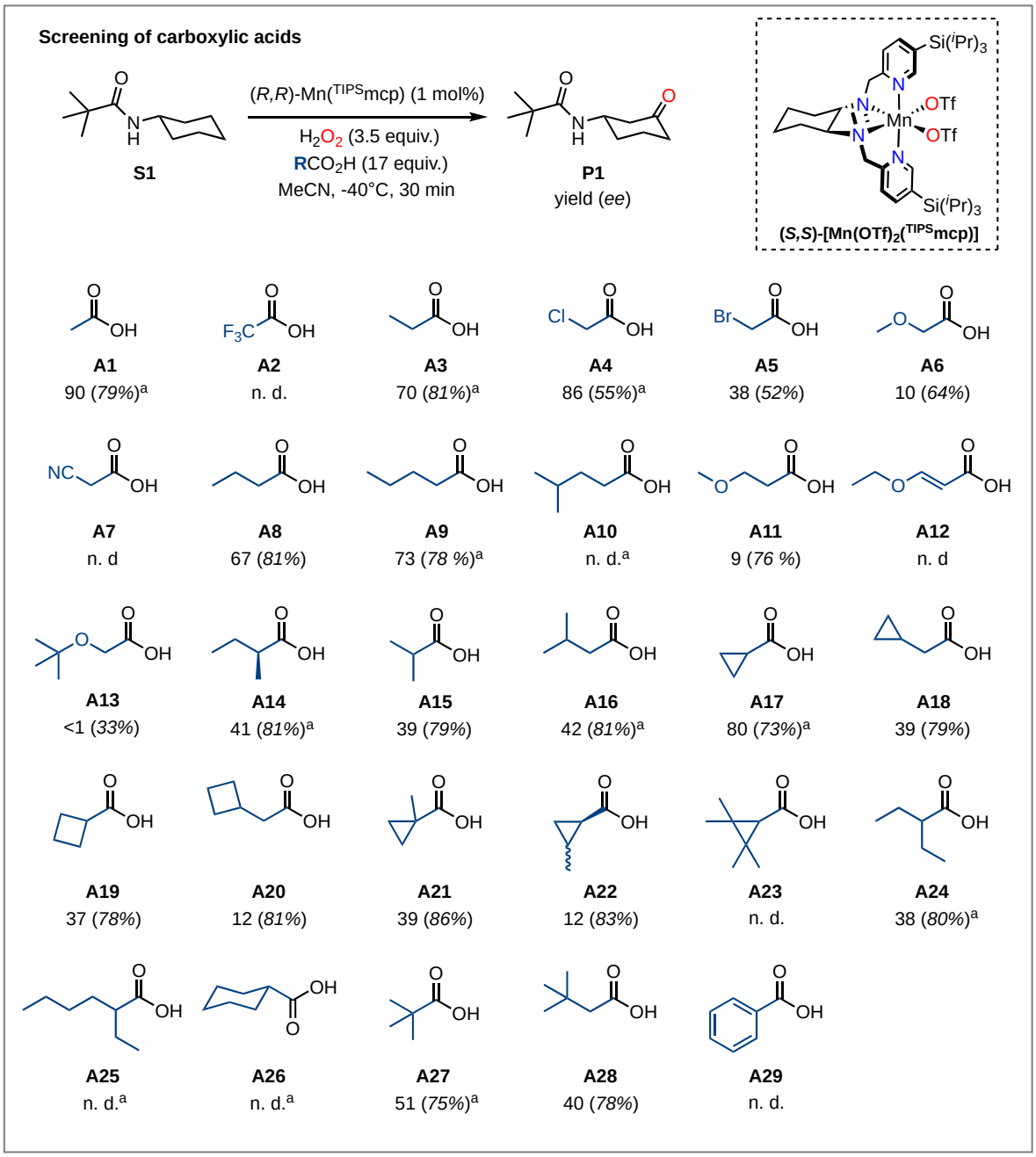

Figure 2. Enantioselective desymmetrization of substrate **S1** using different carboxylic acids. Yields and enantioselectivities (in parenthesis) determined by GC. n.d.: not detected. [a] Previously reported data.[20]

Observing this clustered output, we chose to deploy classification algorithms to identify which features correlate with performance. Towards this goal, molecular features were collected by initially performing a conformational search on each of the carboxylic acids using Macromodel; each

conformer was subsequently optimized at the DFT level (B3LYP/6-31G(d)) with single point energy corrections (M06-2X/cc-pVTZ). To account for the parameters derived from the most populated conformers, all descriptors from DFT-computed conformers were Boltzmann-averaged at 233 K based on their relative energies. Additionally, using the collection of conformers the maximum, minimum, and low-energy-conformer values for each descriptor were extracted, generating a library of over 600 molecular descriptors (See workflow in the Supporting Information, Scheme S1).

We found that the use of logistic regression—a probability-based regression algorithm that can provide insights into binary classifications for moderately sized data sets—allowed identification of the key structural features of the acids enabling high efficiency by defining the probability that a given carboxylic acid will provide a high product yield in the oxidation of **S1** (20% yield threshold).[34-36]

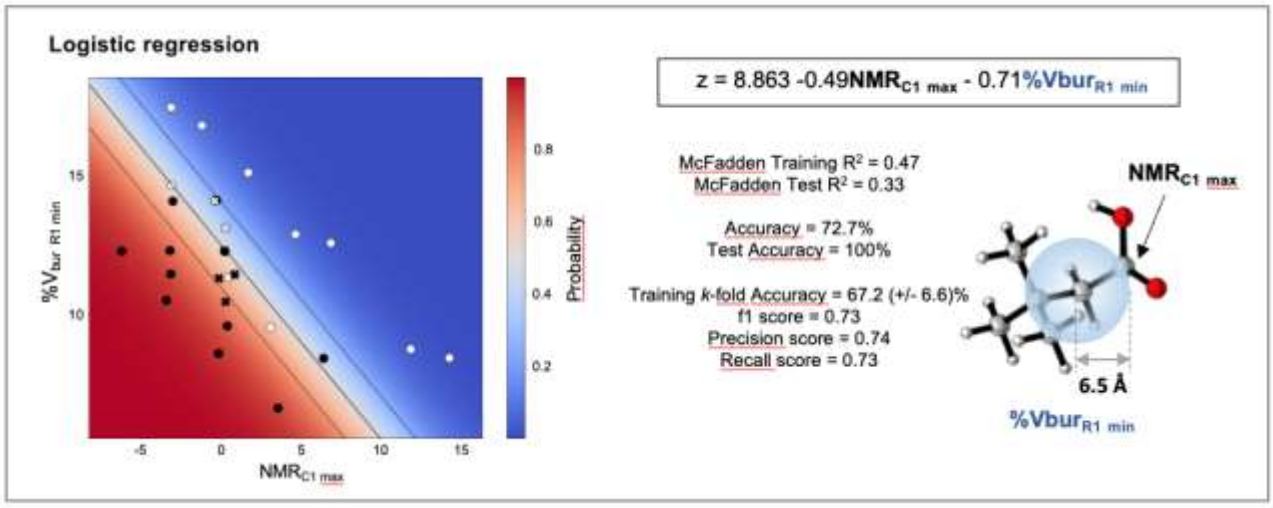

Figure 3. Logistic regression modeling for the enantioselective desymmetrization of substrate **S1** based on the carboxylic acids displayed in Figure 2. The acid data set was partitioned using the Kennard-stone split method into 80:20 training:validation set. Black and white points denote active and inactive acids, respectively (20% yield threshold). Red and blue denote active and inactive chemical space, respectively. Dashed lines denote 25% and 75% probability that a certain acid will be active. The solid black line denotes a 50% probability that the acid will be active. *Probability* = 1/(1 + e^{-z})

