ACS Catalysis

pubs.acs.org/acscatalysis Research Article

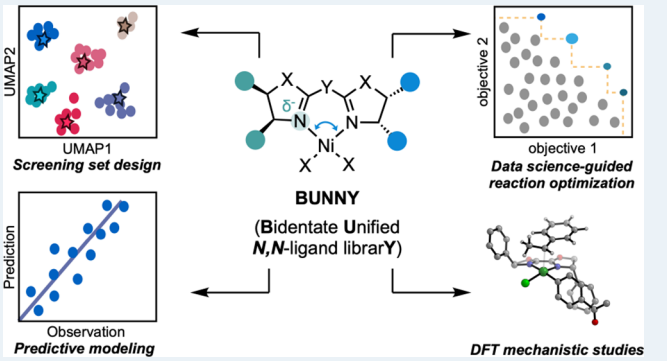
BUNNY: An *N,N*-Bidentate Nitrogen Ligand Descriptor Library—Development and Application to Modeling of Ni-Catalyzed Asymmetric Cross-Electrophile Coupling Reactions

Neyci E. Gutiérrez-Valencia,‡ Jules Schleinitz,‡ Therese H. Wild,‡ Wendy L. Williams, Danielle Mantin, Abigail G. Doyle,* Sarah E. Reisman,* and Matthew S. Sigman*

Cite This: https://doi.org/10.1021/acscatal.6c02585 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: *N,N*-Bidentate ligands are widely employed in Ni-catalyzed cross-electrophile coupling (CEC) reactions; however, it is often difficult to predict *a priori* which scaffold will provide optimal selectivity and yield for a reaction under development. More generally, for a given Ni-catalyzed reaction, models that provide structure-reactivity and structure-selectivity relationships across different *N,N*-bidentate ligand scaffolds remain elusive. Here, we report BUNNY, a density functional theory-based descriptor library of approximately 1100 *N,N*-bidentate ligands designed to support modeling tasks for Ni-catalyzed cross-coupling. Using BUNNY, a screening set of 31 ligands was selected to represent 7 ligand scaffolds, and the value of the screening set was demonstrated by modeling of two Ni-catalyzed CEC case studies. The first case study demonstrates that enantioselectivity can be modeled across different ligand families for two different benzylic electrophiles. The second case study used the screening library to collect ee data for an established Ni-CEC, which was used to guide development of an enantioselective Ni-catalyzed CEC of a different but related electrophile. Differences in descriptors selected by the enantioselectivity models for the two case studies inspired investigation of the radical capture and reductive elimination steps by DFT, which found that either step can be selectivity-determining, depending on the reaction. This work illustrates how BUNNY can be used to model reactivity and selectivity across ligand scaffolds, guiding reaction development and mechanistic understanding.



KEYWORDS: *nickel, cross-electrophile coupling, enantioselectivity, data-science, descriptors*

INTRODUCTION

The identification of ligands that confer high reactivity and selectivity in metal-catalyzed reactions is often a major bottleneck in method development as the structural features responsible for these properties are not well understood, and ligand synthesis can be expensive and time-consuming. Therefore, research into quantifying ligand structure-reactivity and structure-selectivity relationships has intensified to accelerate reaction optimization (Figure 1A). Indeed, the identification of such correlations can be used to guide ligand design and to provide insights on reaction mechanism. Historically, quantification of ligand effects relied on experimentally determined descriptors such as those reported by Tolman.[1,2] More recently, computed ligand descriptors ranging from structure-derived fingerprints to higher level quantum-based descriptors have emerged as advantageous alternatives (Figure 1B);[3,4] the compilation of large virtual libraries of computed descriptors has been used to design ligand training sets prior to an optimization campaign and to construct statistical models for the prediction of improved ligand performance. Among the computed descriptors, DFT-based strategies have provided a platform to define nuanced ligand steric and electronic properties[5−10] and are often preferred for analyzing small, sparse data sets[11,12] involved in many catalyst optimization campaigns. As an example, *Kraken*, an open-source descriptor library of monodentate-phosphorus containing compounds, has enabled widespread use of machine learning tools for retrospective analysis of reaction data and prospective phosphine design.[13]

Given the success of the *Kraken* library, we were interested in developing a similar descriptor library for bidentate nitrogen

Received: April 1, 2026
Revised: May 20, 2026
Accepted: May 20, 2026

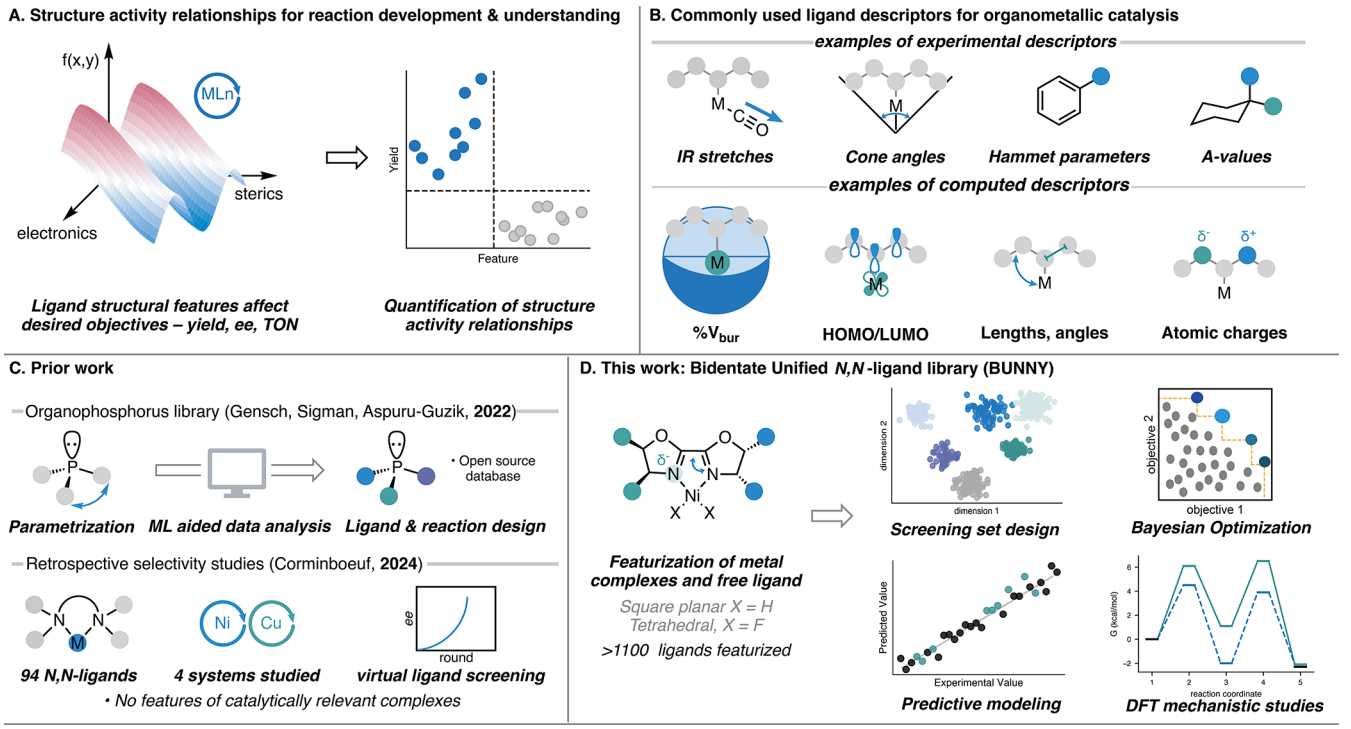


Figure 1. (A) Structure activity relationships can be used to optimize reactions and improve chemical understanding. (B) Examples of commonly used descriptors for homogeneous metal-catalyzed reactions. (C) Prior work on the development and use of ligand descriptor libraries. (D) Design of a library of *N,N*-bidentate ligands and its application for optimization and mechanistic analysis of nickel-catalyzed CEC reactions.

ligands, which have seen extensive use in catalysis. This effort was motivated by our experience in the development of Ni-catalyzed C(sp^2)–C(sp^3) cross-electrophile couplings (CEC),[14] for which new electrophile pairs often require the use of different ligands or ligand classes to achieve optimal yield and selectivity. This empirical observation, coupled to a dearth of understanding as to how ligand properties vary across disparate *N,N*-ligand classes, makes it challenging to anticipate which ligand class will be optimal for a new reaction of interest. In this context, the Liao group demonstrated that black box models trained using high-throughput experimentation data and fingerprint descriptors could help identify chiral ligands for different Ni-catalyzed reactions.[15] Recently, Waser and Corminboeuf showed that description of 312 bidentate ligands enabled prediction of highly selective ligands for three Cu- and one Ni-catalyzed reaction.[16] Although this descriptor library proved helpful for selectivity optimization, we hypothesized that including conformational effects and catalytically relevant structures with Ni-centered descriptors would facilitate the construction of improved statistical models to guide ligand design and provide insights into how *N,N*-bidentate ligand structure and class impact reactivity and selectivity.

Herein, we present the construction of an open-source *N,N*-bidentate ligand library (BUNNY) designed to encompass commercially available and synthetically accessible *N,N*-bidentate ligands. The library includes conformer-dependent DFT-calculated descriptors of 1106 ligands, as well as descriptors of the corresponding [ligand]NiH_2 and [ligand]NiF_2 complexes. The utility of the library is demonstrated in two use cases. First, we designed a ligand screening set to evaluate the response surface of two established Ni-catalyzed CECs, in which the same C(sp^2) electrophile (aryl iodide) is cross-coupled with two different C(sp^3) electrophiles (benzyl chlorides and styrenyl aziridines).[17,18] We demonstrate that the *N,N*-ligand descriptors support modeling reactivity and selectivity trends across disparate ligand classes and transferring these trends to new reaction partners, providing a means to accelerate ligand optimization of new reactions. Then, to probe the generality of the approach, we evaluated the ligand library on a previously developed Ni-CEC between acyl chlorides and benzyl chlorides[19] and leveraged that information to develop a new reaction between acyl chlorides and aziridines, affording enantioenriched β-amino ketones, a pharmaceutically relevant motif.[20]

These case studies showcase the utility of the descriptor library for the optimization of new Ni-CECs. The comparison of the modeling performances in case study I and case study II, supported with a detailed DFT analysis, also provided insights into the mechanistic conditions required to build statistical models unifying ligand classes. Specifically, this work suggests that descriptors extracted from structures resembling those involved in enantio-determining steps improve the modeling of reaction selectivity. Finally, our results demonstrate that the reactivity and selectivity of Ni-CECs are more sensitive to the choice of C(sp^2) electrophile than C(sp^3) electrophile.

■ LIBRARY DEVELOPMENT

Curation and Scope of the Descriptor Library

The objective of this descriptor library is to enable the design of a set of *N,N*-bidentate ligands to screen reactions and train models for reactivity and enantioselectivity prediction. When modeling chemical reactions, there is a trade-off between the extent of the domain of applicability of a model and its accuracy.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX–XXX

Therefore, we limited the chemical space to bidentate *N*-donor ligands and removed any functionality incompatible with practical catalytic applications. An initial set of ∼300 ligands was selected from mining reactions from the Doyle, Reisman, and Sigman labs as well as a limited search of Ni-CEC literature (see Table S34 for details). From this search, we identified seven ligand subclasses most utilized in Ni-catalyzed cross-coupling reactions. Two achiral ligand classes, bipyridine (bpy) and phenanthroline (phen) classes, were included as well as five chiral ligand classes applied in asymmetric Ni-CECs: bi-oxazoline (BiOx), bi-imidazole (BiIm), bisoxazoline (BOX), 2-(2-pyridyl)oxazoline (PyOx), and 2-(2-pyridyl)imidazole (PyIm). Expanding from the initial search space, ligand additions were driven by commercial availability or synthetic accessibility to increase the impact of the library on experimental applications. Thus, commercially available ligands (based on a Reaxys search[21]) were added to the library. The library was further expanded to include virtual ligands that may be accessed through commercially available starting materials and established synthetic routes. As above, these spaces were filtered to limit the compounds to bidentate ligands and remove any ligands with functional groups likely to make them incompatible with Ni-catalyzed cross-couplings. Further details of this curation may be found in the SI. A total of 1106 ligands is included in the descriptor library.

Construction of the DFT Ligand Descriptor Library

Ni-CECs have been proposed to proceed through different mechanisms depending on the identity of the ligand or substrate.[14,22,23] The Ni(II) oxidation state was chosen for descriptor computation because it is generally the most stable oxidation state and therefore likely to be relevant to the broadest scope of reactions. The relative energy between Ni(II) singlet and triplet complexes is highly dependent on the coordination environment; therefore, to develop a set of descriptors enabling downstream modeling without assuming a preferred spin state, we computed two Ni complexes with different spin states (Figure 2A). In practice, descriptors were calculated on a square planar [ligand]NiH_2 complex (singlet), a tetrahedral [ligand]NiF_2 complex (triplet), and on the ligand only. This approach was inspired by the *Kraken* workflow[13] and the development of descriptors for bisphosphine ligands.[24] The computational workflow is detailed in Figure 2A; see Materials and Methods for computational details.

