

Interplay of Computation and Experiment in Enantioselective Catalysis: Rationalization, Prediction, and—Correction?

Michael P. Maloney,[▽] Brock A. Stenfors,[▽] Paul Helquist, Per-Ola Norrby, and Olaf Wiest*



Cite This: *ACS Catal.* 2023, 13, 14285–14299



Read Online

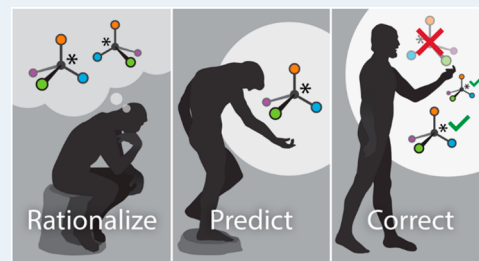
ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: The application of computational methods in enantioselective catalysis has evolved from the rationalization of the observed stereochemical outcome to their prediction and application to the design of chiral ligands. This Perspective provides an overview of the current methods used, ranging from atomistic modeling of the transition structures involved to correlation-based methods with particular emphasis placed on the Q2MM/CatVS method. Using three enantioselective palladium-catalyzed reactions, namely, the conjugate addition of arylboronic acids to enones, the enantioselective redox relay Heck reaction, and the Tsuji–Trost allylic amination as case studies, we argue that computational methods have become truly equal partners to experimental studies in that, in some cases, they are able to correct published stereochemical assignments. Finally, the consequences of this approach to data-driven methods are discussed.

KEYWORDS: Stereoselectivity prediction, Computational chemistry, Enantioselective catalysis, Cross coupling, Force field



INTRODUCTION

The interplay of experimental and computational chemistry has been exceptionally fruitful in all aspects of organic chemistry, including catalysis. Figure 1 shows that the

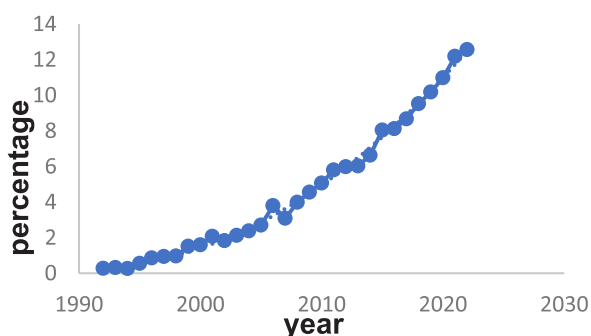


Figure 1. Percentage of publications using computational methods, relative to all publications in the area of catalysis.¹

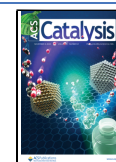
percentage of publications in the area of catalysis that use computational methods as a percentage of a publications in the field¹ continues to grow faster than linear over the last 30 years. This rapid increase, not just in absolute numbers but also as a proportion of the field as a whole, remains close to exponential and demonstrates the value that computational methods bring to the field of catalysis. Although the focus of this Perspective will be on homogeneous catalysis by small molecules, specifically transition-metal complexes, the trend shown in Figure 1 also holds true for the mechanisms of

heterogeneous catalysis^{2,3} and biocatalysis^{4,5} that have been studied extensively using computational methods.

The most common application of computational chemistry to small molecule catalysis is to help understand the mechanism of catalytic reactions and the reaction outcomes at the conceptual and atomistic level.^{6–8} This application has a long history reaching back to the earliest days of computational chemistry⁹ becoming too large and fertile of a field of chemical research to be covered in a single overview. Indeed, many studies of catalytic reaction mechanisms are now partially or completely computational (for some examples from our groups, see, e.g., refs^{10–14}), and computational methods have become an indispensable tool in the study of catalytic reactions.

Driven by the need to make cost- and time-intensive experimental studies more efficient, the focus has increasingly shifted from rationalization to the prediction of reaction outcomes in the last two decades. A key demonstration of this shift toward sometimes counterintuitive predictions that are subsequently confirmed experimentally is the prediction of the stereochemical outcome of electrocyclic ring openings of cyclobutenes that led to the development of the concept of

Received: August 20, 2023
Revised: October 4, 2023
Accepted: October 5, 2023
Published: October 26, 2023



torquoselectivity.¹⁵ In the 35 years since this seminal contribution, many other studies have shown the usefulness of computational methods to predict reaction outcomes¹⁶ up to the point where they are now a common tool in the design and development of chiral ligands for enantioselective synthesis.^{8,15–17} To be useful for such predictive applications, computational methods must not only provide reliable and verifiable predictions but also must be faster than experimental approaches and would ideally also deliver new physical insights. This is particularly important in the area of enantioselective catalysis and the prediction of reaction outcomes where the factors leading to the formation of one stereoisomer over the other are often subtle.

The logical next step in establishing computational methods as a truly equal and synergistic partner to experimental studies is to go beyond rationalization and prediction to the correction of experimental results where needed. Currently, the default assumption in the field of catalysis is that, if there are discrepancies between the experimental and computational findings, the computational study is the less-reliable one. This is in contrast to other areas such as the determination of the structure and stereochemistry of complex molecules including natural products where the use of computational methods, supported by experiments, is widely used to correct proposed structures. Despite significant progress in this area, the assignment of the configuration of stereocenters and sometimes connectivity is still a difficult problem. As a result, there is a large number of misassigned structures of natural products that were often only resolved after extremely time-consuming total synthesis projects that provided an isomer of the originally published structure.^{18,19} Computational methods, especially the combination of DFT calculations with NMR spectroscopy^{20–22} and chiroptical methods,²³ have led to reassignments of both the overall structure²⁴ and stereochemistry of natural products.^{25,26} The prediction of the stereochemical outcome of reactions is more demanding because ideally both the configuration of the major product and the selectivity should be predicted, which requires an accurate calculation of the relative Gibbs free energy of activation ($\Delta\Delta G^\ddagger$).

In this Perspective, we argue that the computational prediction of stereochemical outcomes in enantioselective catalysis is approaching a truly synergistic relationship with experiment whereby the calculations are fast enough to be useful to guide experiments with comparable accuracy, including the ability to correct experimental results where needed. After an overview of the different approaches to the prediction of stereochemistry in enantioselective catalysis, we will discuss three case studies to highlight the potential and challenges for computational predictions of enantioselectivity.

METHODS FOR THE PREDICTION OF THE STEREOCHEMICAL OUTCOME OF REACTIONS

Statistical and Linear Free-Energy Relationship Methods. The prediction of reaction outcomes and rates by linking thermodynamic and kinetic properties of a reaction with the structure of the reactants using linear free-energy relationships (LFERs) has a long history in chemistry. Electronic and steric effects or inherent reaction barriers were studied extensively using Hammett plots,²⁷ Sterimol parameters, Taft/Charton plots,^{28–31} Marcus theory,^{32,33} and other LFERs.^{34–36} The pioneering work in these correlation methods showed that the free-energy barrier for a given

reaction varied linearly (with some exceptions) with changing substituents not directly involved in the reaction. The application of LFERs to the prediction of the stereochemical outcome of a reaction, which, in most cases, depends on the difference in $\Delta\Delta G^\ddagger$ of the pathways leading to the stereoisomeric products, is therefore a logical extension of this approach. Nevertheless, applications of these concepts have been limited in their application to enantioselective reactions. This is in part due to the intrinsic difficulty of striking the correct balance of steric and electronic factors.³⁷ Furthermore, the difference in energy between diastereomeric transition states (TS) is generally small, requiring a high degree of accuracy that can be challenging for correlation methods.

An early example of this approach was inspired by quantitative structure activity relationships (QSAR) and used a combination of force field and QSAR techniques to evaluate the relative importance of different steric influences on regioselectivity and stereoselectivity in Pd-catalyzed allylations.³⁸ A few years later, Kozłowski and co-workers published a series of studies of the asymmetric addition of Et_2Zn to aldehydes.^{39,40} They created linear regression models that correlate structures of chiral catalysts with their corresponding enantioselectivities using three-dimensional (3D) property grids. This method only takes minutes of computing time to generate realistic predictions and was used to design novel chiral amino alcohol ligands for asymmetric additions to aryl aldehydes.⁴¹ Lipkowitz and Pradhan reported a similar prediction strategy for the asymmetric Diels–Alder reaction of *N*-2-alkenyl-1,3-oxazolidine-2-one with cyclopentadiene using catalysts containing seven bisoxazoline or phosphinooxazoline ligands.⁴² Their method is related to Comparative Molecular Field Analysis (CoMFA) and was developed on a set of 23 catalysts containing the aforementioned ligands. Ultimately, they were able to quantitatively predict which catalysts are the most effective at inducing high enantioselectivities, and they furthermore showed that ~70% of the model's variance was due to the steric field while ~30% was attributed to the electrostatic field. This allowed for the quantitative definition of regions in space where steric bulk would influence the selectivity, potentially allowing for the design of novel chiral ligands. In related work, Denmark and co-workers used quantitative structure–selectivity relationships⁴³ to study the addition of thiols to enamides.^{44,45} Their grid-based occupancy model of conformationally averaged structures allowed them to obtain descriptors that they fed into a machine learning (ML) algorithm.

Sigman and co-workers developed a 3D correlation method of steric and electronic free-energy relationships to design and optimize chiral ligands.^{15,46,47} Initially shifting from linear to polynomial fitting and selecting appropriate steric and electronic descriptors allowed for the prediction of stereoselectivity and design of new chiral ligands for the Nozaki–Hiyama–Kishi propargylation of ketones.⁴⁶ Beyond the ability to design new chiral ligands, this method also allows for the quick exclusion of ligand classes that are unlikely to give high enantiomeric excess (ee) values. As with other correlation-based methods, this approach does not require an understanding of the mechanism or stereodetermining step to make enantioselectivity predictions.⁴⁸ Building on this work, the Sigman group expanded their use of multivariate regression models to include a range of electronic and steric descriptors, typically derived from density functional theory (DFT) calculations,⁴⁹ to predict not only enantioselectivity⁵⁰ but a

range of other aspects of reaction development.⁵¹ (See Figure 2.) For example, application of multivariate regression to the

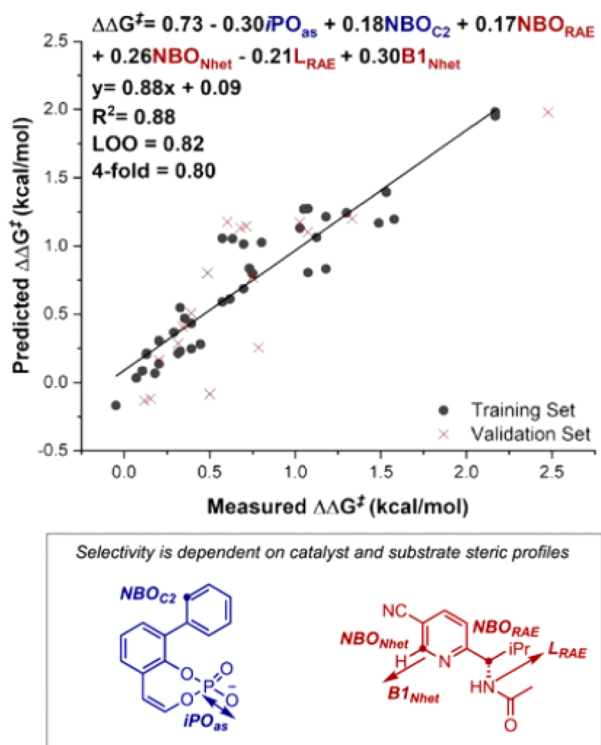


Figure 2. Multivariate regression correlation reveals that enantioselectivity is dependent on catalyst and substrate steric profiles, as represented by various catalyst/product terms. [Reproduced from ref 52. Copyright 2019, American Chemical Society, Washington, DC.]