The resulting logistic regression model with a decision boundary of 50%, along with statistical model performance metrics and selected validation predictions, is shown in Figure 3. The model provides satisfactory levels of accuracy, precision, and f1 score for the training set while maintaining robust performance on the validation set. The coefficients and parameters of the logistic regression model, namely $NMR_{C1\ max}$ (the maximum value of the NMR shift of the carbonyl carbon of the acid) and the $\%V_{bur\ min}$ (the minimum buried volume of the acid substituent at 6.5 Å from the α-carbon) provide useful insights into the influence of the carboxylic acid co-ligand on the reaction outcome. Among electronic descriptors, $NMR_{C1\ max}$ was found to be an appropriate molecular descriptor for predicting the efficiency of the acids. The negative value of the coefficient indicates that the

probability that the **P1** yield will experience erosion increases as the electron-withdrawing character of the acid increases. For instance, the probability that the simple acetic acid provides a **P1** yield lower than 20% is severely reduced by 90% when incorporating strong electron-deficient groups such as CF_3 (**A2**) and CN (**A7**) at the α-carbon. We reasoned that electron-rich acids would better coordinate to the Mn center facilitating the heterolytic cleavage of H_2O_2, while stabilizing the electrophilic $[Mn^{V}O(OCOR)(N_4L)]^{2+}$ species.[37]

The negative coefficient of $\%V_{bur\ min}$ suggests that bulky carboxylic acid co-ligands are deleterious to the reaction. For instance, substituting acetic acid (**A1**, for which $\%V_{bur\ min}$ = 6.6) for cyclohexanecarboxylic acid (**A26**,$\%V_{bur\ min}$ =14.6), the likelihood that the system provides a **P1** yield higher than 20% is reduced by 42%. Likewise, using bulky carboxylic acids such as **A23** ($\%V_{bur\ min}$ = 16.8) and **A25** ($\%V_{bur\ min}$ = 17.4), significantly increases the probability (by 80-84%) of lowering the yield of **P1** below 20%. It is reasonable to consider that the increase in co-ligand size significantly reduces catalyst pocket free space, limiting accessibility to the manganese oxo unit for the incoming substrate. However, it is important to mention that the lack of **P1** formation when using branched carboxylic acids may be also explained in terms of competition between the intermolecular amide substrate oxidation and intramolecular directed oxidation of carboxylic acids.[17] For instance, 4-methylpentanoic acid (**A10**) bearing a tertiary γ-C-H bond, undergoes intramolecular γ-lactonization generating 5,5-dimethyl-dihydro-furan-2-one, and this completely overrides the intermolecular oxidation of **S1**. Although such analysis provides evidence that the size of the acid clearly impacts reaction output, as noted by the model equation, the interplay of both terms is not inconsequential. For instance, the likelihood of obtaining a reaction output of >20% yield increases from 64% to 96% when using a bulky and electron-rich co-ligand such as pivalic acid (**A27**), compared to that predicted for the smaller and electronically deactivated methoxyacetic acid (**A6**). Finally, the steric properties of the acids do not have a clear significant positive impact on enantioinduction, in contrast to some asymmetric C-H oxidation[22] and oxygen transfer reactions,[38] although the electronics of the acid seem to have an important role.

Statistical models of the substrates

Having established the optimal conditions for the desymmetrization of substituted cyclohexanes that comprise the use of (*S*,*S*)-$Mn(OTf)_2(^{TIPS}ecp)$ as the catalyst and cyclopropanecarboxylic acid,[20] we explored a series of substrates differing in electronics and sterics. This combination provides improved yields and enantioselectivities with respect to the (*S*,*S*)-$Mn(OTf)_2(^{TIPS}mcp)$-acetic acid couple. [20] As shown in Figure 4, the yields and enantioselectivities obtained in the desymmetrization of a series of monosubstituted cyclohexanes are strongly dependent

on the nature of the substituent. Considering that substrate conformation may have an important role in determining the selectivity in asymmetric C–H oxidations, the substrate structure was subjected to conformational search followed by geometry optimization by DFT calculations to generate a library of over 130 molecular descriptors. Following the previously described workflow, the parameters derived from the most representative populated conformers at 233 K were extracted.

The measured enantioselectivity was converted into $\Delta\Delta G^{\ddagger}$ using the Gibbs energy equation ($\Delta\Delta G^{\ddagger}$ = -RTln(*e.r.*)) where T represents the reaction temperature, R denotes the gas constant, and *e.r.* the enantiomeric ratio collected from each catalytic assay. Using $\Delta\Delta G^{\ddagger}$ as a reaction outcome allowed us to model the differential energy between the minor and major diastereoisomeric pathways involved in the presumed HAT enantiodetermining step.

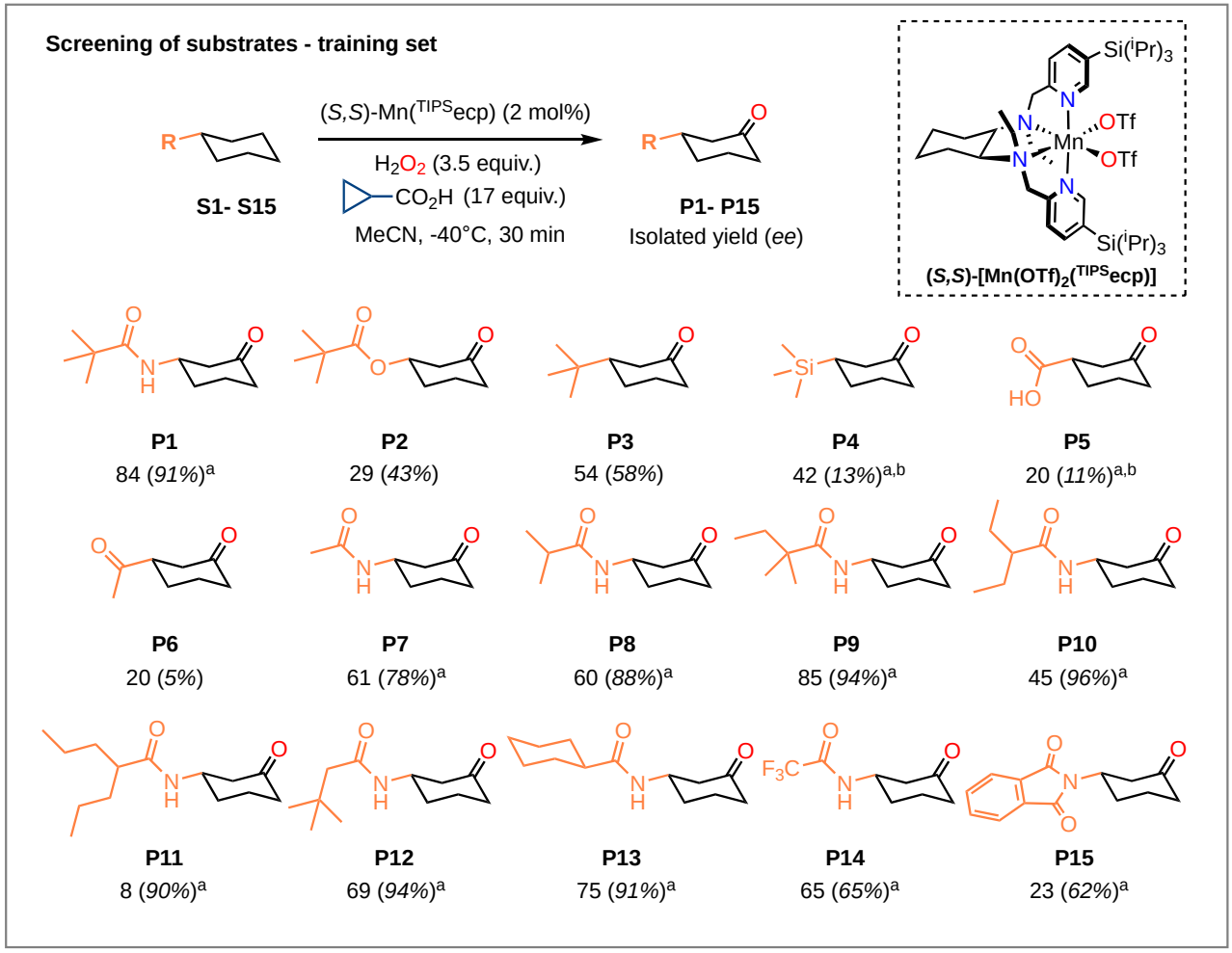

Figure 4. Oxidation of different substituted cyclohexane derivatives selected for the training set. *Ee*'s determined by GC or SFC-HPLC with chiral stationary phase. [a] Previously reported data.[20] [b] Reaction done with (*S,S*)-Mn(OTf)$_2$(^{TIPS}ecp) (1 mol%) and acetic acid (17 equiv.).

