JACS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

Cite This: J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

pubs.acs.org/JACS

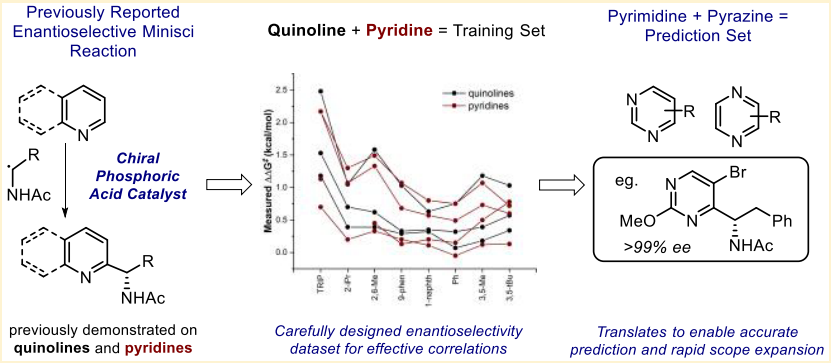
Predictive Multivariate Linear Regression Analysis Guides Successful Catalytic Enantioselective Minisci Reactions of Diazines

Jolene P. Reid,[†,§] Rupert S. J. Proctor,[‡,§] Matthew S. Sigman,[*,†] and Robert J. Phipps[*,‡]

[†]Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112, United States

[‡]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, United Kingdom

Supporting Information



ABSTRACT: The Minisci reaction is one of the most direct and versatile methods for forging new carbon−carbon bonds onto basic heteroarenes: a broad subset of compounds ubiquitous in medicinal chemistry. While many Minisci-type reactions result in new stereocenters, control of the absolute stereochemistry has proved challenging. An asymmetric variant was recently realized using chiral phosphoric acid catalysis, although in that study the substrates were limited to quinolines and pyridines. Mechanistic uncertainties and nonobvious enantioselectivity trends made the task of extending the reaction to important new substrate classes challenging and time-intensive. Herein, we describe an approach to address this problem through rigorous analysis of the reaction landscape guided by a carefully designed reaction data set and facilitated through multivariate linear regression (MLR) analysis. These techniques permitted the development of mechanistically informative correlations providing the basis to transfer enantioselectivity outcomes to new reaction components, ultimately predicting pyrimidines to be particularly amenable to the protocol. The predictions of enantioselectivity outcomes for these valuable, pharmaceutically relevant motifs were remarkably accurate in most cases and resulted in a comprehensive exploration of scope, significantly expanding the utility and versatility of this methodology. This successful outcome is a powerful demonstration of the benefits of utilizing MLR analysis as a predictive platform for effective and efficient reaction scope exploration across substrate classes.

1. INTRODUCTION

First developed into a general synthetic process by Minisci and co-workers in the late 1960s, the addition of nucleophilic radicals to electron-deficient heteroarenes has arguably become the leading method for direct carbon−carbon bond formation onto heteroaromatic scaffolds.[1] The ubiquity of pyridines, quinolines, and the numerous derivatives thereof as structural features in molecules of biological interest has rendered so-called "Minisci-type" chemistry an indispensable tool for medicinal chemists.[2] While the original conditions developed by Minisci for radical generation are still widely applied, the past decade in particular has seen tremendous attention paid to the development of new protocols for Minisci-type reactions.[3] The major emphasis of these advances has been on enhanced approaches for radical generation. Indeed, Minisci-type chemistry has, to some degree, become a testbed for the latest developments in emerging areas such as photoredox catalysis[4] and electrochemistry.[5] However, the Minisci reaction presents several fascinating selectivity challenges to the synthetic chemist. The first is regioselectivity, since on heteroarenes the LUMO coefficients can be very similar at multiple positions.[6] The second is the question of whether a prochiral nucleophilic radical may be coaxed into forming a new stereocenter in an enantiocontrolled manner during the C−C bond-forming process.[7] We recently disclosed a strategy that enabled influence to be exerted over both of these selectivity aspects for the addition of *N*-acyl, *α*-amino radicals to a range of pyridines and quinolines.[8] Jiang and co-workers subsequently demonstrated that this strategy could also be applied to isoquinolines.[9] Our approach was founded on the use of a chiral phosphoric acid catalyst to activate the substrate, which we anticipated was able to subsequently engage in a network of noncovalent interactions (NCIs) with the radical cation