Minisci reaction of *N*-acyl, α -amino radicals to pyridines and quinolones allowed for the guided optimization of pyrimidine substrates.⁵²

A database of steric and electronic descriptors derived from DFT calculations was developed by Fey and co-workers to explore organometallic ligand properties.^{53–56} Several ligand knowledge bases (LKBs) have been constructed to describe many types of common ligand classes, including monodentate phosphine ligands,⁵⁷ bidentate P,P- and P,N-ligands,⁵⁸ other bidentate ligands,⁵⁹ and carbenes.⁶⁰ The LKBs have been explored with multivariate statistical methods, including principal component analysis, which allows for the mapping of chemical space based on steric and electronic descriptors. A ligand's steric and electronic descriptors can then be set against measurable reaction outcomes (yield, rates, selectivity), which allows for statistical models to guide the discovery, optimization, and design of catalysts. However, this strategy is, in most cases, unable to make reliable stereoselectivity predictions due to the limited treatment of conformational space and the lack of ground truth inherent to chemical space approaches.⁵⁴ The lack of data to validate these predictions against renders LKBs, and other chemical space approaches, unreliable in their stereoselectivity predictions. Additionally, LKBs do not take into consideration substrate–catalyst matching, which could lead to erroneous predictions if the catalyst winds up being nonreactive toward the substrate of interest.

Applications of ML in its many variations to predicting stereoselectivity represent a promising direction for the field, as these algorithms can readily learn from datasets and potentially make predictions without information about the mechanism of the reaction. This field has experienced a rapid expansion in recent years, including in the area of stereoselective enzymatic reactions,⁶¹ which are beyond the scope of this Perspective. In the area of small-molecule catalysis, several groups have shown the promise of applications of ML to stereoselectivity predictions. Corminboeuf and co-workers used reaction-based machine learning representations to predict the enantioselectivity of various organocatalysts for the propargylation and allylation of aromatic aldehydes.⁶² Their use of dissimilarity plots allowed the authors to progressively improve reaction-based representations that were mapped to the activation energy of the stereodetermining TS. Additionally, they identified a fundamental limitation of physics-based molecular representations being that neither the ground-state structure before or after the TS is a suitable fingerprint for the TS itself. This limitation can be overcome using a reaction-based representation derived from both structures. Chen and co-workers used molecular field-based regression analysis to perform asymmetric catalyst optimization for divergent control of multiple stereocenters in the α -C-allylation of carboxylic acids.⁶³ Their method used Molecular Field Analysis (MFA), which is a regression analysis between the reaction outcomes and molecular fields calculated from 3D structures. Their approach used intermediate structures in the enantiodetermining step to extract and visualize 3D-structural information, similar to other MFA-based methods for predicting stereoselectivity.^{64,65} The Hong group developed a composite machine learning model that learned from existing stereoselective reactions and was able to accurately and quantitatively predict the activation energies of new reactions.⁶⁶ Their composite model outperformed individual models (LASSO, Overall RF, and nucleophile-focused RF) with a mean absolute errors (MAEs) lower across all reaction types. When using the composite model for 64 reactions not included in the training set, the MAE was 0.39 with a correlation coefficient (r^2) of 0.951, which is on par with other state of the art models. Seeberger and co-workers demonstrated the application of random forest models to the notoriously difficult quantitative prediction of stereoselectivity in glycosylations, achieving a root-mean-square (RMS) error of 6.8% in experimental validations of the ML predictions.⁶⁷ Despite the limited interpretability of ML, which often does not provide insights into the underlying physical reason for the observed effects as well as problems such as data-set quality, scope, and coverage,^{68,69} these examples show that the use of ML methods to stereoselectivity has significant potential for further development and application.

Electronic Structure Calculation of the Transition Structures. In irreversible reactions, the stereochemical outcome of a reaction depends on $\Delta\Delta G^\ddagger$ for the pathway leading to the stereoisomeric products. Consequently, knowledge of the stereodetermining step of the reaction mechanism is a prerequisite for studies of the stereoselectivity. Electronic structure methods, namely, density functional theory (DFT), have been used extensively to rationalize molecular properties, reaction mechanisms, and the origin of stereoselectivity.^{6,7,70,71} Although the prediction of stereoselectivity using DFT has been successful in many cases,^{8,72,73} it must address several challenges. The first is the accuracy of DFT calculations for the

system of interest. Even if “chemical accuracy”, often considered to be ~ 4 kJ/mol, is achieved, this is not sufficient for the accurate prediction of the stereochemical outcome of reactions as the $\Delta\Delta G^\ddagger$ values responsible are of the same magnitude. The assumption of error cancellation between the diastereomeric TS, which could lead to higher accuracies, is not generally correct, especially because many enantioselective reactions rely on differences in noncovalent and dispersion interactions in the diastereomeric transition states. A detailed analysis of the conformational space of the diastereomeric TS for the Pd-catalyzed conjugate addition of arylboronic acids to 2-substituted chromones is shown in Figure 3, which

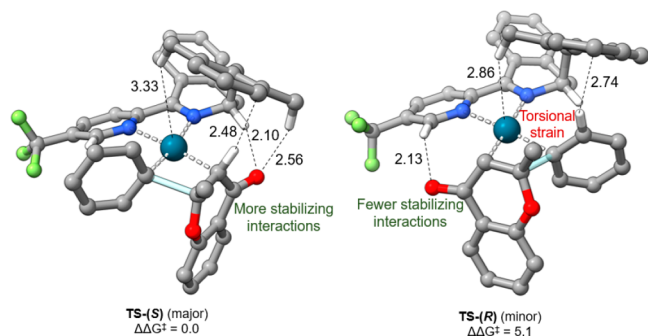


Figure 3. Noncovalent interactions in the diastereomeric TSs of the Pd-catalyzed conjugate addition of aryl boronic acids to 2-substituted chromanones. [Adapted with permission from ref 74. Copyright 2022, Royal Chemical Society, London.]

demonstrates the importance of differential noncovalent interactions in the TS.⁷⁴ To address the problem of short-range noncovalent interactions, different dispersion corrections⁷⁵ and dispersion-aware functionals⁷⁶ have been developed. Their importance in calculating ee values was studied by Paton, who investigated the role of noncovalent interactions in the oxazaborolidinium-catalyzed cycloaddition of maleimides.⁷⁷ Only when including dispersion corrections could the experimentally observed ee values be replicated, as favorable dispersive forces bias the complexation to a specific face of the catalyst. Dispersion corrections are unable to account for medium- to long-range noncovalent interactions due to the semilocal treatment of electron correlation that all functionals use. Johnson et al. developed a method for revealing noncovalent interactions,⁷⁸ which was successfully applied to predicting enantioselectivities by Arbour and co-workers who studied the Ag-catalyzed addition of alcohols and amines to allenes,⁷⁹ where noncovalent interactions between

the substrate and ligand were shown to play a decisive role in predicting stereoselectivity.

The second challenge is to properly account for the conformational ensemble about the TS. The free-energy differences between different conformations of the TS are again of the same magnitude as the $\Delta\Delta G^\ddagger$ values of the diastereomeric TS. Even if the lowest energy conformation of the pathways leading to the stereoisomers is found, conformational sampling is desirable due to the Boltzmann-weighted contribution of higher energy conformations to the overall $\Delta\Delta G^\ddagger$. One approach to the conformational search problems is the AARON toolkit,⁸⁰ developed by Wheeler and co-workers that automates geometry optimizations and builds several conformations of each diastereomeric TS. The toolkit then performs a series of DFT optimizations of the conformations giving a conformational ensemble of each pathway. The applicability of this method was demonstrated in the study of a Rh-catalyzed asymmetric hydrogenation of (*E*)- β -aryl-*N*-acetylenamides⁸¹ and for design of organocatalysts for asymmetric propargylations.^{82,83}

A third challenge is the significant computational resources necessary due to the unfavorable scaling of DFT methods for large ligands and the large number of conformations that need to be calculated. The time needed for these calculations can therefore limit the application to guide experiments in a predictive manner. This is particularly true because the wider availability of high-throughput experimentation (HTE) in recent years, possibly in combination with ML methods such as Bayesian optimization,⁸⁴ greatly accelerated the traditional “trial-and-error” approach to catalyst optimization. Competitive computational methods for prediction therefore need to be both fast and accurate. One approach to accelerate DFT calculations is through hybrid QM/MM approaches where a small subsection of the system, typically where bonds are breaking/forming, is treated with QM methods while the rest of the system is calculated with molecular mechanics (MM) methods. These hybrid calculations are commonly used to study biological systems; however, Goodman and co-workers applied them to the study of the BINOL-derived phosphoric acid-catalyzed asymmetric allylboration of aldehydes.⁸⁵ Their calculations led to the correction of the absolute stereochemistry of a previously misassigned result,⁸⁶ showcasing the utility of the method as a means to proofread experimentally assigned stereoselectivity.

The Q2MM/CatVS Approach. Over the last three decades, we have developed the quantum-guided molecular mechanics (Q2MM) method that combines speed and accuracy with the insight into the physical origin of

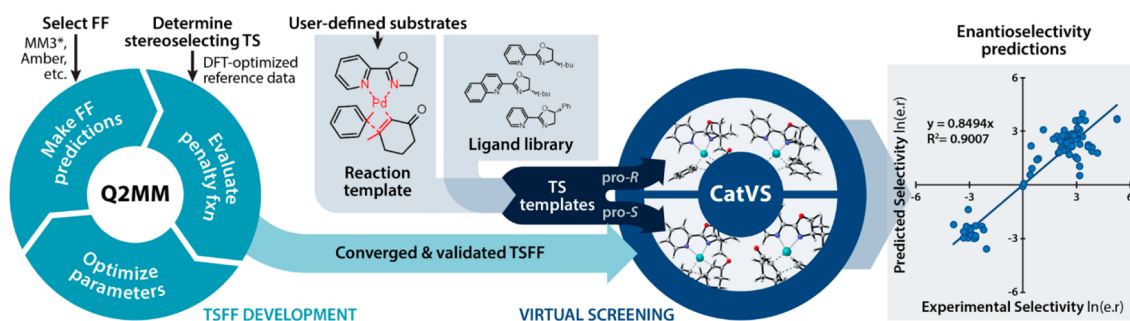


Figure 4. Flow scheme for the Q2MM/CatVS method.

stereoselectivity that is typically provided by atomistic modeling. Q2MM is a surrogate method for the prediction of stereoselectivities by creating parametrized reaction-specific model systems of stereodetermining TS. As outlined in more detail below, Q2MM automates the process of the fitting of a reaction specific force field (FF) of arbitrary functional form to the results of appropriate electronic structure calculations by minimizing an objective function χ^2 . Although Q2MM can also generate FFs for stable molecules,⁸⁷ it is most often used to generate transition-state force fields (TSFFs).⁸⁸ The Q2MM results are utilized in the Catalyst Virtual Screening (CatVS) method⁸⁹ to generate stereoselectivity predictions (see Figure 4).

The use of potential energy functions of the reactants and products such as FF to approximate the connecting TS has a long history in chemistry.^{90,91} Most commonly, this is done by mixing the reactant and product potential energy surfaces (PES) with different weights and corrections.^{92–95} TSFFs differ from these approaches in that the FF aims to describe the TS directly, originally by manual fitting using empirical steric parameters⁹⁶ or toward suitable electronic structure calculations.^{97,98} With few exceptions,⁹⁹ the TS is modeled as an energetic minimum through inversion of the PES and use of the Hessian Matrix, which describes changes in energy with respect to geometric distortions, the key difference in Q2MM, which separates it from traditional FF fitting methods.^{100,101} Inverting the PES is done by decomposing the Hessian (H) into

$$H = VSV^T \quad (1)$$

with corresponding eigenvectors (V) and eigenvalues (S) and replacing the negative S with a large positive value or ignoring the negative S eigenmodes and fitting directly to the eigenmodes and eigenvectors.¹⁰¹

As a starting point, the stereoselecting TS of the reaction for a training set needs to be calculated at an appropriate level of theory. Typical training sets include 8–10 simplified model systems that capture the key steric and electronic factors of the TS in the smallest possible models to minimize convolution of the effects. Next, the user needs to select the functional form of the FF and the atoms to be reparametrized, typically within 3–4 atom distance of the reaction center where bonds are formed or broken. All remaining atoms are treated by standard FF parameters. While other options^{102,103} are available in the Q2MM/CatVS code,¹⁰⁴ the MM3* FF^{105,106} is used for the study of small-molecule catalysts. The FF parameters selected to be fitted are then adjusted in an automated, iterative procedure in which the FF is fitted to the reference values by minimizing an objective function, χ^2 , which depends on the difference between the results from the FF and the reference calculations.

$$\chi^2 = \sum_i w_i^2 (x_i^o - x_i)^2 \quad (2)$$

where x_i^o is the reference data point, x_i is the corresponding FF calculated data point, and w_i is the weight. The function is weighted by the inverse of the acceptable error for a given data type (bond lengths, torsions, angles, etc.)¹⁰⁷ and convergence is achieved when the objective function value is no greater than N , where N is the number of data points in the training set. Gradient-based and simplex optimization methods are used to generate all necessary bonded and nonbonded FF parameters, starting with electrostatic parameters alone. Charges are

calculated using either RESP or CHELPG fitting. Given the abundance of reference data associated with the Hessian matrix, consisting of $3n \times 3n$ data points, where n is the number of atoms, overfitting of parameters in the TSFF is avoided while also containing important energetic information for deviations in equilibrium geometry.¹⁰⁸ Upon convergence, the TSFF is validated by comparing the geometries, relative energies, and Hessian matrix values obtained through electronic structure calculations for a test set of TSs not contained within the training set. Finally, the TSFF is used to calculate the $\Delta\Delta G^\ddagger$ of the diastereomeric TS of experimentally studied reactions, and the predicted ee value is compared to experimental values as an external validation. It should be noted that this is the only use of experimental information in Q2MM, making it a truly predictive method.