The enantioselectivity range observed in the data set varied widely from 5 to 96% *ee*, which enables interrogation of the system using multivariate linear regression (MLR) modeling. A larger

range of $\Delta\Delta G^{\ddagger}$ data improves the modeling outcome, thereby aiding the identification of privileged properties of the ring substituents that predictably impart high levels of enantioselectivity. The full data set comprised of 18 data points (**S1**-**S18**) was partitioned using the y-equidistant split method into 80:20 training:validation set. The library of molecular descriptors extracted from substrates **S1**-**S15** was used for model building (Figure 4), and those from **S16-S20** substrates were used to test the model (Figure 5).

After an iterative process using a forward stepwise linear regression algorithm, potential two-parameter models including steric and electronic molecular descriptors on the cyclohexane substituent were generated. These models were selected based on common statistical metrics such as R^2, leave-one-out (LOO), and *k*-fold. Notably, a two-parameter model with good statistical fidelity was identified (Figure 5). The model contains an electronic term that describes the Boltzmann-weighted partial charge of the atom connected to cyclohexane ($ChelpG_{R\ Boltz}$), and the Sterimol parameter that describes the minimum width of the R group at 2 Å along the C1-R axis among all conformers ($Sterimol_{C1\text{-}R\ Bmin}$).

The $ChelpG_{R\ Boltz}$ term carries the larger coefficient (-0.47) of the equation and suggests that the electronics of the substrate dominate enantioinduction. The negative correlation with selectivity suggests that higher negative charge buildup is directly correlated with increased enantioselectivity. The $ChelpG_{R\ Boltz}$ value calculated on the N of amide **S1** (-0.846) is comparatively different from that calculated on the O of ester analog **S2** (-0.471) and captures the significant variation in enantioinduction determined by the two groups. Analogously, the erosion in *ee* observed when decreasing the Lewis basic character of the amide nitrogen (direct comparison between **S7** (-0.833) and **S14** (-0.720)), or by appending an electrophilic carbon atom (substrates **S5** (0.660) and **S6** (0.536)) is described by the model.

In addition, the $Sterimol_{C1\text{-}R\ Bmin}$ term is directly correlated with enantioselectivity, which highlights the importance of substrates bearing bulky substituents near the cyclohexane ring to increase the interactions with the chiral catalyst pocket. Overall, this predictive model explains well the role of the bulky amide substituents and the reduced enantioselectivities obtained when using smaller and/or stronger electron-withdrawing substituents.

Internal validation techniques (*k*-fold = 0.75 and R^2_{pred} = 0.83) indicated a statistically robust model. The model was predictive with an average prediction error (MAE) of 0.26 kcal/mol. We then evaluated two new out-of-sample substrates not included in the initial data set. The predicted reaction outputs of the bulky ester (**S19**) and aromatic amide (**S20**) were satisfactory, albeit with a maximum MAE of 0.43 kcal/mol (Figure 5).

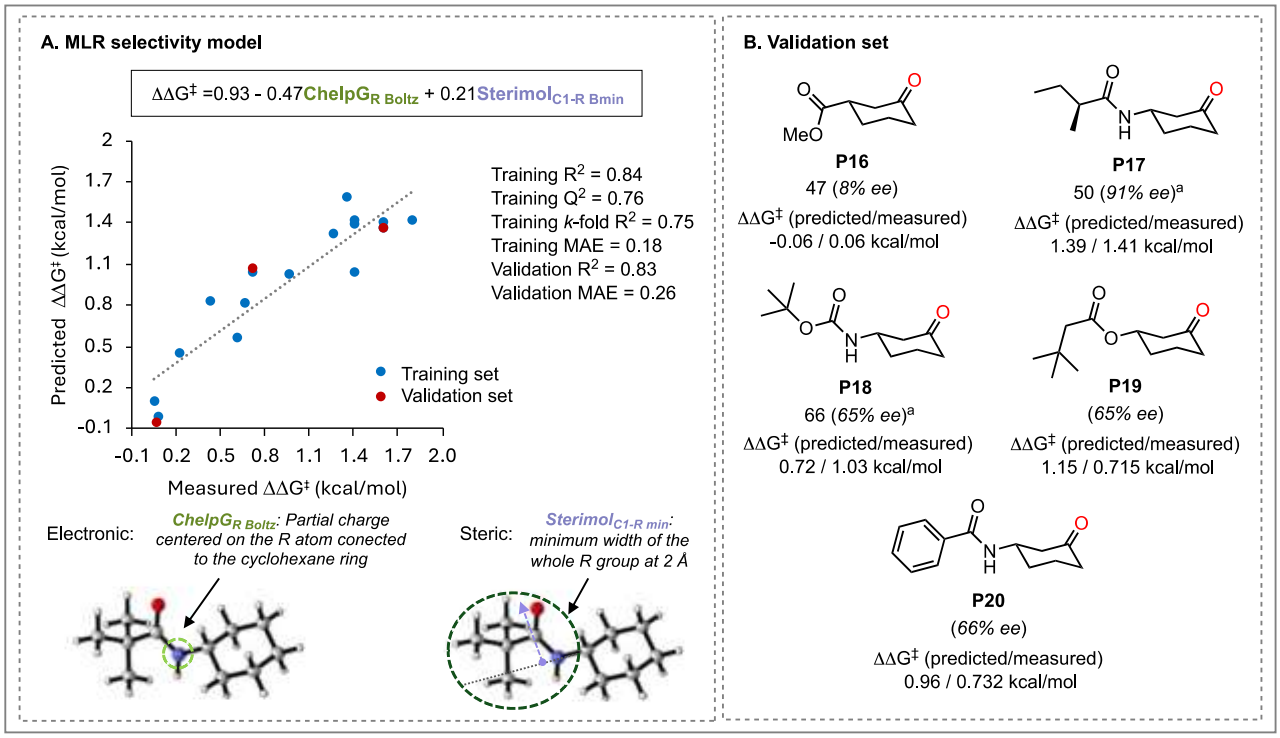

Figure 5. A) Validated multivariate linear regression with out-of-set predictions plotted. MLR model identifies the key features of the substrate for enantioinduction. B) Prediction of substrates not included in the training set. [a] Previously reported data.[20] Despite the clear univariate correlation between the experimentally observed $\Delta\Delta G^{\ddagger}$ values and the specific partial charge (CHELPG) centered on carbon C3 of the substituted cyclohexane substrate, statistical two-parameter models based on CHELPG values were discarded due to their poor predictability capability (See SI section 9). We therefore focused our attention on the molecular description of the cyclohexane substituent.

Statistical models of the catalysts

Intrigued by the catalyst-dependent activity observed in our initial exploration of the reaction space,[20] and driven by the importance of identifying the key catalyst structural features enabling high yields and enantioselectivities, we next sought to apply similar data science tools to understand the structure function relationships imparted by catalyst structure. Molecular descriptors were collected from a single conformer of catalytically active $[(Mn^{V}O(OAc)(N_4L)]^{2+}$ species, where the lone pair of the carbonyl group of the η^1-acetate bound ligand holds a weak O···O interaction with unpaired *p* electron density of the oxo ligand. Previous studies on iron[39] and manganese[40, 41] catalysts demonstrated that such an interaction energetically stabilizes the active species, and thus enables HAT reactivity.[29, 31, 40, 42]

We initially focused on studying the effect of the chiral diamine backbone. For this purpose, different Mn catalysts based on the *N,N'*-dialkyl-*trans*-1,2-cyclohexanediamine backbone but differing in the nature of the *N*-alkyl group were tested in the reaction of **S3**, and product yields correlated with selected molecular descriptors. Analysis of the $\%V_{bur}$ centered on the oxygen atom of the oxo revealed that as the size of the *N*-alkyl group increases, the $\%V_{bur}$ gradually increases and

correlates well with the observed drop in catalytic turnovers and selectivity (Figure 6). As is also evidenced by the optimized geometries, the incorporation of bulky groups in the backbone significantly restricts the accessibility to the Mn=O unit.

