

# Ligand Control of Regioselectivity in Palladium-Catalyzed Heteroannulation Reactions of 1,3-Dienes

Dasha Rodina<sup>‡</sup>, Jakub Vaith<sup>‡</sup>, Shauna M. Paradine<sup>\*</sup>

<sup>‡</sup>These authors contributed equally to this work

Department of Chemistry, University of Rochester, Rochester, New York 14627, United States

## ABSTRACT:

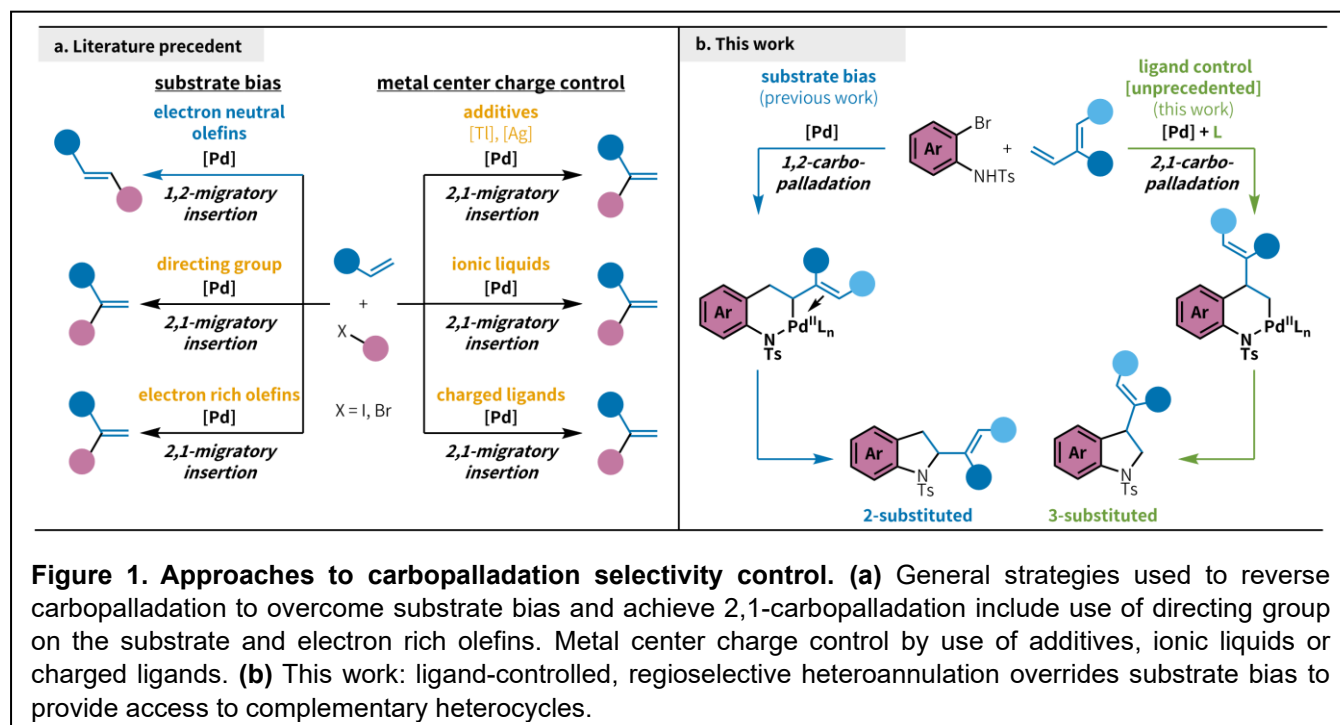
Olefin carbofunctionalization reactions are indispensable tools for constructing diverse, functionalized scaffolds from simple starting materials. However, achieving precise control over regioselectivity in intermolecular reactions remains a formidable challenge. By using  $\text{PAd}_2^{\text{nBu}}$  as a ligand, we have achieved regioselective heteroannulation of *o*-bromoanilines with branched 1,3-dienes through ligand control. This enables regiodivergent access to 3-substituted indolines, showcasing excellent regioselectivity and reactivity across a range of functionalized substrates. To gain further insights into the origin of selectivity control, we employed a data-driven strategy, developing a linear regression model using calculated parameters for phosphorus ligands. This model identified four key parameters governing regioselectivity in this transformation, paving the way for future methodology development. Additionally, density functional theory calculations elucidated key selectivity-determining transition structures along the reaction pathway, corroborating our experimental observations and establishing a solid foundation for future advancements in regioselective olefin difunctionalization reactions.

## Introduction

Olefin carbofunctionalization reactions are powerful scaffold-building transformations that generate diversely functionalized products from simple starting materials.<sup>1,2</sup> These reactions form multiple new C–C and C–heteroatom bonds in a single step, with carbometallation of the olefin being a crucial elementary step that both enables further reactivity and dictates regioselectivity.<sup>3,4</sup> Despite decades of synthetic advances in carbofunctionalization reactions and other related coupling reactions of olefins, controlling regioselectivity of carbometallation in intermolecular reactions remains a significant challenge. Existing strategies for overriding innate substrate bias in reactions involving carbometallation of olefins have primarily relied on substrate control (e.g., directing groups, Figure 1a)<sup>5,6</sup> or the use of stoichiometric additives (Figure 1a).<sup>7–13</sup> Catalyst-mediated control over regioselectivity of olefin addition presents an attractive alternative as it allows for fine-tuning of reaction outcomes, making it a potentially more generalizable approach.