After validation, user-defined substrates and ligands are subjected to screening, utilizing the CatVS method, where ligand/substrate combinations of interest are merged to a TS template with initial geometries corresponding to the associated TSFF, affording unique diastereomeric configurations for the pro-*R* and pro-*S* faces. The CatVS method computationally mimics aspects of high-throughput experimentation (HTE) techniques, a resource-intensive and time-consuming method when done experimentally. The screening can involve virtual libraries of potential ligands, including both known and novel ligands, predicting those with high selectivity for the desired enantiomer for future synthesis. The process has been adapted to support a wide range of ligand/substrate combinations. CatVS automates the process from the initial ligand/substrate combination to the final Boltzmann-averaged prediction. A key assumption is that the mechanism and the stereoselecting TS for which the TSFF was developed will stay constant across the different combinations screened.

The constructed TS models are then subjected to a conformational search using established software packages such as MacroModel. Through an automated selection of conformational search parameters, the resulting TSs are subjected to a combination of Monte Carlo and Low Mode conformational search methods,^{109,110} affording conformational ensembles of each stereoisomeric TS. The stereoselectivity of the reaction is then calculated by Boltzmann averaging, according to

$$\% ee = \frac{er - 1}{er + 1} \times 100 \quad \text{with } er = \frac{\sum_{i \in R} e^{-\Delta\Delta G_i^\ddagger/RT}}{\sum_{j \in R} e^{-\Delta\Delta G_j^\ddagger/RT}} \quad (3)$$

whereby $\Delta\Delta G^\ddagger$ values for the diastereomeric TS leading to the two products i and j are summed over conformational ensembles, relative to the lowest energy conformation.

The Q2MM/CatVS method has shown high correlation between predicted and experimentally determined stereoselectivities, with typical correlation coefficients in the range of $R^2 = 0.8–0.9$ ^{88,111} across a wide range of reactions including transition-metal-catalyzed oxidations,^{70,112,113} hydrogenations,^{114–117} aminations,¹¹⁸ and additions to aldehydes.^{119,120} The automated setup, conformational search, and calculation of ee values performed by CatVS, along with the automated parameter optimization of TSFFs in Q2MM, significantly decreases the need for human intervention. This, together with the high accuracy of the predicted stereoselectivities for various systems, showcase the benefits of computation in organic synthesis, offering a potential preview into the future of small-

molecule synthesis, whereby computation exceeds most aspects of human intuition. Even in cases where the Q2MM/CatVS method performs poorly in comparison to previous results, it would still be useful as a prescreening technique prior to experimental efforts via HTE, saving time and resources by offering starting points for further screening methods.⁸⁹ Increasing accessibility to such computational tools lends credence to more efficient experimental efforts in terms of time, cost, and resources.

While the performance of the Q2MM/CatVS approach across many systems has been impressive, some of its limitations need to be kept in mind: (i) the mechanism of the reaction under study, specifically the stereoselecting step, needs to be well understood and constant across the virtual library of substrate/ligand combinations; (ii) the relevant interactions need to be captured in the training set used; (iii) the performance of Q2MM for strongly solvent dependent reactions involving large changes in partial charges in the TS is often low,¹¹⁹ due to the fixed-charge nature of the currently used FF and the limited accuracy of implicit solvent models;¹²¹ and (iv) even though some degree of error cancellation is expected in the calculation of closely related TS, both the underlying electronic structure methods and the FF used to fit the TSFF have a limited accuracy and (v) if multiple pathways are involved, their relative energy needs to be carefully calibrated. The strengths and weaknesses of the method and their interplay with experiment will be highlighted in the following three case studies.

■ CASE STUDY 1: ENANTIOSELECTIVE 1,4-ADDITION OF ARYLBORONIC ACIDS TO ENONES

The first enantioselective Pd-catalyzed conjugate addition of arylboronic acids to enones was achieved by Minnaard and co-workers, using a Pd(II)/DuPHOS system to form tertiary stereocenters.¹²² This was preceded by significant advancements in the field by Miyaura^{123,124} and Lin and Lu,^{125,126} who contributed to the development of racemic variations. Stoltz and co-workers were able to generate quaternary stereocenters in high yields and enantioselectivities for a variety of enones and arylboronic acids using the (*S*)-*tert*-butylpyridinooxazoline (PyrOx) ligand (Figure 5, top).¹²⁷ Minnaard and co-workers were also able to form quaternary stereocenters using a Pd(II)/

PhBOX system and observed high selectivities but only moderate yields.¹²⁸ The Houk and Stoltz groups undertook a mechanistic study of this system to understand the origins of enantioselectivity. They found that the stereochemistry of the product depends not on two, but four isomeric stereodetermining TSs that differ by the coordination of the aryl group relative to the oxazoline group and the coordination of Pd on the two enantiotopic faces of the enone (Figure 5, bottom).^{129,130} While TS-A was found to be the lowest energy isomer for most systems, the difference in energy between the structures was only 4–8 kJ/mol, well within typical error limits of the computational methodology used. Despite the excellent agreement of the experimental and DFT calculated enantioselectivities,¹²⁹ the fact that multiple TSs and conformations need to be considered renders the computationally expensive DFT an inadequate tool for downstream studies of this reaction, such as high-throughput virtual screening of alternative substrates or chiral ligands. Thus, an alternative method of computationally screening chiral ligand and substrate libraries to rapidly predict enantioselectivities is desirable. Furthermore, the presence of four structurally different but stereomeric TSs, each with its own conformational ensemble, is an interesting test case for the ability of the Q2MM/CatVS method to account for multiple isomeric TSs.

A TSFF was developed for this reaction,¹³¹ and validated by comparing the experimental and calculated enantiomeric ratio (er) for a set of 82 reactions that were not part of the training set. Overall, the TSFF reproduced the experimentally observed enantioselectivities for the 82 diverse examples from the literature with a mean unsigned error (MUE) of 1.8 kJ/mol and an R^2 value of 0.877 between the calculated and experimental er (Figure 6). The points marked in red feature

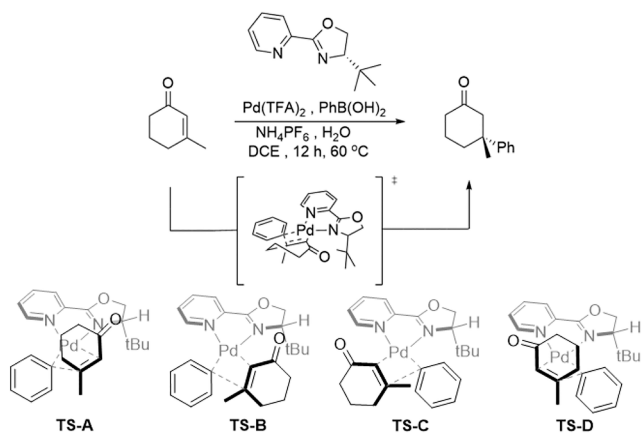


Figure 5. Asymmetric Pd-catalyzed 1,4-addition developed by Stoltz and co-workers (top) and the four isomeric stereodetermining TSs (bottom).^{127,129}

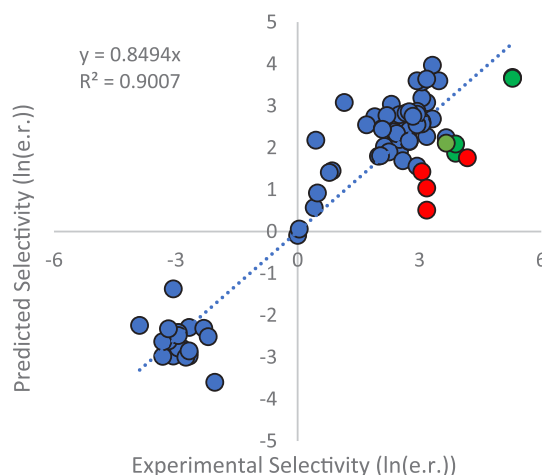
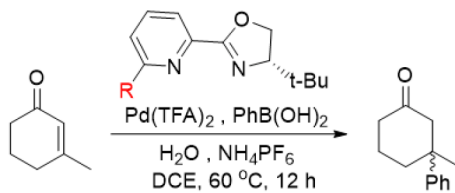


Figure 6. Comparison of experimental and calculated stereo-selectivities.¹³¹

an electron-withdrawing group (EWG) in the meta position, and those in green have an EWG in the para position. Since these data are the largest sources of deviation, it should be noted that all aryl groups were represented by a phenyl ring in the training set, so parameters for substituents come from the published MM3* FF parameters and are likely the origin of the deviation. If these data points are removed from the plot in Figure 6, the MUE improves to 1.4 kJ/mol and R^2 to 0.93 for 73 examples.

After validating the TSFF, it was used to screen a virtual library of (*S*)-*tert*-butyl PyrOx ligands featuring substitution in the 6-position of the pyridine ring (Table 1) that could flip the

Table 1. Predicted er Values for a Virtual Library of 6-Substituted Pyrox Ligands



R-group	% e.r. (pred.)
H	96:4
CH ₃	60.5:39.5
CH(CH ₃) ₂	66.5:33.5
CH ₂ CH ₂ CH ₃	34.5:65.5
CHCH ₂	14:86
CF ₃	6:84
CN	1:99
C(CH ₃) ₃	0.5:99.5

observed stereochemistry of the product while maintaining the stereochemistry of the ligand. Intriguingly, three ligands 6-CF₃, 6-CN, and 6-*tert*-butyl substituted (*S*)-*tert*-butyl PyrOx derivatives, were predicted to give a high er of the opposite stereoisomer.

While it is known that the related quinox ligands give low yields and selectivities,¹³² the use of 6-substituted pyrox ligands for this reaction had not been described in the literature. Experimental validation for the 6-CN and 6-*tert*-butyl PyrOx ligands in the reaction shown in Table 1 did not lead to the desired coupling, yielding instead the homocoupling product, biphenyl, along with unreacted enone. Using an excess (6 equiv) of enone¹³³ and the 6-CF₃ PyrOx ligand, the product of the conjugate addition could be isolated in a yield of 8%, and an er value of 58:42, in favor of the same major enantiomer achieved with the original, nonsubstituted (*S*)-*tert*-butyl PyrOx ligand. While these results do not agree with the computational predictions, they should be interpreted cautiously as the low yield does not exclude the possibility of alternative pathways.

To further understand the poor performance of the ligands featuring substitution in the 6-position of the pyridine, follow-up DFT optimizations on the two or three lowest energy conformations of each of the isomeric TSs were performed for structures including substitution in the 6-position of the pyridine ring. It was discovered that the pyridine(N)–Pd bond

length becomes elongated, causing the ligand to partially dissociate at the TS. Dissociations of sterically encumbered bidentate ligands have been experimentally seen in other systems¹³⁴ and have been hypothesized to be contributing factors in cases in which there were significant differences between experimental and predicted selectivities.¹³⁵ Ultimately, this study emphasizes one of the main limitations of Q2MM and TSFFs in that they are only able to predict selectivities for the stereoselecting TS that they are developed for. Deviations from that TS, such as partial dissociation of the ligand, leads to a disconnect between experiment and prediction. Furthermore, this emphasizes the importance of including suitable structures in the training set to ensure the proper TS is being captured.