Received: October 29, 2019
Published: November 11, 2019

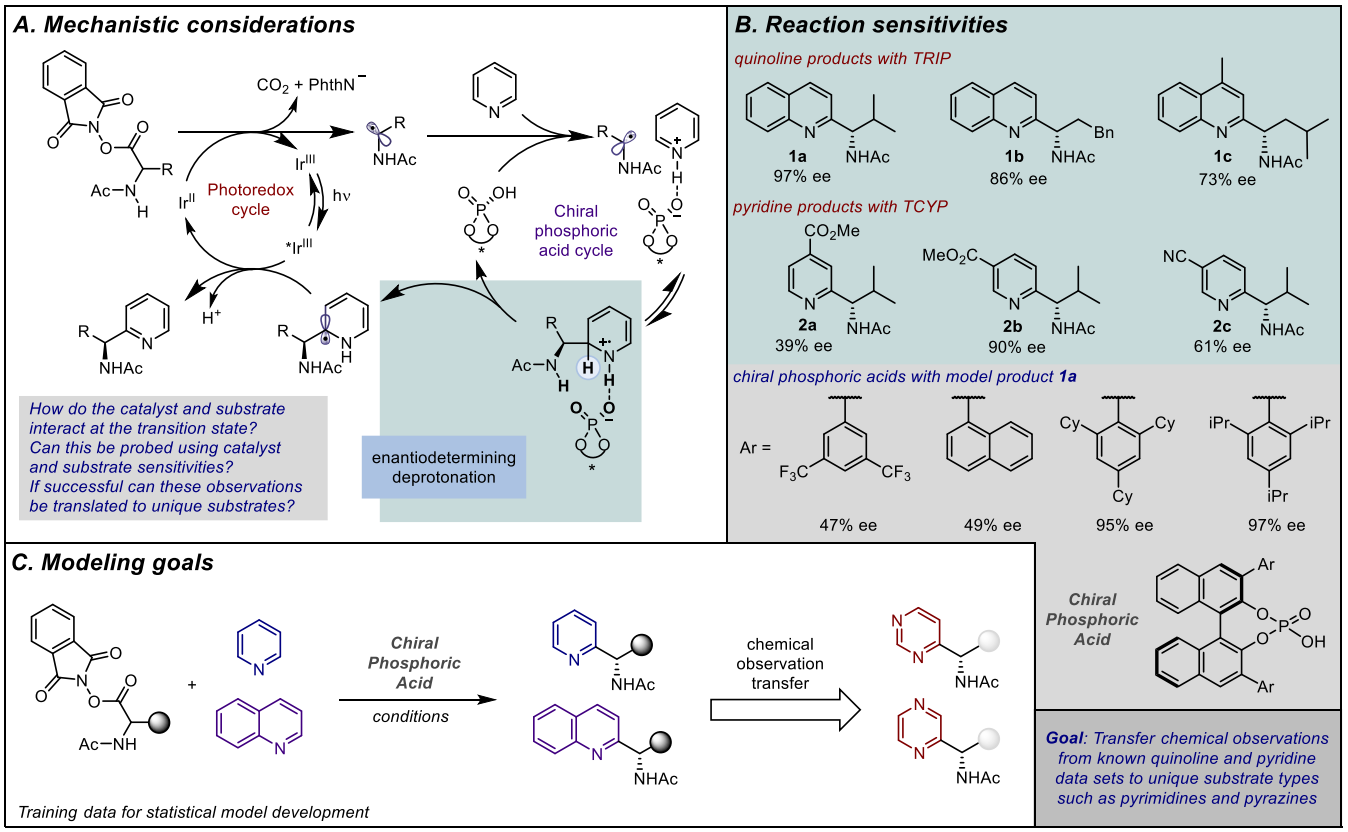


Figure 1. Application of statistical analysis tools to reaction development. (A) Working mechanistic hypothesis for asymmetric radical addition to heteroarenes. (B) Substrate and catalyst sensitivities deployed as a mechanistic probe. (C) The mechanistic principles leading to enantioselective catalysis captured by the statistical models can be transferred to genuinely different structural motifs not contained in the training data set, facilitating reaction development.

intermediate in the transition state (TS) for selectivity-determining deprotonation (Figure 1A).[10] However, the enantioselectivity trends with respect to both catalyst and substrate were not obvious. Even modest structural modifications resulted in substantial differences (Figure 1B), making immediate extension of the protocol to other substrate types challenging. The notion that selectivity could be influenced to such an extent by minor structural modifications to the substrate is intriguing, as it alludes to subtle albeit important molecular features impacting asymmetric catalysis. In targeting the understanding and prediction of substrate efficacy, we approached this problem within the context of a modern physical organic analysis. In this scenario, enantioselectivity[11] or site-selectivity[12] values can report on specific interactions between catalyst and substrate. Specifically, we reasoned that by designing a data set in which the structural features of the catalyst and substrate were appropriately modified, effective correlations could reveal the underlying causal interactions. It was anticipated that such an analysis would not only provide key insights into reaction mechanism but also provide the ability to predict the performance of new substrate types to ultimately expand the scope of the process. With regard to the latter, our initial report had only explored the reaction of pyridines and quinolines. Yet, the prevalence of diverse heterocycles possessing additional heteroatoms in medicinal compounds led us to question the broader applicability of this enantioselective Minisci method.[2b] If the selectivity discriminants were consistent for a range of substrates, it may be possible to quantitatively transfer the insights gained from the correlations to the prediction of unique substrates not included in the training sets (Figure 1C). Moreover, it is widely acknowledged by medicinal chemists that increasing the three-dimensionality of scaffolds in lead molecules enhances the odds of success as a drug candidate.[13] Three-dimensionality inevitably leads to stereoisomers, which often elicit distinct biological activity. As such, a method to predict the viability of directly appending chiral scaffolds to a range of basic heteroarenes with control of absolute stereochemistry would likely have significant impact in pharmaceutical research. To this end, we report a study employing predictive, statistical modeling techniques to relate both catalyst and substrate structures to selectivity outcomes. With statistical models that describe the general mechanistic features of the system, we can quantitatively transfer chemical insights to new substrate components. Furthermore, our model has identified pyrimidines and pyrazines to be amenable to the reaction conditions, successfully predicting protocol extension to the use of these valuable basic heteroarene motifs.

2. RESULTS AND DISCUSSSION

Data Set Design and Modeling. Intrigued by the reduced levels of selectivity observed for certain substrate subsets in the initial exploration of reaction space, and driven by the importance of accessing varied chiral heteroarene building blocks, we initiated a study into the scope and limitations of the enantioselective Minisci protocol. Despite previously reporting a collection of experimental observations for this chemistry, we anticipated that a designed data set in

DOI: 10.1021/jacs.9b11658
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

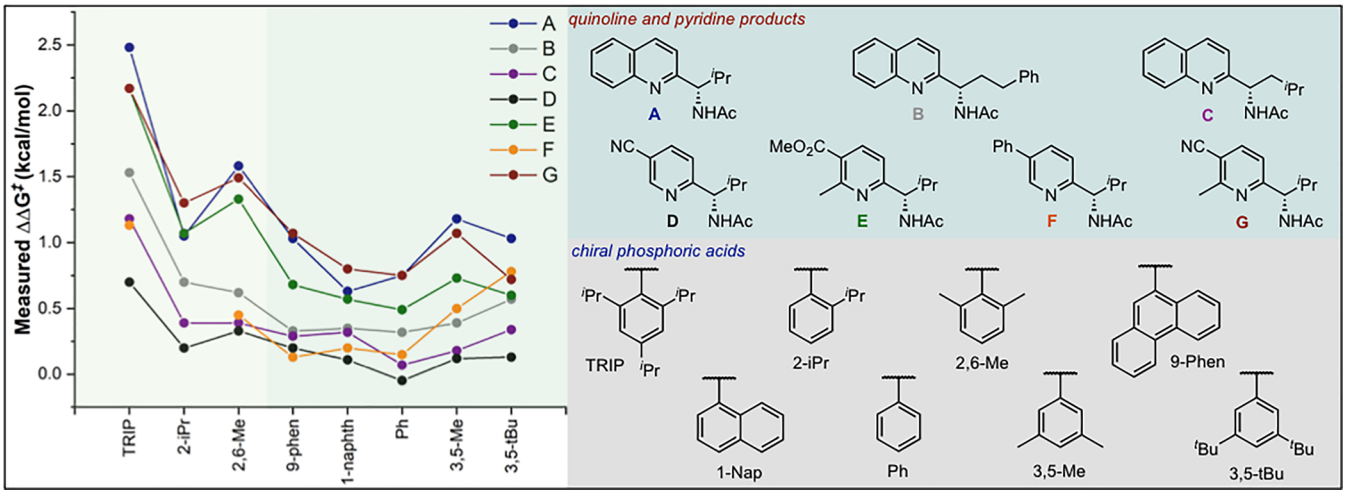


Figure 2. Graphical representation of substrate structure−selectivity trends as a function of catalysts. Colors partition catalysts that have 2- or 2,6-substituents that exhibit a unique response as a function of substrate, compared to other CPAs.

