

Transfer Learning-Enabled Ligand Prediction for Ni-Catalyzed Atroposelective Suzuki–Miyaura Cross-Coupling Based on Mechanistic Similarity: Leveraging Pd Knowledge for Ni Discovery

Xin-Yuan Xu,[§] Li-Gao Liu,[§] Li-Cheng Xu,[§] Shuo-Qing Zhang,* and Xin Hong*



Cite This: *J. Am. Chem. Soc.* 2025, 147, 15318–15328



Read Online

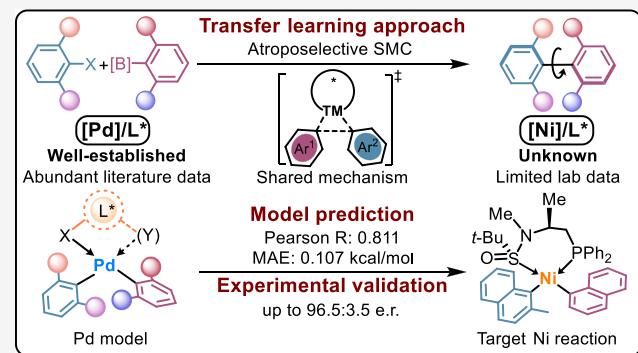
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: The rational design of novel molecular catalysts often confronts challenges due to complex structure–performance relationships. Emerging data-driven approaches provide revolutionary solutions, yet the application of machine learning to new catalyst development inevitably faces a low-data regime with limited effective structure–performance modelings available. In this study, we present a transfer learning strategy to facilitate knowledge transfer from well-documented Pd catalysis to a novel, underexplored Ni system. By synergistically modeling extensive Pd catalysis data with limited Ni/Sadphos data, our approach accurately predicted novel Sadphos ligands, enabling the first atroposelective Ni-catalyzed Suzuki–Miyaura cross-coupling reaction. The synthetic utility of the machine learning-predicted ligand was further demonstrated in its broad synthetic scope, gram-scale synthesis, and precise control of dual axial chiralities in ternaphthalene through the sequential coupling under Ni and Pd catalysis. Additionally, density functional theory calculations were employed to reveal the reaction mechanism and stereochemical model of this new Ni catalyst, validating the proposed mechanistic connection between Ni and Pd. This work demonstrates how machine learning models can effectively leverage mechanistic connectivity, applying extensive structure–performance relationship data from the literature to predict new catalysts, providing a novel strategy for the rational design of molecular catalysts from a few-shot learning perspective.



INTRODUCTION

Catalyst design is an essential technology that drives the efficiency and selectivity of chemical reactions, indispensable for the manufacture and innovation of functional chemicals and advanced materials.¹ However, the cosmic magnitude of molecular structural space,² along with the multitude of controlling factors,³ creates a high-dimensional network of structure–performance relationships (SPRs), rendering rational design and prediction of catalysts profoundly challenging.⁴ Currently, mainstream catalyst design strategies continue to rely on limited modifications at specific sites within existing catalyst frameworks, leveraging only the basic physicochemical properties for dimension reduction and ranking. These empirical catalyst design methods face bottlenecks in precision, efficiency, and, especially, comprehensiveness, struggling to meet the demanding and rapidly evolving needs of frontier synthetic transformations. Consequently, there is a strong desire for novel catalyst design strategies that transcend traditional approaches, accelerating the pace of catalyst development to address the escalating challenges of the synthesis.

With the advancement of machine learning (ML) technologies, chemists are increasingly focusing on harnessing

the value of data to decipher the SPRs of molecular catalysts.^{5–8} Recent exciting developments have demonstrated that supervised learning can effectively constrain models to capture the complexities of catalysis data, thereby providing guided predictions for new catalyst designs.⁹ For example, Doyle et al. utilized parametrization and data analysis of phosphine ligands to uncover the intriguing effects of remote steric hindrance on Ni catalysis, which led to successful predictions of Ni-catalyzed Suzuki coupling of benzylic acetals.¹⁰ Sigman, Mack, and co-workers developed a workflow that combines multiobjective optimization with chemical space mapping, enabling the optimization of bisphosphine ligands across a series of reactions.¹¹ In addition, the field of asymmetric catalysis has also seen substantial advancements in data-driven modeling. Denmark's work illustrated how high-dimensional stereoenvironment encoding and neural network

Received: January 15, 2025

Revised: March 17, 2025

Accepted: March 20, 2025

Published: March 28, 2025



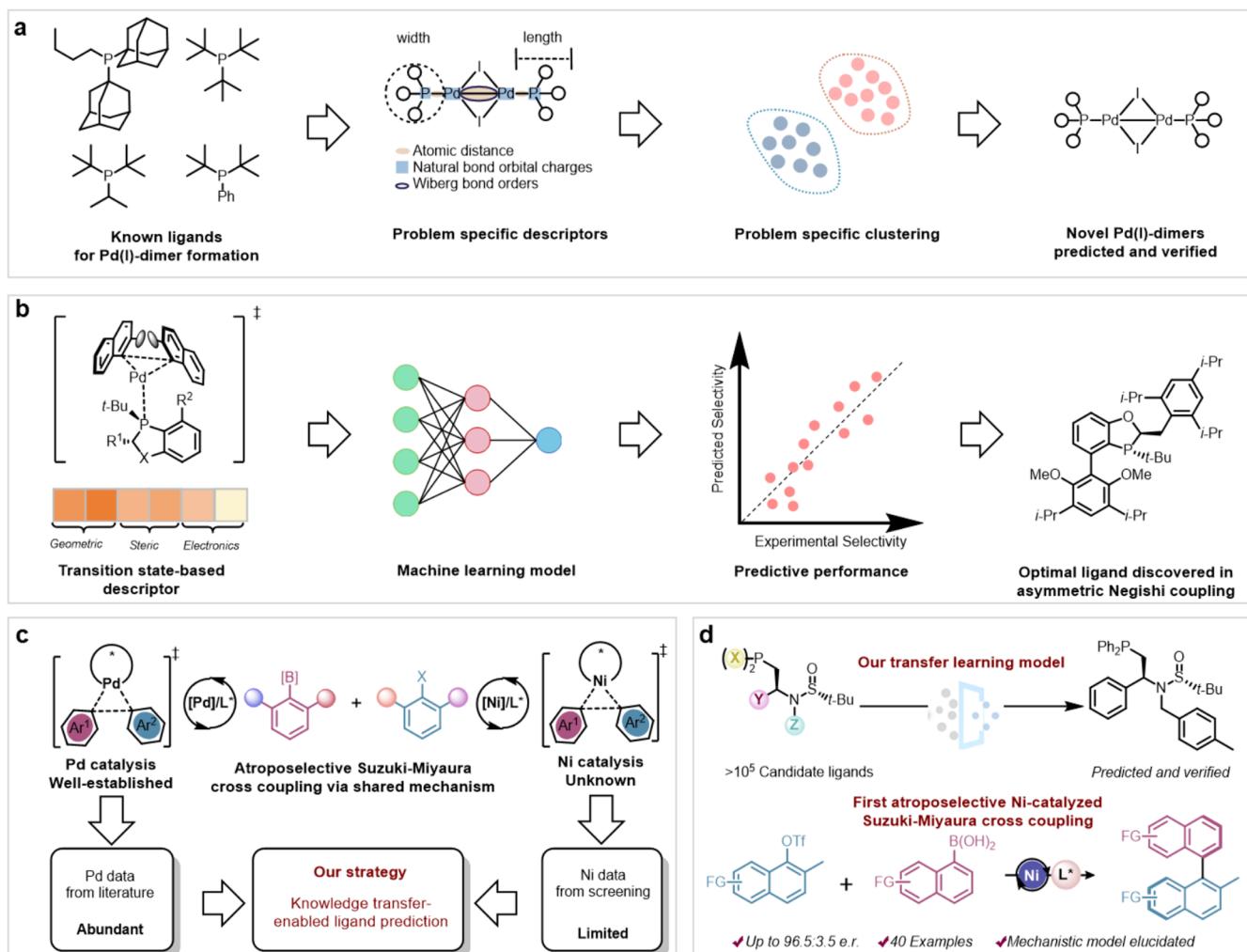


Figure 1. Strategies of machine learning prediction of molecular catalysis under a low-data scenario. (a) Prediction of novel Pd(I)-dimers through an unsupervised learning approach. (b) Transition state-based descriptor design and its machine learning application in the ligand prediction for asymmetric Negishi coupling. (c) Transfer learning strategy that connects the mechanistically related Pd catalysis data from the literature and Ni catalysis data from screening in the atroposelective Suzuki–Miyaura coupling (SMC) reaction. SMC, Suzuki–Miyaura coupling. (d) Machine learning-predicted new Sadphos ligand and its experimental performances in the first atroposelective Ni-catalyzed Suzuki–Miyaura coupling reactions.