CASE STUDY 2: ENANTIOSELECTIVE REDOX-RELAY HECK REACTION

The Sigman group reported the enantioselective Heck arylation of alkenyl alcohols¹³⁶ and, shortly afterward, the enantioselective redox-relay Heck arylations of acyclic alkenyl alcohols that provided remotely functionalized arylated carbonyl products (Figure 7).¹³⁶ This catalyst system showed higher selectivity for the γ -product with excellent er values across a wide scope. Mechanistic studies investigating the chain-walking nature of palladium showed that it proceeded through a series of migratory insertion and β -H eliminations until arriving at the alcohol, leading to a low-energy intermediate.^{10,13,137} The relay process occurs in one direction toward the low-energy intermediate, with the stereochemistry being set during the migratory insertion events based on either *E*- or *Z*-alkene starting material. The stereochemistry is maintained during the chain-walking process, consistent with a previous report wherein a preinstalled stereocenter was maintained during the relay process.¹³⁸ This method was later expanded to the construction of propargylic stereocenters relying on the alkynylation of alkenols.¹³⁹

Due to the wide and evolving scope of alkene, nucleophile, and functional group on the alkene that terminates the chain walking event, as well as the utility of the products, a TSFF was developed to predict the stereoselectivity of redox-relay Heck reactions using Q2MM.¹⁴⁰ This allows for fast stereoselectivity predictions for a wide range of substrates for the redox-relay Heck reaction, provided they proceed through the same stereodetermining migratory insertion TS. The TSFF was validated using published results on five distinct classes of substrates,^{136,138,141–143} resulting in a larger dataset than that typically used with Q2MM. A comparison of the stereoselectivities reported in the literature to the ones obtained from the TSFF calculations is shown in Figure 8.

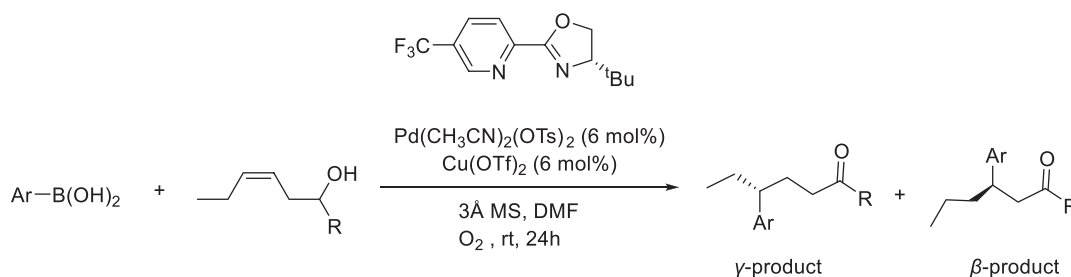


Figure 7. General Redox-Relay Heck reaction of aryl boronic acids and alkenols.

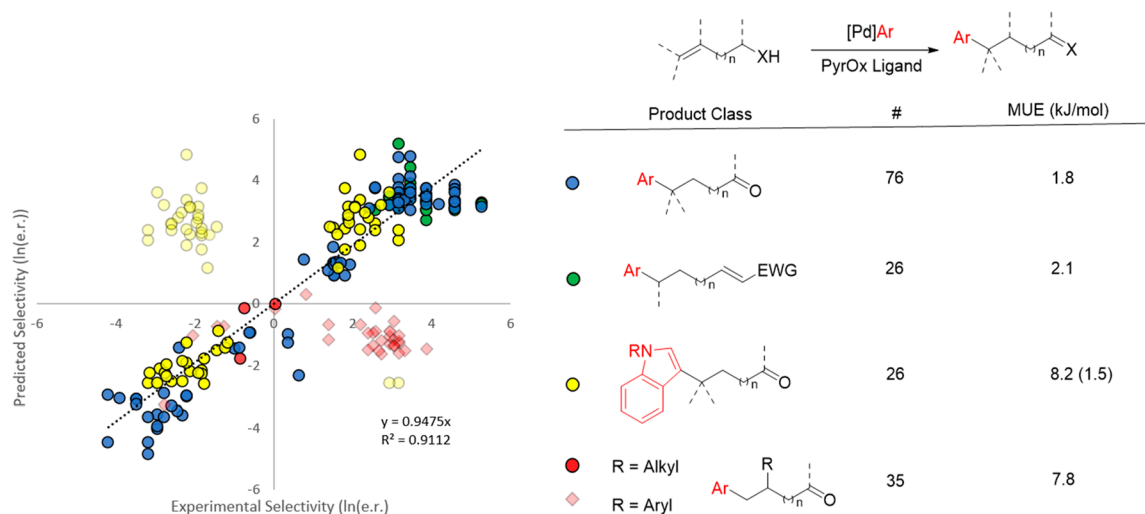


Figure 8. Comparison of 184 predicted and experimental selectivities. Product classes are color coded on the right to show the number of entries in the plot and their corresponding MUE. Light yellow and dark yellow circles indicate C3-alkylated products before and after correction, respectively.

The performance of the TSFF for arylations of internal disubstituted and trisubstituted alkenols resulted in a MUE of 1.8 kJ/mol over 76 examples and contained both cases of high and low enantioselectivity. Additionally, including 26 examples of substrates with a distant EWG, which were not present in the training set, resulted in a MUE of 2.1 kJ/mol. The subset for this reaction class contained relatively few examples of low enantioselectivity, which limits the ability to assess the predictive power of the Q2MM method, as accurately predicting smaller differences in energy between diastereomeric TSs is inherently more difficult. Together, these datasets demonstrate that the TSFF can make accurate selectivity predictions for any disubstituted or trisubstituted alkene substrate, regardless of the nature of the redox acceptor.

There were two sets of substrates that showed poor correlation between the experimental and predicted er: 1,1-disubstituted alkenes and C3-alkylated indoles. These subsets are significantly smaller than the other classes in the validation set, and most examples within the subset have only moderate selectivities. The 1,1-disubstituted alkene set was investigated, and it was noted that previous studies show that the enantioselectivity of the migratory insertion is dictated by the steric difference between the *cis*-oriented alkyl and hydrogen substituents.¹³⁶ However, this analysis does not translate to this class of substrates. An example reaction was further studied by DFT, which calculated a selectivity of 95.5:4.5, in agreement with the experimentally observed er of 98:2. In comparison, the TSFF predicted an er of 20.5:79.5, leading to a discrepancy between the TSFF and the DFT values of 9.4 kJ/mol. Further analysis determined that the origin of this discrepancy is a known inaccuracy in the MM3* FF that had been noted in previous benchmarking studies.^{144,145}

The second substrate class that showed significant discrepancies (MUE = 8.2 kJ/mol) between the TSFF and experimental results is the C3-alkylated indoles. There is a set of outliers, shown in faint yellow in the upper left quadrant of Figure 8, that share the use of PyrOx ligands featuring a (*R*)-CH₂Ph substituent on the oxazoline. The placement of these outliers in the plot with similar values but opposite sign, compared to the reported values, suggests that the degree of selectivity is being accurately reproduced but that the opposite

absolute stereochemistry is predicted. It is noteworthy that reactions using this ligand were reported to have the same absolute stereochemistry as the similar (*S*)-CH₂Ph PyrOx ligand.^{140,141} These contradictory results suggested that a closer evaluation of experimental and computational results was required.

To assess the accuracy of the TSFF prediction, the TSs of the reactions featuring the (*R*)-CH₂Naph, (*S*)-CH₂Ph, and (*R*)-Ph PyrOx ligands were calculated using DFT. The results comparing the experimental, DFT, and TSFF energies are shown in Table 2. For (*R*)-CH₂Naph PyrOx (L1), the DFT

Table 2. Experimental and Calculated Stereoselectivities

ligand	exp.	DFT	TSFF
L1	−7.9 (4:96)	12.5 (0.5:99.5)	4.0 (84:16)
L2	−5.7 (9:91)	−6.8 (94:6)	−4.7 (12.5:87.5)
L3	3.8 (82:18)	9.8 (98.5:1.5)	2.5 (73.5:26.5)

and TSFF results both predict the (*R*)-enantiomer product, while the literature reports the (*S*)-enantiomer. In contrast, the DFT and TSFF energies for the other two ligands were consistent with the results reported in the literature. This discrepancy between the experimental and predicted selectivities suggest that the enantiomeric products were incorrectly assigned for reactions using the (*R*)-CH₂Naph PyrOx ligand. In the initial report, the absolute configurations were not directly determined experimentally but were assigned by comparing to reactions with 3-H and 3-BPin indoles, which proceed through a *syn*-nucleopalladation pathway.¹⁴¹ These results prompted the experimental reinvestigation of the absolute stereochemistry. The attempted crystallization and determination of the absolute stereochemistry of a functionalized product of a C3-alkylated indole was not successful.

Thus, the absolute stereochemistry of these products was determined by comparing the TDDFT-computed (red) and experimental (black) CD spectra (Figure 9) which unambig-

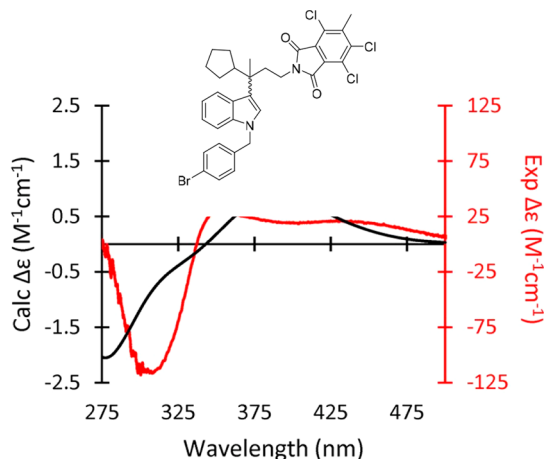


Figure 9. ECD spectra obtained from experiment (black) and TDDFT calculations (red). [Reproduced from ref 140. Copyright 2020, American Chemical Society, Washington, DC.]

uously confirm an absolute stereochemistry of (*R*), in agreement with the computational predictions. Thus, the published stereochemistry for the compounds shown in faint yellow in Figure 8 should be corrected to the deep yellow data points.

Overall, excluding the data for the 1,1-disubstituted alkyl styrenes (for reasons discussed above) and using the correct stereochemistry for C3-alkylation of indoles, the TSFF shows excellent correlation to 151 experimental results (Figure 8). The total MUE over these 151 data points was 1.8 kJ/mol, slightly lower than the MUE obtained with previously developed TSFFs using Q2MM for other reactions.^{88,140} Furthermore, this study showcased a new application of predictive methods for stereochemistry by correcting previously misassigned stereochemistry in the literature. Taking the speed of the TSFF calculation into account, this allows for

high-throughput screening of datasets to ensure the proper stereochemistry has been unambiguously assigned and establishing the computational approach as an equal partner of experiment for the development of enantioselective reactions.

CASE STUDY 3: TSUJI–TROST AMINATION

Asymmetric metal-catalyzed reactions have broadened synthetic accessibility of numerous biologically and biomedically significant compounds. Within this important class of reactions is the Tsuji–Trost allylation, with the ability to generate enantiopure compounds from racemic material.¹⁴⁶ The large number of potential enantiodiscriminate pathways that are dependent on the nature of the reaction components is computationally intriguing for stereochemical prediction and rationalization. Detailed mechanistic examinations of the Tsuji–Trost allylation reaction via DFT have been reported in the literature, giving insight into various stereodetermining factors.^{147,148} Given the complexities involved analyzing the stereodetermination factors, the insight gained from the computational modeling of such a widely used reaction can also aid in validation when experimental methods alone fall short.