which both the substrate and catalyst were systematically modified would allow effective correlation and prediction of substrate performance. In approaching the design of such a data set, we sought to first establish the enantioselectivity range accessible by changing both the reactants and catalysts. This step was used to facilitate rapid identification of the features that most perturbed the enantioselectivity of the process to inform the correct choice of combinations for a matrix. In regard to structural changes, with the generally optimal catalyst for each substrate subset, pyridines were the most sensitive [39−93% ee, with 3,3′-bis(2,4,6-tricyclohexylphenyl)-1,1′-binaphthyl-2,2′-diyl hydrogenphosphate (TCYP)] and quinolines the least [73−97% ee, with 3,3′-bis(2,4,6-triisopropylphenyl)-1,1′-binaphthyl-2,2′-diyl hydrogenphosphate (TRIP)]. Perhaps most notably, small steric profiles on the *N*-heterocycles reduced enantioselectivities; however, electronic effects were much subtler. To probe the effect of the 3- and 3′-substituents on the catalyst, we examined a variety of BINOL-derived phosphoric acids. The screen demonstrated that reasonably large groups at the 3,3′-positions were necessary for high enantioselectivities, a finding common in similar transformations.[14] In targeting the description of general trends, seven substrates (**A**−**G**) were selected that evenly covered the range of enantioselectivities representative of the different substitution patterns presumed to influence the selectivity (Figure 2). Simultaneously, eight phosphoric acid catalysts were prepared with variable substitution at the 3- and 3′-positions of the BINOL backbone. TRIP, 2,6-dimethyl- and 2-^{i}Pr-substituted catalysts were selected to probe proximal sterics, while 3,5-dimethyl- and 3,5-di-^{t}Bu-substituted catalysts were chosen to understand remote steric effects. Two other substituted catalysts, 1-naphthyl and 9-phenanthryl, were prepared to evaluate the possibility of attractive noncovalent interactions, as opposed to repulsive steric ones. Finally, phenyl was intended to serve as a deconstructed derivative of each scenario outlined above to probe any isolated effects. This training set was used not only to provide requisite structural changes as a function of enantioselectivity but also incorporates sufficient overlapping molecular feature space required for the development of comprehensive parameter libraries and statistical analysis. For example, TIPSY [(*S*)-3,3′-bis(triphenylsilyl)-1,1′-binaphthyl-2,2′-diyl hydrogenphosphate], which has large $SiPh_3$ groups at the 3- and 3′-positions, is reasonably effective (product A, 75% ee), but its inclusion in the training set would render the parameter space, which is required to connect changes in structure to selectivity, to be dramatically reduced. For example, aromatic-derived catalysts all contain a six-membered ring with different substituents at these positions; by contrast, nonaromatic derived catalysts do not. Therefore, we only consider BINOL-derived phosphoric acids with aryl substituents at the 3- and 3′-positions, the most commonly used class for asymmetric catalysis. As such, caution should be taken in extrapolating outcomes to other classes, which can prove superior in some situations (see the Supporting Information, SI).[15]

With the appropriate libraries of the substrates and catalysts in hand, the enantioselective outcome of each combination was measured as depicted in Figure 2. From this visual analysis, catalysts with proximal steric bulk (Figure 2, left-hand side) demonstrate a unique response as a function of substrate compared to other Chiral Phosphoric Acids (CPAs). Specifically, a strong dependence of the ee on the heterocycle substituent(s) was observed, resulting in a $\Delta\Delta G^{\ddagger}$ range of ∼1.7 kcal/mol. In contrast, the reaction was less sensitive to these substitution patterns with the remaining catalysts, and the enantioselectivities remained relatively poor. The unique behavior of the 2,6-substituted catalysts is consistent with enhanced stereocontrolling interactions with this catalyst subset.

To truly interrogate the interactions between catalyst and substrate, we sought to employ multivariate linear regression analysis (MLR).[16] In this approach, parameter sets describing the important structural features of the reaction components are related to selectivity outputs expressed as $\Delta\Delta G^{\ddagger}$. The resulting mathematical equation, generally consisting of multiple terms, can be deployed to predict the outcome when features are adjusted. Traditionally, parameter selection is accomplished using candidate structures, which can either be the entire molecule or a simplified structural surrogate (for substrates these are often starting materials to mirror Hammett-type analysis). In this case, we used the product structures, as it combines both reactants while also expanding the features one may extract for aiding correlation development. We viewed this as a simple yet crucial means of describing the molecular features most relevant to the enantio-determining step. To build the parameter set, computation

DOI: 10.1021/jacs.9b11658
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

optimizations were performed on these structures at the M06-2X/def2-TZVP level of theory wherein natural bond orbital (NBO) charges, IR vibrations, and Sterimol values were collected to probe structural effects.

Through an iterative MLR modeling process (see the SI for workflow), the combination of steric and electronic parameters resulted in the model depicted in Figure 3. Both Leave-one-out

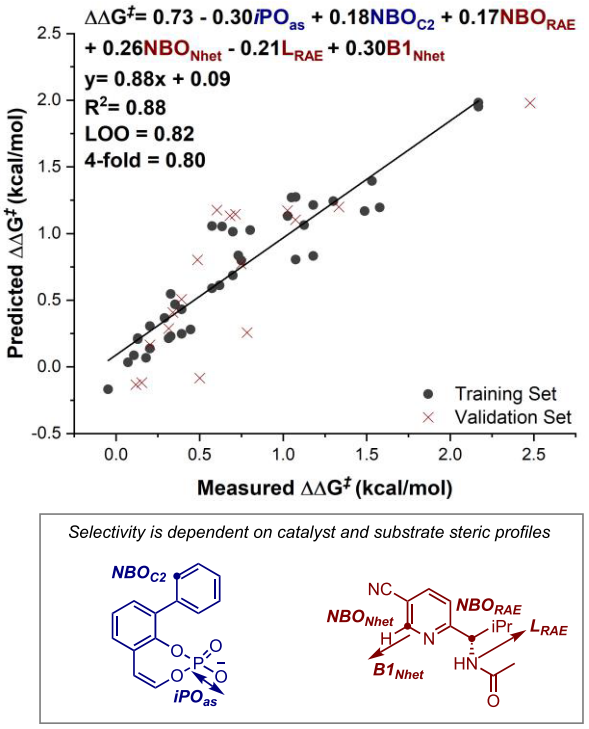


Figure 3. MLR correlation reveals that enantioselectivity is dependent on catalyst and substrate steric profiles as represented by various catalyst/product terms.