model can accurately predict the enantioselectivities of chiral phosphoric acid (CPA) catalysts, thus guiding the selection of highly selective CPAs.¹² Reid's ML modeling has provided valuable insights into the generality of CPA catalysis, enabling the informed decisions of catalyst selection for diverse transformations.¹³ Mlynarski, Grzybowski, and co-workers recently revealed that ML has the capability to design the complete catalytic system including the conditions, which was validated in new catalyst prediction of Mg-catalyzed asymmetric reactions.¹⁴ These studies underscore the transformative potential of ML in revolutionizing the design of molecular catalysts, offering chemists new tools to innovate in the realm of catalysis technologies.

However, in the catalyst prediction for synthetic frontiers, supervised learning inevitably faces the challenge of low-data regime, demanding high predictive capabilities for new structures and enhanced activity/selectivity.¹⁵ Recent advancements in few-shot learning are regarded as a promising solution to this challenge. Schoenebeck et al. innovatively applied unsupervised learning to predict the speciation of dinuclear Pd catalysts, using cluster analysis of phosphine ligands'

physicochemical properties to achieve powerful predictions with merely five data points (Figure 1a).¹⁶ They subsequently extended this method to Ni catalysis, accurately forecasting the formation of Ni(I) dimers.¹⁷ Alternatively, enhancing the descriptor capability has shown effectiveness. Newhouse, Batista, and colleagues reported the usage of computed transition state features to increase the predictive ability for Negishi coupling where limited data is provided (Figure 1b).¹⁸ This transition state-based descriptor design has also been employed in reaction modelings by Grzybowski,¹⁹ Buttar,²⁰ and our group.²¹ However, a traditional catalyst design approach used by chemists—utilizing data from mechanistically related reactions to inspire new catalyst designs, has not been fully integrated into chemical ML. To the best of our knowledge, no ML research has yet achieved the integration of data across different transition metals for knowledge-transfer predictions of new catalysts. Therefore, despite the abundance of catalytic data in the literature, its potential to enhance ML modeling of new catalytic systems remains largely underexploited.

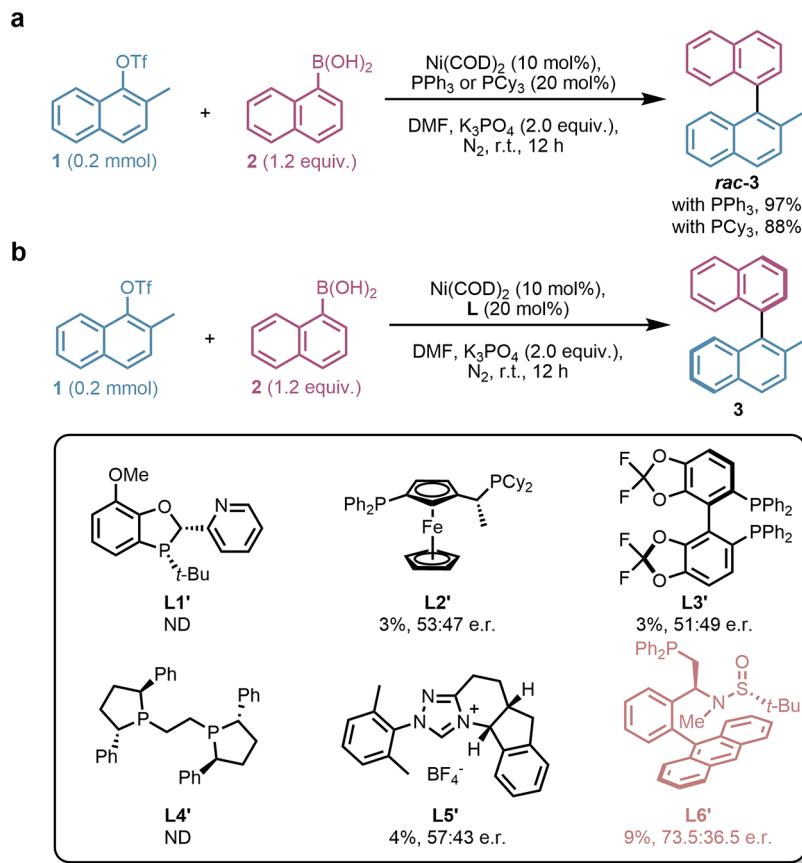


Figure 2. Optimized reaction conditions for the racemic Ni-catalyzed Suzuki–Miyaura coupling and performances of representative chiral ligand scaffolds in the enantioselective version. (a) Yields of the racemic synthesis of 2-methyl-1,1'-binaphthalene under the optimized reaction conditions. rt, room temperature. DMF, dimethylformamide. (b) Yields and enantioselectivities of the stereoselective synthesis of 2-methyl-1,1'-binaphthalene under the representative chiral phosphine and *N*-heterocyclic carbene ligands. er, enantiomeric ratio. ND, not detected. and its effectiveness in gram-scale synthesis, including the precise control of dual axial chiralities in ternaphthalene.

To achieve data-driven knowledge transfer of molecular catalysis, we surmised that mechanistically related SPR data can be synergistically modeled through transfer learning, thereby harnessing literature information to empower catalyst predictions for new reactions with limited data. In this study, we report an ML model that utilized the wealth of Pd catalysis data to guide the ligand prediction for Ni catalyst, which led to the successful prediction of a novel Sadphos ligand that enabled the first atroposelective Ni-catalyzed Suzuki–Miyaura cross-coupling reactions from over ten thousand of candidates (Figure 1c,d). Furthermore, our findings extend beyond the initial prediction, showcasing the catalyst's broad applicability across diverse substrates. In addition, density functional theory (DFT) calculations were performed to verify the proposed mechanistic connection and elucidate the stereochemical model behind the excellent selectivity of this ML-predicted catalyst. This research not only highlights the capability of ML to bridge information gaps between diverse catalytic systems but also establishes a practical framework for transforming extensive catalysis data into actionable insights for new catalyst designs.

RESULTS AND DISCUSSION

Initial Attempts and Optimization of the Target Ni-Suzuki–Miyaura Cross-Coupling. Considering the enantioselectivity-determining reductive elimination step as a critical mechanistic connection between Pd and Ni in the

target atroposelective Suzuki–Miyaura cross-coupling reactions, the key objective of our ML study is focused on enantioselectivity modeling and the prediction of highly selective chiral ligands. Therefore, we first set out to identify the conditions and ligand scaffold with the desired catalytic efficiency. Based on the prior experimental research on Ni/phosphine-catalyzed aryl–aryl Suzuki–Miyaura cross-coupling,²² our condition screenings were initiated from the racemic Suzuki–Miyaura cross-coupling between **1** and **2** using PCy_3 or PPh_3 as ligand. Comprehensive screenings of Ni catalyst precursors, solvents, and bases revealed that the combination of $\text{Ni}(\text{COD})_2$, dimethylformamide (DMF), and K_3PO_4 proved effective, demonstrating good to excellent yields in both alkyl and aryl phosphine systems, with the highest yield reaching 97% (Figure 2a, full details in Tables S1–S3). Consequently, these conditions were adopted for subsequent evaluation of chiral ligands.