Recently, we reported methods for Pd-catalyzed heteroannulation<sup>14</sup> of 1,3-dienes with *o*-bromoanilines and *o*-bromophenols using ureas as sterically-undemanding pro-ligands to overcome longstanding synthetic limitations in the heteroannulative synthesis of 2-substituted indolines<sup>15</sup> and dihydrobenzofurans.<sup>16</sup> During our ligand structure-reactivity investigations in the indoline synthesis, we observed trace formation of a by-product in our model reaction between *N*-tosyl *o*-bromoaniline (**1a**) and myrcene (**2a**), which we later identified as the 3-substituted indoline. We hypothesized that this regioisomer was formed through a divergent 2,1-carbopalladation step, contrasting with the conventional 1,2-carbopalladation that leads to 2-substituted indoline formation (Figure 1b).<sup>17</sup> Our aim was to identify a ligand with the appropriate steric and electronic properties to selectively promote the 2,1-carbopalladation pathway and provide regiodivergent access to 3-substituted indolines.<sup>18</sup> To accomplish this, we needed to overcome the inherent electronic bias for 1,2-carbopalladation, which situates the Pd in the allylic position.<sup>19</sup>

In this work, we report the first example of ligand-controlled regioselectivity in a 1,3-diene carbofunctionalization reaction. Using PdAd<sub>2</sub><sup>n</sup>Bu (CataCXium A, **L1**) as ligand, we have achieved highly regioselective heteroannulation of *o*-bromoanilines and branched 1,3-dienes, offering regiodivergent access to 3-substituted indolines and delivering excellent regioselectivity and good reactivity for a wide range of functionalized substrates. To further our understanding of selectivity control in this transformation, we employed a data-driven strategy that involved devising a linear regression model capable of predicting regioselectivity in the model reaction, using phosphorus ligand parameters from



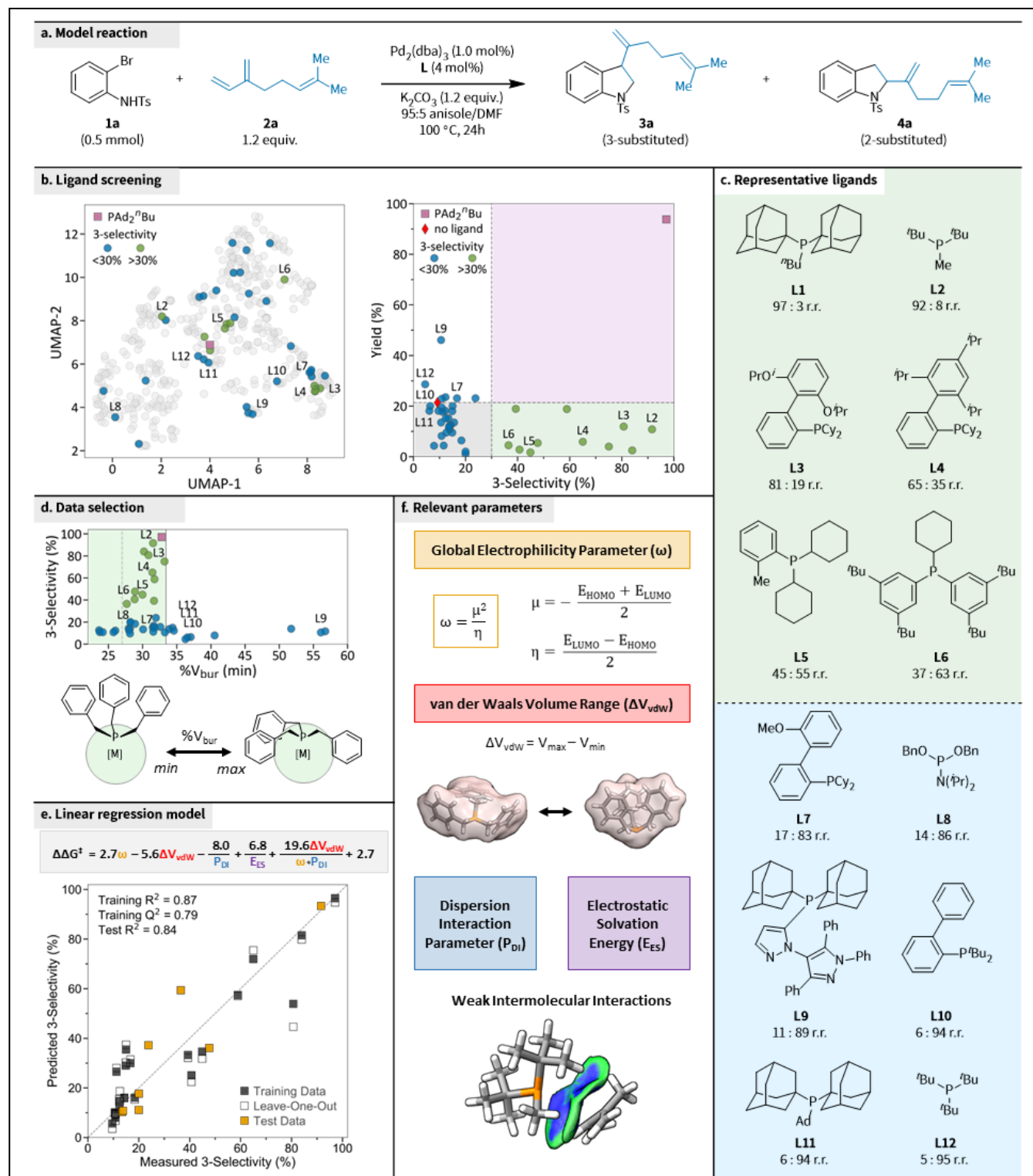
the Kraken database.<sup>20</sup> This predictive model highlighted four crucial parameters that govern regioselectivity control in this transformation. Additionally, we used density functional theory (DFT) calculations to examine the selectivity-determining steps along the reaction pathway with the optimal 3-selective ligand (**L1**). The outcomes of these calculations corroborated our experimental observations; together, these studies substantiated our findings and have establish a foundation for further methodological advancements in regiodivergent olefin difunctionalization reactions.