(LOO) and external validation, in which the data set is partitioned pseudo-randomly into 70:30 training:validation sets, suggest a relatively robust model. The consistency in descriptors of the top 10 models, as determined by their statistical scores and predictability, demonstrates that interpretability of a singular model does not affect the overall analysis (see the SI for full details). Interestingly, the largest coefficients in the depicted normalized model correspond to the product, with the heterocycle and redox-active ester (RAE) represented by seemingly individual components. Variations in *N*-heterocycle component of the product can be described by NBO_{NHet}, $B1_{NHet}$, and NBO_{RAE}. In considering NBO_{RAE}, this term acts as a descriptor for both heterocycle and RAE structural features. This illustrates the advantage of simplifying correlation equations to collective terms through deploying product structures that combine both reactants as the parameter acquisition platform. However, it is likely that the descriptor is reflecting more than one physical effect in the diastereomeric TS structures, making precise interpretations difficult. Ultimately, this analysis implies that the substrate effect on enantioselectivity is mostly additive but suggests there could be some circumstances where correct matching of heterocycle and RAE may be beneficial.

Consistent with other studies, the overall incorporated terms support steric bulk as the major catalyst selectivity discriminant.[17] Specifically, both reasonably large 3,3′-substituents and *N*-heterocycles were important for high levels of enantioselectivity. This is congruent with the hypothesis that TS_{minor} is disfavored as a consequence of energetically penalizing steric repulsions with the catalyst substituents enhanced through large substrate sterics.

The inclusion of L_{RAE} with a negative coefficient suggests that TS_{major} is also sensitive to the substrate molecular features. In other words, longer substituents introduce enhanced steric effects with the catalyst in the TS leading to the observed product, ultimately favoring formation of the opposite enantiomer. Since the Sterimol L term is a conformationally sensitive parameter, it may also describe the role of a preferred geometry.[18] Indeed, surveying the enantioselectivities of the reactions forming **B** and **C**, in which they differ only by RAE, shows that **B** performs better overall despite $-^{i}Pr$ appearing to be shorter than $-CH_2Bn$. However, computation optimizations demonstrate that **B** can adopt more compact arrangements and smaller L values at the RAE than **C**, clarifying this nonintuitive trend. This highlights that substrate dynamics are also important in determining selectivity.

The impact of catalyst and substrate on regioselectivity (2- vs 4-position) was also probed. Since the inherent regioselectivity of the mechanism is masked by the pyridine subset, in which 4-addition does not occur with 3,3′-substituted acids, only the quinoline substrates (**A**–**C**), exhibiting variable regioselectivity as a function of catalyst, were further investigated. Employing the same modeling techniques led only to complex models. This observation is compounded by the training set restriction in terms of data range and structure. A correlation of the C2:C4 isomeric ratio (rr) with the enantioselectivity of the product reveals a linear relationship: as the ee increases, the rr generally increases (see the SI). This suggests that for the quinoline substrate subset, the undesired 4-regioisomer could arise from an unselective pathway. This becomes evident when the enantioselectivity of the 4-isomer product was measured, where possible, resulting in low enantioselectivity values (<15% ee).

Reaction Design. While the obtained model, shown in Figure 3, provides insightful mechanistic information on the transformation, the clear practical utility lies in its ability to predict the performance of unique substrate classes, thereby directing future synthetic efforts. If effective out-of-sample prediction were possible, the model could estimate the impact of a new heterocycle, RAE, and/or catalyst on selectivity, provided that the prediction platforms incorporated sufficient overlap with the training set. Typically, exploration of the synthetic scope of a new enantioselective chemical reaction involves evaluation of a large number of substrates, only a proportion of which yield the desired high levels of enantiomeric excess. This can be a time- and resource-consuming process, particularly when substrates require multistep synthesis. Conversely, in target-driven synthesis, only a single specific substrate is of interest and a number of different synthetic approaches may be considered. A reliable, predictive mathematical model, accessible to bench chemists, has the potential to narrow down the myriad options in the latter scenario and greatly accelerate reaction scope exploration in the former. The workflow for ee prediction is straightforward and is initiated by locating the ground state of the targeted reaction variable by DFT computation, collecting the requisite parameters, and submitting them to the equation (see the SI for a tutorial).

DOI: 10.1021/jacs.9b11658
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

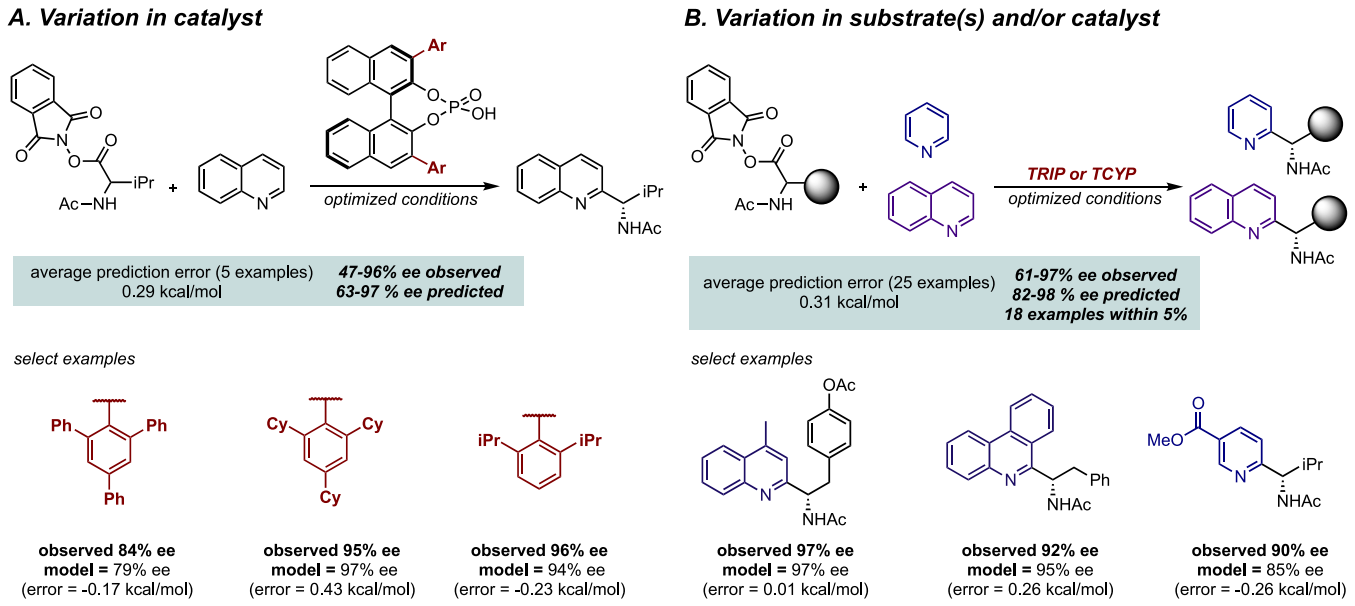


Figure 4. Prediction platforms. (A) Assessing prediction capabilities with various 3,3′-substituted phosphoric acids. (B) Prediction of assorted reaction systems containing substrate and catalyst components not explicitly included in the training set.