Efforts to model the Tsuji–Trost reaction shown in Figure 10, particularly the development of a TSFF for allylic amination using Q2MM/CatVS have proven to be successful. The TSFF was derived from modeling of *endo* and *exo* stereoselecting TSs and subsequent parametrization of a TSFF.¹¹⁸ It was determined that the stereochemistry is set during the nucleophilic attack on the reactive η^3 -allyl palladium intermediate. Phosphite-oxazole ligands were chosen for the training set due to their ability to electronically differentiate atoms within the Pd-allyl intermediate of an unsymmetric allyl palladium complex. This introduces an *endo/exo* isomeric preference in the allylic species formed from 1,3-diphenyl allyl acetate, and directs a distinct nucleophilic attack. Such specificity consolidates potential diastereomeric pathways, essential for accurate determination of a stereoselecting TS and development of a TSFF, especially in the case of the Tsuji–Trost reaction.

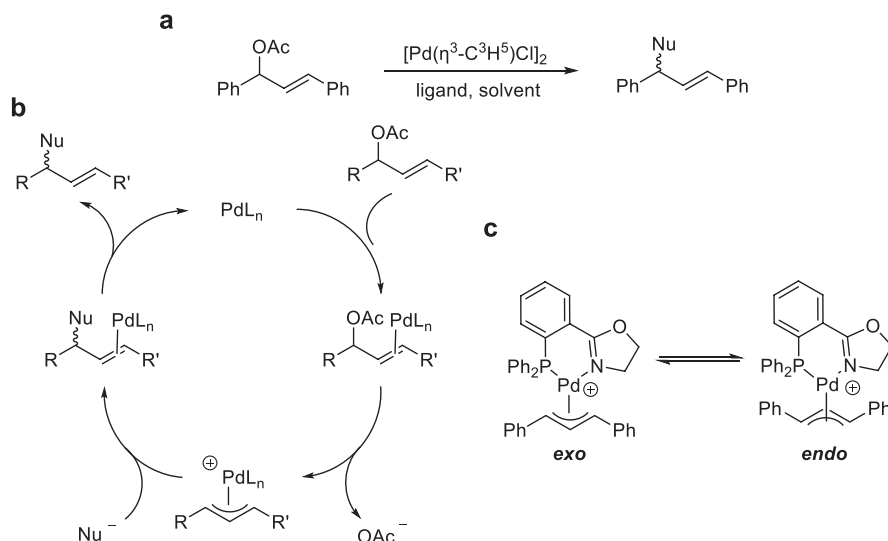


Figure 10. (a) Pd-catalyzed allylation reaction, (b) associated simplified mechanism, and (c) the *exo/endo* isomerization of the η^3 -allyl palladium intermediate.

Application of the TSFF to a dataset of 39 reactions of benzylamine with 1,3-diphenyl allyl complexed to palladium in the presence of 39 previously reported phosphite-oxazole ligands and comparison to the reported experimental results,^{149,150} shown in Figure 11, gave an R^2 value of 0.41

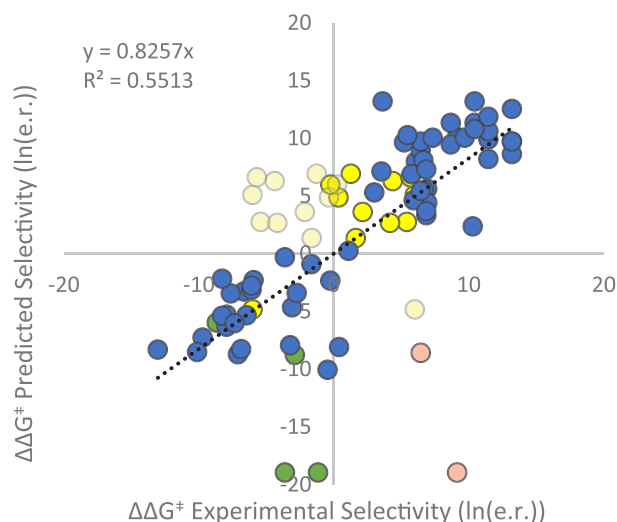


Figure 11. Comparison of relative energies of the experimental values to the calculated MM values after the corrected assignment of 11 data points for the phosphite-oxazole ligands (yellow). Data points before the correction of stereochemistry are shown as faded yellow circles. The plot also contains data corresponding to PHOX ligands (red) and ligands with indole backbone (green).

for the 77 ligand/substrate combinations studied. This value is notably lower than what is observed in previous applications of TSFFs, prompting a detailed analysis of the possible origin of this finding.

A closer analysis of the outliers showed that the majority, shown in yellow in Figure 11, is derived from two studies, including 39 reactions using phosphite-oxazole ligands, 11 of which (shown in yellow in Figure 11) had more similar characteristics than what was discussed in case study 2, i.e., the value of the computed stereoselectivity is similar to the reported value,^{149,150} but the assigned absolute stereochemistry is opposite in the computational and literature results. Experimental studies were done to confirm the absolute configuration in the Pd-catalyzed amination of 1,3-diphenylallyl acetate with benzylamine, using the ligand shown in Figure 12. The major enantiomer of the *N*-benzyl-1,3-diphenylprop-2-en-1-amine product was isolated using chiral HPLC and the configuration confirmed to be (*S*) by comparison with literature data.¹⁵¹ Integration of the HPLC traces determined the *er* value of the reaction to be 92:8. This is opposite to the (*R*) configuration originally reported but is in agreement with the computational prediction by the TSFF,

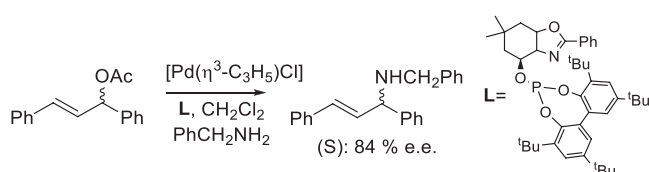


Figure 12. Test reaction for assignment of absolute stereochemistry in the Tsuji–Trost allylic amination.

and therefore the original stereochemical assignment for this series of reactions needs to be changed to the yellow data points, as shown in Figure 11.

Two other data points, shown in light red in Figure 11, also have a significant effect on the calculated R^2 and merit closer analysis. Although the reported¹⁵² (*S*) stereochemistry is based on a crystal structure using anomalous dispersion, closer analysis of the data reveals that the Flack parameter of the structure used for the stereochemistry assignment is 0.1 ± 0.6 , well below the threshold for unambiguous assignment. After the sign of the values shown in light yellow are corrected and the light red data points are excluded due to the uncertain assignment, an R^2 value of 0.71 is obtained. While this is still lower than the results obtained for many other reactions, possibly due to the conformational flexibility of the Pd-allyl complex and the resulting mechanistic complexity of the Tsuji–Trost reaction, the study demonstrates the feasibility of mechanism-based “proofreading” of absolute stereochemical assignments in the literature.

CONCLUSIONS

Over the last decades, the interplay of experimental and computational organic chemistry has moved from rationalization of observations to predictions of reaction outcomes to the point where experiment and computation are now widely seen as synergistic and complementary, each with their own strengths and weaknesses. Because of the ever-increasing speed and accuracy of the computational methods, they are increasingly becoming equal partners to experimental studies. In at least some cases, they can point out problems and errors in the published literature. Therefore, it should not automatically be assumed that, if there is a discrepancy between the computational results and the interpretation of experimental data, the latter is always correct. We argue that while the agreement of experiment and computation is the most common case, detailed investigations of reactions where they disagree is scientifically more fruitful. Although this Perspective is focused on computational studies of the stereochemical outcome of catalytic enantioselective reactions where the interpretation of experimental results often requires detailed analysis, the findings for the case studies are representative to many other applications such as published mechanisms that are corrected based on computational results, which is a topic that goes beyond the scope of the present perspective.

Despite the excellent performance and speed of the Q2MM/CatVS method, the case studies discussed here also highlight some of the caveats that need to be considered in any computational study. When applying surrogate models such as Q2MM-trained TSFFs, it should be noted that changes in mechanism such as those in case study 1 invalidate the underlying model. Case study 2 shows that, despite significant progress, continued development of computational methods such as force fields is needed. Finally, case study 3 shows that complex reaction schemes may lead to lower correlation coefficients between computational and experimental results. Nevertheless, even in those cases, more detailed experimental studies of the computational predictions may lead to the correction of previously published stereochemical assignments.

The advent of data-driven methods such as machine learning in organic synthesis has led to a focus on the quality of published data.^{69,153} The development of accurate methods will require significantly higher data quality that can only be achieved by careful curation of the literature data. The high

speed and physical basis of TSFF calculations is one approach to proofreading for the case of published stereochemistry assignments that is fast enough to be applicable to relatively large datasets and can correct published results where appropriate. While this Perspective has necessarily focused on results from our group, other approaches outlined earlier can provide both complementary approaches to the automated calculated of transition structures and provide a path toward improved correlation methods such as ML-based methods.

■ ASSOCIATED CONTENT

Data Availability Statement

The Q2MM/CatVS code and the associated TSFFs discussed in this contribution are available free of charge at github.com/q2mm.

■ AUTHOR INFORMATION

Corresponding Author

Olaf Wiest – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States; orcid.org/0000-0001-9316-7720; Email: owiest@nd.edu

Authors

Michael P. Maloney – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States; orcid.org/0009-0001-3385-7567

Brock A. Stenfors – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States

Paul Helquist – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States; orcid.org/0000-0003-4380-9566

Per-Ola Norrby – Data Science and Modelling, Pharmaceutical Sciences, R&D, AstraZeneca Gothenburg, SE-431 83 Mölndal, Sweden; orcid.org/0000-0002-2419-0705

Complete contact information is available at: <https://pubs.acs.org/10.1021/acscatal.3c03921>

Author Contributions

[†]M.P.M. and B.A.S. contributed equally. The paper was written through contributions of all authors. All authors have given approval to the final version of the paper.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge the long-term support of the Q2MM/CatVS project by the National Science Foundation, under Grant Nos. CHE-1565669, CHE-1855908 and CHE-2247232. We would like to dedicate the work to Ken Houk, pioneer of predictive computational chemistry, on the occasion of his 80th birthday.