## Results and discussion

### Ligand structure-reactivity and selectivity studies

To investigate the ligand structure-reactivity and selectivity relationship, we used *N*-tosyl *o*-bromoaniline (**1a**) and myrcene (**2a**) as model substrates and examined the impact of various phosphine ligands on selectivity and reactivity in the heteroannulation reaction (Figure 2a). We began with conditions optimized in our previous paper for urea ligands and the synthesis of 2-substituted indolines. Under these conditions, all initially tested phosphine ligands were inhibitory (yields <32%) compared to the reaction without exogenous ligand (32% yield) and yielded 2-substituted indoline as the primary product. However, with **L1**, we observed an enhancement of the combined yield to 56%, and more importantly, **3a** was the dominant product in a 92:8 regioisomeric ratio (r.r.) (**3a:4a**). We then optimized the initial reactivity and selectivity, allowing us to generate the desired product in high yield (88%) with excellent regioselectivity (>95:5 r.r.); this was accomplished by lowering the reaction temperature from 120 °C to 100 °C and switching the [Pd] precursor from Pd(OAc)<sub>2</sub> to Pd<sub>2</sub>(dba)<sub>3</sub>. We hypothesize that the increased regioselectivity is partly due to the reduction of background reactivity arising from phosphine-free [Pd] when starting with a Pd(II) salt.

Under these optimized conditions, we performed a two-round ligand screening, investigating the effects of 20 ligands in the first round and 22 additional ligands in the second round. The first-round ligands broadly covered the steric-electronic monodentate phosphorus ligand space (Figure 2b). Without an exogenous ligand, the combined product yield in the model reaction was 21%, with the 2-substituted indoline product **4a** being strongly favored (9:91 r.r.). Even in this expanded screen, nearly all ligands were inhibitory compared to the reaction without an exogenous ligand, with **L1** and **L9** being the only ligands that led to a yield above 30%. While most first-round ligands still favored the 2-substituted indoline, twelve ligands (e.g., P<sup>t</sup>Bu<sub>2</sub>Me (**L2**), PCy<sub>3</sub>, **L4**, Figure 2c) shifted the selectivity towards the 3-substituted indoline (>30:70 r.r.). Using the ligand parameters from the Kraken database, we were unable to identify a single parameter that would predict selectivity inversion in the model reaction, though we did observe certain trends. Electron-deficient ligands did not invert selectivity, and



**Figure 2. Ligand structure-reactivity and selectivity studies.** (a) Model reaction and conditions used for the ligand screening. (b) 2D visualization of phosphorus ligand space and reaction selectivity results with yield. (c) Selected examples from the ligand screen with selectivity results. (d) Relationship between selectivity and percent buried volume of the ligands, highlighting the ligand cut off at  $V_{bur}(min)$  33%. (e) Linear regression model reveals four key ligand parameters that influence selectivity. (f) Visual representations of key parameters.

regarding steric parameters such as pyramidalization P or minimum of percent buried volume ( $\%V_{\text{bur}}(\text{min})$ ), the ligands that significantly inverted selectivity were neither small nor large ( $\%V_{\text{bur}}(\text{min})$  between 28 and 33).<sup>21</sup> These trends informed our second round of ligand screening, which identified five additional ligands that inverted selectivity and provided sufficient data points to transition from qualitative to quantitative analysis.

### Ligand Selectivity Analysis through Linear Regression

To further explore the factors governing how ligand structure affected the regioselectivity on the model reaction, multivariate linear regression analysis was performed on our experimental data (Figure 2b), using ligand parameters from the Kraken database. The lack of spread in our product yield data prevented us from performing a comparable analysis of ligand effects on yield. We converted the measured regioselectivity into  $\Delta\Delta G^\ddagger$  using the Gibbs free energy equation ( $\Delta\Delta G^\ddagger = -RT\ln(\text{r.r.})$ ), to model the differential energy between the transition states of the two competing carbopalladation pathways (i.e., 1,2-carbopalladation vs. 2,1-carbopalladation). We organized our results into three classes based on the observed 3-selectivity (<30%, 30-75%, >75%), and each group's data was randomly split into a training and validation set, following a 75:25 partition. We used an exhaustive-search linear regression algorithm to identify initial model candidates, which were evaluated based on their  $R^2$  values and also  $Q^2$  values from Leave-One-Out cross-validation (LOOCV) analysis.

Our first attempt was to apply this search to our entire data set, but we were unable to identify any viable model candidates. This led us to hypothesize the existence of two distinct mechanistic pathways within our data set, each with a different regioselectivity control mechanism. Our hypothesis was supported by the high 2-selectivity observed with the most sterically demanding ligands (e.g., **L10-12**) compared to conditions without exogenous ligands and all other ligands. Specifically, we hypothesized that highly sterically-demanding ligands lead to the formation of an underligated, three-coordinate palladium complex in the regioselectivity determining step, while for the remaining ligands, a four-coordinate palladium species would be expected. Density functional theory (DFT) calculations that support our mechanistic hypothesis are presented in the next section of this paper.

