

UC Berkeley

UC Berkeley Previously Published Works

Title

Abstracting Complementarity for the Design of Stereoselective Acid Catalysts

Permalink

<https://escholarship.org/uc/item/75k83334>

Journal

Journal of the American Chemical Society, 148(37)

ISSN

0002-7863

Authors

Smith, AndrewL

Le, Christine M

Ye, Baihua

et al.

Publication Date

2026-09-23

DOI

10.1021/jacs.6c13855

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Abstracting complementarity for the design of stereoselective acid catalysts

Andrew L. Smith¹, Christine Le^{1,2}, Baihua Ye^{1,3}, Jie Zhao^{1,4}, F. Dean Toste^{1*}

¹Department of Chemistry, University of California, Berkeley, California 94720, United States

²Department of Chemistry, York University, North York, Ontario M3J 1P3, Canada

³School of Physical Science and Technology, ShanghaiTech University, Shanghai 201210, China

⁴Key Laboratory for Advanced Materials and Joint International Research Laboratory of Precision Chemistry and Molecular Engineering, Feringa Nobel Prize Scientist Joint Research Center, School of Chemistry and Molecular Engineering, East China University of Science and Technology, Shanghai, 200237, China

asymmetric catalysis, organocatalysis, complementarity, data science, predictive modeling

ABSTRACT: Complementarity is invoked to holistically characterize interactions between a catalyst and its substrate. Although the paradigm maintains utility across many fields of chemistry, it is challenging to proactively deploy, requiring post hoc analyses to demonstrate. Here, we report active site-based buried volume furnishes a quantitative and generalizable representation for complementarity by Brønsted acid organocatalysts. The low-dimensional, geometric abstraction helps construct analytical Gaussian models encoding complementarity. From these footholds, experimental validation realizes new stereoselective solutions across diverse structural motifs and functional modes. Unlike other heuristics for stereoinduction, active site-based buried volume guides the design of stereoselective solutions without requiring detailed attention to the mechanisms of stereoinduction. We anticipate the successful abstraction of high-dimensional, stereoselective interactions by complementarity into a single, low-dimensional descriptor will offer a systematic framework for future catalyst design.

Introduction

Selectivity in catalysis has long been attributed to *complementarity* between catalyst active site and substrate. As the name implies, *complementarity* reflects an *ideal* fit between the two species. Historically, the phenomenon characterized the specificity of enzymes where their activity relates to *binding to the transition state stronger than the ground state*.^{1,2} Yet this idea is not exclusive to homogeneous systems. Even heterogeneous catalysts (ex. zeolites) leverage *complementarity*-like intuition to justify selective outcomes from a collection of possible products.^{3,4} In general, *complementarity* is principally attributed to the selective *stabilization* of matched intermediate(s) and/or *destabilization* of others via an array of non-covalent interactions (NCIs, **Fig. 1A**).^{5,6} Each NCI displays varying dependence on distance, where close interactions are disfavored due to Pauli repulsion while far interactions lose electronic coupling. Interactions at *just the right* distance maximize stabilization (**Fig. 1B**).⁷⁻⁹ The collective action of many interactions culminates in the observed *complementarity*.² As the field has evolved to consider increasingly complex chemical phenomena, the

paradigm has remained a unifying criterion for the design of catalysts achieving chemo-, regio-, and stereoselectivity.^{4,10-13}

Although *complementarity* promises an intuitive platform for catalyst design, molecular representations for the paradigm are not well-formed. Accordingly, construction of *complementary* catalysts is rarely predictive. Instead, the principle is often only inferred by elegant, *post hoc* rationalizations drawn from rigorous mechanistic and computational investigations.¹⁴⁻¹⁸ These challenges are especially persistent in the context of stereoselectivity. Within a single transformation, structurally analogous catalysts can display disparate stereoinduction. Without clear avenues to reflect *complementarity* for stereoselectivity, subsequent structure-activity relationships (SARs) drawn can require high-dimensional models (ex. multivariate regression, machine learning-based models) comprised of case-specific molecular features with only circumstantial relevance.^{6,19-24} Simple, qualitative heuristics like quadrant diagrams or stereochemical models can emerge, but they lack quantitative insight. As a result, their persistent utility is quite rare whereby only a few

instances are routinely engaged by chemists (ex. Sharpless' dihydroxylation,²⁵ Evans' auxiliary,²⁶ Ellman's sulfonamide²⁷).

Identifying a molecular representation that clearly reflects *complementarity* would empower chemists to proactively leverage the principle for catalyst design. Such a representation must generalize across structural features and functional applications of the catalyst, escaping any bespoke relationships (**Fig. 1C**). In the context of stereoselectivity, our group has recently reported active site-based buried volume is a robust and holistic representation for diverse sets of Brønsted acid organocatalysts (BAOCs).²⁸ These catalysts are constructed from myriad structural motifs and consequentially access a wide range of functional modes (ex. acid/base, cooperative transition metal catalysis, etc.) enabling chemists with asymmetric access to ever-expanding regions of chemical space.^{4,29–31} As active site-based buried volume is a low-dimensional geometric abstraction of the catalyst active site, we wondered whether the representation can reflect *complementarity*.^{32,33} If so, models constructed with *active site-based* buried volume alone should anticipate stereoselective BAOCs.

In this work, we show active site-based buried volume quantitatively reflects *complementarity*, transforming the fundamental principle into a practical design tool (**Fig. 1D**). We experimentally demonstrate models constructed from active site-based buried volume can analytically represent catalysts with optimal *complementarity*. Furthermore, we demonstrate the representation can transfer *complementarity* across disparate scaffolds, transcending traditional biases surrounding stereoselection. Altogether, this simple geometric abstraction readily steers chemists to ideal regions of catalyst chemical space, furnishing a framework for the design of selective catalysts.

Results

Modeling complementarity with active site-based buried volume

Complementarity by a stereoselective catalyst suggests the catalyst active site selectively *stabilizes* one diastereomeric transition state and/or *destabilizes* the other.² Such a phenomenon implies a reaction-specific optimum, pairing each substrate with a

potentially unique catalyst. Catalysts approaching *complementarity* are expected to furnish increasingly selective outcomes, while those diverging should perform relatively worse (**Fig. 2A**).