■ REFERENCES

- (1) "Catalysis" and "catalysis and (DFT or computational)", 1992–2022 *Web of Science search*; available via the Internet at: [webofscience.com](https://www.webofscience.com) (accessed April 5, 2023).
- (2) Greeley, J.; Norskov, J. K.; Mavrikakis, M. Electronic structure and catalysis on metal surfaces. *Annu. Rev. Phys. Chem.* **2002**, *53*, 319–348.
- (3) Wang, A. Q.; Li, J.; Zhang, T. Heterogeneous single-atom catalysis. *Nat. Rev. Chem.* **2018**, *2*, 65–81.
- (4) Kamerlin, S. C. L.; Warshel, A. At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis? *Proteins-Struct. Funct. Bioinf.* **2010**, *78*, 1339–1375.
- (5) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (6) Bickelhaupt, F. M.; Houk, K. N. Analyzing reaction rates with the distortion/interaction-activation strain model. *Angew. Chem., Int. Ed.* **2017**, *56*, 10070–10086.
- (7) Houk, K. N.; Gonzalez, J.; Li, Y. Pericyclic reaction transition states—Passions and Punctillios 1935–1995. *Acc. Chem. Res.* **1995**, *28*, 81–90.
- (8) Cheong, P. H. Y.; Legault, C. Y.; Um, J. M.; Celebi-Olcum, N.; Houk, K. N. Quantum Mechanical Investigations of Organocatalysis: Mechanisms, Reactivities, and Selectivities. *Chem. Rev.* **2011**, *111*, 5042–5137.
- (9) Heitler, W.; London, F. Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik. *Z. Phys.* **1927**, *44*, 455–472.
- (10) Xu, L.; Hilton, M. J.; Zhang, X.; Norrby, P.-O.; Wu, Y.-D.; Sigman, M. S.; Wiest, O. Mechanism, reactivity, and selectivity in palladium-catalyzed redox-relay Heck arylations of alkenyl alcohols. *J. Am. Chem. Soc.* **2014**, *136*, 1960–1967.
- (11) Tutkowski, B.; Kerdphon, S.; Lime, E.; Helquist, P.; Andersson, P. G.; Wiest, O.; Norrby, P.-O. Revisiting the stereodetermining step in enantioselective iridium-catalyzed imine hydrogenation. *ACS Catal.* **2018**, *8*, 615–623.
- (12) Bacher, E. P.; Koh, K. J.; Lepore, A. J.; Oliver, A. G.; Wiest, O.; Ashfeld, B. L. A phosphine-mediated dearomative skeletal rearrangement of dianiline squaraine dyes. *Org. Lett.* **2021**, *23*, 2853–2857.
- (13) Hilton, M. J.; Xu, L. P.; Norrby, P. O.; Wu, Y. D.; Wiest, O.; Sigman, M. S. Investigating the nature of palladium chain-walking in the enantioselective redox-relay heck reaction of alkenyl alcohols. *J. Org. Chem.* **2014**, *79*, 11841–11850.
- (14) Tutkowski, B.; Meggers, E.; Wiest, O. Understanding rate acceleration and stereoselection of an asymmetric Giese reaction mediated by a chiral rhodium catalyst. *J. Am. Chem. Soc.* **2017**, *139*, 8062–8065.
- (15) Sigman, M. S.; Harper, K. C.; Bess, E. N.; Milo, A. The Development of Multidimensional Analysis Tools for Asymmetric Catalysis and Beyond. *Acc. Chem. Res.* **2016**, *49*, 1292–1301.
- (16) Meng, S.-S.; Yu, P.; Yu, Y.-Z.; Liang, Y.; Houk, K.; Zheng, W.-H. Computational Design of Enhanced Enantioselectivity in Chiral Phosphoric Acid-Catalyzed Oxidative Desymmetrization of 1, 3-Diol Acetals. *J. Am. Chem. Soc.* **2020**, *142*, 8506–8513.
- (17) Peng, Q.; Duarte, F.; Paton, R. S. Computing organic stereoselectivity—from concepts to quantitative calculations and predictions. *Chem. Soc. Rev.* **2016**, *45*, 6093–6107.
- (18) Nicolaou, K.; Snyder, S. A. Chasing molecules that were never there: misassigned natural products and the role of chemical synthesis in modern structure elucidation. *Angew. Chem., Int. Ed.* **2005**, *44*, 1012–1044.
- (19) Yoo, H.-D.; Nam, S.-J.; Chin, Y.-W.; Kim, M.-S. Misassigned natural products and their revised structures. *Arch. Pharm. Res.* **2016**, *39*, 143–153.
- (20) Bagno, A.; Saielli, G. Addressing the stereochemistry of complex organic molecules by density functional theory-NMR. *WIRE Comput. Mol. Sci.* **2015**, *5*, 228–240.
- (21) Lodewyk, M. W.; Siebert, M. R.; Tantillo, D. J. Computational prediction of ¹H and ¹³C chemical shifts: a useful tool for natural product, mechanistic, and synthetic organic chemistry. *Chem. Rev.* **2012**, *112*, 1839–1862.
- (22) Nazarski, R. B. Summary of DFT calculations coupled with current statistical and/or artificial neural network (ANN) methods to assist experimental NMR data in identifying diastereomeric structures. *Tetrahedron Lett.* **2021**, *71*, 152548.
- (23) Batista, J. M., Jr; Blanch, E. W.; da Silva Bolzani, V. Recent advances in the use of vibrational chiroptical spectroscopic methods

for stereochemical characterization of natural products. *Nat. Prod. Rep.* **2015**, *32*, 1280–1302.

(24) Kutateladze, A. G. Structure Revision of Decurrensides A–E Enabled by the RFF Parametric Calculations of Proton Spin–Spin Coupling Constants. *J. Org. Chem.* **2016**, *81*, 8659–8661.

(25) Chhetri, B. K.; Lavoie, S.; Sweeney-Jones, A. M.; Kubanek, J. Recent trends in the structural revision of natural products. *Nat. Prod. Rep.* **2018**, *35*, 514–531.

(26) Liu, Y.; Holt, T. A.; Kutateladze, A.; Newhouse, T. R. Stereochemical revision of xylogranatin F by GIAO and DU8+ NMR calculations. *Chirality* **2020**, *32*, 515–523.

(27) Hammett, L. P. The Effect of Structure upon the Reactions of Organic Compounds. Benzene Derivatives. *J. Am. Chem. Soc.* **1937**, *59*, 96–103.

(28) Taft, R. W. Linear Steric Energy Relationships. *J. Am. Chem. Soc.* **1953**, *75*, 4538–4539.

(29) Taft, R. W. Linear Free Energy Relationships from Rates of Esterification and Hydrolysis of Aliphatic and Ortho-Substituted Benzoate Esters. *J. Am. Chem. Soc.* **1952**, *74*, 2729–2732.

(30) Charton, M. The Nature of the Ortho Effect. II. Composition of the Taft Steric Parameters. *J. Am. Chem. Soc.* **1969**, *91*, 615–618.

(31) Charton, M.; Steric Effects, I. I. I. Bimolecular Nucleophilic Substitution. *J. Am. Chem. Soc.* **1975**, *97*, 3694–3697.

(32) Marcus, R. A. On the Theory of Oxidation-Reduction Reactions Involving Electron Transfer. *Int. J. Chem. Phys.* **1956**, *24*, 966.

(33) Marcus, R. A. Reflections on Electron Transfer Theory. *J. Chem. Phys.* **2020**, *153*, 210401.

(34) Lewis, E. S. Rate-equilibrium LFER characterization of transition states: The interpretation of α . *J. Phys. Org. Chem.* **1990**, *3*, 1–8.

(35) Mihai, C.; Kravchuk, A. V.; Tsai, M. D.; Bruzik, K. S. Application of Brønsted-Type LFER in the Study of the Phospholipase C Mechanism. *J. Am. Chem. Soc.* **2003**, *125*, 3236–3242.

(36) Carotti, A.; Altomare, C.; Cellamare, S.; Monforte, A.; Bettoni, G.; Loiodice, F.; Tangari, N.; Tortorella, V. LFER and CoMFA Studies on Optical Resolution of α -Alkyl α -Aryloxy Acetic Acid Methyl Esters on DACH-DNB Chiral Stationary Phase. *J. Comput. Aided Des.* **1995**, *9*, 131–138.

(37) 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*, 948–954.

(38) Oslob, J. D.; Åkermark, B.; Helquist, P.; Norrby, P.-O. Steric influences on the selectivity in palladium-catalyzed allylation. *Organometallics* **1997**, *16*, 3015–3021.

(39) Kozlowski, M. C.; Dixon, S. L.; Panda, M.; Lauri, G. Quantum mechanical models correlating structure with selectivity: Predicting the enantioselectivity of β -amino alcohol catalysts in aldehyde alkylation. *J. Am. Chem. Soc.* **2003**, *125*, 6614–6615.

(40) Ianni, J. C.; Annamalai, V.; Phuan, P.-W.; Panda, M.; Kozlowski, M. C. A Priori Theoretical Prediction of Selectivity in Asymmetric Catalysis: Design of Chiral Catalysts by Using Quantum Molecular Interaction Fields. *Angew. Chem., Int. Ed.* **2006**, *45*, 5502–5505.

(41) Huang, J.; Ianni, J. C.; Antoline, J. E.; Hsung, R. P.; Kozlowski, M. C. De Novo Chiral Amino Alcohols in Catalyzing Asymmetric Additions to Aryl Aldehydes. *Org. Lett.* **2006**, *8*, 1565–1568.

(42) Lipkowitz, K. B.; Pradhan, M. Computational studies of chiral catalysts: A Comparative Molecular Field Analysis of an asymmetric Diels-Alder reaction with catalysts containing bisoxazoline or phosphinooxazoline ligands. *J. Org. Chem.* **2003**, *68*, 4648–4656.

(43) Zahrt, A. F.; Athavale, S. V.; Denmark, S. E. Quantitative structure–selectivity relationships in enantioselective catalysis: past, present, and future. *Chem. Rev.* **2020**, *120*, 1620–1689.

(44) Henle, J. J.; Zahrt, A. F.; Rose, B. T.; Darrow, W. T.; Wang, Y.; Denmark, S. E. Development of a Computer-Guided Workflow for

Catalyst Optimization. Descriptor Validation, Subset Selection, and Training Set Analysis. *J. Am. Chem. Soc.* **2020**, *142*, 11578–11592.

(45) Zahrt, A. F.; Henle, J. J.; Rose, B. T.; Wang, Y.; Darrow, W. T.; Denmark, S. E. Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. *Science* **2019**, *363*, 1–11.

(46) Harper, K. C.; Sigman, M. S. Three-Dimensional Correlation of Steric and Electronic Free Energy Relationships Guides Asymmetric Propargylation. *Science* **2011**, *333*, 1875–1878.

(47) Harper, K. C.; Sigman, M. S. Predicting and Optimizing Asymmetric Catalyst Performance Using the Principles of Experimental Design and Steric Parameters. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 2179–2183.

(48) Reid, J. P.; Sigman, M. S. Holistic Prediction of Enantioselectivity in Asymmetric Catalysis. *Nature* **2019**, *571*, 343–348.

(49) Gensch, T.; dos Passos Gomes, G.; Friederich, P.; Peters, E.; Gaudin, T.; Pollice, R.; Jorner, K.; Nigam, A.; Lindner-D'Addario, M.; Sigman, M. S.; Aspuru-Guzik, A. A comprehensive discovery platform for organophosphorus ligands for catalysis. *J. Am. Chem. Soc.* **2022**, *144*, 1205–1217.

(50) Reid, J. P.; Sigman, M. S. Comparing quantitative prediction methods for the discovery of small-molecule chiral catalysts. *Nat. Rev. Chem.* **2018**, *2*, 290–305.

(51) Williams, W. L.; Zeng, L.; Gensch, T.; Sigman, M. S.; Doyle, A. G.; Anslyn, E. V. The evolution of data-driven modeling in organic chemistry. *ACS Central Sci.* **2021**, *7*, 1622–1637.

(52) Reid, J. P.; Proctor, R. S. J.; Sigman, M. S.; Phipps, R. J. Predictive Multivariate Linear Regression Analysis Guides Successful Catalytic Enantioselective Minisci Reactions of Diazines. *J. Am. Chem. Soc.* **2019**, *141*, 19178–19185.

(53) Durand, D. J.; Fey, N. Computational Ligand Descriptors for Catalyst Design. *Chem. Rev.* **2019**, *119*, 6561–6594.

(54) Durand, D. J.; Fey, N. Building a Toolbox for the Analysis and Prediction of Ligand and Catalyst Effects in Organometallic Catalysis. *Acc. Chem. Res.* **2021**, *54*, 837–848.

(55) Fey, N. The contribution of computational studies to organometallic catalysis: descriptors, mechanisms and models. *Dalton Trans.* **2010**, *39*, 296–310.

(56) Fey, N. Lost in chemical space? Maps to support organometallic catalysis. *Chem. Central J.* **2015**, *9*, 38.

(57) Fey, N.; Tsipis, A. C.; Harris, S. E.; Harvey, J. N.; Orpen, A. G.; Mansson, R. A. Development of a ligand knowledge base, Part 1: Computational descriptors for phosphorus donor ligands. *Chem.—Eur. J.* **2006**, *12*, 291–302.

(58) Jover, J.; Fey, N.; Harvey, J. N.; Lloyd-Jones, G. C.; Orpen, A. G.; Owen-Smith, G. J. J.; Murray, P.; Hose, D. R. J.; Osborne, R.; Purdie, M. Expansion of the ligand knowledge base for chelating P₂-donor ligands (LKB-PP). *Organometallics* **2012**, *31*, S302–S306.

(59) Fey, N.; Koumi, A.; Malkov, A. V.; Moseley, J. D.; Nguyen, B. N.; Tyler, S. N. G.; Willans, C. E. Mapping the properties of bidentate ligands with calculated descriptors (LKB-bid). *Dalton Trans.* **2020**, *49*, 8169–8178.

(60) Fey, N.; Haddow, M. F.; Harvey, J. N.; McMullin, C. L.; Orpen, A. G. A ligand knowledge base for carbenes (LKB-C): maps of ligand space. *Dalton Trans.* **2009**, *38*, 8183–8196.

(61) Li, G.; Dong, Y.; Reetz, M. T. Can machine learning revolutionize directed evolution of selective enzymes? *Adv. Synth. Catal.* **2019**, *361*, 2377–2386.

(62) Gallarati, S.; Fabregat, R.; Laplaza, R.; Bhattacharjee, S.; Wodrich, M. D.; Corminboeuf, C. Reaction-based machine learning representations for predicting the enantioselectivity of organocatalysts. *Chem. Sci.* **2021**, *12*, 6879–6889.