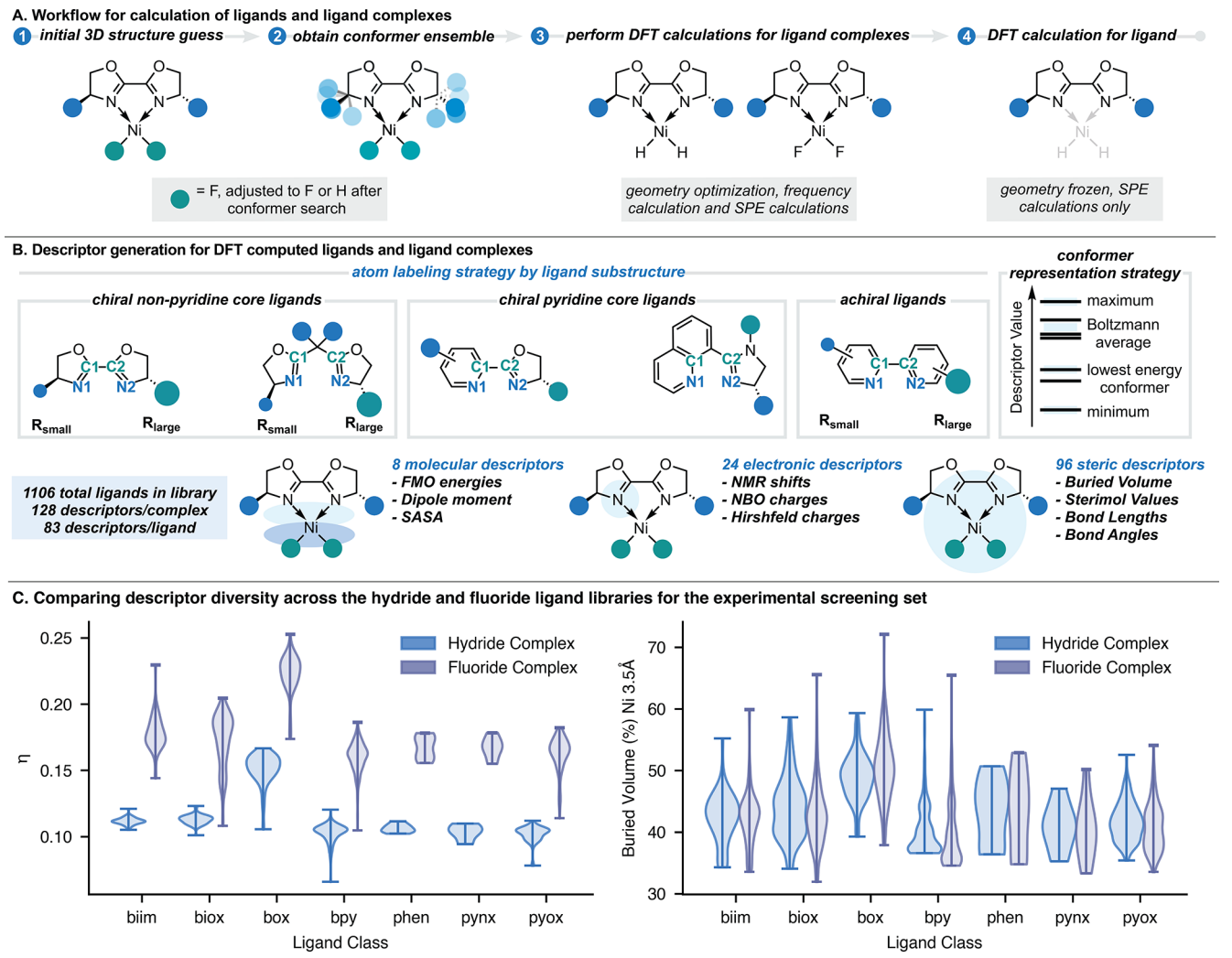


Figure 2. (A) Workflow utilized to construct the ligand library, including two geometries for the Ni complex and the ligand-only structure. (B) Description of the shared structure between ligand classes for descriptor collection; the convention used to attribute C1 and C2 is detailed. (C) Histograms comparing the square planar complex and tetrahedral complex descriptor distributions on two selected descriptors.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX−XXX

Descriptor Collection

Global and atomic steric, electronic, and geometric descriptors were then collected for each of the three structures. To collect comparable atom-based descriptors across libraries, the 5-atom common substructure present in all ligand classes was used, specifically the nickel center, the two binding nitrogen atoms, and their neighboring carbon belonging to the chelate ring (Figure 2B). Since not all ligands are C_2-symmetric, additional considerations were required for consistent labeling of the dative nitrogen and their neighboring carbons (labeled as N1, C1 and N2, C2). For PyOx and PyIm ligand classes, N1 and C1 were assigned to be the pyridine nitrogen and carbon. For all other ligand classes, N1 was assigned as the less sterically hindered side of the ligand. Relative steric hindrance was determined by comparing the buried volume of two hemispheres defined such that the poles of the sphere were set on the N–N axis, and the Ni atom belongs to the equatorial plane. N2 and C2 were attributed to the atoms in the hemisphere with greater buried volume (i.e., containing R_{large}) while N1 and C1 belong to the less hindered side of the ligand (i.e., containing R_{small}). The atom labeling scheme is outlined in Figure 2B.

A broad range of properties was collected, including global molecular descriptors such as volume, solvent-accessible surface area (SASA), and the energy of frontier molecular orbitals. Additionally, atom-based electronic parameters focused on charge or electron distribution, such as NBO charges or Hirshfeld atom dipoles, and bond or atom-based steric descriptors, such as Sterimol[25,26] or buried volume descriptors, were collected. To accurately represent the conformational diversity of these structures, condensed descriptors were utilized to describe the maximum and minimum descriptor values from a ligand conformer ensemble, as well as the descriptor value for the lowest-energy conformer and the Boltzmann-weighted descriptor[27] (respectively denoted with the min, max, low_E, and Boltz suffixes). For the ligand-only descriptors, the lowest-energy conformer and Boltzmann-weighted descriptor were replaced with an averaged descriptor. A total of 128 unique DFT-derived descriptors were collected for each Ni-ligand complex and 83 for ligand-only structures. Further details on this procedure and a complete database of these descriptors can be found in the SI and a complete set of descriptors can be accessed at this link: https://zenodo.org/records/19224101.

Designing a Representative Ligand Screening Set

While several ligand classes are often tested at the initiation of optimization campaigns, once an optimal ligand class is identified, further reaction optimization and modeling efforts tend to focus on this unique ligand class. As such, only a small selection of ligand descriptor space is tested in a reaction, potentially overlooking high-performing ligands and limiting the domain of applicability of generated models to a narrow chemical space. Additionally, traditional ligand sampling curtails a mechanistic understanding of ligand structure-selectivity/reactivity across ligand classes. We envisioned that the shared descriptors used in this library could be employed to optimize ligand space sampling and generate cohesive models across different *N,N*-ligand classes to expand our understanding of ligand effects on reactivity and mechanism.

Due to the high dimensionality of the ligand space, dimensionality reduction was used to visualize the diversity of the library in a two-dimensional space to facilitate the selection of a screening set.[28,29] In this case, Uniform Manifold Approximation and Projection (UMAP)[30] was utilized. Examination of this space

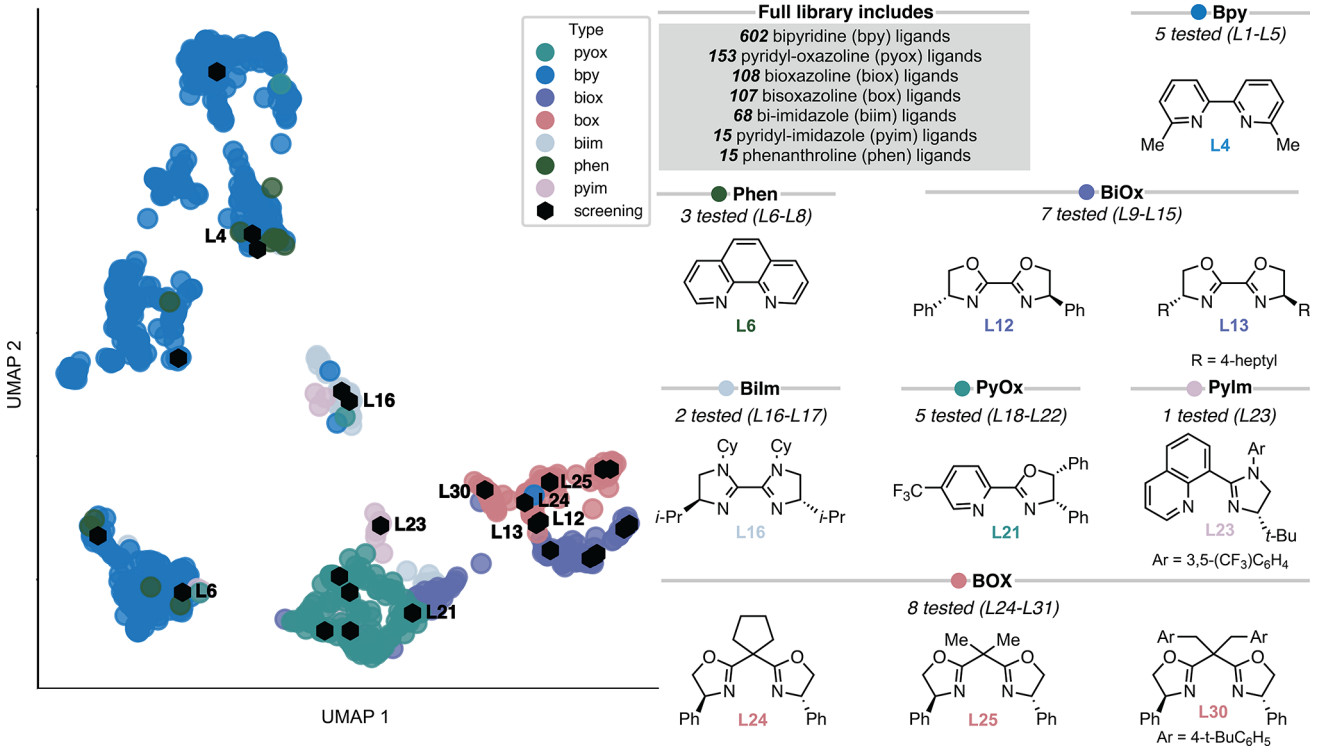


Figure 3. Left panel: Uniform Manifold Approximation and Projection (UMAP) of the 1106 ligands included in the library, computed on 25 descriptors. Right panel: ligand screening set showcasing the seven classes represented and a highlight of specific ligands.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX–XXX

(see Figure 3) suggested that there is little overlap between the different ligand classes in our descriptor space, illustrating its diversity. A set of 31 ligands was selected to act as an experimental screening set. The screening set was selected to cover the chemical space, thereby ensuring descriptor diversity, and each ligand subclass in the library. Chiral ligand classes were overrepresented in the training set due to interest in asymmetric catalysis. To minimize synthetic efforts and cost associated with using this screening set, ligands that were commercially available or had known syntheses were preferentially chosen. The resulting ligand screening set (Figure 3 and Figure S29) includes 31 ligands spanning 7 ligand classes.

Case Study I. To evaluate our ligand selection approach, the screening set was used to evaluate two reported asymmetric Ni-CECs (Figure 4A: reaction **I** and reaction **II**).[17,18] These transformations were chosen because they used the same C(sp^2) electrophile but different C(sp^3) electrophiles and were reported to have the same optimal *N,N*-ligand: 4-heptyl-BiOX (**L13**). Experimental screening shows that the ligand set offered a wide range of outcomes across selectivity and reactivity (Figure 4) consistent with our ligand selection strategy. These yield and enantioselectivity (expressed as $\Delta\Delta G^{\ddagger}$) distributions provide opportunities for applying statistical modeling.

Modeling Reactivity in Case Study I. In a prospective ligand design campaign, most time and resources are dedicated to ligand synthesis. As a first step, we assessed reactivity by evaluating yield as a target for statistical modeling. Yield is a challenging target to model due to a multitude of factors including data sparsity and noise.[31] Simple classification models, which do not require precise predictions, however, have been applied effectively for yield prediction in sparse data regimes.[32] Thus, a single-descriptor classification was utilized to identify regions of ligand space that consistently resulted in lower reactivity. It was important, however, not to remove any possible high-performing ligands and so a conservative approach aimed at maximizing recall, and thus minimizing false negatives, was pursued.