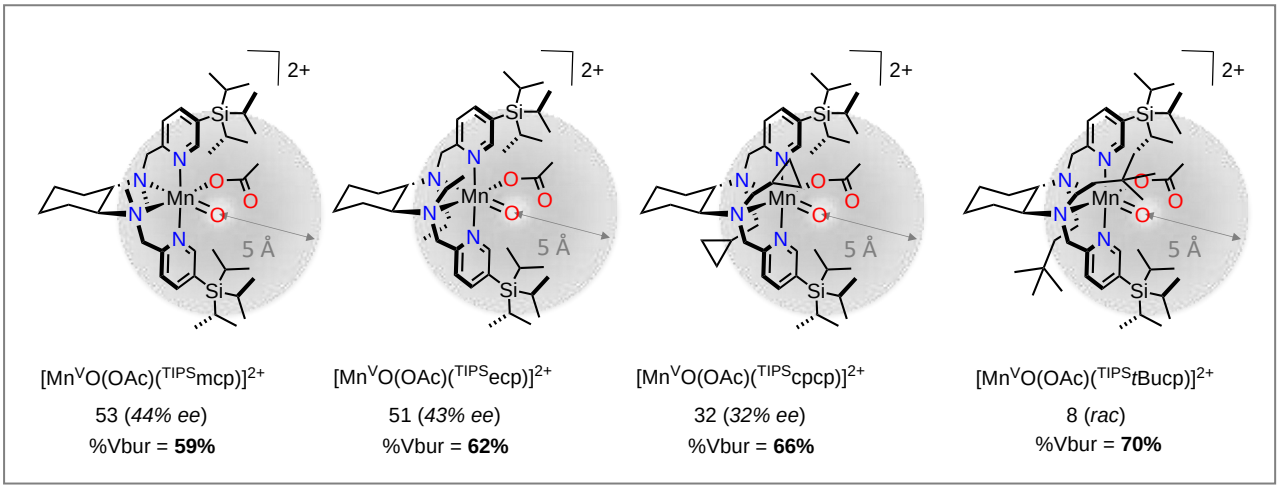

Figure 6: $\%V_{bur}$ values calculated for different catalysts varying the size of the chiral diamine backbone. The product yield and enantioselectivity (in parenthesis) in the oxidation of **S3** are shown for comparison. Reaction conditions: (*S*,*S*)-Mn catalyst (1 mol%), acetic acid (17 equiv.), H_2O_2 (3.5 equiv.) in acetonitrile at 233 K.

We anticipated that an expanded data set where the electronic and steric properties of the catalyst are systematically modified would facilitate effective comprehension and prediction of reaction outcomes. To this aim, a data set was therefore constructed by modifying the pyridine moiety of the tetradentate ligand, using the simple chiral pdp (pdp = *N*,*N*′- bis(2-pyridylmethyl)-2,2′-bipyrrolidine), mcp and ecp backbones. **S1** was employed as the model substrate using cyclopropanecarboxylic acid as co-catalyst under standard reaction conditions. The screen demonstrates that the yield as well as enantioselectivities are sensitive to the nature of the Mn catalyst, providing enantioselectivities ranging from 24 to 90% (Figure 7). Comparatively, catalytic activities using acetic acid follow a similar trend, and the enantioselectivities experience small variations.

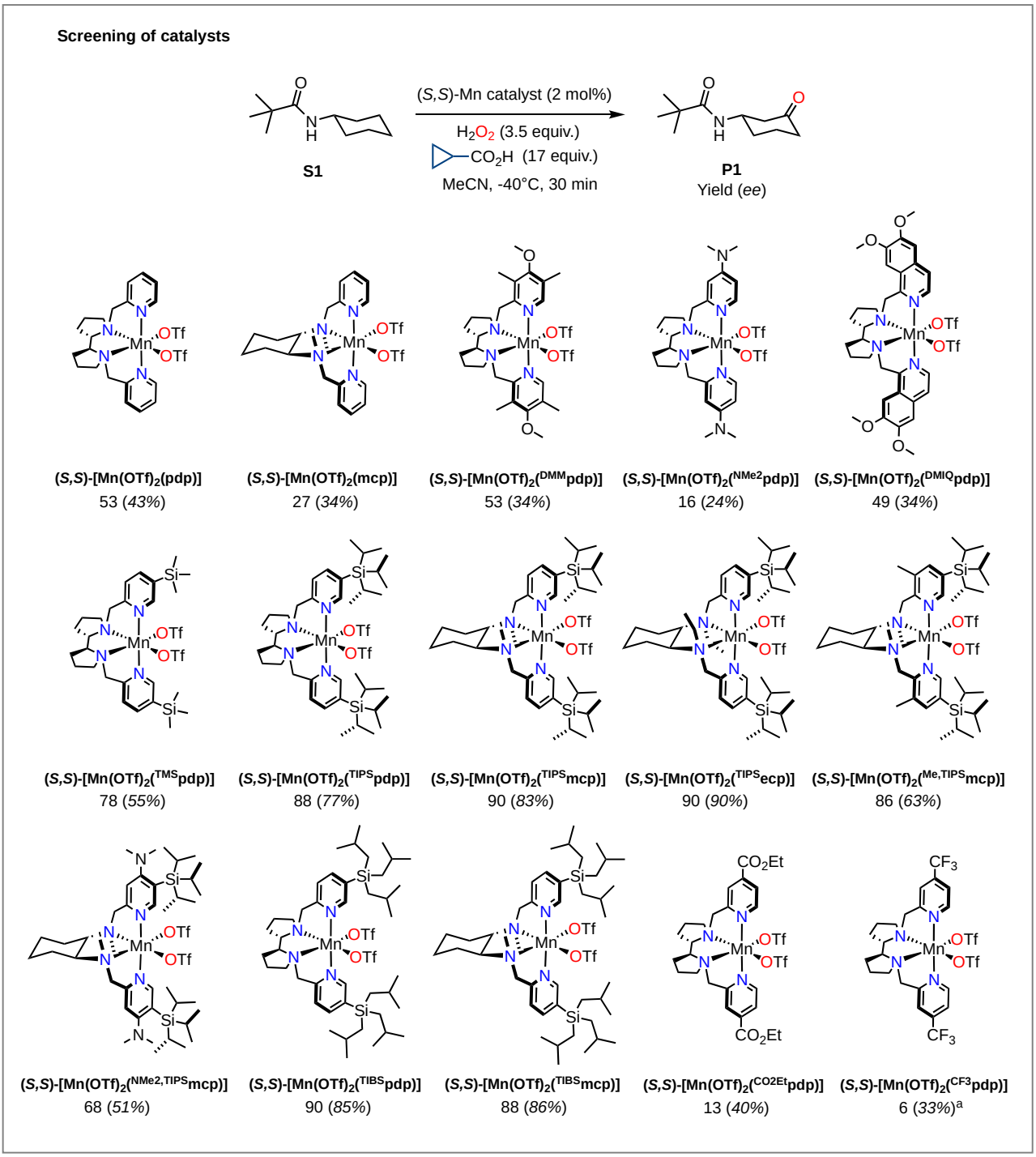

Figure 7. Enantioselective oxidation of **S1** with different catalysts. Reaction scope: substrate (25 mM) and Mn catalyst (2 mol%) using acetonitrile as a solvent and cyclopropanecarboxylic acid as co-ligand at -40 ºC. H_2O_2 (0.9 M in MeCN, 3.5 equiv.) was delivered over 30 min by syringe pump. Product yields were determined by GC from reaction mixtures using biphenyl as an internal standard. *Ee*'s (in parenthesis) have been determined by GC with chiral stationary phase. [a]Reaction performed at 0 ºC.