In the context of the enantioselective Minisci reaction, we sought to expand the scope of the heterocyclic component beyond the pyridines and quinolines that had been included in our initial report but to do so in a rational manner that would not involve "screening" numerous substrates in search of hits with high ee. We envisaged that successful application of MLR analysis to reaction scope expansion would be a very effective showcase of the practical benefits of this approach. However, before progressing to new heterocycle classes, we first sought to evaluate the model's prediction performance on the previously reported data set to consider the feasibility of this endeavor.[8] As a first assessment, we evaluated the ability to predict five additional reactions, involving catalysts with various 3,3′-substituents, with a model substrate contained in our training set (Figure 4A). The ability to predict in this reaction dimension would be particularly useful if the optimal catalyst for a specific substrate combination was not contained in the training set. Treating these as virtual predictions, this set was predicted accurately, with an average absolute $\Delta\Delta G^{\ddagger}$ error of 0.29 kcal/mol.

As a second case study, the model was assessed in the same manner with 25 additional reactions involving various substrate subsets catalyzed by TRIP or TCYP.

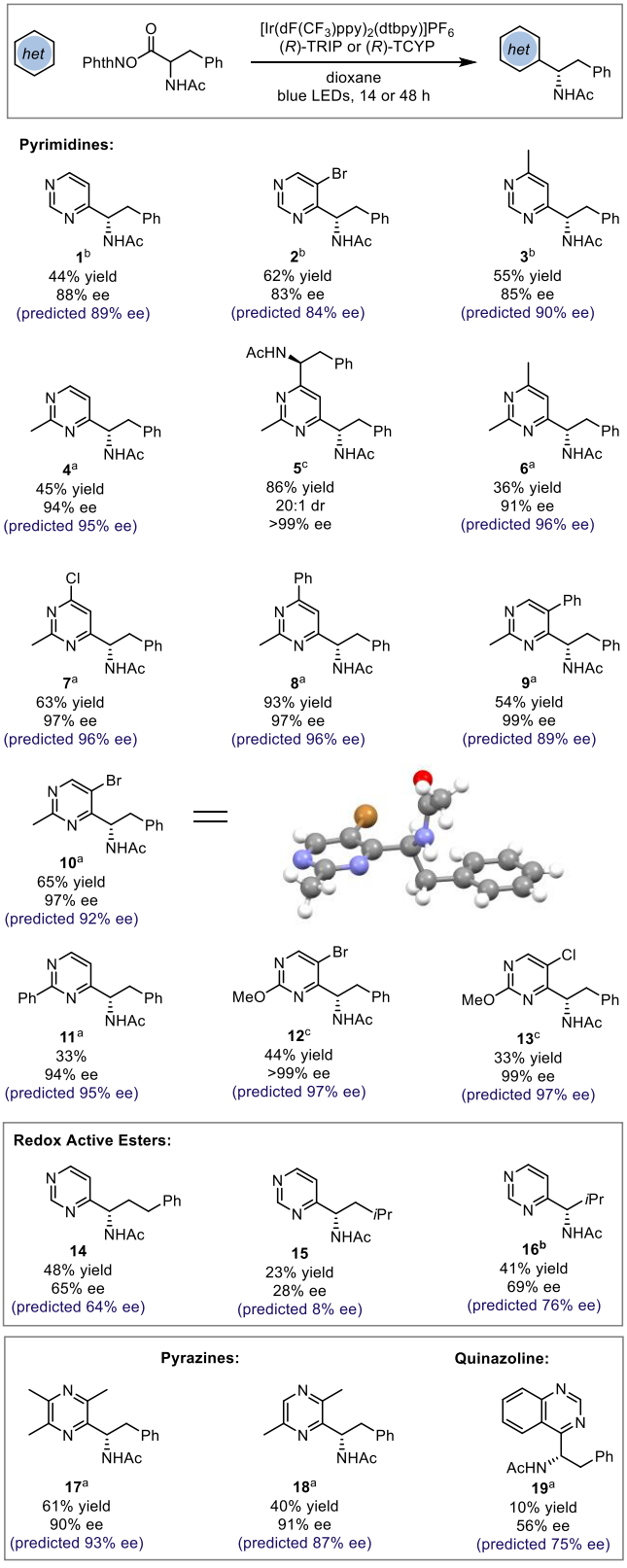
Examples were selected on the basis of a range of enantioselectivity (61−97% ee) and substrate structure (the full list can be found in the SI). This is a more challenging scenario, as some substrate and catalyst components are not explicitly included in the training set. Again, accurate prediction of the outcomes was construed using the model, with an average absolute error of 0.31 kcal/mol and 18 examples predicted within 5% ee (Figure 4B). These results suggest that the ability to effectively extrapolate to new reaction components results from a set of general transition state features that are fundamentally similar across the reaction range.

On the basis of the key parameters in the model, we envisaged that pyrimidines should, in principle, constitute excellent substrates, as the inclusion of the second ring nitrogen would be expected to increase the magnitude of the NBO_{Nhet} term significantly. Pyrimidines are ubiquitous in pharmaceuticals, agrochemicals, and small molecules of medicinal interest, so demonstration of the protocol on this class would be of substantial practical value. Thus, we evaluated a number of reactions involving various electronically and sterically unique pyrimidine and pyrazine substrates, guided by our predictive model. A phenylalanine-derived RAE was selected as the radical precursor, along with either TRIP or TCYP as catalyst. The predictions obtained from the model are shown in Scheme 1 alongside the experimental results that were ultimately obtained, and pleasingly, the agreement was generally excellent. Furthermore, the observed enantioselectivities were typically superior to the use of pyridines, a previously explored subset. Each measured enantioselectivity in Scheme 1 was predicted with an average absolute $\Delta\Delta G^{\ddagger}$ error of 0.39 kcal/mol (13 examples within 5% ee), demonstrating the ability of the model to extrapolate effectively to an entirely new class of substrates.