To test our hypothesis, we identified the  $\%V_{\text{bur}}(\text{min})$  parameter as a potential criterion for separating ligands into two classes, as it has been identified as a good predictor of changes in phosphorus ligand ligation states.<sup>21</sup> Our selectivity data, plotted against this parameter (Figure 2d), clearly shows that no 3-selective ligands are found above  $\%V_{\text{bur}}(\text{min}) = 33$ . However, since the 3-selective ligands are all clustered within the intermediate range for  $\%V_{\text{bur}}(\text{min})$ , we decided to examine the impact of eliminating both the largest and smallest ligands ( $\%V_{\text{bur}}(\text{min}) >33$  and  $<28$ , respectively) on the viability of potential

model candidates. Removal of the smallest ligands did not yield any suitable model candidates, but when we excluded ligands with  $\%V_{\text{bur}}(\text{min}) > 33$ , consistent with our hypothesis, we were able to identify several suitable model candidates. The initial exhaustive search was used to identify four-term models dependent on three parameters and a cross-term. The final model (Figure 2e) was then identified by testing addition of a fourth parameter to the selected model candidates. A five-term model using four parameters and a cross-term [coefficient of determination ( $R^2$ ) = 0.87] was identified to describe the observed regioselectivity. Cross-validation techniques, including LOOCV [ $Q^2$  = 0.79] and external validation [test  $R^2$  = 0.84], were performed to indicate that a statistically robust model was produced.

Our model depends on the global electrophilicity parameter ( $\omega$ ), the van der Waals volume range of the ligand ( $\Delta V_{\text{vdW}}$ ), the dispersion interaction parameter ( $P_{\text{DI}}$ ), the electrostatic component of solvation energy ( $E_{\text{ES}}$ ), and a conformation-dispersion-electrophilicity interaction cross-term (Figure 2f). Overall, the model emphasizes the importance of weak noncovalent interactions, such as electrostatic interactions and London dispersion effects, between the ligand and the reagent fragments. The role of these interactions as factors in regioselectivity control in the model heteroannulation was investigated further through computational analysis.

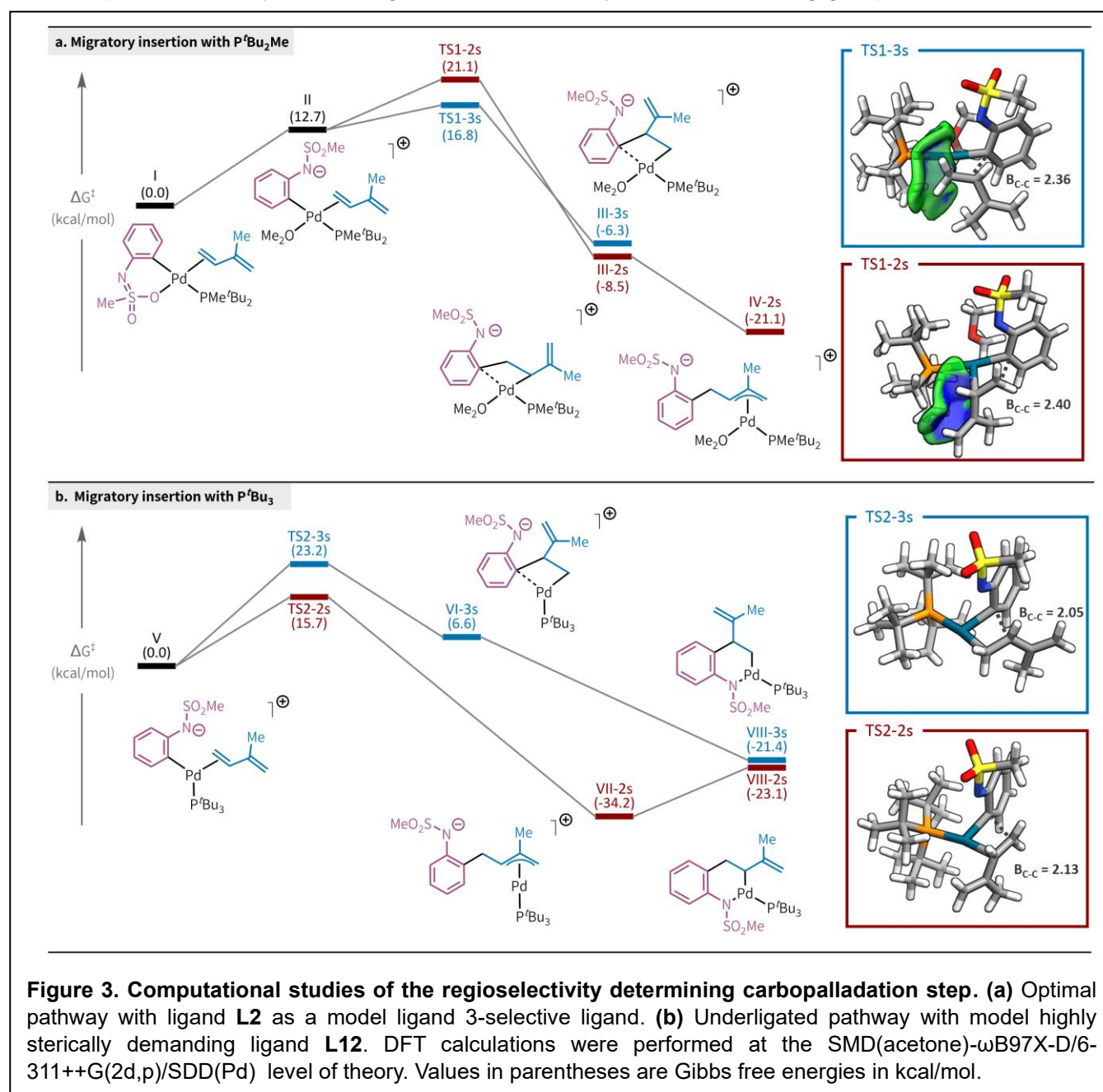
### Computational Analysis of Selectivity-Determining Steps