(63) Chen, H.; Yamaguchi, S.; Morita, Y.; Nakao, H.; Zhai, X.; Shimizu, Y.; Mitsunuma, H.; Kanai, M. Data-driven catalyst optimization for stereodivergent asymmetric synthesis by iridium/boron hybrid catalysis. *Cell Rep. Phys. Sci.* **2021**, *2*, 100679–100679.

(64) Yamaguchi, S. Molecular field analysis for data-driven molecular design in asymmetric catalysis. *Org. Biomol. Chem.* **2022**, *20*, 6057–6071.

- (65) Yamaguchi, S.; Sodeoka, M. Molecular Field Analysis Using Intermediates in Enantio-Determining Steps Can Extract Information for Data-Driven Molecular Design in Asymmetric Catalysis. *Bull. Chem. Soc. Jpn.* **2019**, *92*, 1701–1706.
- (66) Li, J.; Zhang, D.; Wang, Y.; Ye, C.; Xu, H.; Hong, P. Predicting the Stereoselectivity of Chemical Transformations by Machine Learning. *ArXiv* **2021**, DOI: 10.48550/arXiv.2110.05671.
- (67) Moon, S.; Chatterjee, S.; Seeberger, P. H.; Gilmore, K. Predicting glycosylation stereoselectivity using machine learning. *Chem. Sci.* **2021**, *12*, 2931–2939.
- (68) Maloney, M. P.; Coley, C. W.; Genheden, S.; Carson, N.; Helquist, P.; Norrby, P.-O.; Wiest, O. Negative Data in Data Sets for Machine Learning Training. *J. Org. Chem.* **2023**, *88*, 5239–5241.
- (69) Saebi, M.; Nan, B.; Herr, J. E.; Wahlers, J.; Guo, Z.; Zuranski, A. M.; Kogej, T.; Norrby, P.-O.; Doyle, A. G.; Chawla, N. V.; Wiest, O. On the use of real-world datasets for reaction yield prediction. *Chem. Sci.* **2023**, *14*, 4997–5005.
- (70) Norrby, P.-O.; Rasmussen, T.; Haller, J.; Strassner, T.; Houk, K. N. Rationalizing the Stereoselectivity of Osmium Tetroxide Asymmetric Dihydroxylations with Transition State Modeling Using Quantum Mechanics-Guided Molecular Mechanics. *J. Am. Chem. Soc.* **1999**, *121*, 10186–10192.
- (71) Houk, K. N.; Paddon-Row, M. N.; Rondan, N. G.; Wu, Y.-D.; Brown, F. K.; Spellmeyer, D. C.; Metz, J. T.; Li, Y.; Loncharich, R. J. Theory and modeling of stereoselective organic reactions. *Science* **1986**, *231* (4742), 1108–1117.
- (72) Houk, K. N.; Cheong, P. H.-Y. Computational prediction of small-molecule catalysts. *Nature* **2008**, *455*, 309–313.
- (73) Lam, Y.-h.; Grayson, M. N.; Holland, M. C.; Simon, A.; Houk, K. Theory and modeling of asymmetric catalytic reactions. *Acc. Chem. Res.* **2016**, *49*, 750–762.
- (74) Andreola, L. R.; Wheeler, S. E. Importance of favourable non-covalent contacts in the stereoselective synthesis of tetrasubstituted chromanones. *Org. Chem. Front.* **2022**, *9*, 3027–3033.
- (75) Grimme, S. Density functional theory with London dispersion corrections. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1* (2), 211–228.
- (76) Zhao, Y.; Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **2008**, *41*, 157–167.
- (77) Paton, R. S. Dissecting non-covalent interactions in oxazaborolidinium catalyzed cycloadditions of maleimides. *Org. Biomol. Chem.* **2014**, *12*, 1717–1720.
- (78) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. Revealing noncovalent interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498–6506.
- (79) Arbour, J. L.; Rzepa, H. S.; Contreras-García, J.; Adrio, L. A.; Barreiro, E. M.; Hii, K. K. Silver-Catalysed Enantioselective Addition of O-H and N-H Bonds to Allenes: A New Model for Stereoselectivity Based on Noncovalent Interactions. *Chem.—Eur. J.* **2012**, *18*, 11317–11324.
- (80) Guan, Y.; Ingman, V. M.; Rooks, B. J.; Wheeler, S. E. AARON: An Automated Reaction Optimizer for New Catalysts. *J. Chem. Theory Comput.* **2018**, *14*, 5249–5261.
- (81) Guan, Y.; Wheeler, S. E. Automated Quantum Mechanical Predictions of Enantioselectivity in a Rhodium-Catalyzed Asymmetric Hydrogenation. *Angew. Chem., Int. Ed.* **2017**, *56*, 9101–9105.
- (82) Doney, A. C.; Rooks, B. J.; Lu, T.; Wheeler, S. E. Design of organocatalysts for asymmetric propargylations through computational screening. *ACS Catal.* **2016**, *6*, 7948–7955.
- (83) Wheeler, S. E.; Seguin, T. J.; Guan, Y.; Doney, A. C. Noncovalent interactions in organocatalysis and the prospect of computational catalyst design. *Acc. Chem. Res.* **2016**, *49*, 1061–1069.
- (84) Torres, J. A. G.; Lau, S. H.; Anchuri, P.; Stevens, J. M.; Tabora, J. E.; Li, J.; Borovika, A.; Adams, R. P.; Doyle, A. G. A Multi-Objective Active Learning Platform and Web App for Reaction Optimization. *J. Am. Chem. Soc.* **2022**, *144*, 19999–20007.
- (85) Grayson, M. N.; Pellegrinet, S. C.; Goodman, J. M. Mechanistic insights into the BINOL-derived phosphoric acid-catalyzed asymmetric allylboration of aldehydes. *J. Am. Chem. Soc.* **2012**, *134*, 2716–2722.
- (86) Lee, J.-Y.; Miller, J. J.; Hamilton, S. S.; Sigman, M. S. Stereochemical diversity in chiral ligand design: discovery and optimization of catalysts for the enantioselective addition of allylic halides to aldehydes. *Org. Lett.* **2005**, *7*, 1837–1839.
- (87) Wahlers, J.; Rosales, A. R.; Berkel, N.; Forbes, A.; Helquist, P.; Norrby, P.-O.; Wiest, O. A Quantum-Guided Molecular Mechanics Force Field for the Ferrocene Scaffold. *J. Org. Chem.* **2022**, *87*, 12334–12341.
- (88) Hansen, E.; Rosales, A. R.; Tutkowski, B.; Norrby, P. O.; Wiest, O. Prediction of Stereochemistry using Q2MM. *Acc. Chem. Res.* **2016**, *49*, 996–1005.
- (89) Rosales, A. R.; Wahlers, J.; Lime, E.; Meadows, R. E.; Leslie, K. W.; Savin, R.; Bell, F.; Hansen, E.; Helquist, P.; Munday, R. H.; Wiest, O.; Norrby, P. O. Rapid virtual screening of enantioselective catalysts using CatVS. *Nat. Catal.* **2019**, *2*, 41–45.
- (90) London, F. Quantenmechanische Deutung des Vorganges der Aktivierung. *Z. Elektrochem. Angew. Phys. Chem.* **1929**, *35*, 552–555.
- (91) Eyring, H.; Polanyi, M. Zur Berechnung der Aktivierungswärme. *Die Naturwissenschaften* **1930**, *18*, 914–915.
- (92) Aqvist, J.; Warshel, A. Simulation of Enzyme Reactions using Valence Bond Force Fields and other Hybrid Quantum-Classical Approaches. *Chem. Rev.* **1993**, *93*, 2523–2544.
- (93) Truhlar, D. G. Reply to comment on molecular mechanics for chemical reactions. *J. Phys. Chem. A* **2002**, *106*, 5048–5050.
- (94) Florián, J. Comment on Molecular Mechanics for Chemical Reactions. *J. Phys. Chem. A* **2002**, *106*, 5046–5047.
- (95) Weill, N.; Corbeil, C. R.; De Schutter, J. W.; Moitessier, N. Toward a computational tool predicting the stereochemical outcome of asymmetric reactions: Development of the molecular mechanics-based program ACE and application to asymmetric epoxidation reactions. *J. Comput. Chem.* **2011**, *32*, 2878–2889.
- (96) Garbisch, E. W., Jr.; Schildcrout, S. M.; Patterson, D. B.; Sprecher, C. M. Strain Effects II. Diimide Reductions of Olefins. *J. Am. Chem. Soc.* **1965**, *87*, 2932–2944.
- (97) Eksterowicz, J. E.; Houk, K. N. Transition-state modeling with empirical force fields. *Chem. Rev.* **1993**, *93*, 2439–2461.
- (98) Wu, Y. D.; Houk, K. N. Transition structures for hydride transfers. *J. Am. Chem. Soc.* **1987**, *109*, 906–908.
- (99) Madarasz, A.; Berta, D.; Paton, R. S. Development of a true transition state force field from quantum mechanical calculations. *J. Chem. Theor. Comput.* **2016**, *12*, 1833–1844.
- (100) Norrby, P.-O.; Brandt, P.; Rein, T. Rationalization of Product Selectivities in Asymmetric Horner–Wadsworth–Emmons Reactions by Use of a New Method for Transition-State Modeling. *J. Org. Chem.* **1999**, *64*, 5845–5852.
- (101) Limé, E.; Norrby, P.-O. Improving the Q2MM method for transition state force field modeling. *J. Comput. Chem.* **2015**, *36*, 244–250.
- (102) Quinn, T. R.; Patel, H. N.; Koh, K. H.; Haines, B. E.; Norrby, P. O.; Helquist, P.; Wiest, O. Automated fitting of transition state force fields for biomolecular simulations. *PLoS One* **2022**, *17*, No. e0264960.
- (103) Quinn, T. R.; Steussy, C. N.; Haines, B. E.; Lei, J. P.; Wang, W.; Sheong, F.; Stauffacher, C. V.; Huang, X. H.; Norrby, P. O.; Helquist, P.; Wiest, O. Microsecond timescale MD simulations at the transition state of PmHMGR predict remote allosteric residues. *Chem. Sci.* **2021**, *12*, 6413–6418.
- (104) github.com/q2mm.
- (105) Lii, J.-H.; Allinger, N. L. Molecular mechanics. The MM3 force field for hydrocarbons. 2. Vibrational frequencies and thermodynamics. *J. Am. Chem. Soc.* **1989**, *111*, 8566–8575.
- (106) Allinger, N. L.; Yuh, Y.; Lii, J.-H. Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
- (107) Norrby, P.-O.; Brandt, P. Deriving force field parameters for coordination complexes. *Coord. Chem. Rev.* **2001**, *212*, 79–109.