Specifically, unsubstituted pyrimidine reacted with complete regioselectivity at the C4 position to deliver product **1** in 88% ee when using TCYP as catalyst (TRIP gave 78% ee, with the model predicting 83% ee). Given the lack of steric features on unadorned pyrimidine, we regarded this as a highly encouraging result, the moderate yield being due to incomplete conversion rather than deleterious pathways. While superficially surprising that a heteroarene with no steric features should perform well, close examination of key parameters in the model reveals that the more positive NBO values associated with pyrimidine, a result of inclusion of the second ring nitrogen, largely compensate for a lower $B1_{NHet}$ term. Furthermore, the ability of the model to accurately reflect the outcomes with different phosphoric acids highlights its utility for predicting the right catalyst for a particular substrate, obviating the need for extensive catalyst screening for each substrate. A bromide substituent was tolerated at the C5 position with only a slight decrease in ee (**2**, 83% ee), and a methyl was similarly incorporated at C4 (**3**, 85% ee). When the pyrimidine possessed a substituent at C2, enantioselectivity increased significantly. Once again, the model clearly explains

DOI: 10.1021/jacs.9b11658
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Scheme 1. Substrate Scope of Enantioselective Minisci Reaction on Pyrimidines and Pyrazines



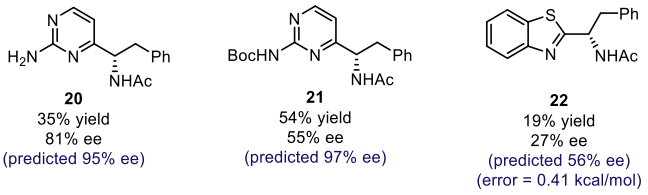
[a]TRIP (5 mol %), 14 h. [b]TCYP (5 mol %), 14 h. [c]TRIP (10 mol %), 48 h

why such substrates should be particularly amenable, since both the NBO and B1 terms are now large and positive. For substrates such as **4**, possessing a substituent only at C2, careful control of the stoichiometry and time of the reaction made it possible to stop at monoalkylation (**4**, 94% ee) or progress all the way to dialkylation [**5**, 20:1 dr, >99% ee (major diastereomer)]. A variety of more complex 2-methylpyrimidine substrates were well-tolerated, including 4-Me (**6**, 91% ee), 4-Cl (**7**, 97% ee), 4-Ph (**8**, 97% ee), 5-Ph (**9**, 99% ee), and 5-Br (**10**, 97% ee). The absolute stereochemistry of the products is predicted to be consistent with that of the original systems, as confirmed by the X-ray crystallographic analysis after recrystallization of **10**. The stereochemistry of the remainder of the entries is assigned by analogy. Aryl substitution was accommodated at the C2 position (**11**, 94% ee), and the model predicted this significant structural perturbation with remarkable accuracy. When progressing to the 2-methoxypyrimidines **12** and **13**, we obtained some of the highest enantioselectivities observed thus far in any enantioselective Minisci reaction (>99% ee for the bromo-functionalized **12** and 99% ee for chloro-functionalized **13**), a result of matching multiple positive structural effects. For these two substrates, moderate conversions led us to raise the catalyst loading to 10 mol% with longer reaction times of 48 h to obtain the yields shown. We sought to test the predictive power of the model on pyrimidine in combination with RAEs other than the phenylalanine-derived variant used thus far. Therefore, RAEs derived from valine, homophenylalanine, and leucine were evaluated. The experimental results were in excellent agreement with the predicted values (Scheme 1, **14**–**16**). Homophenylalanine- (**14**) and leucine-derived (**15**) RAEs gave significantly lower ee than phenylalanine, which are consistent with observations utilizing quinolines and pyridines (Figure 1B). For the valine-derived RAE (**16**), the model accurately predicted moderate enantioselectivity (69% ee), contrasting with the excellent results this RAE had given with quinolines and pyridines. By analyzing terms in the statistical model, the lower enantioselectivity can be attributed to the more negative NBO_{RAE}, which overrides any beneficial impact garnered from the more negative L_{RAE} and positive NBO_{Nhet} terms. Furthermore, this outlines an instance in which correct matching of heterocycle and RAE is beneficial. We also explored several examples of pyrazine substrates and pleasingly observed that various combinations of methyl substitution worked effectively and were accurately predicted (**17**, 90% ee; **18**, 91% ee). One limitation to acknowledge is that the model can only guide users of the methodology toward assessing selectivity outputs and therefore will not be capable of predicting reactivities. This is exemplified by the fact that no reactivity was observed with pyridazines and quinoxalines under our conditions (see the SI). Quinazoline exhibits poor reactivity, giving product **19** in 56% ee. Predicted at 75% ee, this is within the average error of the model ($\Delta\Delta G^{\ddagger}$ error of 0.41 kcal/mol compared with that for the averaged diazine prediction set of 0.39 kcal/mol).

While the observed enantioselectivities for the substrates presented in Scheme 1 generally show good agreement with the prediction for both high and low ee examples, we discovered that substrates incorporating an amino substituent between the two nitrogen atoms (Scheme 2, **20** and **21**) gave results that were rather lower than predicted. We speculate that additional hydrogen bonds formed with these groups are likely interrupting the hydrogen-bonding network leading to stereoinduction, a critical catalyst–substrate interaction expressed by the model terms. Ultimately, these are unique contacts that are

DOI: 10.1021/jacs.9b11658
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Scheme 2. Substrates Revealed as Limitations



not represented in the training set and demonstrate a limitation of the present model.

To test this approach on more structurally disparate bicyclic heteroarenes, the reaction of benzothiazole was probed. The reduction in overlapping features with our training set structures creates a challenge to extending our comprehensible parameter sets to this substrate class (five-membered vs six-membered ring). To address this featurization challenge, we used 0 digits as descriptors for the missing benzothiazole components. By deploying the adapted descriptor set and the training model, the resultant extrapolation of substrate space predicted only modest enantioselectivities, an observation validated by experiment (Scheme 2, **22**). This result is compelling in that we could reach an informed decision about pursuing benzothiazoles as a substrate class. Although the predicted enantioselectivities for **22** were higher than observed, the $\Delta\Delta G^{\ddagger}$ error of 0.41 kcal/mol is comparable to that of the averaged diazine prediction set (0.39 kcal/mol), suggesting that the source of the error may be systematic. Taken together, these examples showcase that the model's predictive capabilities are not limited to classifying published data sets but can be applied to analyze and predict new reactions, even in situations where multiple components are varied. Particular highlights of this protocol are the uniformity of the conditions employed for the diverse set of heteroarenes and the ability to extrapolate to new substrate types in the absence of persuasive mechanistic information.

