Rethinking catalyst design using data science

Sara N. Alektiar and Zachary K. Wickens

Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, United States

In this issue of *Chem*, Sigman, Toste, and coworkers use data science tools to target the development of a novel family of chiral phosphoric acid scaffolds. These conformationally flexible catalysts were designed to induce enantioselectivity through predictable non-covalent interactions, offering an attractive alternative to state-of-the-art chiral catalyst scaffolds that rely on rigid chiral pockets.

The design of novel chiral catalyst scaffolds is fundamental to the development of new asymmetric transformations. Established catalyst design strategies in enantioselective catalysis have typically focused on structures that ensure the desired reaction occurs in a conformationally inflexible chiral environment. The synthesis and evaluation of rigid catalyst structures has led to the identification of a collection of privileged chiral scaffolds that each promote a mechanistically diverse array of asymmetric transformations.[1] However, as the field has progressed, these design principles have also led to increasingly complex and high molecular weight structures within privileged catalyst families while introduction of fundamentally new chiral scaffolds has become relatively uncommon. In part, this situation is a consequence of the linear optimization process typically employed in the development of chiral catalysts to promote specific enantioselective reactions; an iterative cycle of synthesis, evaluation, and modification of catalyst features is performed until a structure that provides suitably high enantioselectivity has been identified. Given this widely employed but laborious protocol, it is no surprise that researchers often aim to maximize the success rate of new catalysts by building upon established catalyst families rather than constructing fundamentally new chiral scaffolds.

In the present study, Sigman, Toste and coworkers put forward an alternative approach to chiral catalyst development that is well-suited to the introduction of new catalyst designs.[2] Rather than aiming to identify the optimal catalyst outright, the authors instead develop a set of catalysts designed to offer a data-rich starting point that informs subsequent optimization efforts.[3] The authors recognized that a structurally diverse set of catalysts should, in principle, result in highly variable enantioselectivity and provide an ideal starting point to identify the origin of selectivity and rationally tailor further catalyst modifications. In principle, this distinct approach should relieve design constraints that hamper catalyst development and are particularly acute for the development of new catalyst scaffolds.

To explore these ideas, Sigman, Toste, and coworkers targeted the development of a novel family of chiral phosphoric acid (CPA) scaffolds designed to be conformationally flexible, rather than rigid,[4–8] and induce enantioselectivity through predictable non-covalent interactions (NCIs). In stark contrast to conventional BINOL-derived CPAs,[9] the basic catalyst design advanced in this study is exceedingly simple. It incorporates a single chiral amino acid derivative into an otherwise achiral biaryl phosphoric acid backbone. A key design feature of this approach is that the catalyst core can then be decorated with a variety of readily installed distinct functional groups that engage in NCIs and dictate the precise shape of the chiral catalyst. The authors employed data science tools to identify a set of 20 catalyst structures to maximize the coverage of chemical space within this new catalyst family. Crucially, no prior knowledge of the precise catalyst interactions necessary to induce enantioselectivity is required to employ this strategy

because the data science tools employed throughout the workflow are intrinsically independent of the mechanisms of enantioinduction.

Following synthesis of this initial catalyst library, each catalyst was assayed for enantioselectivity in an established CPA-catalyzed transfer hydrogenation reaction.[10] These experiments validated the meaningful structural diversity represented within the catalyst training set, as the product enantiomeric excesses ranged from 0–87%. Using these data, linear regression models and DFT analyses identified the key catalyst features that control enantioinduction. Of particular note, two distinct intramolecular hydrogen-bonding networks could be induced to reinforce a well-defined chiral environment within these otherwise flexible structures. Extrapolation from the selectivity model predicted an improved, specific catalyst structure for enantioselective transfer hydrogenation. Synthesis and evaluation of the predicted new catalyst validated the model, with this new structure outperforming not only prior catalysts within the series but also state-of-the-art CPAs based on a rigid BINOL scaffold. Intriguingly, the model provides mechanistically-agnostic insight into what structures would be effective for asymmetric synthesis. As a demonstrative example of this point, the model identified that the partial charge on the catalyst-based amide and sulfonamide oxygen atoms predicted structure-dependent racemization that occurs during catalyst synthesis, an issue fully decoupled from the model transfer hydrogenation reaction mechanism. This observation offers a clear illustration of how such an approach can be particularly valuable in the development of new catalyst scaffolds without well-established routes for their preparation.