We turned to DFT calculations to examine the regioselectivity-determining steps along the reaction pathway and to gain insights into how the parameters from the linear regression model influence interactions between the ligand and the reagent fragments in the carbopalladation transition states, ultimately giving rise to regioselectivity in the heteroannulation reaction. For the computational models, we used mesyl-protected *o*-bromoaniline (**1a-Ms**), which still exhibited high regioselectivity in the reaction with **2a** (see Supporting Information for more details), and isoprene (**2b**), which yielded 75% of **4ab** with a >95:5 r.r. Additionally, we used the second most 3-selective ligand,  $P^t\text{Bu}_2\text{Me}$  (**L2**), which is less computationally demanding than our optimal ligand **L1**, but shares a similar structure. Since the focus of this work is on ligand control of regioselectivity in carbopalladation, we restricted our computational studies to the key migratory insertion/carbopalladation step, starting from the oxidative addition complex I after halide displacement (Figure 3a).

After assessing several potential reaction pathways, we identified the "trans-cationic" pathway to be the only one that accurately predicted the observed 3-selectivity in the experimental reaction (see Supporting Information for analysis of other pathways). This pathway commences from the oxidative addition complex following halide displacement (therefore it is "cationic" and has the phosphine ligand and the aryl group in a trans configuration). The "cationic" pathway being operational in our reaction is

supported by the inhibitory effect of halide anion addition (Table S1, entries 19-21), similar to what is observed in “cationic” Heck reactions.<sup>13, 22</sup>

The 2,1- and 1,2-carbopalladation steps in the “trans-cationic” pathway are both moderately exothermic, with the 2-selective  $\sigma$ -complex **III-2s** being slightly favored by 2.2 kcal/mol. While isomerization of  $\sigma$ -complex **III-2s** into the  $\pi$ -allyl complex **IV-2s** provides significant additional stabilization ( $-12.6$  kcal/mol), no such rearrangement is possible for  $\sigma$ -complex **III-3s**. Consequently, the carbopalladation step is likely irreversible, suggesting that product selectivity is under kinetic control. In a computational study of their regioselective Heck arylation, the Mecking group determined that the



cis arrangement of the ligand and diene is essential for effective control of carbopalladation regioselectivity via ligand-alkene steric interactions.<sup>8</sup> The “trans-cationic” pathway, identified as the probable pathway in our reaction, features the phosphine ligand and diene in this beneficial cis configuration, optimally positioning them for steric interactions that can influence regioselectivity. The 3-selective transition state, **TS1-3s**, exhibits weak repulsive ligand-diene interactions and is significantly favored over **TS1-2s** (−4.2 kcal/mol, >99:1 r.r. at 100 °C), which has stronger repulsive interactions, as evaluated using an independent gradient model analysis.<sup>23</sup> These results are consistent with the observed 3-selectivity in our model reaction (92:8 r.r.). These interactions destabilize the conventionally favored 1,2-carbopalladation transition state, thus overriding the inherent substrate bias and allowing for regioselective synthesis of 3-substituted indolines, and are accounted for in our quantitative linear regression model through the dispersion interaction parameter ( $P_{DI}$ ), the electrostatic component of solvation energy ( $E_{ES}$ ), and the conformation-dispersion-electrophilicity interaction cross-term. It is expected that the conformational flexibility of the ligand, approximated in the model through the van der Waals volume range ( $\Delta V_{vdW}$ ), would also impact the strength of these interactions.