To build the parameter set, structure optimizations of the $[Mn^{V}O(OAc)(N_4L)]^{2+}$ species were performed at the DFT(B3LYP) level with the basis set LANL2DZ and the corresponding pseudopotential for Mn and 6-311+G(d,p) for the remainder of the atoms, taking into account

dispersion corrections using Grimme's GD3BJ algorithm.[43, 44] The energies were further refined by single-point calculations with the basis set SDD and the corresponding pseudopotential for Mn and def2-TZVPP for the remaining atoms including dispersion corrections using Grimme algorithm GD3BJ, wherein a library of 650 molecular descriptors were collected to probe structural properties.

Following the workflow previously described for model building, the iterative process afforded the combination of structural parameters resulting in the model depicted in Figure 8. The catalyst data set was partitioned using the Kennard-stone split method into 80:20 training:validation set. External validation techniques (*k*-fold = 0.85 and R^2_{pred} = 0.81) indicated a statistically robust model. In addition, the change of the training/test splitting method did not impact the model equation (Figure S4). Among the best models obtained in the iteration process, the presented model was selected based on the statistical scores and predictability. The appearance of related parameters among the series of most successful models indicates that we have identified the global operating components of this class of tetradentate chiral manganese catalysts through our workflow (Figure S5). However, since all of them combine structural parameters, the interpretability of a singular model does not affect the overall chemical understanding of the reaction.

After screening all possible univariate correlations, it was found that the angle (φ) comprised by the Mn=O bond and the line connecting the Mn atom and H_6 of the pyridine (coplanar with the Mn=O vector) moiety pointing towards the oxo (Figure 8) correlates with the experimentally observed $\Delta\Delta G^{\ddagger}$ values (R^2 = 0.79) (Figure S3). The selected two-parameter model is composed by φ and by the $Sterimol_{C5\text{-}R}$ term that describes the general steric features of the pyridine unit along the C_5-R axis at 5 Å. The φ descriptor presents the highest coefficient (0.28) and has a positive correlation with enantioselectivity which suggests that an increase in φ has a beneficial effect on enantioselectivity. The $Sterimol_{C5\text{-}R}$ term also has a positive impact on enantioselectivity and carries a smaller coefficient (0.16). Consistent with other studies, this suggests that the positive effect on enantioselectivity associated to an increase in size of the substituent at position 5 of the pyridine moiety results from a restriction of the cone of trajectories available for the incoming substrate to reach the Mn=O unit.[45] A second important consequence of the introduction of bulky substituents on the pyridines is a general improvement in the yield of **P1**, which can be reasonably attributed to an increased steric isolation of the metal center, which prevents or slows degradation pathways (Figure 7).

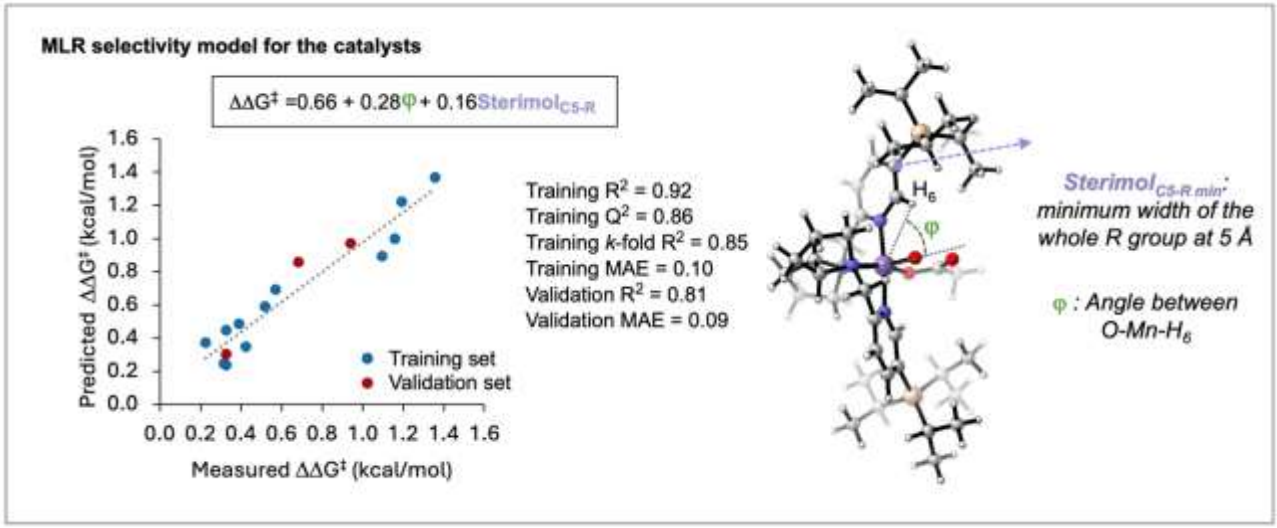

Figure 8. Predictive multivariate linear regression with out-of-set predictions plotted. The statistical regression model from the data displayed in Figure 7 reveals two key features of the catalyst determining enantioselectivity.

The subtle changes in φ when modifying ligand architecture, translated into a significant impact on the observed enantioselectivity. For instance, the slight increase in φ calculated by decreasing the electron-donating character of the pyridine as in $[Mn^{V}O(OAc)(^{NMe2,TIPS}mcp)]^{2+}$ (49.1º) and $[Mn^{V}O(OAc)(^{TIPS}mcp)]^{2+}$ (49.4º), or by changing the nature of the chiral diamine backbone as in $[Mn^{V}O(OAc)(^{TIPS}ecp)]^{2+}$ (49.9º), are in line with the observed increase in *ee*. As a result, the description of the system based on φ values correlates well to the effect of electronic and structural modification of the catalyst with good predictivity (MAE of 0.09 kcal/mol).

The impact of the electronic properties of the catalyst on the enantioselectivity in the currently studied reaction deserves discussion. Previous studies on asymmetric olefin epoxidation demonstrated that the systematic increase in the electron density of the Fe and Mn complexes,[21, 37, 46] led to a notable improvement in enantioselectivity: the less electrophilic the metal-oxo species, the closer the olefin is to the chiral pocket of the catalyst in the transition state (TS) consistent with the Hammond postulate.[47, 48] Moreover, the catalytic activity was also improved, an aspect that was interpreted as a result of a more effective push effect that assists H_2O_2 activation.[49] The electronic dependence observed in our system (Figure 7) contrasts with those observed in oxygen atom transfer reactions. To understand this phenomenon and to complement the interpretation of MLR modelling, we undertook computational studies of the TS structures for the $[Mn^{V}O(OAc)(pdp)]^{2+}$ and $[Mn^{V}O(OAc)(^{NMe2}pdp)]^{2+}$ species undergoing HAT from the axial C_3–H bond of the model substrate **S1**, the enantiodetermining step.[20] Previous studies performed in fluorinated alcohol solvents demonstrated that hydroxylation at this bond dominates over that at the equatorial C_3-H, and that the former is associated to a significantly higher enantioselectivity in the hydroxylated product.[22] Therefore, we proceeded under the assumption that for all catalysts the generation of the chiral axial

C_3-OH hydroxyamide is the primary pathway prior to the formation of the final ketoamide product. All possible conformations of reactant complexes bearing the axial C_3-H bond pointing towards the Mn oxo unit were searched using Macromodel. The most representative 30 geometries were selected based on the minimum Kelly Penalty score. The ensemble of structures generated were optimized at the B3LYP-GD3BJ/6-31G(d,p)(Mn-LANL2DZ) level with B3LYP-GD3BJ/def2-TZVPP/SMD(MeCN) single-point energy corrections. Interestingly, despite starting from different initial structures, the most stable reactant complexes within 2 kcal/mol feature a common weak interaction between the substrate and the diamine backbone. The lone pairs of the amide carbonyl group on **S1** interacts with methylenic α- to N C–H bonds of the pyridine arm and pyrrolidine ring of the catalyst (Figure 9).[22] The acidity at these sites are likely enhanced as a result of the binding of the metal center to the amine N atoms. Such an interaction significantly stabilizes the reactant species and contributes to locating the two axial enantiotopic C_3–H and C_5–H bonds in a privileged position for HAT to the Mn=O unit by positioning the bulky amide group opposite to the carboxylate co-ligand.