Considering the structural diversity of ligand scaffolds, we experimented with over 20 representative chiral phosphine ligands. Phosphine ligands typically used in asymmetric Ni catalysis showed a significant decrease in catalytic activity compared to PCy_3 or PPh_3 (Figure 2b, full details in Table S4). This decline highlights the challenges of catalytic efficiency when sterically demanding aryl moieties are involved in C–C bond formation. However, to our surprise, Zhang's Sadphos ligands,²³ which have rarely been used in Ni catalysis,²⁴ demonstrated promising activity and selectivity (Figure 2b, full

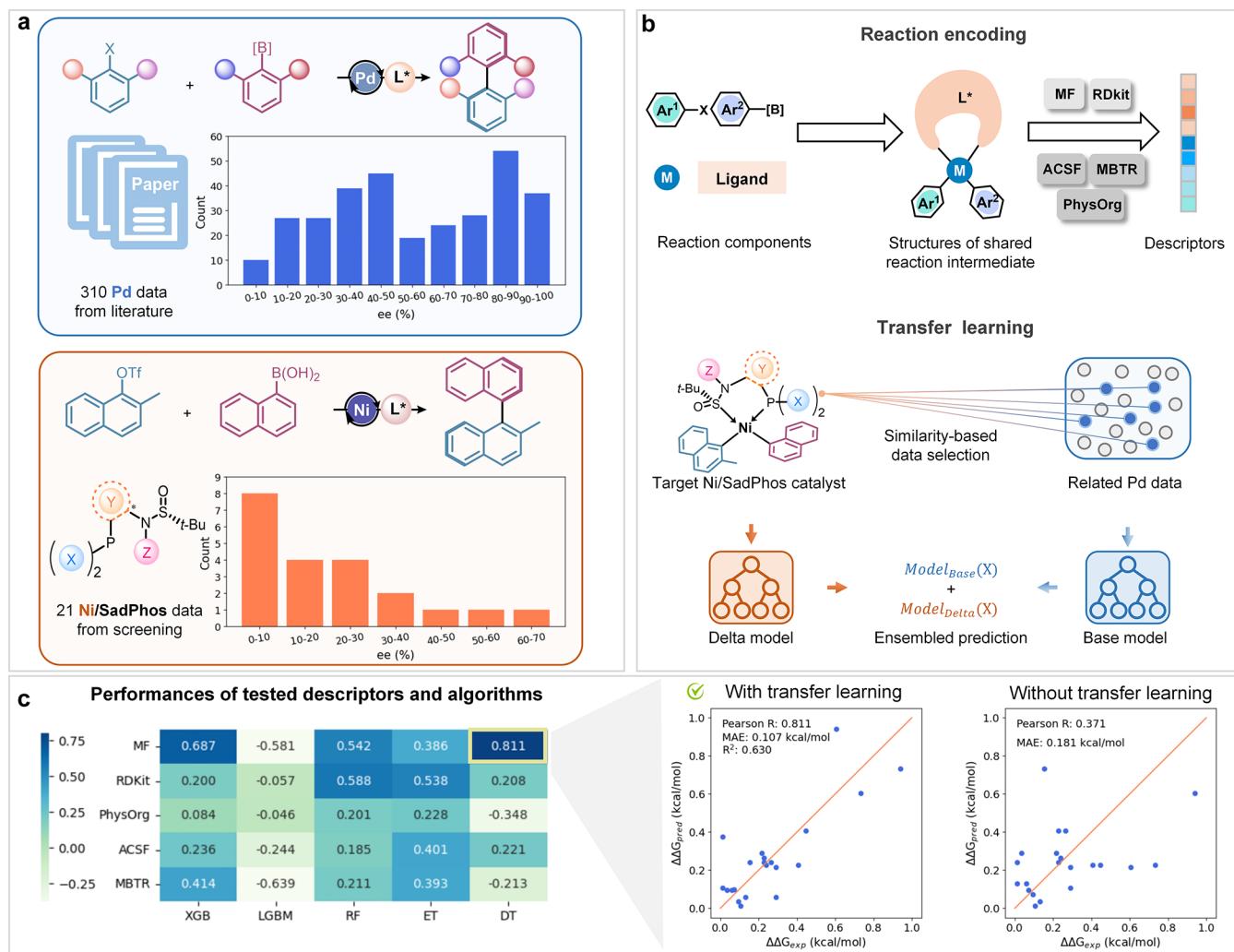


Figure 3. Design of the transfer learning strategy and its modeling performances. (a) Data sources and distributions of the atroposelective Pd- and Ni-catalyzed Suzuki–Miyaura coupling reactions. ee, enantiomeric excess. (b), Mechanism-based reaction encoding using the shared $LM(II)Ar_2$ intermediates and the transfer learning workflow. MF, Morgan molecular fingerprints. ACSF, atom-centered symmetry functions. MBTR, many-body tensor representation. PhysOrg, physical organic descriptors. (c) Enantioselectivity prediction performances of the target Ni/Sadphos system using various combinations of molecular descriptors and regression algorithms, and the comparison of models with or without the transfer learning approach under the optimal combination of Morgan fingerprints and decision tree algorithm. XGB, XGBoost. LGBM, light gradient-boosting machine. RF, random forest. ET, extra-tress. DT, decision tree. Pearson R, Pearson correlation coefficient. MAE, mean absolute error.

details in Table S4). Additionally, given Sadphos’ modular synthetic strategy and versatile derivatization potential, this ligand type emerged as an ideal candidate for our ML prediction and screening efforts.

Machine Learning Prediction of Highly Selective Ligands. For the enantioselectivity modeling of the target reaction, our transfer learning approach leverages data from two sources: documented literature on Pd catalysis and our experimental evaluations of Ni/Sadphos catalysis for the same type of atroposelective Suzuki–Miyaura cross-coupling reactions. For the Pd catalysis data, we curated a data set focused on the catalytic condition screening from related studies published since 2000. This data set comprises 310 reaction entries sourced from 20 publications, involving 28 aryl halides and 15 boronic acids, covering six types of chiral ligands including *N*-heterocyclic carbenes and phosphine ligands (Figure 3a, top, full details in Figure S2). The distribution of the Pd catalysis data is balanced, with 37 catalyst combinations achieving enantioselectivities of 90% enantiomeric excess (ee)

or higher, while 167 reactions recorded ee values below 60%, offering a critical basis for successful modeling in terms of the value range and distribution. Regarding the targeted Ni/Sadphos catalysis, we evaluated the selectivities of commercially available Sadphos ligands in the model Suzuki–Miyaura cross-coupling reaction between substrates 1 and 2, yielding 21 entries (Figure 3a, bottom, full details in Table S5). This data set is limited in number, with values predominantly in the low selectivity range, reflecting a typical, realistic distribution for a catalytic system in need of optimization. It is important to note that such a distribution of Ni/Sadphos data alone is insufficient for ML modeling as it fails to meet the data requirements necessary for generating a predictive model (vide infra).