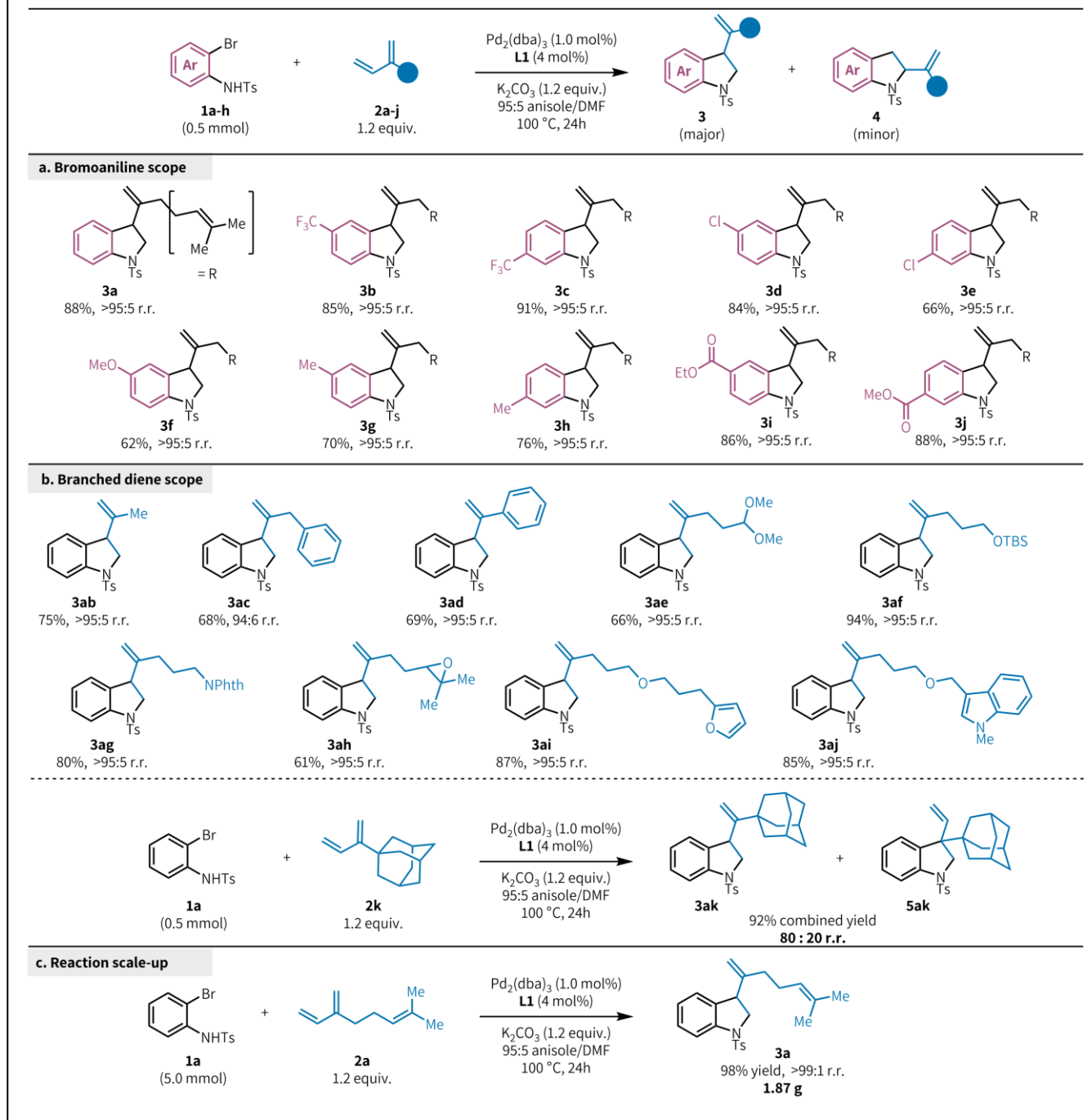
Next, we investigated the carbopalladation step with ligand **L12**, as an example of the more sterically demanding ligands that demonstrated increased 2-selectivity in the model reaction compared to all other ligands and for which we hypothesized an alternative underligated, three-coordinate palladium complex as the active species during the regioselectivity-determining step (Figure 3b). In the transition state optimization calculations of a four-coordinate palladium complex with **L12** for both the “neutral” and “cationic” pathways – where the fourth ligand is either bromide or solvent – no stable stationary point could be localized. These calculations generally resulted in the expulsion of a ligand (diene, halide, or solvent), leading to the formation of an underligated, three-coordinate palladium complex. Starting from the underligated “zwitterionic” complex **V**, the 1,2-carbopalladation pathway is strongly favored (−7.5 kcal/mol) compared to the 2,1-carbopalladation pathway, consistent with the observed 2-selectivity with this ligand in the model reaction (4:96 r.r.). This preference primarily stems from the stabilization of **TS2-2s** through weak coordination of the other olefinic portion of the diene to the underligated palladium center. This pre-coordination leads directly to the formation of the stabilized  $\pi$ -allylpalladium intermediate **VII-2s** during carbopalladation, while no stable  $\sigma$ -complex **VI-2s** could be identified. In the 2,1-carbopalladation pathway, the formation of  $\sigma$ -complex **VI-3s** appears to be endothermic, suggesting potential reversibility. However, there is a strong thermodynamic drive for the formation of chelate complex **VIII-3s**, which can then undergo reductive elimination to produce the 3-substituted indolinic product. Ultimately, the calculated energetics of the carbopalladation pathways involving ligand **L12** support our hypothesis that a distinct mechanism is at play with highly sterically demanding ligands.



## Substrate scope

We explored the generality of our method with various *o*-bromoanilines (Table 1a) and branched 1,3-dienes (Table 1b). Both electron-rich and electron deficient bromoanilines (**3b-3h**) were good substrates under our reaction conditions, producing 3-substituted indolines in 62% to 91% yields with excellent