The irreversible nature of the HAT step, leading to a carbon-centered radical prior to the formation of the hydroxylation product, supports the hypothesis that the reaction enantioselectivity is imparted in this step.[20] Interestingly, the main differences of such structures can be identified by analyzing the positioning of the substrate in the catalyst pocket. In the TS structure of the reaction with $[Mn^{V}O(OAc)(pdp)]^{2+}$ that leads to the formation of the major enantiomer[20] (TS^{major}), the amide group of the substrate interacts with the methylenic α- to N C-H bonds of the ligand, and the targeted axial C_3-H bond points toward the Mn=O, promoting the rotation of the carboxylate co-ligand (O_1-Mn-O_2 angle of 66.1º) (Figure 9). Similarly, in the corresponding TS leading to the minor enantiomer (TS^{minor}), the interaction of the substrate with the carboxylate co-ligand slightly increases (O_1-Mn-O_2 angle of 70.3º), which may result in a modest destabilization of TS^{minor} (0.43 kcal/mol higher in energy than TS^{major}) while maintaining the H-bonding with the amide group. Interestingly, such an energy difference is in good agreement with both experimentally measured ($\Delta\Delta G^{\ddagger}_{exp}$ = 0.42 kcal/mol) and predicted values by the MLR model ($\Delta\Delta G^{\ddagger}_{MLR}$ = 0.35 kcal/mol). The computed lengths of the O–H and C–H bonds that are formed and cleaved in both diastereoisomeric TS structures are 1.75 and 1.13 Å, respectively, indicative of reactant-like TSs.

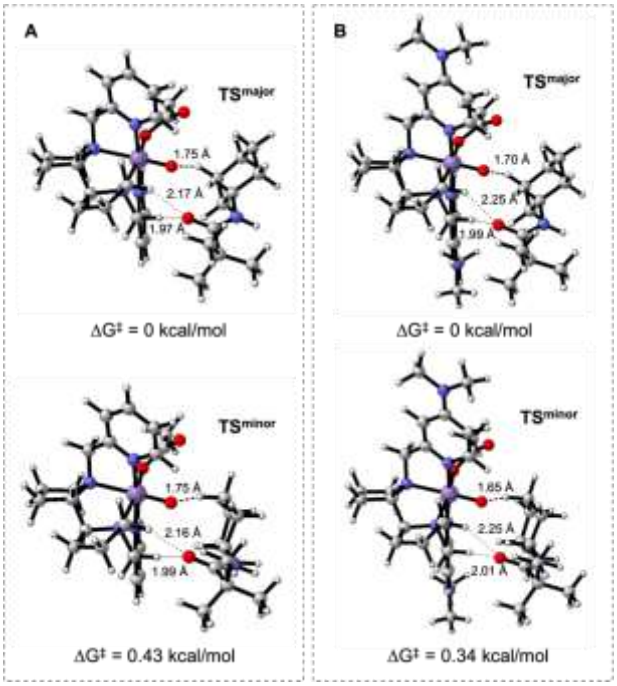

Figure 9. Optimized transition-state (triplet state) geometries of $[Mn^{V}O(OAc)(pdp)]^{2+}$ (A) and $[Mn^{V}O(OAc)(^{NMe2}pdp)]^{2+}$ (B) for axial C_3-H and C_5-H oxidation of substrate **S1**. Geometries at the top correspond to the formation of the major enantiomer (TS^{major}) while those at the bottom refer to the minor enantiomer (TS^{minor}). The CH_2---O=C and Mn=O---H-C bond distances are shown.

Analogously, TS structures computed for the reaction with $[Mn^{V}O(OAc)(^{NMe2}pdp)]^{2+}$ resemble those obtained with $[Mn^{V}O(OAc)(pdp)]^{2+}$. The difference in activation free energy between these systems is 0.34 kcal/mol, close to that measured in the catalytic assays ($\Delta\Delta G^{\ddagger}_{exp}$ = 0.24 kcal/mol) and predicted by the MLR model ($\Delta\Delta G^{\ddagger}_{MLR}$ = 0.23 kcal/mol). Interestingly, compared to the corresponding reaction with $[Mn^{V}O(OAc)(pdp)]^{2+}$, the length of the C–H bond that is cleaved remains virtually unaffected (1.13 Å). Additionally, the O–H bond that is formed is shortened likely due to the lower reactivity of the $[Mn^{V}O(OAc)(^{NMe2}pdp)]^{2+}$ species in executing HAT. Simultaneously, the H-bonding interactions between the carbonyl group of the amide substituent with distances of 2.25 and 1.99 Å (TS^{major}), and 2.25 and 2.01 Å (TS^{minor}) become weaker as a result of the increased electron density determined by the tetradentate ligand that presumably renders the α-N C–H bonds engaging in H-bonding less acidic. As expected, the calculated φ in the TS^{major} of $[Mn^{V}O(OAc)(^{NMe2}pdp)]^{2+}$ (47.8º) is lower than that calculated for the $[Mn^{V}O(OAc)(pdp)]^{2+}$ (48.4º) system, which matches what the MLR model would predict for this less enantioselective catalyst.

Interestingly, analysis of the TS structures suggests that the enantioinduction may not be directly correlated to the position of the TS of the HAT event in the reaction coordinate of the free energy landscape, in contrast to what has been proposed in oxygen atom transfer reactions,[21, 37, 38, 46-48, 50] supporting the hypothesis provided by the MLR model that electronic and steric decoration of

the system translates into subtle structural changes of the catalyst, which are directly connected to enantioinduction.

It should also be considered that the H-bonding between the carbonyl lone pair of the incoming substrate and the relatively acidic sites of the catalyst may also contribute to positioning the substrate to collect such changes. The relative importance of this is expected to increase with increasing catalyst electron deficiency and basic character of the substrate carbonyl group. These observations are reconciled with the erosion of enantioinduction observed in substrates bearing ketones, esters, and electron-poor amide substituents, which are less Lewis basic than the amide analogs.[20] Such an effect is connected to the dominant electronic term of the MLR model for the substrates. Previous results on enantioselective non-directed hydroxylation of tertiary C(sp^3)-H bonds also demonstrated that enantioselectivity is highly sensitive to the position of the carbonyl functionality along the cyclohexane substituent, which affects the establishment of H-bonding interactions.[22] In both scenarios, the increase of the H-bonding capability enforces the interaction of the incoming substrate with the chiral pocket, which likely affects enantiodiscrimination.

It also seems plausible to consider that the enhanced H-bonding behavior derived from the Lewis basic character of the amides plays a directing role in inducing the high site-selectivity for C_3-H oxidation.[20] The beneficial effect of this directing role is complementary to the non-covalent interactions established with the bulky trialkyl silyl groups, which synergistically contribute to the destabilization of TS^{minor}.[22] This effect is well-collected by the $Sterimol_{C3\text{-}R}$ term of the catalysts model, which may be important for enantioinduction in substrates lacking a carbonyl moiety at C1 such as **S3** and **S4**.

Conclusions

In summary, statistical and computational methods enable identification of the key parameters controlling both product yield and enantioselectivity in the desymmetrization of cyclohexane scaffolds via oxidation of methylenic units with chiral manganese catalysts. Based on an extended experimental dataset, internally validated models for carboxylic acids, substrates and catalyst have been developed. The logistic regression model for the acids identifies that small and electron-rich acids maximize the probability of providing a high reaction yield. The predictive MLR model found for the substrates correlates basicity and size of the cyclohexane substituents with enantioselectivity. The predictive model obtained for the catalysts constructed based on parameters of the Mn-oxo carboxylato reactive species shows that enantioinduction responds to structural changes of the catalysts, determined by both electronics of the pyridine moiety and nature of the diamine backbone. We anticipate that this data-driven workflow will provide confidence to extrapolate this methodology

to the analysis of other asymmetric C-H oxidation reactions, representing a valuable predictive platform for guiding reaction optimization and mechanistic interrogation.