Although the yield response surfaces from reaction **I** and **II** are substantially different (Figure 5C), they could both be modeled with a single-node decision tree classifying yields above or below a 30% threshold as a function of the N1–C1–C2 angle with a recall of 100% (Figure 5B). In practice, the model shows that high-yielding ligands require an N1–C1–C2 angle greater than 100° (arbitrary selection within the 98° to 110° range). Analysis of the class labels displayed in Figure 5B illustrates that this angle acts as a binary classifier: ligands with a 6-member metallacycle (e.g., BOX and quinoline imidazoline **L23**) have N1–C1–C2 angles below 100° and consistently give yields below 30% while other ligand classes (N1–C1–C2 angles > 100°) provide more variability in yield. We note that the model displayed in Figure 5B using the [ligand]NiH_2

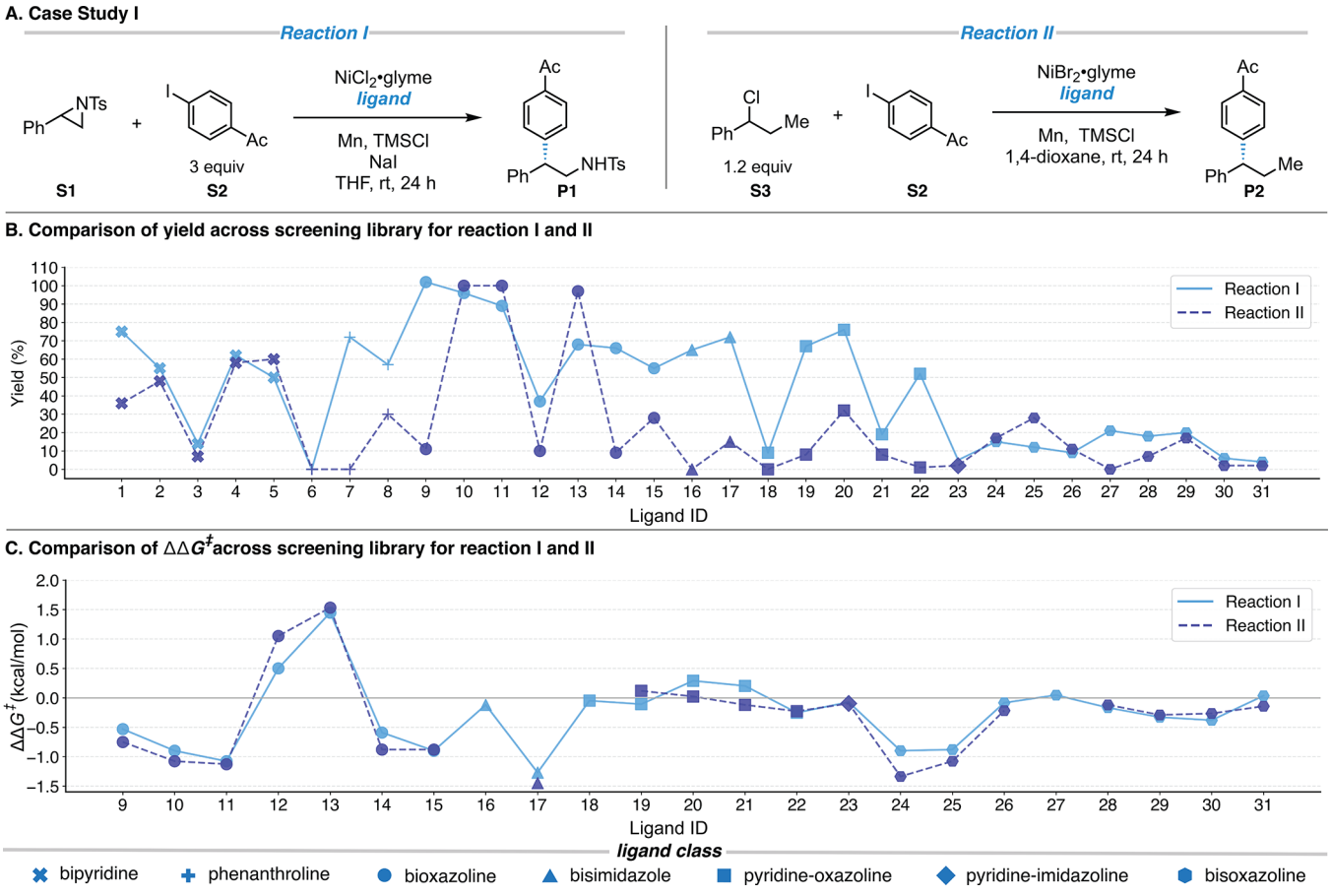


Figure 4. (A) Selected transformations (reaction **I** and **II**) used in case study I. Reaction **I**: $NiCl_2$·glyme (15 mol %) ligand (18 mol %), Mn (3 equiv), NaI (2 equiv), TMSCl (20 mol %), THF (0.1 M), rt, 24 h. Reaction **II**: $NiBr_2$·glyme (11 mol %), ligand (20 mol %), Mn (3 equiv), TMSCl (0.75 equiv), 1,4-dioxane (0.35 M), rt, 18 h. (B) Overlayed yield surface plots for reaction **I** and **II**. (C) Reaction surface plots showing mirrored selectivity ($\Delta\Delta G^{\ddagger}$) trends across ligand identity in reactions **I** and **II**.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX–XXX

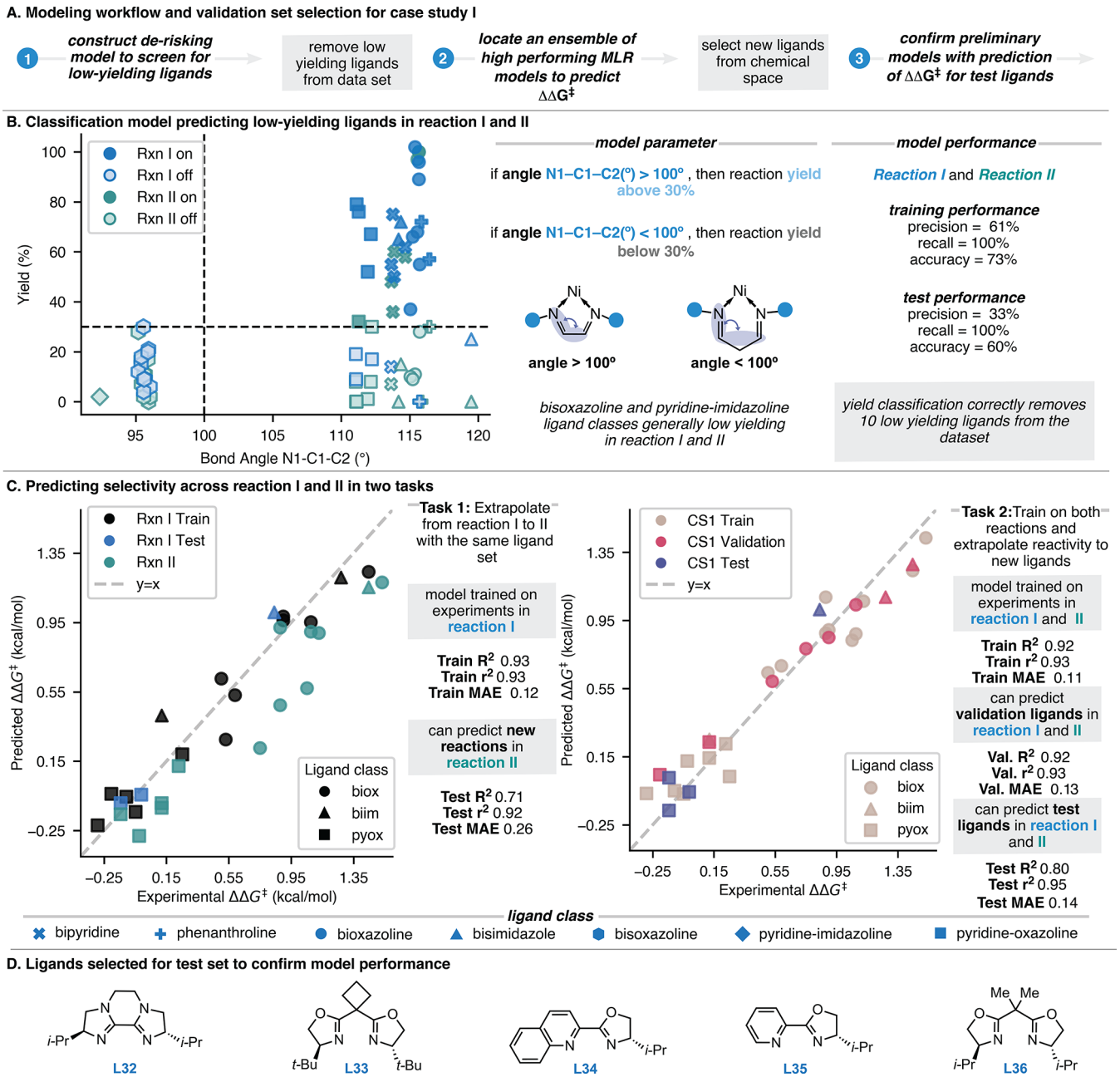


Figure 5. (A) Outline of the modeling workflow applied in case study I to test the suitability of the library for ligand optimization and reaction prediction. (B) Yield classification (below or above 30%) using a single-node decision tree on reaction **I** and **II** data; the N1−C1−C2 angle was extracted from [ligand]NiH_2 structures. (C) Regression models for the prediction of $\Delta\Delta G^{\ddagger}$. On the left (task I), descriptor selection was performed on reaction **I** using a y-equidistant train/validation split and then the model was tested for the prediction of reaction **II**. The regression models displayed are the average of two MLR models with independent variables ($\Delta\Delta G^{\ddagger}$ = 0.44 + 0.43·NMR_C2_min + 0.24·Vbur_Ni_3.5Å_min and $\Delta\Delta G^{\ddagger}$ = 0.44 + 0.52•SASA_max − 0.55·Sterimol_B5_N2–Ni_min). On the right (task II) the descriptor selection was performed on reaction **I** using a y-equidistant train/validation split and then the model was tested on a set of 5 external ligands selected to cover low/medium and high ee responses. For the modeling, the sign of $\Delta\Delta G^{\ddagger}$ measured for the "R" ligands was inverted as descriptors were computed on the "S" ligands; thus, $\Delta\Delta G^{\ddagger}$ for **L12**, **L13** has a different sign than in Figure 4C. (D) Ligands selected (**L32**−**L36**) for external test.

descriptors can also be constructed when the same descriptor is extracted from the [ligand]NiF_2 and ligand-only structures. This type of classification utilizing the descriptor library enables prospective ligand design campaigns to immediately remove non-productive ligand classes without wasting resources.

Modeling Selectivity in Case Study I. After removing the low-yielding ligands identified by the classifier, we investigated the performance of BUNNY molecular descriptors in two enantioselectivity prediction tasks – a common goal in reaction optimization. First, we sought to evaluate whether the selectivity response from reactions using the same C(sp^2) electrophile could identify promising ligands for a new reaction involving a different C(sp^3) electrophile. Second, we evaluated how models trained on the ligand screening set could predict selectivity for

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX−XXX

unseen ligands from the entire ligand search space. In each case, models were interpreted to rationalize reactivity trends across ligand classes.

To test how reaction **I** could inform reaction **II** selectivity (task I), data from reaction **I** was divided into training and validation sets using a y-equidistant split (70 and 30% of data respectively). This split was used to search for high-performing multivariate linear regression (MLR) models using the library descriptors (see Supporting Information section 6.1.2 for details). The search revealed two distinct and high-performing MLR models (MAE of 0.11 and 0.29 kcal/mol, respectively, on the validation set). Since both models performed desirably and offered distinct mechanistic insight, ensemble predictions from both models were utilized going forward in task I and II. The ensemble was then refit to the training and validation set from reaction **I** and utilized to predict outcomes for reaction **II**. The performance on reaction **II** gave a 0.26 kcal/mol MAE: while the model underestimates the selectivity of several ligands in reaction **II**, leading to a higher MAE, it accurately ranks ligand enantioselectivity (correlation coefficient $r^2 = 0.92$). This suggests that transfer learning from previous couplings to optimize new transformations is achievable. The correlation between the enantioselectivity response of reaction **I** and **II** across the ligands screened suggests that the modification of the C(sp^3) electrophile from benzyl chloride to aziridine has a minor impact on the magnitude of selectivity; moreover, the correlation, combined with the ability to model both reactions together in a cohesive linear model, could indicate they involve analogous selectivity-determining steps.