Conclusion. We have described the development of a predictive, mathematical model for the enantioselective Minisci addition of *N*-acyl, *α*-amino radicals to pyridines and quinolines through careful evaluation of catalyst/substrate training sets and parameter acquisition platforms. The model describes the general transition-state features important for the reaction class, which ultimately provided the basis for the transfer of experimental observations from one substrate subset to another. The model parameters suggested that pyrimidines, with typically larger NBO values than pyridines, should be particularly amenable to the same reaction conditions. The specific predictions produced by the model prompted us to explore a range of substituted pyrimidines, as well as several pyrazines. The accurate predictive power avoided the need to assess a large number of substrates in order to discover those most compatible with the method—we were guided there directly, saving valuable time and resources. This should provide confidence to synthetic chemists looking to extrapolate this methodology further, to other diverse heterocyclic classes. More broadly, this successful outcome is a powerful demonstration of the benefits of utilizing MLR analysis as a predictive platform for effective and efficient reaction scope exploration in asymmetric catalysis.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.9b11658.

Experimental details, procedures, compound characterization data, computational details, and copies of ^{1}H and ^{13}C NMR spectra of new compounds (PDF)

X-ray crystallographic data for **10** in CIF format (CIF)

■ AUTHOR INFORMATION

Corresponding Authors

*matt.sigman@utah.edu

*rjp71@cam.ac.uk

ORCID

Rupert S. J. Proctor: 0000-0002-2296-448X

Matthew S. Sigman: 0000-0002-5746-8830

Robert J. Phipps: 0000-0002-7383-5469

Author Contributions

§J.P.R. and R.S.J.P. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

R.S.J.P. is grateful to GlaxoSmithKline and the EPSRC for a CASE Ph.D. studentship. R.J.P. is grateful to the Royal Society for a University Research Fellowship and the ERC (StG 757381) and EPSRC (EP/N005422/1) for funding. J.P.R. thanks the EU Horizon 2020 Marie Skłodowska-Curie Fellowship (grant no. 792144) and M.S.S thanks the NIH (1 R01 GM121383) for support of this work. Computational resources were provided from the Center for High Performance Computing (CHPC) at the University of Utah and the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by the NSF (ACI-1548562) and provided through allocation TG-CHE180003. We are grateful to Dr. Andrew Bond (University of Cambridge) for solving and refining the X-ray crystal structure, Dr. Holly J. Davis for initial experiments, and Jonathan Taylor (GSK) for useful discussion.

■ REFERENCES

(1) (a) Minisci, F.; Vismara, E.; Fontana, F. Recent Developments of Free-Radical Substitutions of Heteroaromatic Bases. *Heterocycles* **1989**, *28*, 489−519. (b) Minisci, F.; Fontana, F.; Vismara, E. Substitutions by nucleophilic free radicals: A new general reaction of heteroaromatic bases. *J. Heterocycl. Chem.* **1990**, *27*, 79−96.

(2) (a) Duncton, M. A. J. Minisci reactions: Versatile CH-functionalizations for medicinal chemists. *MedChemComm* **2011**, *2*, 1135−1161. (b) Vitaku, E.; Smith, D. T.; Njardarson, J. T. Analysis of the Structural Diversity, Substitution Patterns, and Frequency of Nitrogen Heterocycles among U.S. FDA Approved Pharmaceuticals. *J. Med. Chem.* **2014**, *57*, 10257−10274.

(3) Proctor, R. S. J.; Phipps, R. J. Recent Advances in Minisci-Type Reactions. *Angew. Chem., Int. Ed.* **2019**, *58*, 13666−13699.

(4) Sun, A. C.; McAtee, R. C.; McClain, E. J.; Stephenson, C. R. J. Advancements in Visible-Light-Enabled Radical C(sp)2−H Alkylation of (Hetero)arenes. *Synthesis* **2019**, *51*, 1063−1072.

(5) Yan, M.; Kawamata, Y.; Baran, P. S. Synthetic Organic Electrochemical Methods Since 2000: On the Verge of a Renaissance. *Chem. Rev.* **2017**, *117*, 13230−13319.

(6) (a) Tauber, J.; Imbri, D.; Opatz, T. Radical Addition to Iminium Ions and Cationic Heterocycles. *Molecules* **2014**, *19*, 16190. (b) O'Hara, F.; Blackmond, D. G.; Baran, P. S. Radical-Based