Acknowledgments:

M. C. acknowledges economic support from European Research Council, (AdvG 883922 to MC.) and the Spain Ministry of Science, (MINECO, PRE2019-090149 to A.P., PID2021-129036NB-I00 to M.C. and PID2022-140666NB-C22 to J.M.L). A. C. acknowledge support from MINECO (IJC2020-046115-I and RYC2022-038031-I). Financial support in the Sigman group was supported by the National Science Foundation (CHE-2154502) and the National Institutes of Health (in the form of a postdoctoral fellowship to J.A.R., F32GM128351). Computational contributions to this work were supported by the University of Utah Center for High Performance Computing (CHPC).

Supporting Information

Materials and methods describing the preparation of complexes and substrates, characterization, and experimental procedures for the catalytic reactions. NMR spectra, SCF and GC traces. Additional modeling data, DFT methods and cartesian coordinates.

References

1. Ye, P.; Feng, A.; Wang, L.; Cao, M.; Zhu, R.; Liu, L., Kinetic resolution of cyclic benzylic azides enabled by site- and enantioselective C(sp^3)-H oxidation. *Nat. Commun.* **2022,** *13* (1), 1621.
2. Sun, Q.; Sun, W., Catalytic Enantioselective Methylene C(sp^3)-H Hydroxylation Using a Chiral Manganese Complex/Carboxylic Acid System. *Org. Lett.* **2020,** *22* (24), 9529-9533.
3. Serrano-Plana, J.; Rumo, C.; Rebelein, J. G.; Peterson, R. L.; Barnet, M.; Ward, T. R., Enantioselective Hydroxylation of Benzylic C(sp^3)-H Bonds by an Artificial Iron Hydroxylase Based on the Biotin-Streptavidin Technology. *J. Am. Chem. Soc.* **2020,** *142* (24), 10617-10623.
4. Ottenbacher, R. V.; Talsi, E. P.; Bryliakov, K. P., Highly enantioselective undirected catalytic hydroxylation of benzylic CH_2 groups with H_2O_2. *J. Catal.* **2020,** *390*, 170-177.
5. Milan, M.; Bietti, M.; Costas, M., Enantioselective aliphatic C–H bond oxidation catalyzed by bioinspired complexes. *Chem. Commun.* **2018,** *54* (69), 9559-9570.
6. Qiu, B.; Xu, D.; Sun, Q.; Miao, C.; Lee, Y.-M.; Li, X.-X.; Nam, W.; Sun, W., Highly Enantioselective Oxidation of Spirocyclic Hydrocarbons by Bioinspired Manganese Catalysts and Hydrogen Peroxide. *ACS Catal.* **2018,** *8* (3), 2479-2487.
7. Ottenbacher, R. V.; Talsi, E. P.; Rybalova, T. V.; Bryliakov, K. P., Enantioselective Benzylic Hydroxylation of Arylalkanes with H_2O_2 in Fluorinated Alcohols in the Presence of Chiral Mn Aminopyridine Complexes. *Chemcatchem* **2018,** *10* (22), 5323-5330.

8. Burg, F.; Gicquel, M.; Breitenlechner, S.; Pothig, A.; Bach, T., Site- and Enantioselective C-H Oxygenation Catalyzed by a Chiral Manganese Porphyrin Complex with a Remote Binding Site. *Angew. Chem. Int. Ed.* **2018,** *57* (11), 2953-2957.
9. Talsi, E. P.; Samsonenko, D. G.; Ottenbacher, R. V.; Bryliakov, K. P., Highly Enantioselective C−H Oxidation of Arylalkanes with H_2O_2 in the Presence of Chiral Mn-Aminopyridine Complexes. *Chemcatchem* **2017,** *9* (24), 4580-4586.
10. Frost, J. R.; Huber, S. M.; Breitenlechner, S.; Bannwarth, C.; Bach, T., Enantiotopos-selective C-H oxygenation catalyzed by a supramolecular ruthenium complex. *Angew. Chem. Int. Ed.* **2015,** *54* (2), 691-5.
11. Groves, J. T.; Viski, P., Asymmetric Hydroxylation by a Chiral Iron Porphyrin. *J. Am. Chem. Soc.* **1989,** *111* (22), 8573-8538.
12. Guan, H.; Tung, C. H.; Liu, L., Methane Monooxygenase Mimic Asymmetric Oxidation: Self-Assembling m-Hydroxo, Carboxylate-Bridged Diiron(III)-Catalyzed Enantioselective Dehydrogenation. *J. Am. Chem. Soc.* **2022,** *144* (13), 5976-5984.
13. Sun, S.; Ma, Y.; Liu, Z.; Liu, L., Oxidative Kinetic Resolution of Cyclic Benzylic Ethers. *Angew. Chem. Int. Ed.* **2021,** *60* (1), 176-180.
14. Sun, S.; Yang, Y.; Zhao, R.; Zhang, D.; Liu, L., Site- and Enantiodifferentiating C(sp^3)-H Oxidation Enables Asymmetric Access to Structurally and Stereochemically Diverse Saturated Cyclic Ethers. *J. Am. Chem. Soc.* **2020,** *142* (45), 19346-19353.
15. Lu, R.; Cao, L.; Guan, H.; Liu, L., Iron-Catalyzed Aerobic Dehydrogenative Kinetic Resolution of Cyclic Secondary Amines. *J. Am. Chem. Soc.* **2019,** *141* (15), 6318-6324.
16. Meng, Z.; Sun, S.; Yuan, H.; Lou, H.; Liu, L., Catalytic enantioselective oxidative cross-coupling of benzylic ethers with aldehydes. *Angew. Chem. Int. Ed.* **2014,** *53* (2), 543-7.
17. Call, A.; Capocasa, G.; Palone, A.; Vicens, L.; Aparicio, E.; Choukairi Afailal, N.; Siakavaras, N.; López Saló, M. E.; Bietti, M.; Costas, M., Highly Enantioselective Catalytic Lactonization at Nonactivated Primary and Secondary γ-C–H Bonds. *J. Am. Chem. Soc.* **2023,** *145* (32), 18094-18103.
18. Cianfanelli, M.; Olivo, G.; Milan, M.; Klein Gebbink, R. J. M.; Ribas, X.; Bietti, M.; Costas, M., Enantioselective C-H Lactonization of Unactivated Methylenes Directed by Carboxylic Acids. *J. Am. Chem. Soc.* **2020,** *142* (3), 1584-1593.
19. Vicens, L.; Bietti, M.; Costas, M., General Access to Modified α-Amino Acids by Bioinspired Stereoselective γ-C−H Bond Lactonization. *Angew. Chem. Int. Ed.* **2021,** *60* (9), 4740-4746.
20. Milan, M.; Bietti, M.; Costas, M., Highly Enantioselective Oxidation of Nonactivated Aliphatic C-H Bonds with Hydrogen Peroxide Catalyzed by Manganese Complexes. *ACS Cent. Sci.* **2017,** *3* (3), 196-204.
21. Cussó, O.; Ribas, X.; Costas, M., Biologically inspired non-heme iron-catalysts for asymmetric epoxidation; design principles and perspectives. *Chem. Commun.* **2015,** *51* (76), 14285-98.
22. Palone, A.; Casadevall, G.; Ruiz-Barragan, S.; Call, A.; Osuna, S.; Bietti, M.; Costas, M., C–H Bonds as Functional Groups: Simultaneous Generation of Multiple Stereocenters by Enantioselective Hydroxylation at Unactivated Tertiary C–H Bonds. *J. Am. Chem. Soc.* **2023,** *145* (29), 15742-15753.
23. Bietti, M., Activation and Deactivation Strategies Promoted by Medium Effects for Selective Aliphatic C-H Bond Functionalization. *Angew. Chem. Int. Ed.* **2018,** *57* (51), 16618-16637.
24. Chen, M. S.; White, M. C., A Predictably Selective Aliphatic C–H Oxidation Reaction for Complex Molecule Synthesis. *Science* **2007,** *318* (5851), 783-787.
25. Maloney, M. P.; Stenfors, B. A.; Helquist, P.; Norrby, P.-O.; Wiest, O., Interplay of Computation and Experiment in Enantioselective Catalysis: Rationalization, Prediction, and–Correction? *ACS Catal.* **2023,** *13* (21), 14285-14299.