Subsequently, we leveraged the mechanistic connection between Pd and Ni catalysis to encode the two data sources into one unified format (Figure 3b, top). Given that both catalytic systems share the enantioselectivity-determining reductive elimination step in the target Suzuki–Miyaura cross-coupling reactions, the corresponding pre-intermediate

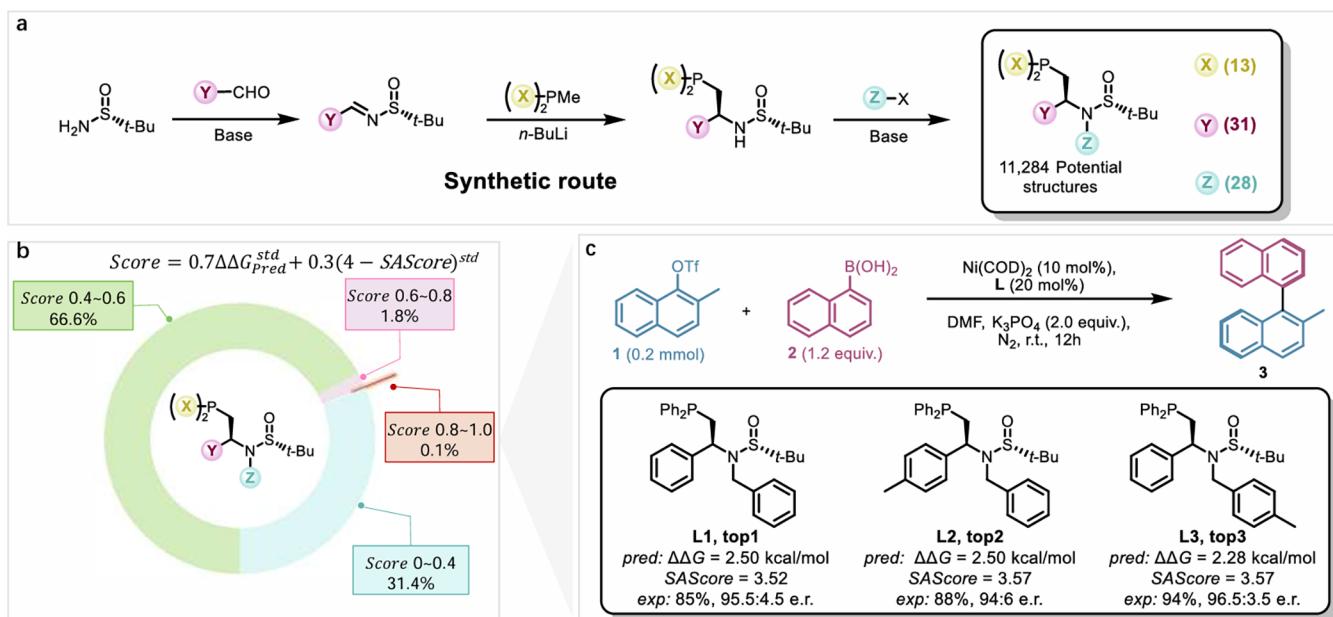


Figure 4. Machine learning predictions of new Sadphos ligands and their experimental verifications. (a) Library generation of the candidate Sadphos ligands based on its synthetic route. (b) Scoring function based on the predicted selectivity, the synthetic accessibility, and the score distribution of the explored Sadphos ligands. SAScore, synthetic accessibility score. (c) Experimental validations of the top three predicted Sadphos ligands in the atroposelective Ni-catalyzed Suzuki–Miyaura coupling reaction of 2-methyl-1,1'-binaphthalene synthesis.

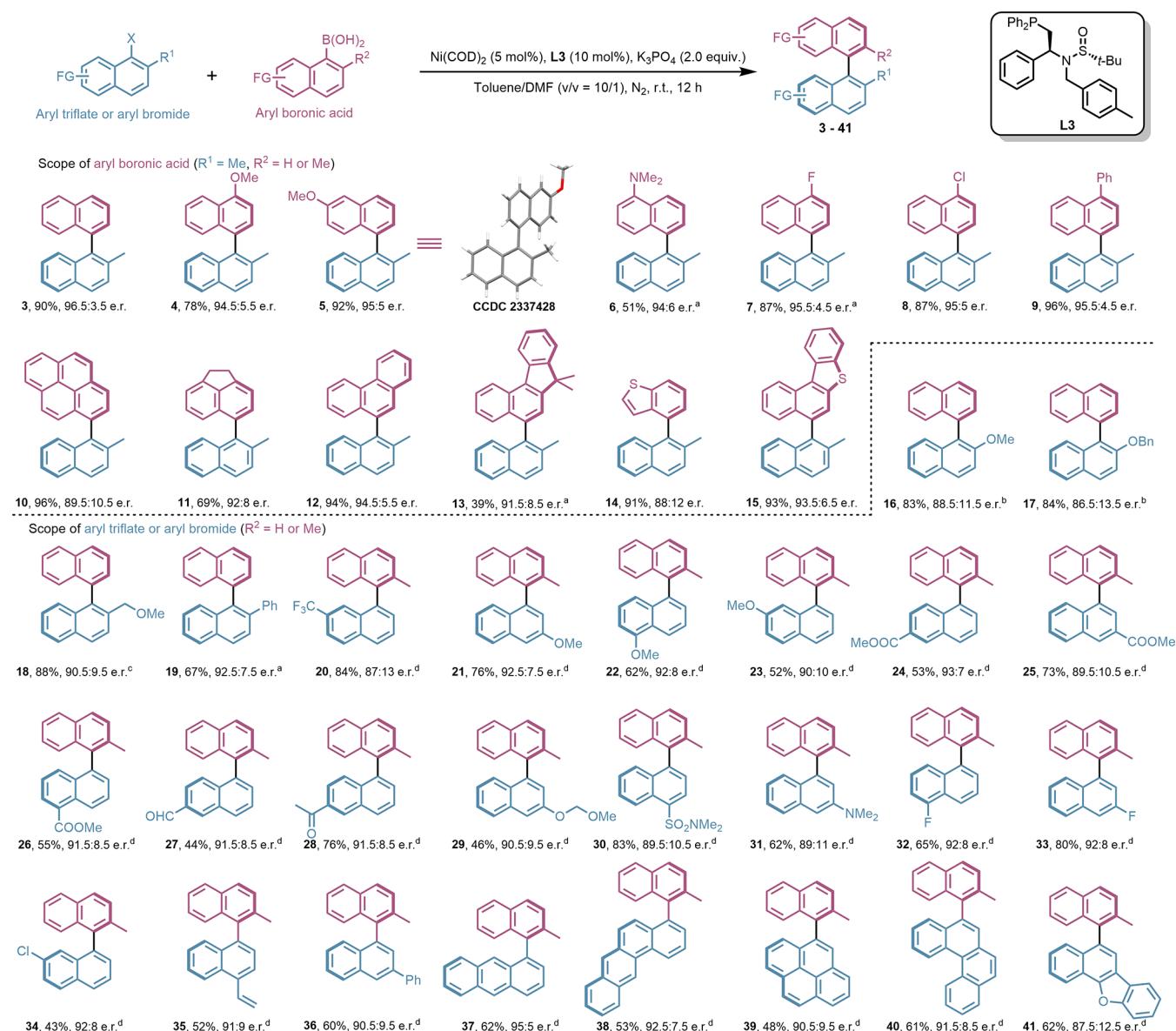
LMAr₂ was used as the structural basis for encoding. Utilizing the metal, ligand, and substrate information from the data set, we generated the two-dimensional structures for the 171 LMAr₂ intermediates, yielding Morgan molecular fingerprints (MF) and RDKit topological descriptors. Moreover, geometry optimizations at the GFN2-xTB level were conducted for the active catalysts and substrates to produce three-dimensional descriptors of atom-centered symmetry functions (ACSF) and many-body tensor representation (MBTR). Physical organic (PhysOrg) descriptors were also incorporated to account for the key substructures of the catalytic center and the reacting carbon in the substrates, enhancing the comprehensiveness of the encoding process. Eventually, each reaction was encoded using five approaches (MF, RDKit, ACSF, MBTR, and PhysOrg) for ML modeling.

In our modeling framework, the transfer learning approach was implemented by extracting relevant data and learning the Pd-to-Ni SPR perturbations (Figure 3b, bottom). Given the structural diversity of the 21 Ni/Sadphos catalysts for target predictions, we believe that their mechanistic connections to the Pd catalysis data are not equal. To better leverage the embedded knowledge, we curated relevant Pd data for each Ni/Sadphos catalyst using the Euclidean distance of MF. This led to the individually tailored Pd models for each Ni/Sadphos catalyst. These Pd models cannot perfectly predict the Ni selectivities due to the obvious perturbations between the Pd and Ni SPRs. However, these perturbations are learnable and provided opportunities for model refinement under a data-limited scenario. To fine-tune the predictions from the Pd model, we trained a separate delta learning model that captures the Pd-to-Ni SPR perturbations. Consequently, the final predictions for each Ni catalyst were generated by combining outputs from both the Pd model and the Pd-to-Ni delta learning model, thereby enabling ML to harness the extensive knowledge from the Pd catalysis literature and its deviations when applied to Ni catalysis. A step-by-step workflow with

details of data splitting is provided in the Supporting Information (Figure S7).