Given that the catalyst training set was selected independently from the model transformation, the highly variable enantioselectivity necessary for effective modeling should be observed in other CPA-catalyzed reactions using this initial library of structurally-diverse catalysts. Indeed, evaluation of the same 20 catalysts in two additional transformations again resulted in a range in enantioselectivity, albeit with a lower maximum e.r. in both cases. Intriguingly, the same catalyst features that predicted selectivity for the transfer hydrogenation model system could again be used to construct a predictive model for these mechanistically distinct reactions. These data highlight the power of thoughtfully developing a catalyst training set designed to provide a strong starting point for further optimization and suggest that the workflow developed in this study could, in principle, be applied across a wide range of enantioselective transformations. Furthermore, the parameters that control enantioselectivity across each unique reaction suggests that understanding the key catalyst NCIs is likely to facilitate *de novo* catalyst optimization.

Overall, the authors have advanced a promising workflow to introduce new catalyst scaffolds for enantioselective catalysis. The catalyst family advanced through this study not only improves enantioselectivity relative to state-of-the-art CPAs in a model test case, but their flexible design is also well-matched to enable a data-science-driven approach for the development of enantioselective reactions. Rather than relying on rigid chiral pockets, the authors intentionally design conformationally flexible catalysts that induce enantioselectivity through predictable NCIs. This study provides a compelling case study in how data science tools can be integrated throughout a catalyst development effort and free researchers from design constraints that require extensive prior knowledge of key selectivity dictating interactions. While this workflow was illustrated within the context of asymmetric catalysis, we anticipate this study unlocks a distinct approach to catalyst development that reduces the risks associated with the development of fundamentally new catalyst scaffolds and is likely to have wide-ranging implications for catalyst design.