All parameters for the best models arise from the [ligand]NiH_2 descriptors. Analysis of the ensemble parameters suggests selectivity is primarily dominated by steric effects. Two descriptors utilized in the ensemble, percent buried volume (3.5 Å, from Ni) and solvent-accessible surface area (SASA), suggest that an increase in steric bulk of the complex is associated with higher $\Delta\Delta G^{\ddagger}$ and thus better enantioselectivity. A decrease in the magnitude of the third steric term, the B5 Sterimol for the N2–Ni bond, however, is associated with an increase in $\Delta\Delta G^{\ddagger}$. While this appears in conflict with the steric trend of the first two terms, this term may serve to identify the impact of more localized steric bulk on selectivity. The fourth term utilized in the ensemble is the NMR shielding for the C2 carbon. This term serves as a classifier distinguishing PyOx ligands, which typically give poorer selectivity and lower $\Delta\Delta G^{\ddagger}$ values, from all other ligand classes (see Figure S31).

To test how the descriptor library could be used to predict the enantioselectivity of unseen ligands, the same train/validation split strategy was applied. Unlike in task I, models were trained using data from reaction **I** and **II** (70% of ligands), and validated using all data, from reaction **I** and **II**, for the unseen ligands in the screening set (30% of ligands). The same ensemble process described for task I was applied here; however, since training and validation data now incorporated data from both reactions, an additional, one-hot encoding parameter was added to distinguish between reaction **I** and **II**. The models were then refit to the combined training data from reaction **I** and **II** and used to predict the validation and test sets. The average model resulted in an MAE of 0.13 kcal/mol on the validation set. To confirm the possibility of using these models for ligand design, we selected an external test set from the virtual library. The external test set included five ligands spanning three *N*,*N*-bidentate classes (**L32**−**36**, Figure 5D) and covering a wide range of predicted ee (from 1 to 58%). This choice of ligands tested the model's ability to predict across ligand classes and identify both low and highly enantioselective ligands. We first applied the yield classification step described above, which predicted **L33** and **L36** as being low performing. The recall statistics for the test set remained at 100%. The remaining test set was carried forward and gave a desirable predictive performance with an MAE of 0.14 kcal/mol, demonstrating the model's ability to extrapolate. The two tasks within this case study demonstrated the ability to use this library to predict selectivity in unseen reactions and to predict unseen ligands in multiple reactions. Ultimately, this case study highlights the possibility of using the ligand library to guide optimization of new enantioselective Ni-catalyzed CEC reactions of similar electrophiles.

Case Study II. Case study I suggested that with aryl iodide electrophiles, the enantioselectivity trend across ligands was conserved using different C(sp^3) electrophile partners (respectively benzyl chlorides and aziridines). We thus investigated whether this remains true using a different C(sp^2) electrophile. As the Reisman group had previously reported the cross-coupling of benzylic chlorides with acyl chlorides (reaction **IV**),[19] we focused on the development of an analogous stereoconvergent cross-coupling between aziridines **S4** and acyl chlorides **S5** (reaction **III**). This reaction afforded **P3,** an enantioenriched β-aminoketone, which is a relevant drug-like motif.[20] Under slightly modified reaction conditions to those initially reported in reaction **III**, we found that product **P3** could be accessed in low yields but with good enantioselectivity (up to 24% yield and 83% ee across the screening library, see Table S7 for details).

The enantioselectivity measured across all chiral ligands tested (**L9**−**L31**) for the new acyl-aziridine coupling reaction **III** strongly correlated with reaction **IV** (Figure 6B), validating the hypothesis that optimal ligands are conserved across different types of C(sp^3) electrophiles. Interestingly, there was little correlation between the enantioselectivity of CECs using the same C(sp^3) but different C(sp^2) electrophiles; BiOX ligands are the best-performing scaffolds for reactions **I** and **II** (aryl iodide), whereas BOX ligands are optimal for reactions **III** and **IV** (acyl chloride). More specifically, the three highest performing ligands for both reaction **III** and **IV** were all BOX ligands: **L24** (**III**: −83% ee, **IV**: −90% ee), **L25** (**III**: −82% ee, **IV**: −91% ee), and **L30** (**III**: −82% ee, **IV**: −85% ee). As observed in case study I, the yield trends were not conserved upon changing C(sp^3) electrophile, and the most selective ligands provided the β-aminoketone **P3** in low yield; for example, **L24** provided **P3** in 83% ee but 15% yield, thus requiring optimization.

Bayesian Optimization (EDBO+) was leveraged to optimize the reaction yield.[33] The optimization search space was designed to include the three most enantioselective ligands (**L24**, **L25**, and **L30** outperformed all other ligands by at least 30% ee), common cross-electrophile coupling solvents, additives, precatalysts, various concentrations, and ligand to metal ratios (Figure 6C, and Tables S9 and S10).[34] The non-numerical parameters of the search space were one-hot encoded and temperature and concentrations were numerically encoded. The first four rounds of EDBO+ optimization led to a modest improvement in yield at the expense of selectivity (31% yield for 43% ee) but enabled us to reduce the search space: Zn and $NiCl_2$·glyme were identified as the optimal reductant and precatalyst, respectively. Moreover, MeCN emerged as a solvent of interest for increasing reactivity of the system.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX−XXX

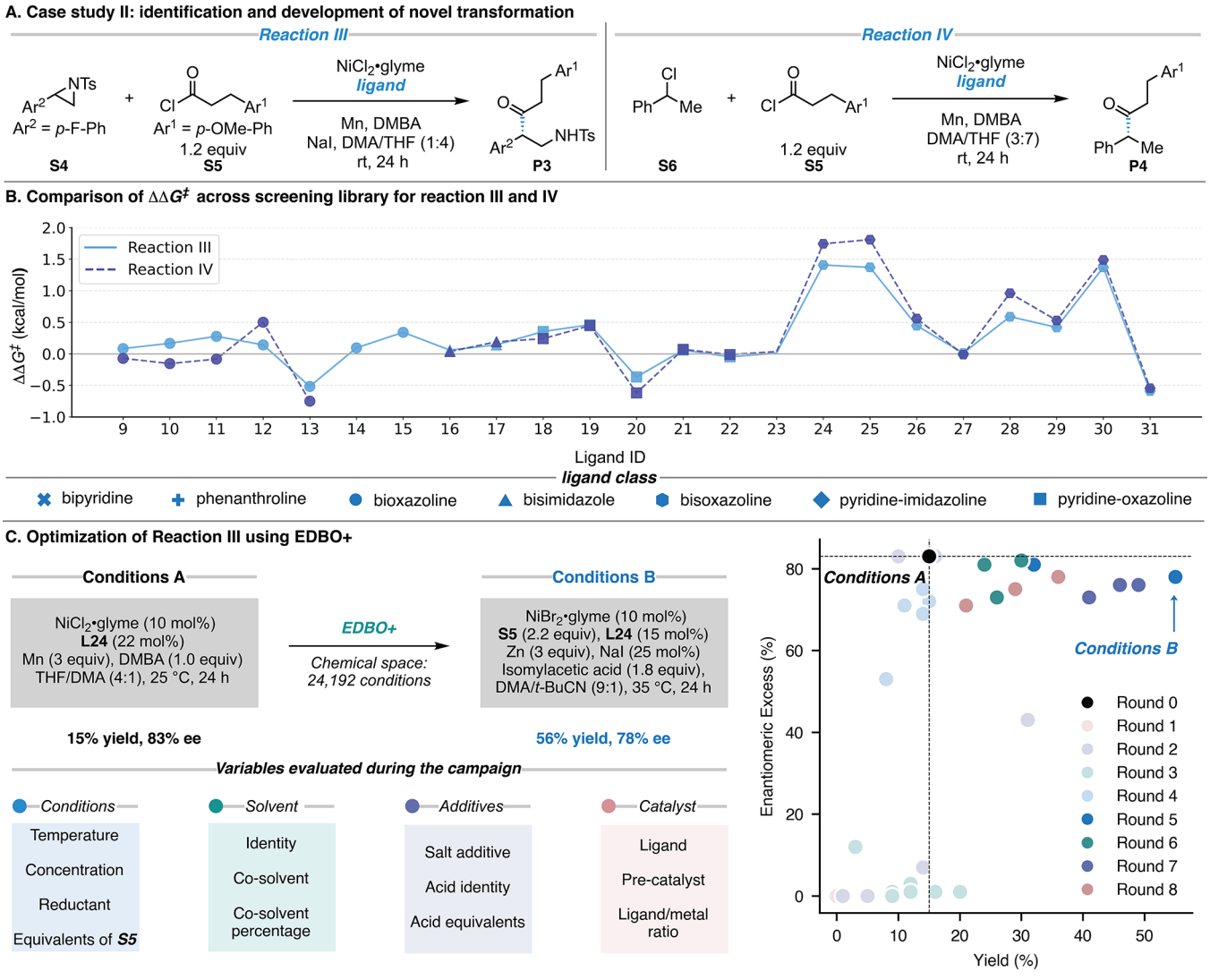


Figure 6. (A) Reaction conditions for reaction **III**: $NiCl_2$·glyme (10 mol %), ligand (22 mol %), Mn (3 equiv), 2,5-dimethylbenzoic acid (DMBA) (1 equiv), NaI (25 mol %), THF/DMA (4:1), 0.15 M, rt, 24 h, and reaction **IV**: $NiCl_2$·glyme (10 mol %), ligand (22 mol %), Mn (3 equiv), DMBA (0.75 equiv), THF/DMA (7:3), 0.37 M, rt, 24 h. (B) Reaction surface plots showing similar selectivity ($\Delta\Delta G^{\ddagger}$) trends in reaction **III** and **IV** across the screening set of asymmetric ligands (**L9**–**L31**). (C) Bayesian optimization campaign for reaction **III**. Initial conditions and optimized conditions after campaign (top left); chemical space surveyed during campaign (bottom left); enantiomeric excess vs yield plot (right) indicated by circle as indicated by color bar. Original conditions (A) correspond to round 0 (see Supporting Information Tables S9–S12 for the full list of experiments).

Using these optimal parameters and the promising results from MeCN screening, we expanded our condition exploration to solvent mixtures and temperature, as well as the identity and concentration of an acid additive (see Tables S11 and S12). The decision to incorporate the latter variable was determined after observing that trace acid was beneficial to the reaction performance, a behavior also observed in the original report for reaction **IV**.[19] After five subsequent EDBO+ rounds, enhanced conditions for the reaction were identified. These conditions represented a considerable improvement in yield (56%) while maintaining good selectivity (78% ee). This showcases the advantages of the ligand library descriptor strategy to identify competent chiral ligands and the use of Bayesian optimization to reach favorable results efficiently (31 ligands screened for one condition, then 41 reactions for condition optimization). Indeed, while eight rounds of EDBO+ were performed, promising conditions were identified during the fifth round of experiments, highlighting the efficiency of this strategy. This case study demonstrates that knowledge of a CEC reaction's response surface can expedite the development of a similar transformation if they share a $C(sp^2)$ electrophile.

Modeling Selectivity in Case Study II. Motivated by the successful identification of classification and regression models for case study I, we sought to use similar techniques to understand and explain the enantioselectivity trends observed with the ligand screening set in reactions **III** and **IV**. While in case study I, we took the practical step of removing low-yielding reactions before modeling enantioselectivity, no accurate yield classification models for reactions **III** and **IV** were found. Unlike case study I, limited success was found in building linear models for selectivity prediction. To understand why this may be the case, we evaluated univariate correlations within each ligand

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX–XXX

class. The best univariate correlations for each of the BiOx, BOX, and PyOX/PyIm ligand classes feature different types of descriptors (Figure 7A). Correlations were not constructed for BiIm ligands due to a lack of data (3 ligands). BiOX ligand selectivity was best correlated with a global descriptor (LUMO), while BOX and PyOX ligands had better correlations with a steric descriptor (N1–Ni B5 Buried Sterimol). Although the correlations are modest, these differences led us to hypothesize that the enantio-determining step might not be the same across the different ligand classes, which might explain the lack of success in building a singular linear model. To account for a scenario in which the descriptors controlling enantioselectivity depend on the ligand class, we turned to nonlinear models to predict $\Delta\Delta G^{\ddagger}$ values.