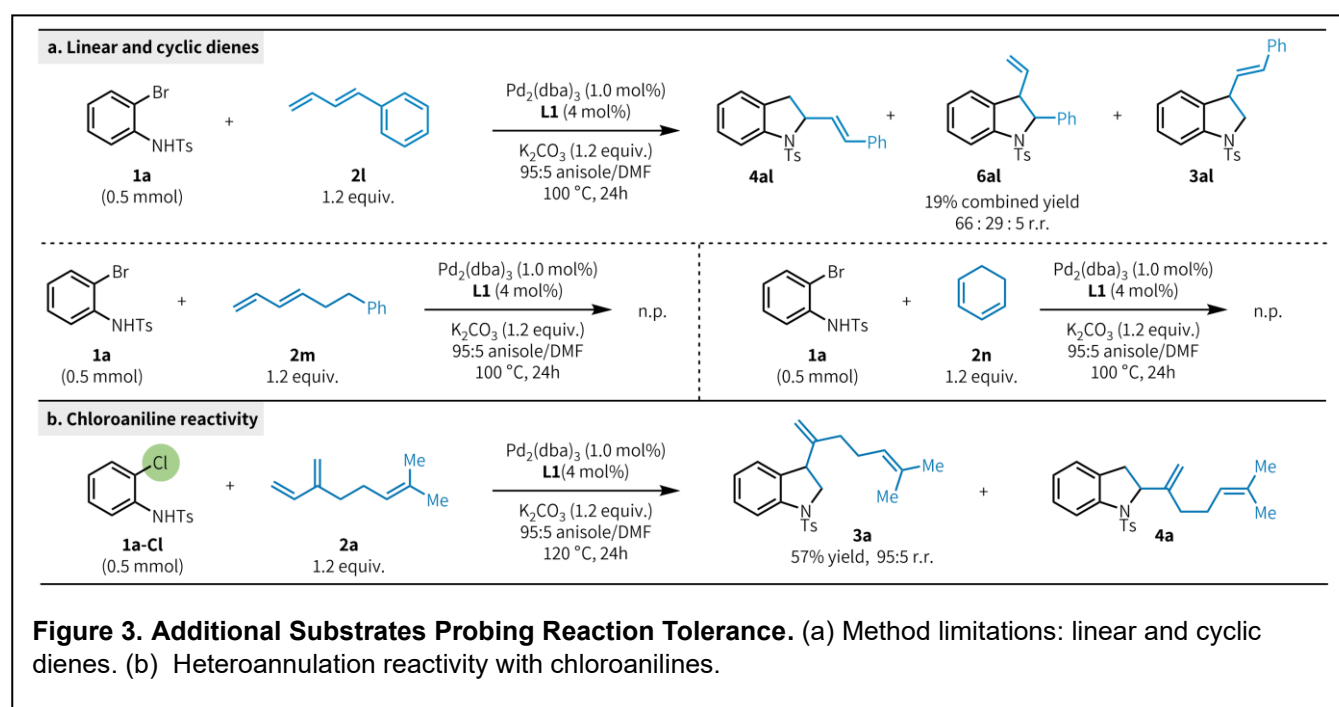
**Table 1. Substrate Scope of *o*-bromoanilines and 1,3-dienes.**



regioselectivity (>95:5 r.r.). Our method was also effective with ester-containing substrates (**3i**, 86% yield and **3j**, 88% yield) that closely resemble bioactive alkaloids such as benzastatins E and G.<sup>24</sup>

Our method accommodates a diverse scope of branched dienes. Both alkyl-substituted dienes (**3ab** and **3ac**) and aryl-substituted dienes (**3ad**) afforded 3-substituted indolines in good yield and excellent regioselectivity (>95:5 r.r.). Furthermore, 3-substituted indolines containing dimethyl acetal **3ae** (66%), silyl-protected alcohol **3af** (94%), and phthalimide-protected amine **3ag** (80%) were readily obtained under our reaction conditions. Branched dienes bearing heterocycles such as epoxides, furans, and indoles, were also suitable substrates (**3ah-3aj**). While examining the tolerance of our system to the steric properties of the diene substrate, we observed excellent yields with adamantyl diene **2k**, yet different selectivity compared to other branched dienes (80:20 r.r.). Interestingly, the minor product **5ak** resulted from the annulation of the internal alkene rather than 1,2-carbopalladation of the terminal alkene. The heteroannulation reaction is readily scaled up 10-fold to afford **3a** in quantitative yields and excellent regioselectivity (Table 1c).

We expanded our investigation to include other related substrates. First, we explored other diene classes (Figure 4a). When we tested a commonly-used linear conjugated diene (**2l**), the selectivity reverted to 2-substituted indoline (**4al**) as the major product and the yield diminished (19% combined yield). Additionally, linear diene **2m** as well as cyclohexadiene (**2n**) yielded no heteroannulation product. Finally, we turned our focus on the ambiphile. By increasing the reaction temperature to 120 °C, we successfully engaged *o*-chloroaniline **1a-Cl** with myrcene **2a**, obtaining indoline **3a** in moderate yield



(57%) and high regioselectivity (Figure 4b). To our knowledge, this represents the first instance of successful heteroannulation of *o*-chloroanilines with dienes,<sup>25</sup> demonstrating the potential for expanding the scope of this synthetic approach to include a wider range of substrate classes.

## Conclusion

Through the exertion of ligand control over carbopalladation, we have successfully achieved a highly regioselective heteroannulation reaction of *o*-bromoanilines and branched 1,3-dienes. This strategy enables the efficient synthesis of 3-substituted indolines and displays exceptional selectivity and reactivity across a diverse range of functionalized substrates. Furthermore, our data-driven approach that employs a linear regression model, using phosphorus ligand parameters from the Kraken database, has provided valuable insights into the key parameters governing selectivity in this transformation. Our findings have highlighted the key role of weak noncovalent interactions between the ligand surface and substrates in influencing reaction outcomes. Additionally, DFT calculations have corroborated our experimental observations and shed light on the selectivity-determining steps in the catalytic cycle. We anticipate that the demonstrated application of ligand control over carbometallation in palladium catalysis, along with predictive models for reaction selectivity, will have a transformative impact on the field of olefin carbofunctionalization.