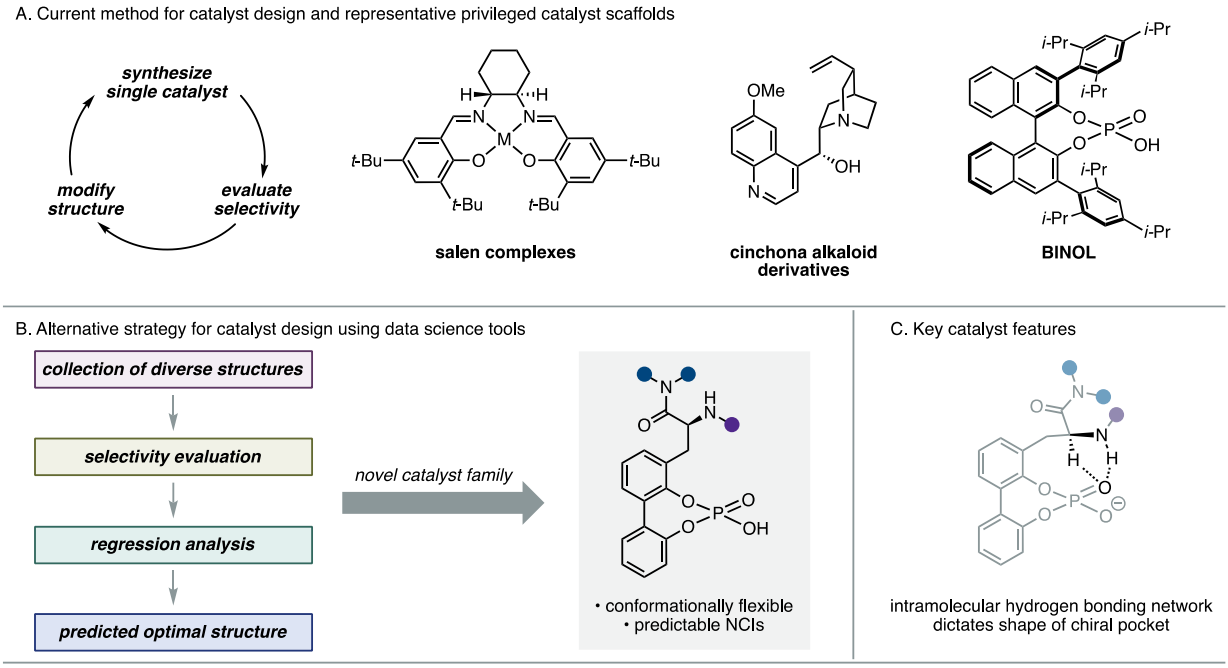

Figure 1. Overview of chiral catalyst design strategies

References

(1) Yoon, T. P.; Jacobsen, E. N. Privileged Chiral Catalysts. *Science* **2003**, *299* (5613), 1691–1693. https://doi.org/10.1126/science.1083622.

(2) Liles, J. P.; Rouget-Virbel, C.; Wahlman, J. L. H.; Rahimoff, R.; Crawford, J. M.; Medlin, A.; O'Connor, V. S.; Li, J.; Roytman, V. A.; Toste, F. D.; Sigman, M. S. Data Science Enables the Development of a New Class of Chiral Phosphoric Acid Catalysts. *Chem* **2023**. https://doi.org/10.1016/j.chempr.2023.02.020.

(3) For a special issue on data science in chemistry, see: Kulik, H. J.; Sigman, M. S. Advancing Discovery in Chemistry with Artificial Intelligence: From Reaction Outcomes to New Materials and Catalysts. *Acc. Chem. Res.* **2021**, *54* (10), 2335–2336. https://doi.org/10.1021/acs.accounts.1c00232.

(4) Birman, V. B.; L. Rheingold, A.; Lam, K.-C. 1,1′-Spirobiindane-7,7′-Diol: A Novel, C2-Symmetric Chiral Ligand. *Tetrahedron: Asymmetry* **1999**, *10* (1), 125–131. https://doi.org/10.1016/S0957-4166(98)00481-9.

(5) Guo, Q.-S.; Du, D.-M.; Xu, J. The Development of Double Axially Chiral Phosphoric Acids and Their Catalytic Transfer Hydrogenation of Quinolines. *Angewandte Chemie International Edition* **2008**, *47* (4), 759–762. https://doi.org/10.1002/anie.200703925.

(6) Xu, F.; Huang, D.; Han, C.; Shen, W.; Lin, X.; Wang, Y. SPINOL-Derived Phosphoric Acids: Synthesis and Application in Enantioselective Friedel−Crafts Reaction of Indoles with Imines. *J. Org. Chem.* **2010**, *75* (24), 8677–8680. https://doi.org/10.1021/jo101640z.

(7) Čorić, I.; List, B. Asymmetric Spiroacetalization Catalysed by Confined Brønsted Acids. *Nature* **2012**, *483* (7389), 315–319. https://doi.org/10.1038/nature10932.

(8) Schreyer, L.; Properzi, R.; List, B. IDPi Catalysis. *Angewandte Chemie International Edition* **2019**, *58* (37), 12761–12777. https://doi.org/10.1002/anie.201900932.

(9) 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* (18), 9047–9153. https://doi.org/10.1021/cr5001496.

(10) Shugrue, C. R.; Miller, S. J. Phosphothreonine as a Catalytic Residue in Peptide-Mediated Asymmetric Transfer Hydrogenations of 8-Aminoquinolines. *Angewandte Chemie International Edition* **2015**, *54* (38), 11173–11176. https://doi.org/10.1002/anie.201505898.