Several nonlinear models were tested, and it was found that kernel-ridge regression models (KRR) performed best. Descriptors were down-selected using recursive feature elimination (RFE) with a ridge estimator. Further details of this procedure can be found in the Supporting Information section 6.1.4. The y-equidistant algorithm was applied to divide data from reaction **III** and **IV** into a 70/30 train-validation ligand split. Since enantioselectivity could be measured for each ligand in

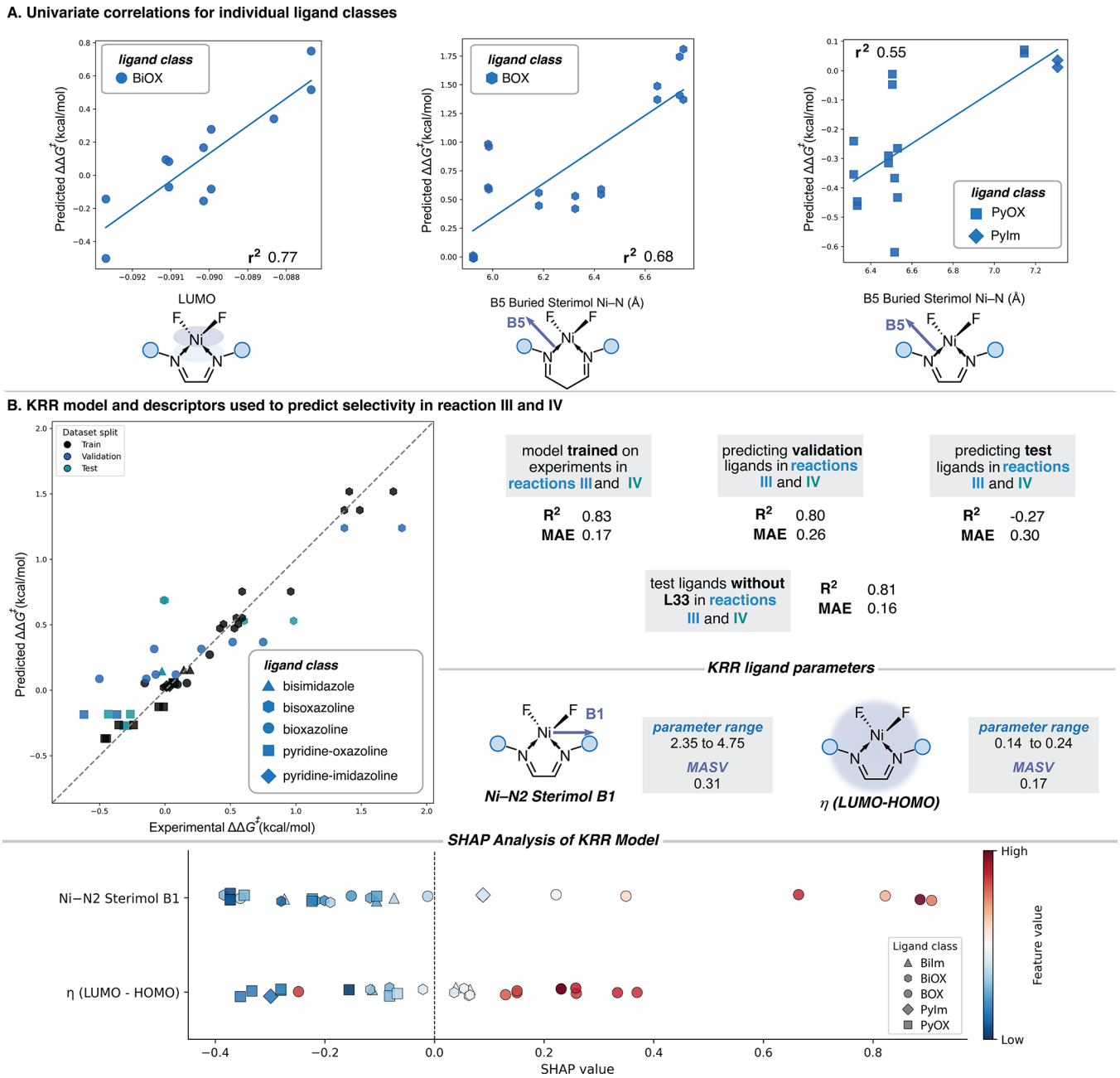


Figure 7. Modeling efforts for reactions **III** and **IV**. (A) Univariate correlations for individual ligand classes using data from reactions **III** and **IV**. (B) Left panel: parity plot for the optimal kernel-ridge regressor (KRR) model; a first descriptor selection was performed using a y-equidistant 70/30 train/validation split on the screening ligand set (**L9–L31**); the best model was tested on **L32–L36**. Right panel: analysis of the KRR model used for prediction. Performance metrics on train/validation and external test sets. Ligand library descriptors selected and SHapley Additive exPlanation (SHAP) analysis, mean absolute SHAP values are denoted as MASV.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX–XXX

reaction **III** but not in reaction **IV** (due to poor yielding ligands), a ligand split was performed using the y-equidistant algorithm on reaction **III** data. This ligand split was used to generate a 70/30 train-validation split that also incorporated reaction **IV** data. The best KRR relies on two descriptors from the [ligand]NiF_2 structures (Figure 7). The selected model gave an R^2 of 0.80 and an MAE of 0.26 kcal/mol for the validation set. When the model was tested on the external ligand set (**L32**−**L36**), an MAE of 0.30 kcal/mol and an R^2 of −0.27 were obtained, showing a significant decrease compared to the validation performance. The reduction in statistics is likely due to a single ligand, *t*-Bu-BOX (**L33**), for which the selectivity is overestimated (∼50% ee predicted). The low selectivity observed (−1 and 0% ee for reaction **III** and **IV**, respectively) likely arises from poor coordination of **L33** to the Ni center due to steric hindrance, consistent with observations in prior reports.[35] This coordination issue is most likely not encountered in the training set and is thus out of the domain of applicability of the model. When the test ligand set is evaluated without the inclusion of **L33**, metrics improve significantly, with an R^2 of 0.81 and MAE of 0.16, comparable to the train and validation set.

The model contains a global electronic descriptor: η, which is calculated by subtracting the HOMO energy from the LUMO energy and is thought to describe the hardness of the complex. The other descriptor is a steric term, the B1 Sterimol for the Ni–N2 bond, quantifying the bulk on the oxazoline or imidazoline moieties across ligand classes. Unlike linear models, understanding and interpreting nonlinear models is less straightforward. Shapley additive explanations (SHAP) analysis[36] was used to determine how ligand descriptors were used by the model to predict selectivity (Figure 7). Higher mean absolute SHAP values (MASV) indicate a heavier reliance on the descriptor in model decision making. The Sterimol term was found to be the most important descriptor (MASV of 0.31). The η parameter carries less weight at a MASV of 0.17. Plotting the SHAP values for each data point (Figure 7), a higher η value (greater LUMO-HOMO gap) is associated with higher $\Delta\Delta G^{\ddagger}$ values and thus greater selectivity. The distribution of descriptor vs. SHAP values also indicates that this descriptor acts as a classifier, likely to distinguish BOX ligands, which typically give good selectivity, from all other ligand classes (see Figure 2C). The analysis for the B1 Sterimol term is directly correlated with $\Delta\Delta G^{\ddagger}$ such that higher descriptor values give higher selectivity. This indicates that increasing ligand steric hindrance enhances selectivity. The points which have especially high SHAP values for the B1 Sterimol term represent the two reactions performed with **L30**, an especially bulky and flexible BOX ligand (Figure 3). In summary, the Sterimol term indicates that bulkier ligand complexes lead to improved selectivity. The η term suggests that more stable complexes lead to greater selectivity in these reactions. This case study demonstrates the utility of nonlinear models in a low-data regime as an alternative to MLR. The need for these nonlinear models led to the hypothesis that the selectivity-determining step may not be conserved across the ligands surveyed.

DFT Analysis of Proposed Reaction Mechanisms

Interestingly, for both models in case study II, all ligand descriptors are derived from the [ligand]NiF_2 library. The origin of these descriptors for the C(sp^2) acyl-couplings contrasts with the linear models obtained for the aryl-couplings, in which the best models exclusively used the [ligand]NiH_2 library. In low-data regimes, the use of DFT descriptors derived from structures that more closely resemble reactive intermediates tends to enable more accurate modeling.[11,37,38] Thus, we investigated whether descriptor selection from the tetrahedral or the square planar complexes could reflect a change in the enantio-determining step occurring when the C(sp^2) partner of the cross-electrophile is modified.

The resting state of Ni-CEC couplings involving aryl halides as the C(sp^2) electrophile (case study I) has been characterized as a [ligand]Ni(Ar)X species.[39] The subsequent steps, radical capture and reductive elimination, have both been proposed to be enantio-determining depending on the nature of the ligand used. Specifically, Reisman and co-workers suggested an enantio-determining radical capture for a cyclopropyl-IndaBOX ligand (**L31**) with the sign of $\Delta\Delta G^{\ddagger}$ being conserved for the reductive elimination.[39] More recently, Shu and co-workers demonstrated that in the case of a BiIm ligand, the $\Delta\Delta G^{\ddagger}$ computed for the radical capture matched the experimental enantioselectivity.[40] Nonetheless, the singlet square planar Ni(II) intermediate invoked in the Shu study is inconsistent with previous calculations establishing that the triplet tetrahedral Ni(II) was more stable for a similar BiIm ligand.[41] In the case of BiOX ligands, Kozlowski and co-workers also proposed that radical capture occurred at a tetrahedral [ligand]Ni(Ar)X complex when a benzylic radical was used.[42] Altogether, prior work suggests that depending on the ligand class, Ni(II) intermediates can be square planar or tetrahedral. In case study II, this could suggest that (a) the nonlinear modeling arises from a change in resting state structure or in the selectivity-determining TS across ligands and/or (b) the change of the C(sp^2) electrophile from aryl iodide to acyl chloride affects the structure of a Ni(II) aryl or acyl intermediate. A change in intermediate structure and subsequent enantio-determining steps could explain why case studies I and II are best modeled by [ligand]NiH_2 or [ligand]NiF_2 descriptors, respectively. To investigate these two hypotheses, two ligands were selected: **L12** (BiOX) and **L24** (BOX). This choice was motivated by the observation that while C(sp^2) aryl-couplings using these two ligands resulted in comparable ee's (−71 and −79% respectively for reaction **II** when corrected to the *S*,*S* enantiomeric series), the C(sp^2) acyl-couplings with the same ligands resulted in opposite ee's (+38 and −90% respectively for reaction **IV** when corrected to the *S*,*S* enantiomeric series). **L12** showing unexpected selectivity for its enantiomeric series.

To understand whether the origin of the best descriptors for case study I and II was due to a change in spin states of the Ni(II) intermediates, the relative energies of the Ni(II) square planar singlet and tetrahedral triplet complexes for the combinations of ligands and C(sp^2) electrophiles were computed. In all four cases, the singlet square planar structures were found to be more stable than their tetrahedral triplet counterpart (see Supporting Information section 3.5.1), which contrasts with the hypothesis that descriptor selection originates from the spin state of the relevant Ni(II) intermediate (**I**).

Given the spin state of the intermediate did not seem to be responsible for the observed discrepancies, a change of enantio-determining step between case study I and case study II was considered. The DFT study was limited to the benzyl chloride C(sp^3) electrophile (reaction **II** and **IV**), as similar enantioselectivity trends were observed for the aziridine throughout the screening library. Diao and co-workers recently proposed a concerted radical capture/reductive elimination mechanism in the case of pyBOX tridentate ligands.[49] This

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX−XXX

mechanism was computed for the less-studied acyl coupling using **L12** and **L24**. The DFT results demonstrate that for these ligands, a concerted mechanism is unlikely to occur (see Supporting Information section 3.5.3), which is evidence against this specific change of mechanism leading to different descriptors being used in the modeling endeavors.

At this stage, the energy surfaces for the reaction pathways involving radical capture and reductive elimination in the acyl and aryl-couplings were computed using **L12** and **L24** (Figure 8). Interestingly, conformational sampling of the Ni(III) intermediates showed that the radical capture intermediate, a square planar pyramid intermediate (**III-a** see Figure S1), can easily reorganize to give a more stable pseudo-octahedral structure (**III-b**). The reductive elimination transition states (TSs) were computed from the **III-a** and **III-b** structures; a comparison of the TSs energies shows that reductive elimination from **III-b** is favored in all cases (SI Supporting Information sections 3.4 and 3.5), and therefore, a comparison of the energy surfaces was performed using the corresponding TSs (as depicted in Figure 8). The energy surfaces of the aryl coupling show that the formation of the Ni(III) intermediate is slightly uphill in energy (+2.3 to +4.4 kcal/mol for the minimum energy found of **L12** and **L24**, respectively). This corroborates results from previous studies on the use of BOX ligands.[49] The computations suggest that the differences between the energies of radical capture reductive elimination TSs are small for the BiOX ligand **L12** (–0.1 kcal/mol blue surface Figure 8) and more significant for the BOX ligand **L24** (+3.1 kcal/mol green surface Figure 8). The comparisons of the surfaces leading to the enantiomeric products (detailed in Supporting Information section 3.4) suggest the enantioselectivity is controlled by the radical capture for the BiOX ligand **L12** ($\Delta G^{\ddagger}$ = 7.7 kcal/mol and $\Delta\Delta G^{\ddagger}$ 1.4 kcal/mol see Figure S3), whereas for the BOX ligand, the highest energy barriers are observed for reductive elimination ($\Delta G^{\ddagger}$ = 11.9 kcal/mol and $\Delta\Delta G^{\ddagger}$ 1.1 kcal/mol; see Figure S7). In principle, the rate-determining states, the TOF-determining intermediate (TDI), and the TOF-determining transition state (TDTS), calculated according to the energetic span model,[50] involve the reductive elimination TS for the BOX and the radical capture for the BiOX ligands. For both ligands, the TDI is the Ni(II) square planar structure.