The designed transfer learning strategy successfully integrated Pd and Ni catalysis data, enabling the desired prediction of the target Ni/Sadphos catalysts. Figure 3c presents the Pearson correlation coefficients of the ML models utilizing the aforementioned encoding methods combined with typical regression algorithms including XGBoost (XGB), light gradient-boosting machine (LGBM), random forest (RF), decision tree (DT), and extra-trees (ET). To our delight, the MF-DT combination exhibited a Pearson correlation coefficient of 0.811 and an R² of 0.630 (full details in Figure S8). Figure 3c details the performance of this model, displaying a mean absolute error (MAE) of 0.107 kcal/mol. Importantly, this model successfully identified the Sadphos ligands with higher selectivities from the 21 structurally and functionally diversified candidates—a task that poses significant challenges even for experienced experts in asymmetric catalysis. Additionally, we explored the effect of different distance metrics and molecular fingerprint encodings on the transfer learning performance. Our results showed that these factors can influence the transfer effectiveness in this few-shot learning scenario, with the combination of Euclidean distance and Morgan molecular fingerprints producing the best outcome (full details in Figure S9). This indicates the critical role of chemical similarity projection and measurement in the successful knowledge transfer in SPR.

To further verify the effectiveness of the transfer learning strategy, we also evaluated the ML model trained solely on the 21 Ni/Sadphos data, which yielded a Pearson coefficient of only 0.371 and the inability to differentiate the ligands' chiral discrimination abilities (Figure 3c, right), highlighting the inherent SPR complexity of the target Ni/Sadphos system. Moreover, if a tailored Pd model was not constructed for each Sadphos ligand and instead a static Pd model was generated using all Pd catalysis data, the transfer learning goal could not

Table 1. Substrate Scope of the Atroposelective Ni/L3-Catalyzed Suzuki–Miyaura Cross-Coupling Reaction^e

^aReactions were performed using $\text{Ni}(\text{COD})_2$ (10 mol %) and **L3** (20 mol %). ^bReactions were performed on 0.4 mmol scale using $\text{Ni}(\text{COD})_2$ (1 mol %) and **L3** (2 mol %). ^cReactions were performed on 0.1 mmol scale using THF (1 M), $\text{Ni}(\text{COD})_2$ (10 mol %), and **L3** (20 mol %).

^dReactions were performed at 60 °C using $\text{Ni}(\text{COD})_2$ (10 mol %) and **L3** (20 mol %). v/v, volume/volume. ^eUnless otherwise noted, reactions were performed using aryl triflate or bromide (0.2 mmol), aryl boronic acid (1.2 equiv), $\text{Ni}(\text{COD})_2$ (5 mol %), **L3** (10 mol %), and K_3PO_4 (2.0 equiv) under nitrogen atmosphere at room temperature. Separated yields.

be achieved (Figure S12). This emphasizes the necessity of dynamic data extraction, considering the diverse positions of the Sadphos targets within the SPR space.

Utilizing the trained transfer learning models, we investigated a vast array of Sadphos ligands. Based on the synthetic route of Sadphos ligands reported by Zhang et al.²⁵ and the commercial availability of related building blocks, we targeted three key modifiable sites on its scaffold, considering a wide array of functional groups at each position (Figure 4a, full details in Figure S4). This resulted in 11,284 distinctive Sadphos ligands, covering a broad range of structural and functional diversity. Each candidate underwent the ML prediction of enantioselectivity, and its synthetic accessibility was also evaluated using the SAScore developed by Ertl et al.²⁶ We then ranked the candidates by integrating the predicted

enantioselectivity and SAScore to identify the leading selections (Figure 4b). The top three predicted Sadphos ligands not only demonstrated superior predicted selectivities but also featured simple and synthesizable structures due to SAScore constraints. The predicted er values for **L1**, **L2**, and **L3** were 98.5:1.5, 98.5:1.5, and 98.2 (Figure 4c), respectively. Subsequent synthesis and evaluations of these three ligands revealed exciting results. These new members of the Sadphos ligands all demonstrated excellent enantioselectivities and yields in the Suzuki–Miyaura cross-coupling reaction between substrates **1** and **2**, with **L3** achieving the highest enantioselectivity of 96.5:3.5 er. These validations not only provided convincing experimental support for our transfer learning strategy but also demonstrated that this strategy can endow the model with extrapolative capabilities in both the

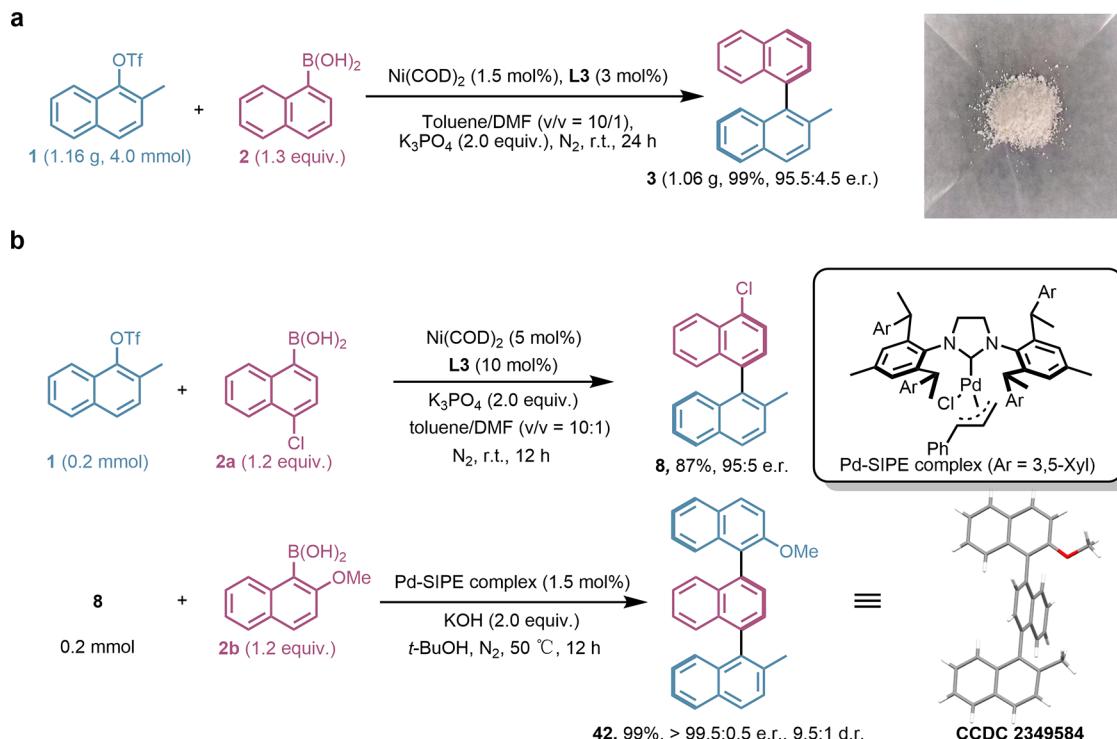


Figure 5. Synthetic applications. (a) Gram-scale reaction. (b) Enantio- and diastereoselective synthesis of unsymmetric ternaphthalene atropisomers using the atroposelective Ni/Sadphos-catalyzed Suzuki–Miyaura cross-coupling reaction and the reported Pd/SIPE conditions.²⁷ SIPE, 1,3-bis(2,6-diisopropyl-4-methylphenyl)imidazolidine.

structure and selectivity spaces, which is crucially needed in the few-shot learning scenario when designing new catalysts.

Notably, all three predicted ligands displayed a distinctive *N*-benzyl substitution, which is not present in any of the Sadphos ligands in the training set or found in any commercially available Sadphos ligands. Examination of the Pd literature data used for training revealed the two diastereoisomeric phosphoramidite ligands that contain the *N*-benzyl substitution (Figure S3), which are likely the source of knowledge that led to the prediction of the top three Sadphos ligands. This finding further indicates the importance of Pd literature data and highlights the ability of our transfer learning model to integrate and leverage external catalysis knowledge for innovative molecular design.