Previously, we have reported active site-based buried volume generally furnishes robust univariate linear regressions to stereoselectivity by BAOCs (50+ reaction datasets). The original report described the representation as a reflection of *BAOC confinement* – a quality describing the relative constriction of an active site and typically attributed to stereoselectivity. Based on the theoretical implications of *complementarity* discussed above, one single *confinement* cannot be uniquely stereoselective across *all* reactions. To validate this, the distribution of the most stereoselective BAOCs was plotted against all BAOCs calculated in the 50+ datasets considered herein (see **Supplementary Information** for dataset details). Immediately, it can be recognized that no specific BAOC confinement is uniquely stereoselective (**Fig. 2B**, max. Kolmogorov–Smirnov statistic = 0.23, $p = 0.08$).

This analysis implies there are reaction-specific regions of chemical space occupied by *complementary* stereoselective catalysts. As we approach such regions, BAOCs should be increasingly stereoselective. In our original report, we detected these gradients with linear regression. However, if BAOCs deviate from

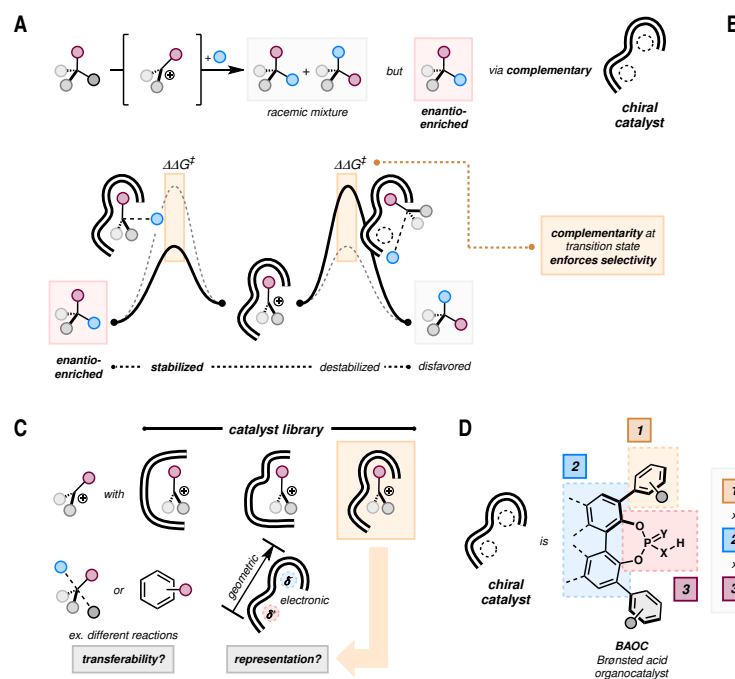


Figure 1. Complementarity between catalyst and substrate (a) Complementarity at the transition state to drive the selective stereoselectivity. (b) Complementarity is the cumulative effect of (c) Complementarity is often recognized *post hoc*. Rendering the descriptor that readily characterizes each catalyst (ex. representation) (ex. transferrable). (d) *This work* – active site-based buried volume – intuitively encodes and reflects complementarity of these catalysts

the *complementary* solution, both more and less confined catalysts should perform worse. This behavior critically implies a nonlinear and non-monotonic modality. Although many of the datasets previously considered could be meaningfully modeled by simple linear regression, physically meaningful fits accommodating this *complementary* behavior would be necessarily advantageous (**Fig. 2C-F**).^{34,35} In fact, based on R^2 , mean absolute error (MAE), leave-one-out cross validation (LOO-CV), and Akaike information criteria (AIC), Gaussian regression furnishes the most meaningful fits across all datasets relative to other linear and non-linear regression-based analyses (Gaussian, avg. $R^2 = 0.72$, avg. MAE = 0.16 kcal/mol; see **Supplementary Information** for numerical fitting methods and statistical analyses, **Supplementary Fig. 2-56** and **Supplementary Table 3-7**).

Well-populated datasets exhibiting the hallmarks of *complementarity* showcase the unique utility of Gaussian regression. We highlight two such datasets here – one reflecting a stereoselective condensation and one reporting access to chiral-at-*N* molecules. **R-I** describes a stereoselective condensation furnishing a chiral spirocycle ($n = 11$, **Fig. 2G**).³⁶ After screening a diverse array of BAOCs, the authors realized only a moderate degree of confinement was required to confer stereoselectivity in the condensation (ref. **CPA-1**). Further increases to confinement only eroded the observed selectivity (ref. **Anth-NTP**). This sensitive evolution of *complementarity* across the wide net of BAOCs screened is excellently captured by Gaussian regression with active site-based buried volume (Gaussian, $R^2 = 0.94$, MAE = 0.15 kcal/mol).

R-II describes access to chiral-at-*N* stereocenters from BAOC-mediated oxidation and substitution ($n = 12$, **Fig. 2H**).³⁷ The authors realized the stereoselectivity from two distinct BAOC scaffolds – chiral phosphoric acids (CPA) and spiroposphoric acids (SPA). The former is a traditional catalyst, but familiar analogs only furnished the *N*-heterocycle in low enantioenrichment (ex. **TCYP-CPA**, obs. 7% *e.e.*). Instead, fine increases to confinement at the active site of SPAs were key to realize stereoselectivity. While sterically bulky CPA-based substituents harmed selectivity, sterically bulky SPA-based substituents were essential (ex. **TCYP-SPA**, obs. 78% *e.e.*). Gaussian regression with active site-based buried volume readily accommodates this emergent *complementarity* (Gaussian, $R^2 = 0.80$, MAE = 0.11 kcal/mol).