For the acyl coupling, we found that the **III-b** intermediates are lower in energy than the Ni(II) structure (–2.3 and –2.0 kcal/mol for the minimum energy of **L12** and **L24** respectively). According to the energy span model, **III-b** becomes the TDI

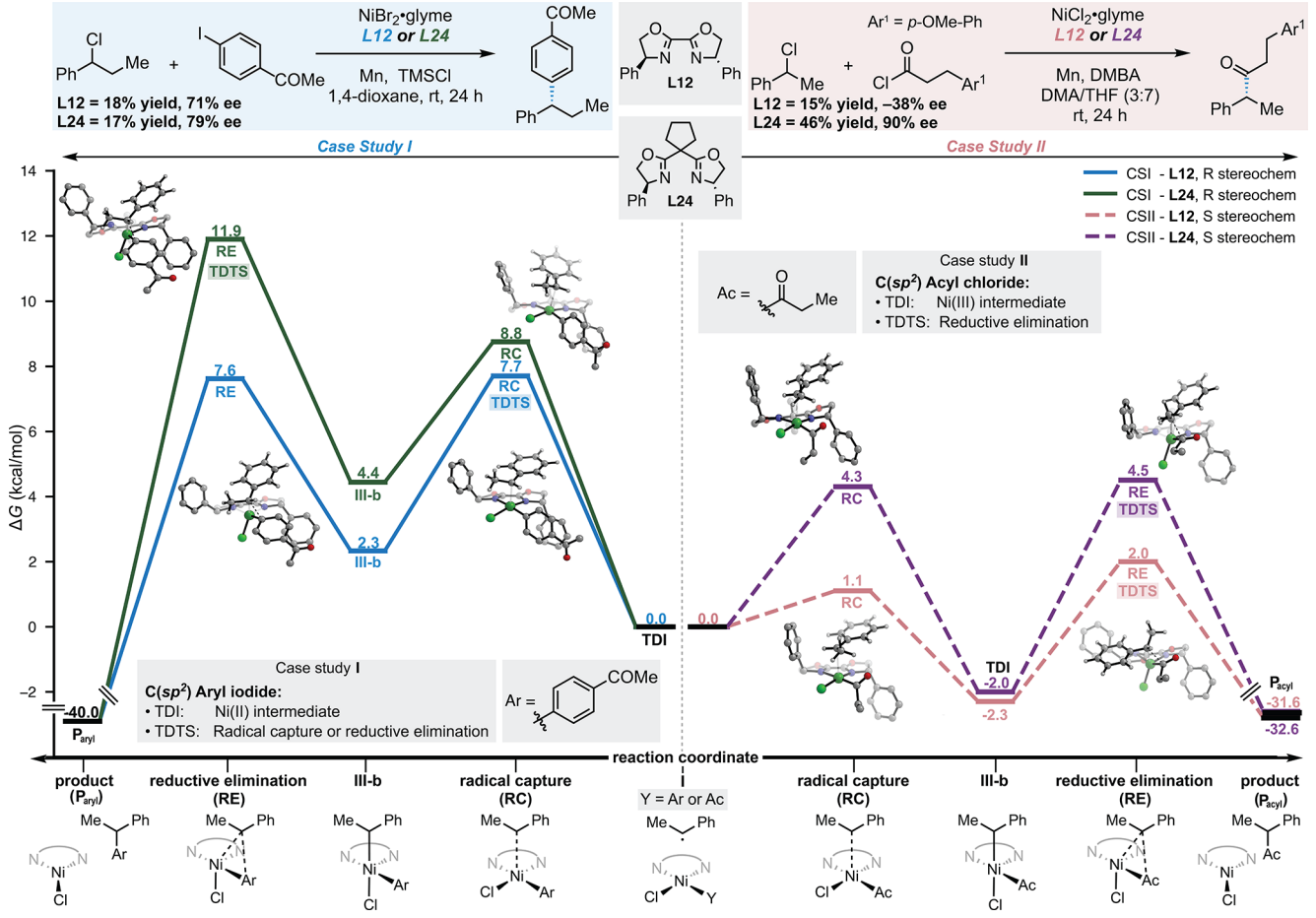


Figure 8. The surfaces represented for each ligand and case study correspond to the most energetically favorable pathway. For each ligand and case study, the surfaces rising from the four diastereo-approaches of the radical on the Ni(II) complexes can be found in supporting information. For each surface, the energy reference was set to the energy of the corresponding species **I** in which the ligand can be **L12** or **L24** and Y depends on the case study. Calculations were performed at the PBE0-D3/cc-pVDZ(H,C,N,O,Cl)/cc-pVTZ(Ni)//M06/cc-pVTZ(Ni) level of theory.[43−47] All structures reported (TS and intermediates) are the minimum energy of a conformer ensemble obtained with CREST[48] and refined by DFT. The turnover frequency (TOF) determining intermediate (TDI) and the TOF-determining transition state (TDTS) are highlighted. We note that the % ee of **L12** depicted in this figure was corrected for the *S,S* enantiomeric series used for the modeling, while the *R,R* ligand was used experimentally.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX–XXX

and the reductive elimination becomes the TDTS for the most favorable surfaces of both **L12** and **L24**. These results support a change in the enantio-determining pathway between the aryl and acyl-couplings. The change of the TDI from a square planar Ni(II) to a pseudo-octahedral Ni(III) intermediate corresponds to the shift in descriptors used by the model (from the square planar [ligand]NiH_2 to the tetrahedral [ligand]NiF_2 descriptors, respectively). Although more in-depth studies are required to understand how and when the descriptor analysis obtained from modeling campaigns can inform changes in mechanism, this example suggests that it could help generate mechanistic hypotheses. We anticipate that the increasing accessibility of accurate electronic structure calculations (i.e., machine learning potentials[51]) will facilitate these analyses.[52]

Overall, the calculations on high-performing ligands (**L12** BiOX and **L24** BOX) and the enantioselectivity trends measured across nitrogen bidentate ligands suggest C(sp^2) electrophiles exert the most control on the mechanism of the reaction. Additionally, two independent studies of Ni-CEC using a third C(sp^2) electrophile (vinyl bromide) with benzyl chloride[39] and aziridine,[53] respectively, report the same optimal ligand: the cyclopropyl-IndaBOX (**L31**). These observations define an interesting paradigm for the development of enantioselective cross-electrophile couplings: optimal ligands reported for established Ni-CECs are likely to be selective for an extended scope of C(sp^3) electrophiles. Provided that reaction yield optimization is successful, this reduces the need for *de novo* ligand design. We hope that this information will help guide the development of cross-electrophile coupling reactions for which precedents using the same C(sp^2) electrophile have been reported.

CONCLUSIONS

In conclusion, we have developed a publicly available library of bidentate *N*,*N*-ligands including DFT level description of the ligand-only, [ligand]NiH_2, and [ligand]NiF_2 structures. The utility of the library was demonstrated in two cases studies. The first case study featured two Ni-CECs sharing the same C(sp^2) electrophile (aryl iodide) with two different C(sp^3) electrophiles (benzyl chloride and aziridine) and highlighted how the library could help identify thresholds and MLR relationships across ligand classes to predict the selectivity and reactivity of unseen ligands. We also showed how the library could facilitate transfer learning between two reactions with different C(sp^3) electrophiles. The second case study was designed to prospectively verify these claims. Starting from an established Ni-CEC between acyl chloride and benzyl chloride, we used the library to screen for successful ligands and accelerate the early-stage development of a new enantioselective acyl coupling using an aziridine as C(sp^3) electrophile. In contrast to the aryl-couplings, MLR modeling of the acyl coupling was unsuccessful; instead, we found that kernel-ridge regressor was effective at predicting new ligand selectivity accurately. The aryl coupling was effectively modeled by descriptors computed from square planar structures while the acyl coupling relied on tetrahedral structures. To investigate whether the change in models and descriptors could be indicative of a change in reaction mechanism, a DFT study of two ligands (BOX and BiOX) for the acyl and aryl coupling with a common benzylic chloride was undertaken. This study revealed a change in TOF-determining intermediates for the different C(sp^2) electrophiles. Interestingly, the reaction with a square planar TDI was best modeled by the square planar descriptors and the one with the pseudo-octahedral TDI was best modeled by tetrahedral descriptors. From a broader perspective, this study suggests that the design and modeling of experimental datasets using high-level descriptors can help identify changes in mechanistic regime and generate mechanistic hypotheses.

METHODS

Materials and General Procedure for Reaction I

Materials. Aziridine **S1** was synthesized according to literature precedent and was recrystallized from EtOAc/hexanes before use.[53] Aryl iodide **S2** was purchased from Oakwood Chemical and used as received. $NiCl_2$·glyme was purchased from Strem Chemicals and used as received. Sodium iodide, manganese (powder, ~325 mesh, >99% trace metal basis), and chlorotrimethylsilane (purified by redistillation, >99%, Sure/Seal™) were purchased from Sigma-Aldrich and used as received. Tetrahydrofuran (THF) was distilled over Na/benzophenone, degassed by sparging with nitrogen, and stored over 4 Å sieves in a nitrogen-filled glovebox.

Procedure. The following screening procedure is modified from the originally reported cross-coupling.[18] All reaction preparation was done inside a nitrogen-filled glovebox. Vials were oven-dried (140 °C) and were allowed to cool under vacuum prior to use in the glovebox. Reactions were run on a 0.05 mmol scale relative to **S1**. *Stock Solution 1* was prepared in a 1 or 2 mL volumetric flask such that the final concentration contained **S1** (13.67 mg, 0.05 mmol, 1 equiv) and 4-iodoacetophenone **S2** (36.90 mg, 0.15 mmol, 3 equiv) in 225 μL of THF. The following stock solutions were prepared in 1-dram vials: *Stock Solution 2:* TMSCl (1.08 mg, 10.0 μmol, 20 mol %) and 50 μL of THF. *Stock Solution 3:* $NiCl_2$·glyme (1.65 mg, 7.5 μmol, 15 mol %), ligand (9.0 μmol, 18 mol %), and 225 μL of THF. This vial was equipped with a Teflon stir bar, capped with a septum cap, and allowed to stir for 30 min.

To a 1-dram vial equipped with a Teflon stir bar was added NaI (14.98 mg, 0.10 mmol, 2 equiv) and Mn (8.24 mg, 0.15 mmol, 3 equiv). Stock solution 1 (225 μL), stock solution 2 (50 μL), and stock solution 3 (225 μL) were then sequentially added. The reaction vial was capped with a septum cap, wrapped with electrical tape, and brought outside of the glovebox. The reaction was set to stir (700 rpm) at room temperature for 24 h. Upon completion, a trimethoxy benzene stock solution (84.1 mg in EtOAc) was prepared, and an aliquot (100 μL, 1 equiv) was added as an internal standard. Yields were determined by GC-FID of the crude reaction mixture. After purifying the product by preparative thin-layer chromatography (50% EtOAc/hexanes), % ee was determined by SFC.

Materials and General Procedure for Reaction II

Materials. 4-iodoacetophenone (**S2**) and 1-chloropropylbenzene (**S3**) were purchased from Combi-Blocks. $NiBr_2$(diglyme), manganese (powder, ~325 mesh, ≥99% trace metal basis), and chlorotrimethylsilane (≥99%, Sure/Seal™) were purchased from Sigma-Aldrich and used as received. 1,4-dioxane was dried by passing through activated alumina columns and stored in a nitrogen-filled glovebox.

Procedure. The following screening procedure is modified from the originally reported cross-coupling.[17] All reactions were set up inside a nitrogen-filled glovebox. Vials were oven-dried (140 °C) and were allowed to cool under vacuum prior to use in the glovebox. Reactions were run on a 0.1 mmol scale.