DOI: 10.1021/jacs.9b11658
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

Regioselective C−H Functionalization of Electron-Deficient Heteroarenes: Scope, Tunability, and Predictability. *J. Am. Chem. Soc.* **2013**, *135*, 12122−12134.
(7) For relevent reviews on enantioselective radical chemistry, see the following: (a) Sibi, M. P.; Manyem, S.; Zimmerman, J. Enantioselective Radical Processes. *Chem. Rev.* **2003**, *103*, 3263−3296. (b) Brimioulle, R.; Lenhart, D.; Maturi, M. M.; Bach, T. Enantioselective Catalysis of Photochemical Reactions. *Angew. Chem., Int. Ed.* **2015**, *54*, 3872−3890. (c) Silvi, M.; Melchiorre, P. Enhancing the potential of enantioselective organocatalysis with light. *Nature* **2018**, *554*, 41−49.
(8) Proctor, R. S. J.; Davis, H. J.; Phipps, R. J. Catalytic enantioselective Minisci-type addition to heteroarenes. *Science* **2018**, *360*, 419−422.
(9) Liu, X.; Liu, Y.; Chai, G.; Qiao, B.; Zhao, X.; Jiang, Z. Organocatalytic Enantioselective Addition of α-Aminoalkyl Radicals to Isoquinolines. *Org. Lett.* **2018**, *20*, 6298−6301.
(10) For selected examples of application of noncovalent catalysis in enantioselective radical chemistry, see the following: (a) Bauer, A.; Westkamper, F.; Grimme, S.; Bach, T. Catalytic enantioselective reactions driven by photoinduced electron transfer. *Nature* **2005**, *436*, 1139−1140. (b) Müller, C.; Bauer, A.; Maturi, M. M.; Cuquerella, M. C.; Miranda, M. A.; Bach, T. Enantioselective Intramolecular [2 + 2]-Photocycloaddition Reactions of 4-Substituted Quinolones Catalyzed by a Chiral Sensitizer with a Hydrogen-Bonding Motif. *J. Am. Chem. Soc.* **2011**, *133*, 16689−16697. (c) Rono, L. J.; Yayla, H. G.; Wang, D. Y.; Armstrong, M. F.; Knowles, R. R. Enantioselective Photoredox Catalysis Enabled by Proton-Coupled Electron Transfer: Development of an Asymmetric Aza-Pinacol Cyclization. *J. Am. Chem. Soc.* **2013**, *135*, 17735−17738. (d) Alonso, R.; Bach, T. A Chiral Thioxanthone as an Organocatalyst for Enantioselective [2 + 2] Photocycloaddition Reactions Induced by Visible Light. *Angew. Chem., Int. Ed.* **2014**, *53*, 4368−4371. (e) Uraguchi, D.; Kinoshita, N.; Kizu, T.; Ooi, T. Synergistic Catalysis of Ionic Brønsted Acid and Photosensitizer for a Redox Neutral Asymmetric α-Coupling of N-Arylaminomethanes with Aldimines. *J. Am. Chem. Soc.* **2015**, *137*, 13768−13771. (f) Tröster, A.; Alonso, R.; Bauer, A.; Bach, T. Enantioselective Intermolecular [2 + 2] Photocycloaddition Reactions of 2(1H)-Quinolones Induced by Visible Light Irradiation. *J. Am. Chem. Soc.* **2016**, *138*, 7808−7811. (g) Lin, J.-S.; Dong, X.-Y.; Li, T.-T.; Jiang, N.-C.; Tan, B.; Liu, X.-Y. A Dual-Catalytic Strategy To Direct Asymmetric Radical Aminotrifluoromethylation of Alkenes. *J. Am. Chem. Soc.* **2016**, *138*, 9357−9360. (h) Lin, L.; Bai, X.; Ye, X.; Zhao, X.; Tan, C.-H.; Jiang, Z. Organocatalytic Enantioselective Protonation for Photoreduction of Activated Ketones and Ketimines Induced by Visible Light. *Angew. Chem., Int. Ed.* **2017**, *56*, 13842−13846. (i) Liu, Y.; Liu, X.; Li, J.; Zhao, X.; Qiao, B.; Jiang, Z. Catalytic enantioselective radical coupling of activated ketones with N-aryl glycines. *Chem. Sci.* **2018**, *9*, 8094−8098. (j) Gentry, E. C.; Rono, L. J.; Hale, M. E.; Matsuura, R.; Knowles, R. R. Enantioselective Synthesis of Pyrroloindolines via Noncovalent Stabilization of Indole Radical Cations and Applications to the Synthesis of Alkaloid Natural Products. *J. Am. Chem. Soc.* **2018**, *140*, 3394−3402. (k) Cao, K.; Tan, S. M.; Lee, R.; Yang, S.; Jia, H.; Zhao, X.; Qiao, B.; Jiang, Z. Catalytic Enantioselective Addition of Prochiral Radicals to Vinylpyridines. *J. Am. Chem. Soc.* **2019**, *141*, 5437−5443. (l) Zheng, J.; Swords, W. B.; Jung, H.; Skubi, K. L.; Kidd, J. B.; Meyer, G. J.; Baik, M.-H.; Yoon, T. P. Enantioselective Intermolecular Excited-State Photoreactions Using a Chiral Ir Triplet Sensitizer: Separating Association from Energy Transfer in Asymmetric Photocatalysis. *J. Am. Chem. Soc.* **2019**, *141*, 13625−13634.
(11) (a) Milo, A.; Neel, A. J.; Toste, F. D.; Sigman, M. S. A data-intensive approach to mechanistic elucidation applied to chiral anion catalysis. *Science* **2015**, *347*, 737−743. (b) Zhao, S.; Gensch, T.; Murray, B.; Niemeyer, Z. L.; Sigman, M. S.; Biscoe, M. R. Enantiodivergent Pd-catalyzed C−C bond formation enabled through ligand parameterization. *Science* **2018**, *362*, 670−674.
(12) (a) Keylor, M. H.; Niemeyer, Z. L.; Sigman, M. S.; Tan, K. L. Inverting Conventional Chemoselectivity in Pd-Catalyzed Amine Arylations with Multiply Halogenated Pyridines. *J. Am. Chem. Soc.* **2017**, *139*, 10613−10616. (b) Toste, F. D.; Sigman, M. S.; Miller, S. J. Pursuit of Noncovalent Interactions for Strategic Site-Selective Catalysis. *Acc. Chem. Res.* **2017**, *50*, 609−615.
(13) (a) Lovering, F.; Bikker, J.; Humblet, C. Escape from Flatland: Increasing Saturation as an Approach to Improving Clinical Success. *J. Med. Chem.* **2009**, *52*, 6752−6756. (b) Lovering, F. Escape from Flatland 2: complexity and promiscuity. *MedChemComm* **2013**, *4*, 515−519.
(14) Reid, J. P.; Goodman, J. M. Goldilocks Catalysts: Computational Insights into the Role of the 3,3′ Substituents on the Selectivity of BINOL-Derived Phosphoric Acid Catalysts. *J. Am. Chem. Soc.* **2016**, *138*, 7910−7917.
(15) Parmar, D.; Sugiono, E.; Raja, S.; Rueping, M. Complete Field Guide to Asymmetric BINOL-Phosphate Derived Brønsted Acid and Metal Catalysis: History and Classification by Mode of Activation; Brønsted Acidity, Hydrogen Bonding, Ion Pairing, and Metal Phosphates. *Chem. Rev.* **2014**, *114*, 9047−9153.
(16) (a) 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. (b) Santiago, C. B.; Guo, J.-Y.; Sigman, M. S. Predictive and mechanistic multivariate linear regression models for reaction development. *Chem. Sci.* **2018**, *9*, 2398−2412. (c) 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.
(17) Kwon, Y.; Li, J.; Reid, J. P.; Crawford, J. M.; Jacob, R.; Sigman, M. S.; Toste, F. D.; Miller, S. J. Disparate Catalytic Scaffolds for Atroposelective Cyclodehydration. *J. Am. Chem. Soc.* **2019**, *141*, 6698−6705.
(18) Brethomé, A. V.; Fletcher, S. P.; Paton, R. S. Conformational Effects on Physical-Organic Descriptors: The Case of Sterimol Steric Parameters. *ACS Catal.* **2019**, *9*, 2313−2323.

DOI: 10.1021/jacs.9b11658
J. Am. Chem. Soc. XXXX, XXX, XXX−XXX