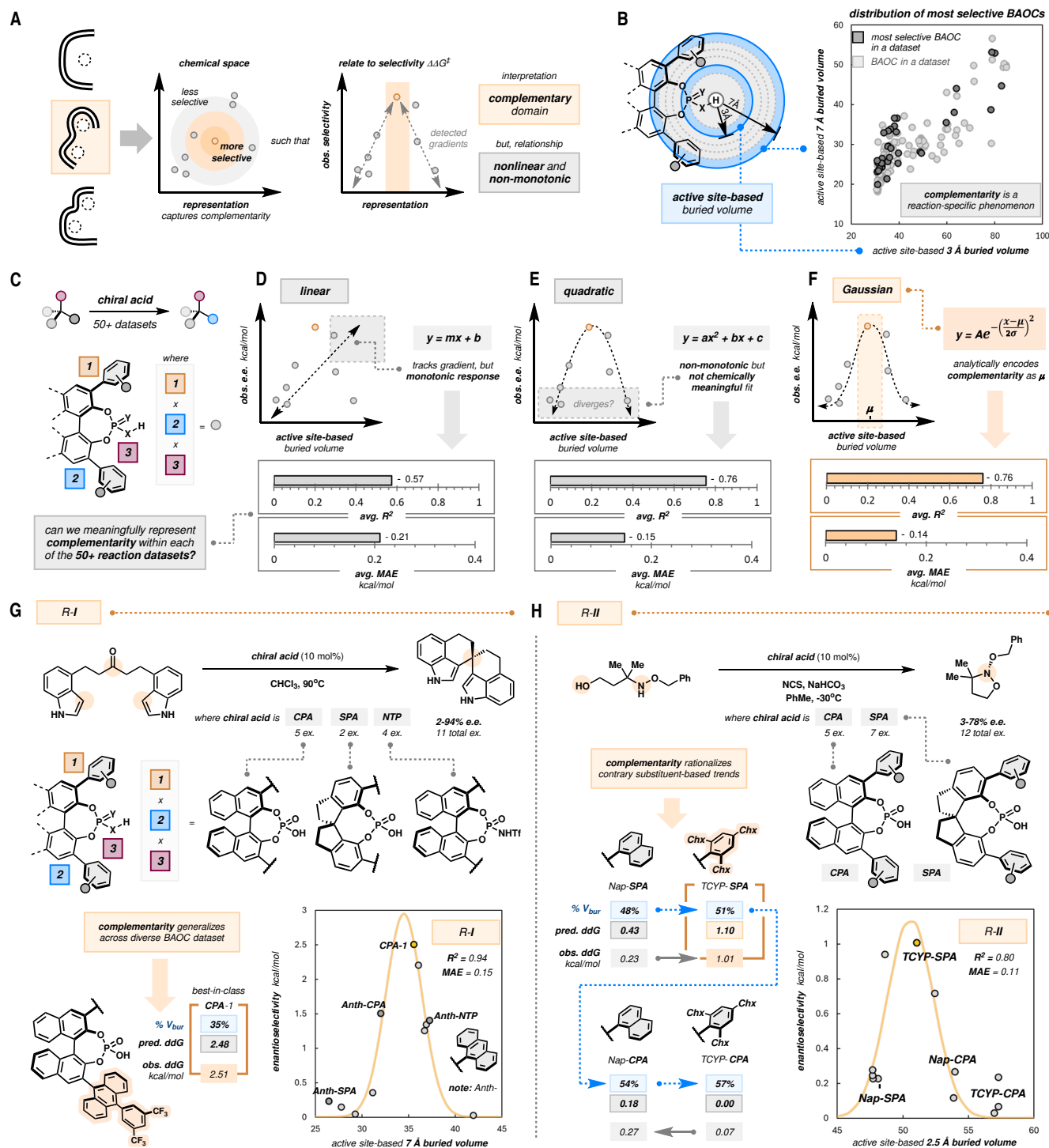


Figure 2. Complementarity is analytically represented by Gaussian regression. (a) Molecular representations for complementarity must capture its nonlinear and non-monotonic nature. (b) No single BAOC confinement universally confers stereoselectivity, aligning with the reaction-specific basis of complementarity. (c) The average performance (avg. R^2 , MAE) by active site-based buried volume of myriad BAOCs in 50+ reaction datasets shows that (d) linear regression cannot reflect complementarity as well as (e) non-physical, quadratic regression or (f) chemically meaningful, Gaussian regression. (g-h) Datasets R-I³⁶ and R-II³⁷ notably exhibit hallmarks of complementarity and are well-represented by Gaussian regression to active site buried volume.

The success of *complementarity* to inform stereoselective solutions above supports the expectation that this criterion conveys the collective consequence of each structural element on stereoinduction. However, each of the above datasets considered only explores stereoselective solutions within analogous scaffolds, where stereoselective NCIs are primarily modulated by altering the NCIs from 3,3'-substituents decorating analogous backbones (**Fig. 3A**).⁶ This design is so common that empirical discovery often exclusively relies on evaluation of BAOCs solely modulated at this position. After all, simple heuristics like quadrant diagrams are frequently deployed for qualitative stereochemical predictions (**Fig. 3B**). Yet, these patterns require specific catalyst architectures to maintain utility. Does the low-dimensional abstraction of *complementarity* continue to reflect selectivity even when disparate catalysts (ex. without 3,3'-substituents) are under consideration?

To investigate whether the representation of *complementarity* transcends structural biases of canonical BAOCs, we considered an alternative architecture critically lacking 3,3'-substituents along the chiral backbone. This new BAOC is constructed from a 5-aryl sulfonamide (SIA) with two flanking phosphoryl groups (denoted as P2-SIA). Now, modular substitution of the aryl group at 5 furnishes a diverse array of P2-SIAs. We report divergent access to 13 P2-SIAs prepared in three steps from trityl NSO (**Fig. 3C**).

Equipped with the previously unreported P2-SIAs, we first validated whether relationships to active site-based buried volume and stereoselectivity could be observed with datasets including traditional BAOCs like CPAs and NTPs (R-III through R-VII, see **Supplementary Information** for full benchmark details). In total, five benchmark transformations were evaluated with P2-SIAs, reflecting the reaction diversity present across all datasets considered, including: thermal or photochemical reactions, additions or eliminations, and inter- or intramolecular reactions (avg. benchmark dataset size = 18 ex. with up to 13 ex. of P2-SIA, see **Supplementary Information** for experimental observations across datasets R-III through R-VII).³⁸⁻⁴² Indeed, across these benchmarks, we observe the P2-SIA-containing datasets are well represented by Gaussian regression constructed with active site-based buried volume (**Fig. 3D**; Gaussian, avg. $R^2 = 0.72$, avg. MAE = 0.17 kcal/mol,

see **Supplementary Information** for regression and statistics).