The following stock solutions were prepared in a 1 or 2 mL volumetric flask: *Stock Solution 1:* $NiBr_2$·diglyme (3.53 mg, 10 μmol, 10 mol %), 4-iodoacetophenone **S2** (24.6 mg, 0.1 mmol,1 equiv), and 228 μL of 1,4-dioxane. *Stock Solution 2:* TMSCl (8.15 mg, 0.075 mmol, 0.75 equiv), 1-chloropropylbenzene **S3** (18.6 mg, 0.12 mmol, 1.2 equiv), and 50 μL of 1,4-dioxane.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX–XXX

The reaction vial was equipped with a Teflon stir bar cut with cut edges, Mn (16.5 mg, 0.3 mmol, 3 equiv), and ligand (x mg, 20 mol %). Stock solution 1 (228 μL) was added, and the mixture was stirred for 10 min at 300 rpm. Then, stock solution 2 (50 μL) was added. The reaction vial was capped with a septum cap, wrapped with electrical tape, and brought outside the glovebox. The reaction was stirred (420 rpm) at room temperature for 18 h. Upon completion, dodecane was added as an external standard. Yields were determined by GC-FID of the crude reaction mixture, and % ee was determined by SFC-MS.

Materials and General Procedure for Reaction III

Materials. Aziridine **S4** was synthesized according to literature precedent and was recrystallized from EtOAc/hexanes before use.[54] Acid chloride **S5** was purchased from Ambeed and distilled before use. $NiCl_2$·glyme was purchased from Strem Chemicals and used as received. Sodium iodide and manganese (powder, ~325 mesh, >99% trace metal basis) were purchased from Sigma-Aldrich and used as received. DMBA was purchased from Oakwood Chemical and used as received. THF was distilled over Na/benzophenone, degassed by sparging with nitrogen, and stored over 4 Å sieves in a nitrogen-filled glovebox. Dimethylacetamide was purified by passing through an activated alumina column under an argon atmosphere, degassed by sparging with nitrogen, and stored over 4 Å sieves.

Procedure. All reactions were set up inside a nitrogen-filled glovebox. Vials were oven-dried (140 °C) and were allowed to cool under vacuum prior to use in the glovebox. Reactions were run on a 0.05 mmol scale relative to **S4**.*Stock Solution 1:* The stock solution was prepared in a 1 mL volumetric flask such that the final concentration contained **S4** (14.57 mg, 0.05 mmol, 1 equiv) and **S5** (11.92 mg, 0.06 mmol, 1.2 equiv) in 100 μL of THF/DMA (4:1). The following stock solutions were prepared in 1-dram vials: *Stock Solution 2:* NaI (1.87 mg, 12.5 μmol, 25 mol %), 2,6-dimethylbenzoic acid (7.51 mg, 0.05 mmol, 1 equiv), and 116.7 μL of THF/DMA (4:1) and *Stock Solution 3:* $NiCl_2$·glyme (1.10 mg, 5.0 μmol, 10 mol %), ligand (11.0 μmol, 22 mol %), and 116.7 μL of THF/DMA (4:1). This vial was equipped with a Teflon stir bar, capped with a septum cap, and allowed to stir for 30 min.

To a 1/2-dram vial equipped with a Teflon stir bar was added Mn (8.24 mg, 0.15 mmol, 3 equiv). Stock solution 2 (116.7 μL), stock solution 1 (100.0 μL), and stock solution 3 (116.7 μL) were then sequentially added. The reaction vial was capped with a septum cap, wrapped with electrical tape, and brought outside of the glovebox. The reaction was set to stir (700 rpm) at room temperature for 24 h. Upon completion, a stock solution of 1-fluoronapthalene (146 mg in 2 mL of EtOAc) was prepared in a volumetric flask. An aliquot (100 μL) was added to the reaction mixture as an internal standard. The resulting solution was shaken, passed through a silica plug and analyzed *via* ^{19}F NMR. Yields were determined by ^{19}F NMR of the crude reaction mixture. After purifying the product (**P3**) by preparative thin-layer chromatography (30% EtOAc/hexanes), % ee was determined by SFC.

Materials and General Procedure for Reaction IV

Materials. (1-chloroethyl) benzene **S6** and 3-(4-methoxyphenyl)-propanoyl chloride **S5** were purchased from Combi-Blocks. $NiCl_2$·dme, manganese (powder, ~325 mesh, ≥99% trace metal basis), DMA (≥99%, Sure/Seal™), and 2,5-dimethylbenzoic acid (DMBA) were purchased from Sigma-Aldrich and used as received. THF was dried by passing through activated alumina columns and stored in a nitrogen-filled glovebox.

Procedure. The following screening procedure is adapted from the originally reported acyl coupling.[19] All reactions were set up inside a nitrogen-filled glovebox. Vials were oven-dried (140 °C) and were allowed to cool under vacuum prior to use in the glovebox. Reactions were run on a 0.1 mmol scale. The following stock solution was prepared in a 1 or 2 mL volumetric flask: *Stock Solution 1:* $NiCl_2$·dme 2.20 mg, 10 μmol, 10 mol %), DMBA (11.3 mg, 0.75 mmol, 0.75 equiv), and 200 μL of DMA/THF (3:7) and *Stock Solution 2:* (1-chloroethyl) benzene **S6** (14.1 mg, 0.1 mmol, 1 eq), 3-(4-methoxyphenyl)propanoyl chloride **S5** (23.8 mg, 0.12 mmol, 1.2 equiv), and 65 μL of DMA/THF (3:7).

To a 1/2-dram vial equipped with a Teflon stir bar, 16.5 mg (0.3 mmol, 3 equiv) of Mn was added. Stock solution 2 (65 μL) and stock solution 1 (200.0 μL) were then sequentially added. The reaction vial was capped with a septum cap, wrapped with electrical tape, and brought outside the glovebox. The reaction was stirred (400 rpm) at room temperature for 18 h. Upon completion, dodecane was added as an external standard. Yields were determined by GC-FID of the crude reaction mixture, and % ee was determined by SFC-MS.

Computational Methods

Library Computation Workflow. RDKit[55] and Open Babel[56] were used to obtain an initial guess for the geometry of the free ligand from a SMILES string using a Python script.[57] A script developed from the Schrodinger python API was then used to constrain the free ligand geometry through addition of copper, formation of bonds between the nitrogen donor atoms, and the metal center and addition of two standard ligands to generate four-coordinate [ligand]$Cu(F)_2$ complexes. Copper was chosen due to force field compatibility. A conformational search was performed on the [ligand]CuF_2 complex using MacroModel's maestro tool that employs molecular mechanics (OPLS4 force field).[58] If a ligand had more than 20 unique conformers, representative conformers were selected. The geometry of the metal center and additional ligands, either hydride or fluoride, was then adjusted to achieve the requisite tetrahedral and square planar geometries, respectively. These were subsequently submitted to DFT structure optimization and single-point energy (SPE) calculations. The optimized geometries were checked with a frequency calculation to ensure they were true minima. Geometry optimizations and frequency calculations of these complexes were performed at the B3LYP-D3BJ/LANL2TZ(Ni), 6-31G(d) (other atoms) level of theory[43,59−63] with an ultrafine integration grid. An effective core potential (ECP) was applied to nickel. These optimized geometries were used for subsequent single points at the M06-D3/Def2-TZVP level of theory.[43,44,64] NBO analysis was performed using the NBO program (version 7.0).[65] To obtain data for the unbound ligand descriptors, the optimized geometry of the bound [ligand]NiH_2 was used with the Ni center and hydride ligands removed. This frozen geometry was carried forward for SPE calculations.

DFT Study. All stationary points (minima and transition states) were subjected to conformational sampling with CREST[48] using an energy window of 6 kcal/mol. The number of conformers considered was capped at 25. When CREST generated more than 25 candidates, the built-in clustering algorithm was applied to reduce redundancy. In cases where clustering still yielded more than 25 structures, the 25 lowest-energy conformers were selected and subsequently refined at the DFT level. All calculations were performed using the Gaussian 16 package.[66] Geometry optimizations were carried out at the PBE0-D3/cc-pVDZ(H,C,N,O,Cl)/cc-pVTZ(Ni) level of theory,[46,47,67] including empirical dispersion corrections using Grimme's D3 scheme[43] with Becke–Johnson damping (D3BJ),[61] and frequency calculations were performed at the same level to confirm that the optimized structures correspond to true minima (no imaginary frequencies) or transition states (one imaginary frequency). All minimum energy transition states were confirmed by intrinsic reaction coordinate (IRC) calculations, which verified that each connects the expected reactant and product minima. Single-point energy refinements were obtained using the M06-D3/cc-pVTZ level of theory.[44,47] All reported free energies include thermal corrections to Gibbs-free energy at 298 K.

CYLView20 was used to generate the structure images.[68]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.6c02585.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX−XXX

Experimental procedures, results, and characterization data for case studies I and II; DFT reaction coordinate diagram calculations; Ni(II) nitrogen ligand library; screening set selection; modeling; NMR spectra (PDF)

AUTHOR INFORMATION

Corresponding Authors

Abigail G. Doyle – *Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, California 90095, United States; Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States;* orcid.org/0000-0002-6641-0833; Email: abigaildoyle@g.ucla.edu

Sarah E. Reisman – *The Warren and Katharine Schlinger Laboratory for Chemistry and Chemical Engineering, Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States;* orcid.org/0000-0001-8244-9300; Email: reisman@caltech.edu

Matthew S. Sigman – *Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, United States;* orcid.org/0000-0002-5746-8830; Email: matt.sigman@utah.edu

Authors

Neyci E. Gutiérrez-Valencia – *Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, California 90095, United States;* orcid.org/0000-0003-1283-1656

Jules Schleinitz – *The Warren and Katharine Schlinger Laboratory for Chemistry and Chemical Engineering, Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States;* orcid.org/0000-0001-7116-4772

Therese H. Wild – *Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, United States*

Wendy L. Williams – *Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, California 90095, United States; Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States;* orcid.org/0000-0003-2227-0856

Danielle Mantin – *The Warren and Katharine Schlinger Laboratory for Chemistry and Chemical Engineering, Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, United States*

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acscatal.6c02585

Author Contributions

‡Neyci E. Gutiérrez-Valencia, Jules Schleinitz, and Therese H. Wild contributed equally and are listed alphabetically.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We acknowledge the financial support from the NSF under the CCI Center for Computer Assisted Synthesis (C-CAS) (CHE-2202693). We gratefully acknowledge Dr. Scott Virgil and the Caltech Center for Catalysis and Chemical Synthesis for access to analytical equipment. We thank the Dow Next Generation Educator Funds and Instrumentation Grants for their support of the Caltech CCE NMR facility and Multiuser Mass Spectrometry Laboratory. These studies were supported by shared instrumentation grants from the National Science Foundation under equipment grant number CHE-1048804 and the NIH Office of Research Infrastructure Programs under grant number S10OD028644. The UCLA Institute of Digital Research and Education and Expanse at SDSC, the Resnick High Performance Computing Center, a facility supported by Resnick Sustainability Institute at the California Institute of Technology, and the Center for High Performance Computing (CHPC) at the University of Utah are gratefully acknowledged for their support and resources used for the computations presented here. This work was supported by the Israel Scholarship Education Foundation's International Fellowship Program.

REFERENCES

(1) Tolman, C. A. Phosphorus Ligand Exchange Equilibriums on Zerovalent Nickel. Dominant Role for Steric Effects. *J. Am. Chem. Soc.* **1970**, *92* (10), 2956–2965.

(2) Tolman, C. A. Electron Donor-Acceptor Properties of Phosphorus Ligands. Substituent Additivity. *J. Am. Chem. Soc.* **1970**, *92* (10), 2953–2956.

(3) Williams, W. L.; Zeng, L.; Gensch, T.; Sigman, M. S.; Doyle, A. G.; Anslyn, E. V. The Evolution of Data-Driven Modeling in Organic Chemistry. *ACS Cent. Sci.* **2021**, *7* (10), 1622–1637.

(4) Matsuoka, W.; Harabuchi, Y.; Maeda, S. Virtual Ligand Strategy in Transition Metal Catalysis Toward Highly Efficient Elucidation of Reaction Mechanisms and Computational Catalyst Design. *ACS Catal.* **2023**, *13* (8), 5697–5711.

(5) 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.

(6) Durand, D. J.; Fey, N. Computational Ligand Descriptors for Catalyst Design. *Chem. Rev.* **2019**, *119* (11), 6561–6594.

(7) Durand, D. J.; Fey, N. Building a Toolbox for the Analysis and Prediction of Ligand and Catalyst Effects in Organometallic Catalysis. *Acc. Chem. Res.* **2021**, *54* (4), 837–848.

(8) Crawford, J. M.; Gensch, T.; Sigman, M. S.; Elward, J. M.; Steves, J. E. Impact of Phosphine Featurization Methods in Process Development. *Org. Process Res. Dev.* **2022**, *26* (4), 1115–1123.

(9) 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* (6), 1292–1301.