Substrate Scopes and Synthetic Applications. With the predicted Ni/L3 catalyst at hand, we further explored its substrate scope with both coupling partners (Table 1). This catalyst exhibited satisfactory performances on aryl boronic acids, accommodating both electron-withdrawing and -donating substituents without compromising enantioselectivity (4–8). Extended conjugated ring systems were also well tolerated, showing good to excellent enantioselectivities in compounds 9 to 13. The effective transformation of benzothiophene derivatives 14 and 15 demonstrated the catalyst's adaptability with heteroatom-containing substrates. In addition, this catalytic condition demonstrated versatility in applications involving nucleophilic components. Introducing moderately bulky or coordinating groups at the ortho-position did not compromise the catalytic performances (16–19). Compounds featuring common electron-donating and -withdrawing substituents were also prepared with high selectivities (20–33). Notably, reactive functional groups including aryl chloride and terminal unsubstituted olefins 34 and 35, as well as extended

conjugated arenes 36 to 41, were well accommodated, providing opportunities for the synthesis of complex or functional molecules.

The gram-scale synthesis and controlled formation of axially chiral ternaphthalenes further demonstrated the application potential of our Ni/L3 catalyst. The atroposelective Suzuki–Miyaura cross-coupling reaction between substrates 1 and 2 was successfully scaled up to gram-scale under optimal conditions, achieving a 99% yield with an enantiomeric ratio of 95.5:4.5 (Figure 5a). Particularly, the catalyst loading for this gram-scale synthesis was low at 1.5 mol %, highlighting the superior catalytic efficiency of the predicted Sadphos ligand L3. Moreover, the differential chemoselectivity between the designed Ni catalysis and the previously reported Pd catalysis²⁷ allowed for the enantioselective and diastereoselective synthesis of unsymmetric ternaphthalene atropisomers. Leveraging the tolerance for aryl chlorides in the Ni/L3 catalyst, aryl triflate 1 was selectively coupled with chlorine-containing 2a to synthesize axially chiral product 8 with an er of 95:5. Subsequently, 8 was transformed using Shi's Pd/NHC catalyst,²⁷ constructing the second axial chirality to precisely synthesize one of the four axially chiral isomers of ternaphthalene 42, with an outstanding er of over 99.5:0.5 and a favorable dr of 9.5:1 (Figure 5b). This application showcases that the knowledge transfer from Pd to Ni and the expansion of mechanistically connected catalysis not only hold theoretical value but also offer promising synthetic opportunities.

Mechanism Investigations. To further confirm the proposed mechanistic consistency between Ni and Pd catalysis in the aryl–aryl Suzuki–Miyaura cross-coupling and to understand the stereochemical model of the predicted Sadphos ligand, we further employed DFT calculations to study the

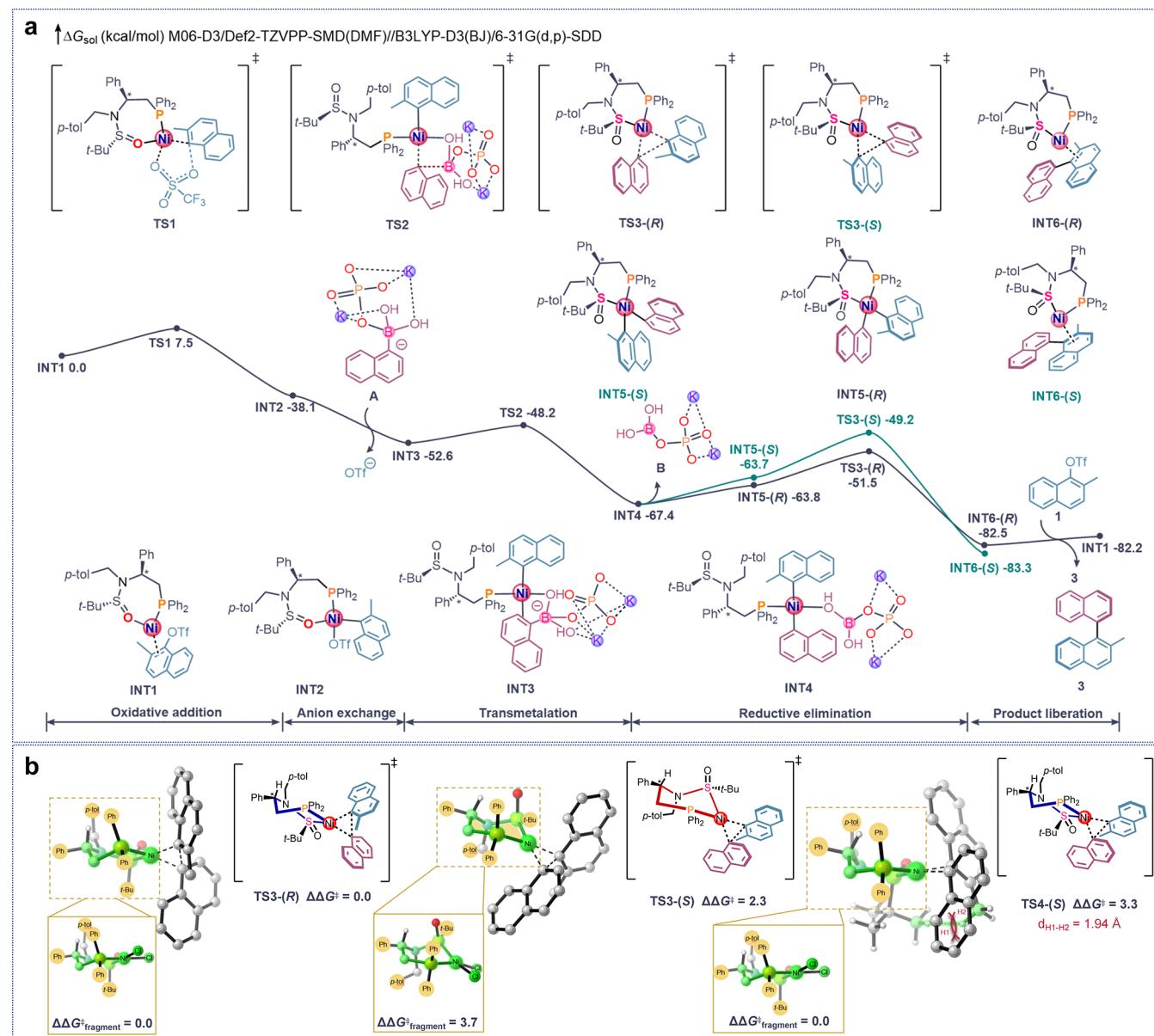


Figure 6. DFT-computed free energy diagram and origins of enantioselectivity. (a) Free energy changes of the complete catalytic cycle of Ni/L3-catalyzed Suzuki–Miyaura cross-coupling reaction using the synthesis of 2-methyl-1,1'-binaphthalene as a model. (b) Optimized structures and relative free energies of the enantioselectivity-determining reductive elimination transition states. $\Delta \Delta G_{\text{fragment}}^{\ddagger}$ refers to the relative free energies of the L3NiCl₂ complex which share the same ring conformation of the active catalyst fragment as the corresponding transition states.

catalytic cycle and the origins of enantioselectivity. We found that this transformation follows a typical M(0)/M(II) cross-coupling mechanism, and Figure 6a presents the DFT-computed free energy diagram of the Ni/L3-catalyzed Suzuki–Miyaura cross-coupling reaction between substrates **1** and **2**. Starting from the Ni(0)-aryl triflate complex INT1, the oxidative addition via TS1 is facile and irreversible,²⁸ generating the arylNi(II) intermediate INT2. Subsequently, the anion exchange between K₂PO₄⁻ and the triflate anion produces the pretransmetalation intermediate INT3. INT3 has L3 coordinating as a monodentate phosphine ligand, which allows the four-membered ring transmetalation through TS2. After the transmetalation, INT4 dissociates the boric acid-K₂PO₄⁻ complex to form the trans-configured Ni(II) intermediate INT5. INT5 undergoes irreversible C–C reductive elimination through TS3-(R) to generate the

Ni(0)-product complex INT6, which eventually liberates the binaphthalene product with *R*-axial chirality. Based on the computed free energy profile, the rate- and enantioselectivity-determining step is identified as the C–C reductive elimination, where the competition between TS3-(R) and TS3-(S) leads to atroposelectivity. A previous study by Lin et al. also indicates that reductive elimination can be the rate-determining step in the Suzuki–Miyaura cross-coupling reaction.²⁹ The computed free energy differences between TS3-(R) and TS3-(S) are 2.3 kcal/mol, which is in excellent agreement with the observed enantioselectivity of 96.5:3.5 er. This mechanistic model is consistent with those of the Pd catalysts,³⁰ further validating our hypothesized mechanistic connection. Throughout the catalytic cycle, the coordination of the Sadphos ligand alters between the monodentate P and bidentate P/S. This P/S bidentate coordination mode has also

been observed in Pd(0) complex by Zhang et al., with crystallographic evidence supporting its existence.²³ Other coordination modes involving oxygen were also considered but were found to be less favorable (full details in Figures S13–S15).