To assess if *complementarity* truly generalizes across each scaffold, we revisited whether active site-based buried volume encodes relationships between P2-SIA and more familiar BAOCs like CPA and NTP. On inspection, P2-SIAs occupy a complementary region of chemical space unpopulated relative to traditional scaffolds (**Fig. 3E**). As such, we expect in some reactions that P2-SIAs are more *complementary* compared to one scaffold or another. By extension, these P2-SIAs would be advantageously stereoselective.

This is critically the case in R-III. Here, Yamamoto, et al. observed the privileged analog **TRIP-NTP** furnishes an enantioselective protonation of silyl enol ether in modest stereoselectivity (rep. 54% *e.e.*, obs. 60% *e.e.*), while the privileged **TRIP-CPA** is unsuitable for the transformation (reported n.d., obs. 8% *e.e.*).³⁸ First, we probed whether Yamamoto's **TRIP-NTP** was *complementary* by screening an additional five NTPs populating the surrounding regions. It was observed both more and less confined BAOCs were less stereoselective compared to Yamamoto's reported solution (Gaussian, *train* $R^2 = 0.77$, *train* MAE = 0.10 kcal/mol). Over this same domain, P2-SIAs populate the chemical space occupied by less confined NTPs but still modestly approach **TRIP-NTP**. This analysis anticipates that the closest P2-SIA (P2-5) is the *next* most stereoselective (pred. 36% *e.e.*), exceeding all other BAOCs. Consistent with the model, P2-5 affords the stereoselective protonation in superior stereoselectivity to the remaining dataset (obs. 36% *e.e.*, next most stereoselective obs. 30% *e.e.*). Moreover, the performance of the remaining seven P2-SIAs is very well-aligned with predictions made by the NTP-and-CPA-trained model (*test* MAE = 0.05 kcal/mol), highlighting the transferability of the representation for *complementarity*, even across disparate catalytic scaffolds lacking familiar structural motifs.

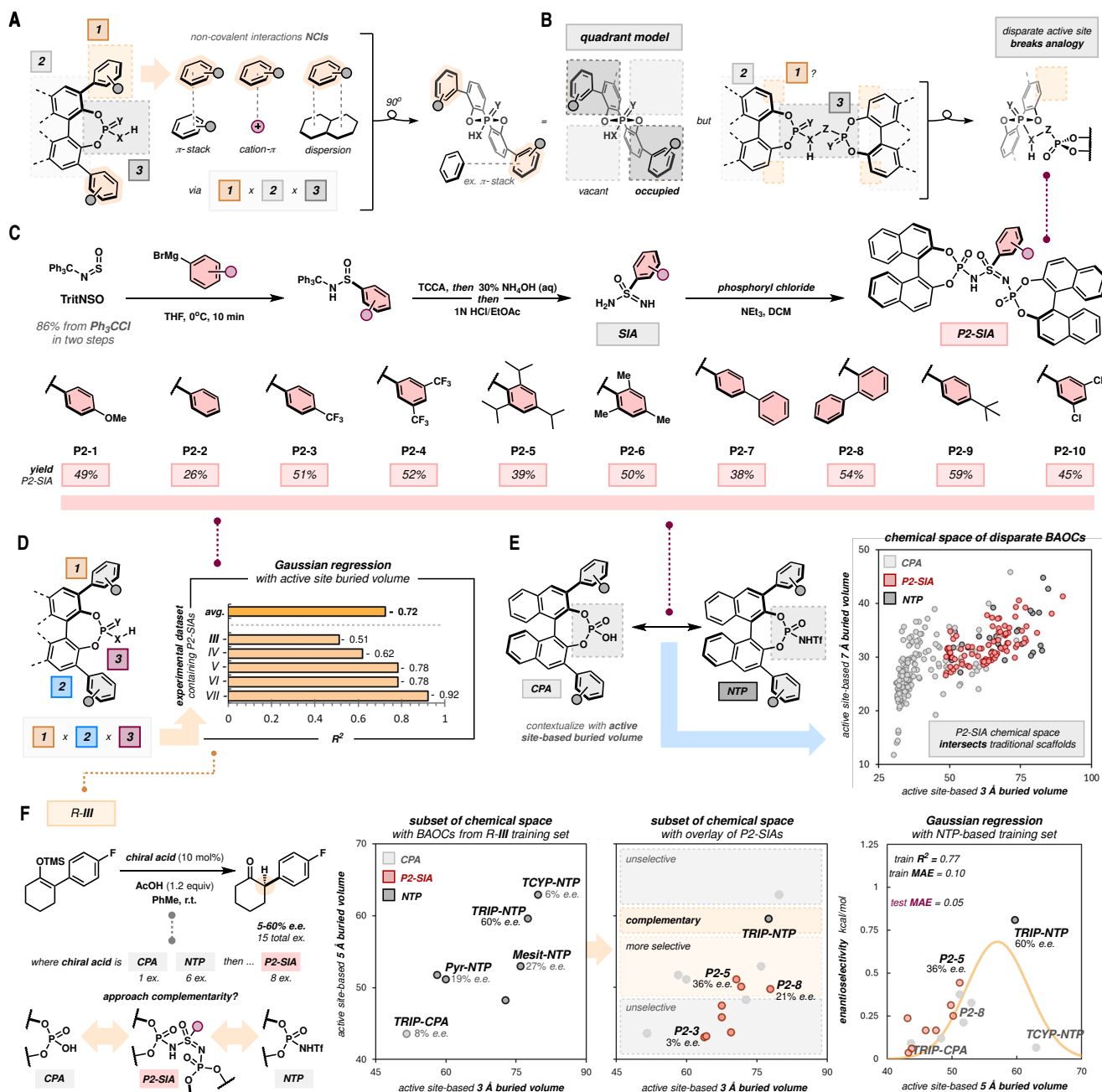


Figure 3. Active site-based buried volume coherently reflects complementarity across disparate catalyst scaffolds. (a) Stereoselective NCIs are traditionally conferred by modular 3,3'-substituents decorating chiral backbones. (b) Simple stereochemical models reflecting complementarity require analogous sense of enantioinduction. (c) Synthetic access into an unreported, disparate catalyst scaffold – P2-SIA (see **Supplementary Information** for full synthetic details). (d) Stereoselectivity by BAOCs in five experimental benchmarks (R-III through R-VII) including P2-SIAs are well-represented by Gaussian regression with active site-based buried volume (see **Supplementary Information** for benchmark reaction details and datasets). (e) P2-SIAs intersect the region of chemical space between traditional BAOCs like CPAs and NTPs. (f) Dataset R-III³⁸ showcases how *near-complementary* P2-SIAs (e.g. P2-5) are advantageously stereoselective relative to traditional BAOCs like CPAs and NTPs.