(10) Crawford, J. M.; Kingston, C.; Toste, F. D.; Sigman, M. S. Data Science Meets Physical Organic Chemistry. *Acc. Chem. Res.* **2021**, *54* (16), 3136–3148.

(11) Jorner, K.; Brinck, T.; Norrby, P.-O.; Buttar, D. Machine Learning Meets Mechanistic Modelling for Accurate Prediction of Experimental Activation Energies. *Chem. Sci.* **2021**, *12* (3), 1163–1175.

(12) Haas, B. C.; Kalyani, D.; Sigman, M. S. Applying Statistical Modeling Strategies to Sparse Datasets in Synthetic Chemistry . *Sci. Adv.* **2025**, *11* (1), eadt3013.

(13) Gensch, T.; Passos Gomes, G.dos; Friederich, P.; Peters, E.; Gaudin, T.; Pollice, R.; Jorner, K.; Nigam, A.; Lindner-D'Addario, M.; Sigman, M. S.; Aspuru-Guzik, A. A Comprehensive Discovery Platform for Organophosphorus Ligands for Catalysis. *J. Am. Chem. Soc.* **2022**, *144* (3), 1205–1217.

(14) Chen, L.-M.; Reisman, S. E. Enantioselective C(Sp2)–C(Sp3) Bond Construction by Ni Catalysis. *Acc. Chem. Res.* **2024**, *57* (5), 751–762.

(15) Gao, Y.; Hu, K.; Rao, J.; Zhu, Q.; Liao, K. Artificial Intelligence-Driven Development of Nickel-Catalyzed Enantioselective Cross-Coupling Reactions. *ACS Catal.* **2024**, *14* (24), 18457–18468.

(16) Schoepfer, A. A.; Laplaza, R.; Wodrich, M. D.; Waser, J.; Corminboeuf, C. Reaction-Agnostic Featurization of Bidentate Ligands for Bayesian Ridge Regression of Enantioselectivity. *ACS Catal.* **2024**, *14* (12), 9302–9312.
(17) Poremba, K. E.; Dibrell, S. E.; Reisman, S. E. Nickel-Catalyzed Enantioselective Reductive Cross-Coupling Reactions. *ACS Catal.* **2020**, *10* (15), 8237–8246.
(18) Woods, B. P.; Orlandi, M.; Huang, C.-Y.; Sigman, M. S.; Doyle, A. G. Nickel-Catalyzed Enantioselective Reductive Cross-Coupling of Styrenyl Aziridines. *J. Am. Chem. Soc.* **2017**, *139* (16), 5688–5691.
(19) Cherney, A. H.; Kadunce, N. T.; Reisman, S. E. Catalytic Asymmetric Reductive Acyl Cross-Coupling: Synthesis of Enantioenriched Acyclic α,α-Disubstituted Ketones. *J. Am. Chem. Soc.* **2013**, *135* (20), 7442–7445.
(20) Roman, G. Mannich Bases in Medicinal Chemistry and Drug Design. *Eur. J. Med. Chem.* **2015**, *89*, 743–816.
(21) Reaxys search from November 2022 of commercially available ligands with the BiOX, BOX and PyOX scaffolds.
(22) Weix, D. J. Methods and Mechanisms for Cross-Electrophile Coupling of Csp2 Halides with Alkyl Electrophiles. *Acc. Chem. Res.* **2015**, *48* (6), 1767–1775.
(23) Dawson, G. A.; Spielvogel, E. H.; Diao, T. Nickel-Catalyzed Radical Mechanisms: Informing Cross-Coupling for Synthesizing Non-Canonical Biomolecules. *Acc. Chem. Res.* **2023**, *56* (24), 3640–3653.
(24) Dotson, J. J.; Dijk, L.; Timmerman, J. C.; Grosslight, S.; Walroth, R. C.; Gosselin, F.; Püntener, K.; Mack, K. A.; Sigman, M. S. Data-Driven Multi-Objective Optimization Tactics for Catalytic Asymmetric Reactions Using Bisphosphine Ligands. *J. Am. Chem. Soc.* **2023**, *145* (1), 110–121.
(25) Development and Application of New Steric Substituent Parameters in Drug Design. In *Drug Design*; Academic Press: **1976**, pp 165–207.
(26) Harper, K. C.; Bess, E. N.; Sigman, M. S. Multidimensional Steric Parameters in the Analysis of Asymmetric Catalytic Reactions. *Nat. Chem.* **2012**, *4* (5), 366–374.
(27) 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* (3), 2313–2323.
(28) Kariofillis, S. K.; Jiang, S.; Żurański, A. M.; Gandhi, S. S.; Martinez Alvarado, J. I.; Doyle, A. G. Using Data Science To Guide Aryl Bromide Substrate Scope Analysis in a Ni/Photoredox-Catalyzed Cross-Coupling with Acetals as Alcohol-Derived Radical Sources. *J. Am. Chem. Soc.* **2022**, *144* (2), 1045–1055.
(29) Zahrt, A. F.; Rose, B. T.; Darrow, W. T.; Henle, J. J.; Denmark, S. E. Computational Methods for Training Set Selection and Error Assessment Applied to Catalyst Design: Guidelines for Deciding Which Reactions to Run First and Which to Run Next. *React. Chem. Eng.* **2021**, *6* (4), 694–708.
(30) McInnes, L.; Healy, J.; Saul, N.; Großberger, L. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Software* **2018**, *3* (29), 861.
(31) Raghavan, P.; Haas, B. C.; Ruos, M. E.; Schleinitz, J.; Doyle, A. G.; Reisman, S. E.; Sigman, M. S.; Coley, C. W. Dataset Design for Building Models of Chemical Reactivity. *ACS Cent. Sci.* **2023**, *9* (12), 2196–2204.
(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) Torres, J. A. G.; Lau, S. H.; Anchuri, P.; Stevens, J. M.; Tabora, J. E.; Li, J.; Borovika, A.; Adams, R. P.; Doyle, A. G. A Multi-Objective Active Learning Platform and Web App for Reaction Optimization. *J. Am. Chem. Soc.* **2022**, *144* (43), 19999–20007.
(34) Ehehalt, L. E.; Beleh, O. M.; Priest, I. C.; Mouat, J. M.; Olszewski, A. K.; Ahern, B. N.; Cruz, A. R.; Chi, B. K.; Castro, A. J.; Kang, K.; Wang, J.; Weix, D. J. Cross-Electrophile Coupling: Principles, Methods, and Applications in Synthesis. *Chem. Rev.* **2024**, *124* (23), 13397–13569.
(35) Denmark, S. E.; Stiff, C. M. Effect of Ligand Structure in the Bisoxazoline Mediated Asymmetric Addition of Methyllithium to Imines. *J. Org. Chem.* **2000**, *65* (18), 5875–5878.
(36) Lundberg, S. M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: **2017**, Vol. 30.
(37) Xu, L.-C.; Frey, J.; Hou, X.; Zhang, S.-Q.; Li, Y.-Y.; Oliveira, J. C. A.; Li, S.-W.; Ackermann, L.; Hong, X. Enantioselectivity Prediction of Pallada-Electrocatalysed C–H Activation Using Transition State Knowledge in Machine Learning. *Nat. Synth.* **2023**, *2* (4), 321–330.
(38) Gallarati, S.; Bucci, E. M.; Doyle, A. G.; Sigman, M. S. Transferable Enantioselectivity Models from Sparse Data. *Nature* **2026**, *651* (8106), 637–646.
(39) Turro, R. F.; Wahlman, J. L. H.; Tong, Z. J.; Chen, X.; Yang, M.; Chen, E. P.; Hong, X.; Hadt, R. G.; Houk, K. N.; Yang, Y.-F.; Reisman, S. E. Mechanistic Investigation of Ni-Catalyzed Reductive Cross-Coupling of Alkenyl and Benzyl Electrophiles. *J. Am. Chem. Soc.* **2023**, *145* (27), 14705–14715.
(40) Fang, J.-D.; Pang, X.; Shu, X.-Z. Nickel-Catalyzed Enantioconvergent Cross-Electrophile Coupling of Benzylic Alcohols with Alkenyl Electrophiles. *J. Am. Chem. Soc.* **2025**, *147* (31), 28313–28321.
(41) Lau, S. H.; Borden, M. A.; Steiman, T. J.; Wang, L. S.; Parasram, M.; Doyle, A. G. Ni/Photoredox-Catalyzed Enantioselective Cross-Electrophile Coupling of Styrene Oxides with Aryl Iodides. *J. Am. Chem. Soc.* **2021**, *143* (38), 15873–15881.
(42) Gutierrez, O.; Tellis, J. C.; Primer, D. N.; Molander, G. A.; Kozlowski, M. C. Nickel-Catalyzed Cross-Coupling of Photoredox-Generated Radicals: Uncovering a General Manifold for Stereoconvergence in Nickel-Catalyzed Cross-Couplings. *J. Am. Chem. Soc.* **2015**, *137* (15), 4896–4899.
(43) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), No. 154104.
(44) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120* (1–3), 215–241.
(45) Balabanov, N. B.; Peterson, K. A. Systematically Convergent Basis Sets for Transition Metals. I. All-Electron Correlation Consistent Basis Sets for the 3d Elements Sc-Zn. *J. Chem. Phys.* **2005**, *123* (6), 64107.
(46) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, *110* (13), 6158–6170.
(47) Dunning, T. H. Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90* (2), 1007–1023.
(48) Pracht, P.; Bohle, F.; Grimme, S. Automated Exploration of the Low-Energy Chemical Space with Fast Quantum Chemical Methods. *Phys. Chem. Chem. Phys.* **2020**, *22* (14), 7169–7192.
(49) Spielvogel, E. H.; Yuan, J.; Hoffmann, N. M.; Diao, T. Nickel-Mediated Radical Capture: Evidence for a Concerted Inner-Sphere Mechanism. *J. Am. Chem. Soc.* **2025**, *147* (23), 19632–19642.
(50) Kozuch, S.; Shaik, S. How to Conceptualize Catalytic Cycles? The Energetic Span Model. *Acc. Chem. Res.* **2011**, *44* (2), 101–110.
(51) Xia, J.; Zhang, Y.; Jiang, B. The Evolution of Machine Learning Potentials for Molecules, Reactions and Materials. *Chem. Soc. Rev.* **2025**, *54* (10), 4790–4821.
(52) Rubin, H.; Cockrell, J.; Morgan, J. B. Scalable Synthesis of N-Acylaziridines from N-Tosylaziridines. *J. Org. Chem.* **2013**, *78* (17), 8865–8871.
(53) Hu, X.; Cheng-Sánchez, I.; Cuesta-Galisteo, S.; Nevado, C. Nickel-catalyzed enantioselective electrochemical reductive cross-coupling of aryl aziridines with alkenyl bromides. *J. Am. Chem. Soc.* **2023**, *145* (11), 6270–6279.

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX–XXX

(54) Dongbang, S.; Doyle, A. G. Ni/Photoredox-Catalyzed C(Sp3)–C(Sp3) Coupling between Aziridines and Acetals as Alcohol-Derived Alkyl Radical Precursors. *J. Am. Chem. Soc.* **2022**, *144* (43), 20067–20077.
(55) *RDKit*. https://www.rdkit.org/ (accessed 2022-11-08).
(56) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An Open Chemical Toolbox. *J. Cheminf.* **2011**, *3* (1), 33.
(57) *Schrödinger Python API*, **2023**. https://www.schrodinger.com/pythonapi.
(58) Lu, C.; Wu, C.; Ghoreishi, D.; Chen, W.; Wang, L.; Damm, W.; Ross, G. A.; Dahlgren, M. K.; Russell, E.; Von Bargen, C. D.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS4: Improving Force Field Accuracy on Challenging Regimes of Chemical Space. *J. Chem. Theory Comput.* **2021**, *17* (7), 4291–4300.
(59) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.
(60) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789.
(61) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465.
(62) Hay, P. J.; Wadt, W. R. Ab Initio Effective Core Potentials for Molecular Calculations. Potentials for K to Au Including the Outermost Core Orbitals. *J. Chem. Phys.* **1985**, *82* (1), 299–310.
(63) Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. 6-31G* Basis Set for Atoms K through Zn. *J. Chem. Phys.* **1998**, *109* (4), 1223–1229.
(64) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297–3305.
(65) Glendening, E. D.; Landis, C. R.; Weinhold, F. NBO 7.0: New Vistas in Localized and Delocalized Chemical Bonding Theory. *J. Comput. Chem.* **2019**, *40* (25), 2234–2241.
(66) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A. Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16* Revision C.01, **2016**.
(67) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.
(68) Legault, C. Y., U. de S. *CYLview Visualization Software*; https://www.cylview.org/ (accessed 2026-02-10).

https://doi.org/10.1021/acscatal.6c02585
ACS Catal. XXXX, XXX, XXX–XXX