26. Santiago, C. B.; Guo, J.-Y.; Sigman, M. S., Predictive and mechanistic multivariate linear regression models for reaction development. *Chem. Sci.* **2018,** *9* (9), 2389-2412.
27. Lampe, J. W.; Alford, J. S.; Boriak-Sjodin, P. A.; Brach, D.; Cosmopoulos, K.; Duncan, K. W.; Eckley, S. T.; Foley, M. A.; Harvey, D. M.; Motwani, V.; Munchhof, M. J.; Raimondi, A.; Riera, T. V.; Tang, C.; Thomenius, M. J.; Totman, J.; Farrow, N. A., Discovery of a First-in-Class Inhibitor of the Histone Methyltransferase SETD2 Suitable for Preclinical Studies. *ACS Med. Chem. Lett.* **2021,** *12* (10), 1539-1545.
28. Vicens, L.; Olivo, G.; Costas, M., Rational Design of Bioinspired Catalysts for Selective Oxidations. *ACS Catal.* **2020,** *10* (15), 8611-8631.
29. Kal, S.; Xu, S.; Que, L., Jr., Bio-inspired Nonheme Iron Oxidation Catalysis: Involvement of Oxoiron(V) Oxidants in Cleaving Strong C-H Bonds. *Angew. Chem. Int. Ed.* **2020,** *59* (19), 7332-7349.
30. Huang, X.; Groves, J. T., Oxygen Activation and Radical Transformations in Heme Proteins and Metalloporphyrins. *Chem. Rev.* **2018,** *118* (5), 2491-2553.
31. Bryliakov, K. P.; Talsi, E. P., Active sites and mechanisms of bioinspired oxidation with H_2O_2, catalyzed by non-heme Fe and related Mn complexes. *Coord. Chem. Rev.* **2014,** *276*, 73-96.
32. Sun, W.; Sun, Q., Bioinspired Manganese and Iron Complexes for Enantioselective Oxidation Reactions: Ligand Design, Catalytic Activity, and Beyond. *Acc. Chem. Res.* **2019,** *52* (8), 2370-2381.
33. Bryliakov, K. P., Catalytic Asymmetric Oxygenations with the Environmentally Benign Oxidants H_2O_2 and O_2. *Chem. Rev.* **2017,** *117* (17), 11406-11459.
34. Dotson, J. J.; van Dijk, L.; Timmerman, J. C.; Grosslight, S.; Walroth, R. C.; Gosselin, F.; Puntener, K.; Mack, K. A.; Sigman, M. S., Data-Driven Multi-Objective Optimization Tactics for Catalytic Asymmetric Reactions Using Bisphosphine Ligands. *J. Am. Chem. Soc.* **2023,** *145* (1), 110-121.
35. Boni, Y. T.; Cammarota, R. C.; Liao, K.; Sigman, M. S.; Davies, H. M. L., Leveraging Regio- and Stereoselective C(sp^3)-H Functionalization of Silyl Ethers to Train a Logistic Regression Classification Model for Predicting Site-Selectivity Bias. *J. Am. Chem. Soc.* **2022,** *144* (34), 15549-15561.
36. Oja, M.; Sild, S.; Maran, U., Logistic Classification Models for pH-Permeability Profile: Predicting Permeability Classes for the Biopharmaceutical Classification System. *J. Chem. Inf. Model.* **2019,** *59* (5), 2442-2455.
37. Cussó, O.; Garcia-Bosch, I.; Ribas, X.; Lloret-Fillol, J.; Costas, M., Asymmetric epoxidation with H_2O_2 by manipulating the electronic properties of non-heme iron catalysts. *J. Am. Chem. Soc.* **2013,** *135* (39), 14871-8.
38. Lyakin, O. Y.; Ottenbacher, R. V.; Bryliakov, K. P.; Talsi, E. P., Asymmetric Epoxidations with H_2O_2 on Fe and Mn Aminopyridine Catalysts: Probing the Nature of Active Species by Combined Electron Paramagnetic Resonance and Enantioselectivity Study. *ACS Catal.* **2012,** *2* (6), 1196-1202.
39. Fan, R.; Serrano-Plana, J.; Oloo, W. N.; Draksharapu, A.; Delgado-Pinar, E.; Company, A.; Martin-Diaconescu, V.; Borrell, M.; Lloret-Fillol, J.; García-España, E.; Guo, Y.; Bominaar, E. L.; Que, L., Jr.; Costas, M.; Munck, E., Spectroscopic and DFT Characterization of a Highly Reactive Nonheme Fe(V)-Oxo Intermediate. *J. Am. Chem. Soc.* **2018,** *140* (11), 3916-3928.
40. Feng, A.; Liu, Y.; Yang, Y.; Zhu, R.; Zhang, D., Theoretical Insight into the Mechanism and Selectivity in Manganese-Catalyzed Oxidative C(sp^3)–H Methylation. *ACS Catal.* **2022,** *12* (4), 2290-2301.
41. Call, A.; Cianfanelli, M.; Besalú-Sala, P.; Olivo, G.; Palone, A.; Vicens, L.; Ribas, X.; Luis, J. M.; Bietti, M.; Costas, M., Carboxylic Acid Directed *γ*-Lactonization of Unactivated Primary C–H Bonds Catalyzed by Mn Complexes: Application to Stereoselective Natural Product Diversification. *J. Am. Chem. Soc.* **2022,** *144* (42), 19542-19558.

42. Li, X. X.; Guo, M.; Qiu, B.; Cho, K. B.; Sun, W.; Nam, W., High-Spin Mn(V)-Oxo Intermediate in Nonheme Manganese Complex-Catalyzed Alkane Hydroxylation Reaction: Experimental and Theoretical Approach. *Inorg. Chem.* **2019,** *58* (21), 14842-14852.
43. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011,** *32* (7), 1456-1465.
44. 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* (15), 154104.
45. Gormisky, P. E.; White, M. C., Catalyst-controlled aliphatic C-H oxidations with a predictive model for site-selectivity. *J. Am. Chem. Soc.* **2013,** *135* (38), 14052-5.
46. Cussó, O.; Garcia-Bosch, I.; Font, D.; Ribas, X.; Lloret-Fillol, J.; Costas, M., Highly Stereoselective Epoxidation with H_2O_2 Catalyzed by Electron-Rich Aminopyridine Manganese Catalysts. *Org. Lett.* **2013,** *15* (24), 6158-6161.
47. Cavallo, L.; Jacobsen, H., Electronic Effects in (salen)Mn-Based Epoxidation Catalysts. *J. Org. Chem.* **2003,** *68* (16), 6202-6207.
48. Palucki, M.; Finney, N. S.; Pospisil, P. J.; Güler, M. L.; Ishida, T.; Jacobsen, E. N., The Mechanistic Basis for Electronic Effects on Enantioselectivity in the (salen)Mn(III)-Catalyzed Epoxidation Reaction. *J. Am. Chem. Soc.* **1998,** *120* (5), 948-954.
49. Khavrutskii, I. V.; Musaev, D. G.; Morokuma, K., Cooperative Pull and Push Effects on the O−O Bond Cleavage in Acylperoxo Complexes of [(Salen)Mn^{III}L]: Ensuring Formation of Manganese(V) Oxo Species. *Inorg. Chem.* **2005,** *44* (2), 306-315.
50. Jacobsen, E. N.; Zhang, W.; Muci, A.; Ecker, J. R.; Deng, L., Highly Enantioselective Epoxidation Catalysts Derived from 1,2-Diaminocyclohexane. *J. Am. Chem. Soc.* **1991,** *113* (18), 7063-7064.