Complementarity anticipates new stereoselective solutions

The application of Gaussian regression models constructed with *active site-based* buried volume promises the discovery of *complementary* BAOCs. If *complementarity* is to be a practical tool, then it must be proactively leveraged to predict previously unknown stereoselective solutions (**Fig. 4A**). Yet, at the outset of catalyst discovery, chemists turn to *privileged* solutions. In the context of BAOCs, *privileged* solutions drastically minimize the number of 3,3'-substituents under consideration. In fact, across the 50+ datasets considered herein, 2,4,6-triisopropylphenyl (TRIP), 9-anthracenyl, and 3,5-bistrifluoromethyl are the most popular, deployed in nearly 80% of all the reactions (note – the average substituent is present in less than 10%, **Fig. 4B**). We wondered whether Gaussian regression trained using only BAOCs from these *empirical* choices accurately anticipates if other, less popular BAOCs are more stereoselective. To this end, we deployed Gaussian regression from empirical data to nominate *complementary* regions of chemical space. Then, we leveraged this insight to identify new stereoselective BAOCs.

As proof-of-concept, we revisited R-I from **Fig. 2F**. In this dataset, a unique anthracene-derived substituent was designed to furnish the condensation in high stereoselectivity (**CPA-1**, rep. 94% *e.e.*), exceeding other popular, anthracenyl substituted BAOCs (**Anth-CPA**, **Anth-NTP**, and **Anth-SPA**, rep. up to 78% *e.e.*, 1.51 kcal/mol). We trained a Gaussian regression model solely on the *empirically popular* BAOCs – those constructed with the above, high-frequency 3,3'-substituents ($n = 5$, $\text{train } R^2 = 0.97$, $\text{train MAE} = 0.06$ kcal/mol, **Fig. 4C**). Still, among the other 6 BAOCs, the model predicts **CPA-1** to be the most stereoselective (pred. 89% *e.e.*, 2.05 kcal/mol), expecting to exceed the training set's best instance by +0.54 kcal/mol. This prediction is in strong agreement with the reported outcome (rep. 94% *e.e.*, 2.54 kcal/mol). Altogether, the test set is well-represented by the *empirical* model ($\text{test } R^2 = 0.64$, $\text{test MAE} = 0.48$ kcal/mol).

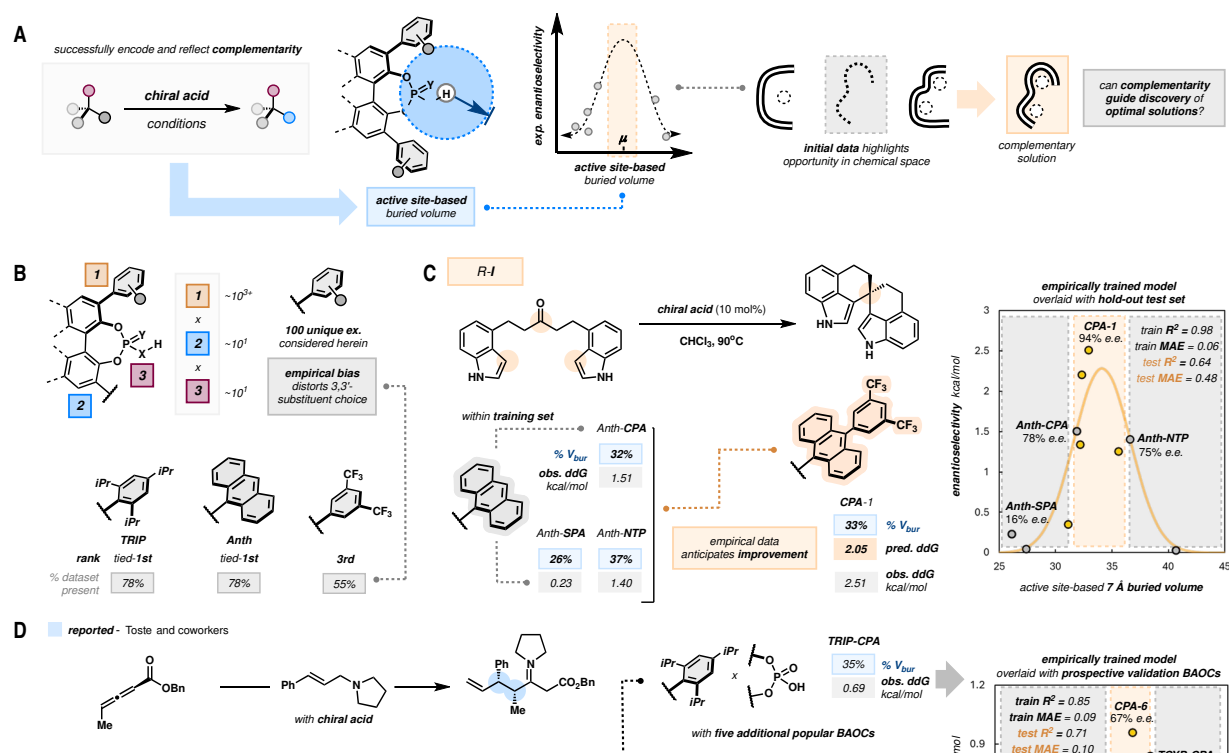
Motivated by this initial success, we sought to apply this approach to identify stereoselective candidates in a catalyst optimization campaign. Building on our prior report of an enantioselective allenolate-Claisen rearrangement,⁴³ we hypothesized