We next examined TS3-(R) and TS3-(S) to elucidate the origins of the enantioselectivity for this ML-designed ligand. Thorough conformational searches were performed, and the key transition state structures are shown in Figure 6b (further details are provided in Figure S19). We found that in the most favorable transition states TS3-(R) and TS3-(S), the Ni/L3 catalyst adopts two distinctive six-membered ring conformations. Fragment analysis indicated that the ring conformation in TS3-(R) is more stable by 3.7 kcal/mol, which is the major contribution to the energy difference between TS3-(R) and TS3-(S). The reason TS3-(S) adopts a less stable catalyst conformation is to mitigate the steric repulsions between the methyl substituent of the substrate and the bulky *t*-butyl group on the chiral sulfur center of the ligand. To demonstrate this, we were able to locate TS4-(S), which has the identical ring conformation of catalyst as TS3-(R). Such a ring conformation in TS4-(S) leads to the steric clash between the highlighted methyl group and the bulky *t*-butyl group on the chiral sulfur center of the ligand, with the closest H–H distance being only 1.94 Å. Consequently, TS4-(S) is less favorable by 3.3 kcal/mol compared to TS3-(R). To avoid this steric clash in TS4-(S), TS3-(S) was forced to adopt a less advantageous catalyst conformation, resulting in barrier differences during reductive elimination between the atropisomer formations. These analyses provide a molecular-level stereochemical model for this new ligand and reveal the role of ring conformation change in the chiral induction of the Sadphos ligands.

CONCLUSIONS

In summary, we have successfully achieved knowledge transfer from Pd to Ni catalysis through machine learning, constructing a structure–performance relationship model that accurately predicted novel chiral ligands and subsequently enabling the first atroposelective Ni-catalyzed Suzuki–Miyaura cross-coupling reaction. Our approach exploits the mechanistic similarities between Pd and Ni catalysts, allowing synergistic modeling of extensive Pd catalysis data with limited Ni/Sadphos information. By integrating sampling and delta learning techniques, we refined the model using structurally similar Pd data for each Ni/Sadphos catalyst and adjusted for discrepancies between Pd model predictions and actual Ni performances. The constructed transfer learning model, trained on only 21 Sadphos ligand data points, predicted enantioselectivity with a Pearson regression coefficient of 0.811 and a mean absolute error of 0.107 kcal/mol, significantly outperforming the training results from the small-sample Ni/Sadphos data set itself, as well as the direct application of the Pd catalysis model in a cross-metal scenario.

This transfer learning model provided the opportunity to virtually screen a large array of candidate Sadphos ligands. Based on the Sadphos scaffold's readily derivatizable positions and the potential groups for modification, we constructed up to 11,284 distinctive Sadphos ligands with diverse structures and properties. The top three predicted Sadphos ligands were synthesized and evaluated experimentally for their compatibility with asymmetric catalysis. These new Sadphos ligands exhibited exciting activity and selectivity in the target atroposelective coupling reaction, providing compelling

experimental validation for our transfer learning strategy. Moreover, the synthetic utility of the predicted Ni/Sadphos catalyst was demonstrated. The catalyst presented broad applicability with coupling partners, compatibility with a wide array of functional groups, and effectiveness in transforming substrates containing extended conjugated systems or certain heterocycles. The catalyst was also successfully scaled up to the gram scale, maintaining near-quantitative yields and excellent enantioselectivity with a low catalyst loading. Furthermore, by leveraging the differential chemoselectivity between our Ni catalyst and previously reported Pd systems, precise control over the double axial chiralities of ternaphthalene was achieved.

Additionally, density functional theory calculations were employed to investigate the catalytic mechanism and the origins of enantioselectivity. The catalytic cycle, which follows a Ni(0)/Ni(II) cross-coupling pathway, pinpointed the aryl–aryl reductive elimination as the rate- and enantioselectivity-determining step, thus confirming the mechanistic connection between Ni and Pd proposed in our study. The effective chiral induction by the new Sadphos ligand was attributed to the steric effects of the chiral sulfur center and the rigidity of the Ni/Sadphos ring conformation. Overall, this work provides an effective synergistic modeling strategy for molecular catalysis data with mechanistic linkage, illustrating that machine learning models can harness extensive catalytic knowledge from the literature to empower new reaction predictions under low-data scenarios. We envision this method to significantly accelerate the mechanism-based discovery of new catalysts by leveraging the digitized representation of mechanistic connections and the effective utilization of literature knowledge through machine learning.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/jacs.Sc00838>.

Experimental procedures, machine learning methods, computational data, and copies of NMR spectra and HPLC spectra ([PDF](#))

Accession Codes

Deposition Numbers [2337428](#) and [2349584](#) contain the supplementary crystallographic data for this paper. Deposition Numbers 2337432 contain the supplementary crystallographic data for the Supporting Information. These data can be obtained free of charge via the joint Cambridge Crystallographic Data Centre (CCDC) and Fachinformationszentrum Karlsruhe [Access Structures service](#).

AUTHOR INFORMATION

Corresponding Authors

Shuo-Qing Zhang – Center of Chemistry for Frontier Technologies, Department of Chemistry, State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, P. R. China;  [orcid.org/0000-0002-7617-3042](#); Email: angellasty@zju.edu.cn

Xin Hong – Center of Chemistry for Frontier Technologies, Department of Chemistry, State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, P. R. China; School of Chemistry and Chemical Engineering, Henan Normal University, Xinxiang 453007, P. R. China;

 orcid.org/0000-0003-4717-2814; Email: hxchem@zju.edu.cn

Authors

Xin-Yuan Xu — *Center of Chemistry for Frontier Technologies, Department of Chemistry, State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, P. R. China*

Li-Gao Liu — *Center of Chemistry for Frontier Technologies, Department of Chemistry, State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, P. R. China*

Li-Cheng Xu — *Center of Chemistry for Frontier Technologies, Department of Chemistry, State Key Laboratory of Clean Energy Utilization, Zhejiang University, Hangzhou 310027, P. R. China*

Complete contact information is available at:

<https://pubs.acs.org/10.1021/jacs.Sc00838>

Author Contributions

[§]X.-Y.X., L.-G.L., and L.-C.X. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

National Key R&D Program of China (2022YFA1504301, X.H.), National Natural Science Foundation of China (22122109, 22271253, and W2512004, X.H.; 22103070, S.-Q.Z.), the Starry Night Science Fund of Zhejiang University Shanghai Institute for Advanced Study (SN-ZJU-SIAS-006, X.H.), CAS Youth Interdisciplinary Team (JCTD-2021-11, X.H.), Fundamental Research Funds for the Central Universities (226-2022-00140, 226-2022-00224, 226-2023-00115, and 226-2024-00003, X.H.), the State Key Laboratory of Clean Energy Utilization (ZJUCEU2020007, X.H.), the State Key Laboratory of Physical Chemistry of Solid Surfaces (202210, X.H.), the Leading Innovation Team grant from Department of Science and Technology of Zhejiang Province (2022R01005, X.H.), and Open Research Fund of School of Chemistry and Chemical Engineering of Henan Normal University (2024Z01, X.H.) are gratefully acknowledged. Calculations were performed on the high-performance computing system at Department of Chemistry, Zhejiang University. We thank Jiyong Liu for the support with X-ray single-crystal diffraction analyses, Dr. Qiaohong He and Xiaoping Li for ESI-HRMS and EI-HRMS analyses, and Dr. Yuanjiang Pan's group for matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) analyses.