- (108) Sherrod, M. J.; Menger, F. M. "Transition-state modeling" does not always model transition states. *J. Am. Chem. Soc.* **1989**, *111*, 2611–2613.
- (109) Chang, G.; Guida, W. C.; Still, W. C. An internal-coordinate Monte Carlo method for searching conformational space. *J. Am. Chem. Soc.* **1989**, *111*, 4379–4386.
- (110) Kolossváry, I.; Guida, W. C. Low Mode Search. An Efficient, Automated Computational Method for Conformational Analysis: Application to Cyclic and Acyclic Alkanes and Cyclic Peptides. *J. Am. Chem. Soc.* **1996**, *118*, 5011–5019.
- (111) Rosales, A. R.; Quinn, T. R.; Wahlers, J.; Tomberg, A.; Zhang, X.; Helquist, P.; Wiest, O.; Norrby, P. O. Application of Q2MM to predictions in stereoselective synthesis. *Chem. Commun.* **2018**, *54*, 8294–8311.
- (112) Fristrup, P.; Jensen, G. H.; Andersen, M. L. N.; Tanner, D.; Norrby, P.-O. Combining Q2MM modeling and kinetic studies for refinement of the osmium-catalyzed asymmetric dihydroxylation (AD) mnemonic. *J. Organomet. Chem.* **2006**, *691*, 2182–2198.
- (113) Fristrup, P.; Tanner, D.; Norrby, P.-O. Updating the asymmetric osmium-catalyzed dihydroxylation (AD) mnemonic: Q2MM modeling and new kinetic measurements. *Chirality* **2003**, *15*, 360–368.
- (114) Donoghue, P. J.; Helquist, P.; Norrby, P.-O.; Wiest, O. Prediction of Enantioselectivity in Rhodium Catalyzed Hydrogenations. *J. Am. Chem. Soc.* **2009**, *131*, 410–411.
- (115) Donoghue, P. J.; Helquist, P.; Norrby, P. O.; Wiest, O. Development of a Q2MM force field for the asymmetric rhodium catalyzed hydrogenation of enamides. *J. Chem. Theor. Comput.* **2008**, *4*, 1313–1323.
- (116) Le, D. N.; Hansen, E.; Khan, H. A.; Kim, B.; Wiest, O.; Dong, V. M. Hydrogenation catalyst generates cyclic peptide stereocenters in sequence. *Nat. Chem.* **2018**, *10*, 968–973.
- (117) Lime, E.; Lundholm, M. D.; Forbes, A.; Wiest, O.; Helquist, P.; Norrby, P. O. Stereoselectivity in Asymmetric Catalysis: The Case of Ruthenium-Catalyzed Ketone Hydrogenation. *J. Chem. Theor. Comput.* **2014**, *10*, 2427–2435.
- (118) Wahlers, J.; Margalef, J.; Hansen, E.; Bayesteh, A.; Helquist, P.; Dieguez, M.; Pamies, O.; Wiest, O.; Norrby, P. O. Proofreading experimentally assigned stereochemistry through Q2MM predictions in Pd-catalyzed allylic aminations. *Nat. Commun.* **2021**, *12*, 6719.
- (119) Lee, J. M.; Zhang, X.; Norrby, P. O.; Helquist, P.; Wiest, O. Stereoselectivity in (Acyloxy)borane-Catalyzed Mukaiyama Aldol Reactions. *J. Org. Chem.* **2016**, *81*, 5314–5321.
- (120) Rasmussen, T.; Norrby, P.-O. Modeling the Stereoselectivity of the β -Amino Alcohol-Promoted Addition of Dialkylzinc to Aldehydes. *J. Am. Chem. Soc.* **2003**, *125*, 5130–5138.
- (121) Plata, R. E.; Singleton, D. A. A case study of the mechanism of alcohol-mediated Morita Baylis–Hillman reactions. The importance of experimental observations. *J. Am. Chem. Soc.* **2015**, *137*, 3811–3826.
- (122) Gini, F.; Hessen, B.; Minnaard, A. J. Palladium-catalyzed enantioselective conjugate addition of arylboronic acids. *Org. Lett.* **2005**, *7*, 5309–5312.
- (123) Nishikata, T.; Yamamoto, Y.; Miyaura, N. Conjugate Addition of Aryl Boronic Acids to Enones Catalyzed by Cationic Palladium-(II)–Phosphane Complexes. *Angew. Chem., Int. Ed.* **2003**, *42*, 2768–2770.
- (124) Nishikata, T.; Yamamoto, Y.; Miyaura, N. 1,4-Addition of Arylboronic Acids and Arylsiloxanes to α,β -Unsaturated Carbonyl Compounds via Transmetalation to Dicationic Palladium(II) Complexes. *Organometallics* **2004**, *23*, 4317–4324.
- (125) Lu, X.; Lin, S. Pd(II)-Bipyridine Catalyzed Conjugate Addition of Arylboronic Acid to α,β -Unsaturated Carbonyl Compounds. *J. Org. Chem.* **2005**, *70*, 9651–9653.
- (126) Lin, S.; Lu, X. Palladium-bipyridine catalyzed conjugate addition of arylboronic acids to α,β -unsaturated carbonyl compounds in aqueous media. *Tetrahedron Lett.* **2006**, *47*, 7167–7170.
- (127) Kikushima, K.; Holder, J. C.; Gatti, M.; Stoltz, B. M. Palladium-Catalyzed Asymmetric Conjugate Addition of Arylboronic Acids to Five-, Six-, and Seven-Membered β -Substituted Cyclic Enones: Enantioselective Construction of All-Carbon Quaternary Stereocenters. *J. Am. Chem. Soc.* **2011**, *133*, 6902–6905.
- (128) Gottumukkala, A. L.; Matcha, K.; Lutz, M.; De Vries, J. G.; Minnaard, A. J. Palladium-Catalyzed Asymmetric Quaternary Stereocenter Formation. *Chem.—Eur. J.* **2012**, *18*, 6907–6914.
- (129) Holder, J. C.; Zou, L.; Marziale, A. N.; Liu, P.; Lan, Y.; Gatti, M.; Kikushima, K.; Houk, K. N.; Stoltz, B. M. Mechanism and enantioselectivity in palladium-catalyzed conjugate addition of arylboronic acids to β -substituted cyclic enones: Insights from computation and experiment. *J. Am. Chem. Soc.* **2013**, *135*, 14996–15007.
- (130) Boeser, C. L.; Holder, J. C.; Taylor, B. L. H.; Houk, K. N.; Stoltz, B. M.; Zare, R. N. Mechanistic analysis of an asymmetric palladium-catalyzed conjugate addition of arylboronic acids to β -substituted cyclic enones. *Chem. Sci.* **2015**, *6*, 1917–1922.
- (131) Wahlers, J.; Maloney, M.; Salahi, F.; Rosales, A. R.; Helquist, P.; Norrby, P. O.; Wiest, O. Stereoselectivity Predictions for the Pd-Catalyzed 1,4-Conjugate Addition Using Quantum-Guided Molecular Mechanics. *J. Org. Chem.* **2021**, *86*, 5660–5667.
- (132) Holder, J. C.; Goodman, E. D.; Kikushima, K.; Gatti, M.; Marziale, A. N.; Stoltz, B. M. Synthesis of diverse β -quaternary ketones via palladium-catalyzed asymmetric conjugate addition of arylboronic acids to cyclic enones. *Tetrahedron* **2015**, *71*, 5781–5792.
- (133) Buter, J.; Moezelaar, R.; Minnaard, A. J. Enantioselective palladium catalyzed conjugate additions of ortho-substituted arylboronic acids to β,β -disubstituted cyclic enones: Total synthesis of herbertenediol, enokipodin A and enokipodin B. *Org. Biomol. Chem.* **2014**, *12*, 5883–5890.
- (134) Gridnev, I. D.; Imamoto, T.; Hoge, G.; Kouchi, M.; Takahashi, H. Asymmetric Hydrogenation Catalyzed by a Rhodium Complex of (R)-(tert-Butylmethylphosphino)(di-tert-butylphosphino)methane: Scope of Enantioselectivity and Mechanistic Study. *J. Am. Chem. Soc.* **2008**, *130*, 2560–2572.
- (135) Donoghue, P. J.; Helquist, P.; Norrby, P. O.; Wiest, O. Development of a Q2MM force field for the asymmetric rhodium catalyzed hydrogenation of enamides. *J. Chem. Theory Comput.* **2008**, *4*, 1313–1323.
- (136) Werner, E. W.; Mei, T. S.; Burckle, A. J.; Sigman, M. S. Enantioselective heck arylations of acyclic alkenyl alcohols using a redox-relay strategy. *Science* **2012**, *338*, 1455–1458.
- (137) Hilton, M. J.; Cheng, B.; Buckley, B. R.; Xu, L. P.; Wiest, O.; Sigman, M. S. Relative reactivity of alkenyl alcohols in the palladium-catalyzed redox-relay Heck reaction. *Tetrahedron* **2015**, *71*, 6513–6518.
- (138) Mei, T. S.; Patel, H. H.; Sigman, M. S. Enantioselective construction of remote quaternary stereocenters. *Nature* **2014**, *508*, 340–344.
- (139) Chen, Z. M.; Nervig, C. S.; DeLuca, R. J.; Sigman, M. S. Palladium-Catalyzed Enantioselective Redox-Relay Heck Alkynylation of Alkenols To Access Propargylic Stereocenters. *Angew. Chem., Int. Ed.* **2017**, *56*, 6651–6654.
- (140) Rosales, A. R.; Ross, S. P.; Helquist, P.; Norrby, P. O.; Sigman, M. S.; Wiest, O. Transition State Force Field for the Asymmetric Redox-Relay Heck Reaction. *J. Am. Chem. Soc.* **2020**, *142*, 9700–9707.
- (141) Zhang, C.; Santiago, C. B.; Crawford, J. M.; Sigman, M. S. Enantioselective Dehydrogenative Heck Arylations of Trisubstituted Alkenes with Indoles to Construct Quaternary Stereocenters. *J. Am. Chem. Soc.* **2015**, *137*, 15668–15671.
- (142) Zhang, C.; Santiago, C. B.; Kou, L.; Sigman, M. S. Alkenyl Carbonyl Derivatives in Enantioselective Redox Relay Heck Reactions: Accessing α,β -Unsaturated Systems. *J. Am. Chem. Soc.* **2015**, *137*, 7290–7293.
- (143) Mei, T. S.; Werner, E. W.; Burckle, A. J.; Sigman, M. S. Enantioselective redox-relay oxidative heck arylations of acyclic alkenyl alcohols using boronic acids. *J. Am. Chem. Soc.* **2013**, *135*, 6830–6833.

- (144) Paton, R. S.; Goodman, J. M. Hydrogen bonding and π -stacking: How reliable are force fields? A critical evaluation of force field descriptions of nonbonded interactions. *J. Chem. Inf. Model.* **2009**, *49*, 944–955.
- (145) Sherrill, C. D.; Sumpter, B. G.; Sinnokrot, M. O.; Marshall, M. S.; Hohenstein, E. G.; Walker, R. C.; Gould, I. R. Assessment of standard force field models against high-quality ab initio potential curves for prototypes of π - π , CH/ π , and SH/ π interactions. *J. Comput. Chem.* **2009**, *30*, 2187–2193.
- (146) Trost, B. M.; Crawley, M. L. Asymmetric Transition-Metal-Catalyzed Allylic Alkylations: Applications in Total Synthesis. *Chem. Rev.* **2003**, *103*, 2921–2944.
- (147) Cusumano, A. Q.; Stoltz, B. M.; Goddard, W. A., III Reaction mechanism, origins of enantioselectivity, and reactivity trends in asymmetric allylic alkylation: a comprehensive quantum mechanics investigation of a C(sp₃)-C(sp₃) cross-coupling. *J. Am. Chem. Soc.* **2020**, *142*, 13917–13933.
- (148) McPherson, K. E.; Croatt, M. P.; Morehead, A. T.; Sargent, A. L. DFT Mechanistic Investigation of an Enantioselective Tsuji–Trost Allylation Reaction. *Organometallics* **2018**, *37*, 3791–3802.
- (149) Mazuela, J.; Pàmies, O.; Diéguez, M. A New Modular Phosphite-Pyridine Ligand Library for Asymmetric Pd-Catalyzed Allylic Substitution Reactions: A Study of the Key Pd- π -Allyl Intermediates. *Chem.—Eur. J.* **2013**, *19*, 2416–2432.
- (150) Diéguez, M.; Pàmies, O. Modular Phosphite–Oxazoline/Oxazine Ligand Library for Asymmetric Pd-Catalyzed Allylic Substitution Reactions: Scope and Limitations—Origin of Enantioselectivity. *Chem.—Eur. J.* **2008**, *14*, 3653–3669.
- (151) Popa, D.; Marcos, R.; Sayalero, S.; Vidal-Ferran, A.; Pericàs, M. A. Towards Continuous Flow, Highly Enantioselective Allylic Amination: Ligand Design, Optimization and Supporting. *Adv. Synth. Catal.* **2009**, *351*, 1539–1556.
- (152) Zhao, Q.; Zhuo, C.-X.; You, S.-L. Enantioselective synthesis of N-allylindoles via palladium-catalyzed allylic amination/oxidation of indolines. *RSC Adv.* **2014**, *4*, 10875–10878.
- (153) Maloney, M. P.; Coley, C. W.; Genheden, S.; Carson, N.; Helquist, P.; Norrby, P.-O.; Wiest, O. Negative Data in Data Sets for Machine Learning Training. *Org. Lett.* **2023**, *25*, 2945–2947.