that propargylic esters could function as conjugate acceptors with allyl amines to generate analogous reactive intermediates (**Fig. 4D**, see **Supplementary Information** for experimental conditions). Initial evaluation revealed that the privileged **TRIP-CPA** catalyzed the desired transformation with high diastereoselectivity and moderate enantioselectivity (obs. >20:1 *d.r.* and 53% *e.e.*). However, screening of other empirically successful catalysts, such as **TCYP-CPA**, resulted in only marginal improvements in selectivity. To guide further optimization, we trained a Gaussian regression model on the compiled dataset ($\text{train } R^2 = 0.85$, $\text{train MAE} = 0.093$ kcal/mol). The model predicted a narrow region of opportunity between **TRIP-CPA** and **TCYP-CPA**. Systematic evaluation of candidate 3,3'-substituents, informed by computed buried volume parameters, identified a single promising CPA substituent – 2,4,6-tricyclopentylphenyl (pred. 61% *e.e.*). Notably, this motif appeared in only two examples within the 50+ datasets considered herein. Synthesis of the BAO (**CPA-6**) and experimental evaluation confirmed the prediction, delivering an optimal result (obs. 67% *e.e.*, +0.12 kcal/mol over training set). In contrast, the analogous 2,4,6-tricycloheptylphenyl substituted (**CPA-7**) only displayed attenuated selectivity as predicted by the model (pred. 38% *e.e.*, obs. 41% *e.e.*). Additional derivatives resembling **TRIP-CPA** were designed and prepared, but none were expected nor observed to improve over **CPA-6**. Across the BAOs prepared, good agreement with the trained Gaussian model was observed ($\text{test } R^2 = 0.71$, $\text{test MAE} = 0.103$ kcal/mol).

We further demonstrate the utility of *complementarity* in the data-driven optimization campaign to discover a stereoselective solution for another previously unreported transformation – the stereoselective reduction of propargyl alcohols to enantioenriched allenes (**Fig. 4E**, see **Supplementary Information** for experimental conditions). Initial reaction discovery relied on empirical insights, whereby popular BAOs only furnished enantioenrichment of the model allene with an avg. 47% *e.e.*, with the privileged **TRIP-CPA** providing the best initial result (75% *e.e.*, see **Supplementary Fig. 71**). Insights based on *confinement* naively would suggest more confined chiral acids should be more stereoselective, but empirical evaluation with of more confined CPAs did not furnish marked improvements over the privileged **TRIP-CPA** (ref. **CPA-S16**, **Supplementary Fig. 72**). Instead, by leveraging the

empirical training set, we instead constructed a Gaussian regression model using active site-based buried volume representation, revealing a promising, but critically unpopulated *complementary* domain (*train* $R^2 = 0.78$, *train* MAE = 0.14 kcal/mol, see **Supplementary Fig. 72**). Guided by the model, we designed, prepared, and evaluated three CPAs to realize an improved stereoselective solution within the anticipated *complementary* domain. The model predicted **CPA-S23**, the most *complementary* instance, would furnish the allene with *best-in-class* enrichment (pred. 84% e.e.). Indeed, experimental evaluation validated this prediction, obs. 83% e.e. (+0.25 kcal/mol over **TRIP-CPA**). Overall, the model predictions were in strong agreement with the observations across the prospective test set examples (*test* $R^2 = 0.85$, *test* MAE = 0.09 kcal/mol, see **Supplementary Fig. 73**).

Gaussian regression with active site buried volume accurately anticipates stereoselective solutions, highlighting experimental opportunities to reveal optimized solutions over training data. Whereas three examples showcasing this are highlighted above, an additional seven analyses are detailed in Supplementary Information (**Supplementary Fig. 64-70**). To summarize, we observed the Gaussian regression models solely trained on empirically popular catalysts readily identified more stereoselective solutions from substantially less popular choices (avg. improvement over training data = +0.45 kcal/mol). Beyond prediction of *best-in-class* stereoselective choices, the trained models also generalized well across the less popular, test set catalysts (avg. *test* $R^2 = 0.73$, avg. *test* MAE = 0.25 kcal/mol).



Conclusion

Complementarity has promised a fundamental paradigm for the design of stereoselective catalysts, yet its realization has long been restricted to retrospective analyses. Here, we have shown that a low-dimensional geometric abstraction of catalyst active sites can reflect *complementarity* in a data-driven manner and render it operational. Across a wide array of acid-catalyzed benchmarks, the representation was deployed in Gaussian regression models to anticipate stereoselective solutions from both familiar and previously unreported BAOC scaffolds (ex. P2-SIA). By reframing *complementarity* from a qualitative principle into a quantifiable workflow, the discovery of new stereoselective catalysts can be approached systematically. Still, some outstanding challenges remain. During data-driven optimization campaigns, Gaussian regression not only helps to prioritize the most *complementary* regions of chemical space but also quantify the maximum selectivity expected to be observed. As a result, fitting and error associated with the amplitude of the Gaussian regression model reports a statistical, upper limit anticipated for stereoselectivity by the BAOCs (within a given dataset). In cases where this predicted limit remains unsatisfactory to the experimentalist, the approach reported offers little guidance to anticipate alternative reaction mechanisms where the desired selectivity may be realized. Accordingly, the present framework most effectively guides optimization of catalyst properties with respect to one view of *complementarity*. Anticipating which alternative mechanisms to engage for more satisfactory stereocontrol remains an important direction for future research.

ASSOCIATED CONTENT

Supporting Information.

This material is available free of charge via the Internet at <http://pubs.acs.org>.

General materials and methods, detailed statistical analyses, experimental benchmarks and datasets, synthetic methods, and NMR spectra. (PDF)
Table of *active site-based* descriptors (CSV)

AUTHOR INFORMATION

Corresponding Author

*fdtoste@berkeley.edu.

Present Addresses

¹Department of Chemistry, University of California, Berkeley, California 94720, United States

²Department of Chemistry, York University, North York, Ontario M3J 1P3, Canada

³School of Physical Science and Technology, ShanghaiTech University, Shanghai 201210, China

⁴Key Laboratory for Advanced Materials and Joint International Research Laboratory of Precision Chemistry and Molecular Engineering, Feringa Nobel Prize Scientist Joint Research Center, School of Chemistry and Molecular Engineering, East China University of Science and Technology, Shanghai, 200237, China

Author Contributions

The manuscript was written through contributions of all authors. / All authors have given approval to the final version of the manuscript. /

Funding Sources

CCI Center for Computer Assisted Synthesis, CHE-2202693 (ALS).