REFERENCES

- (1) Ahn, S.; Hong, M.; Sundararajan, M.; Ess, D. H.; Baik, M.-H. Design and Optimization of Catalysts Based on Mechanistic Insights Derived from Quantum Chemical Reaction Modeling. *Chem. Rev.* **2019**, *119*, 6509.
- (2) Coley, C. W. Defining and Exploring Chemical Spaces. *Trends Chem.* **2021**, *3*, 133–145.
- (3) Cheong, P. H.-Y.; Legault, C. Y.; Um, J. M.; Çelebi-Ölçüm, N.; Houk, K. M. Quantum Mechanical Investigations of Organocatalysis: Mechanisms, Reactivities, and Selectivities. *Chem. Rev.* **2011**, *111*, 5042.
- (4) Poree, C.; Schoenebeck, F. A Holy Grail in Chemistry: Computational Catalyst Design: Feasible or Fiction? *Acc. Chem. Res.* **2017**, *50*, 605.

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

(6) Sigman, M. S.; Harper, K. C.; Bess, E. N.; Milo, A. The Development of Multidimensional Analysis Tools for Asymmetric Catalysis and Beyond. *Acc. Chem. Res.* **2016**, *49*, 1292.

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

(8) 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*, 827.

(9) Żurański, A. M.; Alvarado, J. I. M.; Shields, B. J.; Doyle, A. G. Predicting Reaction Yields via Supervised Learning. *Acc. Chem. Res.* **2021**, *54*, 1856.

(10) Wu, K.; Doyle, A. G. Parameterization of phosphine ligands demonstrates enhancement of Ni catalysis via remote steric effects. *Nat. Chem.* **2017**, *9*, 779.

(11) Dotson, J. J.; Dijk, L. V.; 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*, 110.

(12) 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*, No. eaau5631.

(13) Betinol, I. O.; Lai, J.; Thakur, S.; Reid, J. P. A Data-Driven Workflow for Assigning and Predicting Generality in Asymmetric Catalysis. *J. Am. Chem. Soc.* **2023**, *145*, 12870.

(14) Baczewska, P.; Kulczykowski, M.; Zambróń, B.; Jaszczebska-Adamczak, J.; Pakulski, Z.; Roszak, R.; Grzybowski, B. A.; Mlynarski, J. Machine Learning Algorithm Guides Catalyst Choices for Magnesium-Catalyzed Asymmetric Reactions. *Angew. Chem., Int. Ed.* **2024**, *63*, No. e202318487.

(15) Schrader, M. L.; Schäfer, F. R.; Schäfers, F.; Glorius, F. Bridging the information gap in organic chemical reactions. *Nat. Chem.* **2024**, *16*, 491.

(16) Hueffel, J. A.; Sperger, T.; Funes-Ardoiz, I.; Ward, J. S.; Rissanen, K.; Schoenebeck, F. Accelerated dinuclear Pd catalyst identification through unsupervised machine learning. *Science* **2021**, *374*, 1134.

(17) Karl, T. M.; Bouayad-Gervais, S.; Hueffel, J. A.; Sperger, T.; Wellig, S.; Kaldas, S. J.; Dabrankaya, U.; Ward, J. S.; Rissanen, K.; Tizzard, G. J.; Schoenebeck, F. Machine Learning-Guided Development of Trialkylphosphine Ni(I) Dimers and Applications in Site-Selective Catalysis. *J. Am. Chem. Soc.* **2023**, *145*, 15414.

(18) Cuomo, A. E.; Ibarraran, S.; Sreekumar, S.; Li, H.; Eun, J.; Menzel, J. P.; Zhang, P.; Buono, F.; Song, J. J.; Crabtree, R. H.; Batista, V. S.; Newhouse, T. R. Feed-Forward Neural Network for Predicting Enantioselectivity of the Asymmetric Negishi Reaction. *ACS Cent. Sci.* **2023**, *9*, 1768.

(19) Moskal, M.; Beker, W.; Szymkuc, S.; Grzybowski, B. A. Scaffold directed face selectivity machine-learned from vectors of non-covalent interactions. *Angew. Chem., Int. Ed.* **2021**, *60*, 15230.

(20) 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*, 1163.

(21) 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*, 321.

(22) Rosen, B. M.; Quasdorf, K. W.; Wilson, D. A.; Zhang, N.; Resmerita, A.-M.; Garg, N. K.; Percec, V. Ni-Catalyzed cross-couplings Involving Carbon-Oxygen Bonds. *Chem. Rev.* **2011**, *111*, 1346.

(23) Li, W.-B.; Zhang, J.-L. Sadphos as Adaptive Ligands in Asymmetric Palladium Catalysis. *Acc. Chem. Res.* **2024**, *57*, 489.

- (24) Pan, Q.; Wang, K.; Xu, W.-P.; Ai, Y.-Q.; Ping, Y.-Y.; Liu, C.-H.; Wang, M.-Y.; Zhang, J.-L.; Kong, W.-Q. Ligand-Controlled, Ni-Catalyzed Stereodivergent Construction of 1,3-Nonadjacent Stereo-centers. *J. Am. Chem. Soc.* **2024**, *146*, 15453.
- (25) Dai, Q.; Li, W.; Li, Z.; Zhang, J. P-Chiral Phosphines Enabled by Pd/Xiao-Phos Catalyzed Asymmetric P–C cross-coupling of Secondary Phosphine Oxides and Aryl Bromides. *J. Am. Chem. Soc.* **2019**, *141*, 20556.
- (26) Ertl, P.; Schuffenhauer, A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J. Cheminf.* **2009**, *1*, No. 8.
- (27) Shen, D.; Xu, Y.-J.; Shi, S.-L. A Bulky Chiral N-Heterocyclic Carbene Pd Catalyst Enables Highly Enantioselective Suzuki–Miyaura cross-coupling Reactions for the Synthesis of Biaryl Atropisomers. *J. Am. Chem. Soc.* **2019**, *141*, 14938.
- (28) Previous computational studies on Ni(0)-mediated aryl C–O bond activation have also indicated that this process is irreversible. For related references, see: (a) Quasdorf, K. W.; Antoft-Finch, A.; Liu, P.; Silberstein, A. L.; Komaromi, A.; Blackburn, T.; Ramgren, S. D.; Houk, K. N.; Snieckus, V.; Garg, N. K. Suzuki–Miyaura Cross-Coupling of Aryl Carbamates and Sulfamates: Experimental and Computational Studies. *J. Am. Chem. Soc.* **2011**, *133*, 6352. (b) Zhang, S.-Q.; Hong, X. Mechanism and Selectivity Control in Ni- and Pd-Catalyzed Cross-Couplings Involving Carbon–Oxygen Bond Activation. *Acc. Chem. Res.* **2021**, *54*, 2158.
- (29) He, X.-S.; Zhang, S.-S.; Guo, Y.-L.; Wang, H.-Y.; Lin, G.-Q. Mechanistic Investigations of a Palladium-Diene Catalyzed Suzuki–Miyaura Cross-Coupling Reaction. *Organometallics* **2012**, *31*, 2945.
- (30) Shen, X.-Q.; Jones, G. O.; Watson, D. A.; Bhayana, B.; Buchwald, S. L. Enantioselective Synthesis of Axially Chiral Biaryls by the Pd-Catalyzed Suzuki-Miyaura Reaction: Substrate Scope and Quantum Mechanical Investigations. *J. Am. Chem. Soc.* **2010**, *132*, 11278.