ACKNOWLEDGMENT

The authors thank CCI Center for Computer Assisted Synthesis (CHE-2202693) for funding. The authors thank Dr. Hasan Celik and UC Berkeley's NMR facility in the College of Chemistry (CoC NMR) for spectroscopic assistance. Instruments in the CoC-NMR are supported in part by National Institute of Health grant S10OD024998. The authors thank Dr. Kathleen Durkin, Dr. Azhagiya Singam, Dr. Dave Small, and the Molecular Graphics and Computation Facility (MGCF) at the University of California, Berkeley for calculation and workflow assistance, as supported by National Institute of Health grant S10OD034382. Drs. C.L., B.Y. and J. Z. performed experiments at UC Berkeley supported by NIGMS (R35 GM118190). B.Y. is Grateful to the Swiss National Science Foundation (SNSF) for a postdoctoral mobility fellowship (P2ELP2_158881).

ABBREVIATIONS

BAOC, Brønsted acid organocatalyst
NCI, non-covalent interaction
SAR, structure-activity relationship
MAE, mean absolute error
CPA, chiral phosphoric acid
NTP, *N*-triflylphosphoramidate
SPA, spirophosphoric acid
P2-SIA, diphosphoryl sulfonimidamide

REFERENCES

- (1) Nelson, D. L.; Cox, M., M. *Lehninger Principles of Biochemistry*, 4th ed.; W.H. Freeman & Co., 2005.
- (2) Anslyn, E.; Dougherty, D. *Modern Physical Organic Chemistry*; University Science Books, 2006.
- (3) Kaeding, W. W.; Chu, C.; Young, L. B.; Weinstein, B.; Butter, S. A. Selective Alkylation of Toluene with Methanol to Produce Para-Xylene. *Journal of Catalysis* **1981**, 67 (1), 159–174.
- (4) Mitschke, B.; Turberg, M.; List, B. Confinement as a Unifying Element in Selective Catalysis. *Chem* **2020**, 6 (10), 2515–2532.
- (5) Sak, M. H.; Jacobsen, E. N. Selective Noncovalent Catalysis with Small Molecules. *Chem. Rev.* **2025**, 125 (17), 8319–8369.
- (6) Betinol, I. O.; Kuang, Y.; Mulley, B. P.; Reid, J. P. Controlling Stereoselectivity with Noncovalent Interactions in Chiral Phosphoric Acid Organocatalysis. *Chem. Rev.* **2025**, 125 (8), 4184–4286.
- (7) Orlandi, M.; Coelho, J. A. S.; Hilton, M. J.; Toste, F. D.; Sigman, M. S. Parametrization of Non-Covalent Interactions for Transition State Interrogation Applied to Asymmetric Catalysis. *J. Am. Chem. Soc.* **2017**, 139 (20), 6803–6806.
- (8) Neel, A. J.; Hilton, M. J.; Sigman, M. S.; Toste, F. D. Exploiting Non-Covalent π Interactions for Catalyst Design. *Nature* **2017**, 543 (7647), 637–646.
- (9) Treacy, S. M.; Smith, A. L.; Bergman, R. G.; Raymond, K. N.; Toste, F. D. Supramolecular Catalyzed Cascade Reduction of Azaarenes Interrogated via Data Science. *J. Am. Chem. Soc.* **2024**, 146 (43), 29792–29800.
- (10) Mouarrawis, V.; Plessius, R.; van der Vlugt, J. I.; Reek, J. N. H. Confinement Effects in Catalysis Using Well-Defined Materials and Cages. *Front. Chem.* **2018**, 6.
- (11) Morimoto, M.; Bierschenk, S. M.; Xia, K. T.; Bergman, R. G.; Raymond, K. N.; Toste, F. D. Advances in Supramolecular Host-Mediated Reactivity. *Nat Catal* **2020**, 3 (12), 969–984.
- (12) Grimm, J. A. A.; Zhou, H.; Properzi, R.; Leutzsch, M.; Bistoni, G.; Nienhaus, J.; List, B. Catalytic Asymmetric Synthesis of Cannabinoids and Menthol from Neral. *Nature* **2023**, 615 (7953), 634–639.
- (13) Luo, N.; Turberg, M.; Leutzsch, M.; Mitschke, B.; Brunen, S.; Wakchaure, V. N.; Nöthling, N.; Schelwies, M.; Pelzer, R.; List, B. The Catalytic Asymmetric Polyene Cyclization of Homofarnesol to Ambrox. *Nature* **2024**, 632 (8026), 795–801.
- (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 (25), 7910–7917.
- (15) Champagne, P. A.; Houk, K. N. Origins of Selectivity and General Model for Chiral Phosphoric Acid-Catalyzed Oxetane Desymmetrizations. *J. Am. Chem. Soc.* **2016**, 138 (38), 12356–12359.
- (16) Alexandre Champagne, P. Identifying the True Origins of Selectivity in Chiral Phosphoric Acid Catalyzed N -Acyl-Azetidine Desymmetrizations. *Chemical Science* **2021**, 12 (47), 15662–15672.
- (17) Duarte, F.; Paton, R. S. Molecular Recognition in Asymmetric Counteranion Catalysis: Understanding Chiral Phosphate-Mediated Desymmetrization. *J. Am. Chem. Soc.* **2017**, 139 (26), 8886–8896.
- (18) Reid, J. P.; Goodman, J. M. Selecting Chiral BINOL-Derived Phosphoric Acid Catalysts: General Model To Identify Steric Features Essential for Enantioselectivity. *Chemistry – A European Journal* **2017**, 23 (57), 14248–14260.
- (19) B. Santiago, C.; Guo, J.-Y.; S. Sigman, M. Predictive and Mechanistic Multivariate Linear Regression Models for Reaction Development. *Chemical Science* **2018**, 9 (9), 2398–2412.
- (20) Gallegos, L. C.; Luchini, G.; St. John, P. C.; Kim, S.; Paton, R. S. Importance of Engineered and Learned Molecular Representations in Predicting Organic Reactivity, Selectivity, and Chemical Properties. *Acc. Chem. Res.* **2021**, 54 (4), 827–836.
- (21) 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 (6223), 737–743.
- (22) 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 (6424).

(23) Zahrt, A. F.; Athavale, S. V.; Denmark, S. E. Quantitative Structure–Selectivity Relationships in Enantioselective Catalysis: Past, Present, and Future. *Chem. Rev.* **2020**, 120 (3), 1620–1689.

(24) Tsuji, N.; Sidorov, P.; Zhu, C.; Nagata, Y.; Gimadiev, T.; Varnek, A.; List, B. Predicting Highly Enantioselective Catalysts Using Tunable Fragment Descriptors. *Angewandte Chemie International Edition* **2023**, 62 (11), e202218659.

(25) Kolb, H. C.; VanNieuwenhze, M. S.; Sharpless, K. B. Catalytic Asymmetric Dihydroxylation. *Chem. Rev.* **1994**, 94 (8), 2483–2547.

(26) Evans, D. A.; Bartroli, J.; Shih, T. L. Enantioselective Aldol Condensations. 2. Erythro-Selective Chiral Aldol Condensations via Boron Enolates. *J. Am. Chem. Soc.* **1981**, 103 (8), 2127–2129.

(27) Liu, G.; Cogan, D. A.; Ellman, J. A. Catalytic Asymmetric Synthesis of Tert-Butanesulfinamide. Application to the Asymmetric Synthesis of Amines. *J. Am. Chem. Soc.* **1997**, 119 (41), 9913–9914.

(28) Smith, A. L.; Toste, F. D. Stereoselective Generalizations over Diverse Sets of Chiral Acids Enabled by Buried Volume. *J. Am. Chem. Soc.* **2026**,

(29) Akiyama, T. Stronger Brønsted Acids. *Chem. Rev.* **2007**, 107 (12), 5744–5758.

(30) Handjaya, J. P.; Patankar, N.; Reid, J. P. The Diversity and Evolution of Chiral Brønsted Acid Structures. *Chemistry – A European Journal* **2024**, 30 (46), e202400921.

(31) Hamilton, G. L.; Kang, E. J.; Mba, M.; Toste, F. D. A Powerful Chiral Counterion Strategy for Asymmetric Transition Metal Catalysis. *Science* **2007**, 317.

(32) Newman-Stonebraker, S. H.; Smith, S. R.; Borowski, J. E.; Peters, E.; Gensch, T.; Johnson, H. C.; Sigman, M. S.; Doyle, A. G. Univariate Classification of Phosphine Ligation State and Reactivity in Cross-Coupling Catalysis. *Science* **2021**, 374 (6565), 301–308.

(33) Antinucci, G.; Dereli, B.; Vittoria, A.; Budzelaar, P. H. M.; Cipullo, R.; Goryunov, G. P.;

Kulyabin, P. S.; Uborsky, D. V.; Cavallo, L.; Ehm, C.; Voskoboynikov, A. Z.; Busico, V. Selection of Low-Dimensional 3-D Geometric Descriptors for Accurate Enantioselectivity Prediction. *ACS Catal.* **2022**, 12 (12), 6934–6945.

(34) Reid, J. P.; Sigman, M. S. Comparing Quantitative Prediction Methods for the Discovery of Small-Molecule Chiral Catalysts. *Nat Rev Chem* **2018**, 2 (10), 290–305.

(35) Ardkhean, R.; Roth, P. M. C.; Maksymowicz, R. M.; Curran, A.; Peng, Q.; Paton, R. S.; Fletcher, S. P. Enantioselective Conjugate Addition Catalyzed by a Copper Phosphoramidite Complex: Computational and Experimental Exploration of Asymmetric Induction. *ACS Catal.* **2017**, 7 (10), 6729–6737.

(36) Zhao, H.-W.; Jiang, F.; Chen, S.; Hu, J.; Xiang, S.-H.; Ding, W.-Y.; Lu, W.; Tan, B. Organocatalytic Asymmetric Construction and Application of Axially Chiral Spiro-Bisindoles. *Angewandte Chemie International Edition* **2025**, 64 (12).

(37) Wu, S.; Chen, P.; Duan, M.; Jiang, P.-Y.; Zhou, Q.; Xiang, S.-H.; Houk, K. N.; Tan, B. Controlling Pyramidal Nitrogen Chirality by Asymmetric Organocatalysis. *Nature* **2025**, 647 (8091), 897–905.

(38) Cheon, C. H.; Yamamoto, H. A Brønsted Acid Catalyst for the Enantioselective Protonation Reaction. *J. Am. Chem. Soc.* **2008**, 130 (29), 9246–9247.

(39) Saito, K.; Moriya, Y.; Akiyama, T. Chiral Phosphoric Acid Catalyzed Asymmetric Synthesis of 2-Substituted 2,3-Dihydro-4-Quinolones by a Protecting-Group-Free Approach. *Org. Lett.* **2015**, 17 (13), 3202–3205.

(40) Liao, S.; Čorić, I.; Wang, Q.; List, B. Activation of H₂O₂ by Chiral Confined Brønsted Acids: A Highly Enantioselective Catalytic Sulfoxidation. *J. Am. Chem. Soc.* **2012**, 134 (26), 10765–10768.

(41) Sherbrook, E. M.; Genzink, M. J.; Park, B.; Guzei, I. A.; Baik, M.-H.; Yoon, T. P. Chiral Brønsted Acid-Controlled Intermolecular Asymmetric [2 + 2] Photocycloadditions. *Nat Commun* **2021**, 12 (1), 5735.

(42) 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* (16), 6698–6705.

(43) Miró, J.; Gensch, T.; Ellwart, M.; Han, S.-J.; Lin, H.-H.; Sigman, M. S.; Toste, F. D. Enantioselective Allenolate-Claisen Rearrangement Using Chiral Phosphate Catalysts. *J. Am. Chem. Soc.* **2020**, *142* (13), 6390–6399.

(44) Sherbrook, E. M.; Jung, H.; Cho, D.; Baik, M.-H.; Yoon, T. P. Brønsted Acid Catalysis of Photosensitized Cycloadditions. *Chem. Sci.* **2020**, *11* (3), 856–861.