## References

1. McDonald, R. I.; Liu, G.; Stahl, S. S., Palladium(II)-Catalyzed Alkene Functionalization via Nucleopalladation: Stereochemical Pathways and Enantioselective Catalytic Applications. *Chemical Reviews* **2011**, *111*, 2981-3019.
2. Patel, M.; Desai, B.; Sheth, A.; Dholakiya, B. Z.; Naveen, T. Recent Advances in Mono- and Difunctionalization of Unactivated Olefins. *Asian Journal of Organic Chemistry* **2021**, *10*, 3201-3232.
3. Chen, X.; Xiao, F.; He, W.-M. Recent developments in the difunctionalization of alkenes with C–N bond formation. *Organic Chemistry Frontiers* **2021**, *8*, 5206-5228.
4. Tsuji, J. Pd(0)-Catalyzed Reactions of Allylic Compounds via  $\pi$ -Allylpalladium Complexes. In *Palladium Reagents and Catalysts*, 2004; pp 431-517.
5. Xu, L.; Hilton, M. J.; Zhang, X.; Norrby, P.-O.; Wu, Y.-D.; Sigman, M. S.; Wiest, O. Mechanism, Reactivity, and Selectivity in Palladium-Catalyzed Redox-Relay Heck Arylations of Alkenyl Alcohols. *Journal of the American Chemical Society* **2014**, *136*, 1960-1967.
6. Larock, R. C.; Guo, L. Palladium-catalyzed Annulation of Oxygen-Substituted 1,3-Dienes. *Synlett* **1995**, 465-466.
7. von Schenck, H.; Åkermark, B.; Svensson, M. Electronic Control of the Regiochemistry in the Heck Reaction. *Journal of the American Chemical Society* **2003**, *125*, 3503-3508.
8. Wucher, P.; Caporaso, L.; Roesle, P.; Ragone, F.; Cavallo, L.; Mecking, S.; Göttker-Schnetmann, I. Breaking the regioselectivity rule for acrylate insertion in the Mizoroki–Heck reaction. *Proceedings of the National Academy of Sciences* **2011**, *108*, 8955-8959.
9. Garnes-Portolés, F.; Greco, R.; Oliver-Meseguer, J.; Castellanos-Soriano, J.; Consuelo Jiménez, M.; López-Haro, M.; Hernández-Garrido, J. C.; Boronat, M.; Pérez-Ruiz, R.; Leyva-Pérez, A. Regioirregular and catalytic Mizoroki–Heck reactions. *Nature Catalysis* **2021**, *4*, 293-303.

10. Zheng, C.; Wang, D.; Stahl, S. S. Catalyst-Controlled Regioselectivity in the Synthesis of Branched Conjugated Dienes via Aerobic Oxidative Heck Reactions. *Journal of the American Chemical Society* **2012**, *134*, 16496-16499.
11. Beccaria, L.; Deagostino, A.; Prandi, C.; Zavattaro, C.; Venturello, P. Heck Reaction on 1-Alkoxy-1,3-dienes in Ionic Liquids: A Superior Medium for the Regioselective Arylation of the Conjugated Dienic System. *Synlett* **2006**, 2989-2992.
12. Qin, L.; Hirao, H.; Zhou, J. Regioselective Heck reaction of aliphatic olefins and aryl halides. *Chemical Communications* **2013**, *49*, 10236-10238.
13. Cabri, W.; Candiani, I.; Bedeschi, A.; Penco, S.; Santi, R.  $\alpha$ -Regioselectivity in palladium-catalyzed arylation of acyclic enol ethers. *The Journal of Organic Chemistry* **1992**, *57*, 1481-1486.
14. Ni, H.-Q.; Cooper, P.; Engle, K. M. Recent advances in palladium-catalyzed (hetero)annulation of C–C bonds with ambiphilic organo(pseudo)halides. *Chemical Communications* **2021**, *57*, 7610-7624.
15. Vaith, J.; Rodina, D.; Spaulding, G. C.; Paradine, S. M. Pd-Catalyzed Heteroannulation Using N-Arylureas as a Sterically Undemanding Ligand Platform. *Journal of the American Chemical Society* **2022**, *144*, 6667-6673.
16. Houghtling, K. E.; Canfield, A. M.; Paradine, S. M. Convergent Synthesis of Dihydrobenzofurans via Urea Ligand-Enabled Heteroannulation of 2-Bromophenols with 1,3-Dienes. *Organic Letters* **2022**, *24*, 5787-5790.
17. Beletskaya, I. P.; Cheprakov, A. V. The Heck Reaction as a Sharpening Stone of Palladium Catalysis. *Chemical Reviews* **2000**, *100*, 3009-3066.
18. Tasker, S. Z.; Jamison, T. F. Highly Regioselective Indoline Synthesis under Nickel/Photoredox Dual Catalysis. *Journal of the American Chemical Society* **2015**, *137*, 9531-9534.
19. Trost, B. M.; Van Vranken, D. L. Asymmetric Transition Metal-Catalyzed Allylic Alkylations. *Chemical Reviews* **1996**, *96*, 395-422.
20. Gensch, T.; dos Passos Gomes, G.; 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. *Journal of the American Chemical Society* **2022**, *144*, 1205-1217.
21. 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*, 301-308.
22. Xu, L.; Chen, W.; Ross, J.; Xiao, J. Palladium-Catalyzed Regioselective Arylation of an Electron-Rich Olefin by Aryl Halides in Ionic Liquids. *Organic Letters* **2001**, *3*, 295-297.
23. Lefebvre, C.; Khartabil, H.; Boisson, J.-C.; Contreras-García, J.; Piquemal, J.-P.; Hénon, E. The Independent Gradient Model: A New Approach for Probing Strong and Weak Interactions in Molecules from Wave Function Calculations. *ChemPhysChem* **2018**, *19*, 724-735.
24. Kim, W.-G.; Kim, J.-P.; Koshino, H.; Shin-Ya, K.; Seto, H.; Yoo, I.-D. Benzastatins E, F, and G: New indoline alkaloids with neuronal cell protecting activity from *Streptomyces nitrosporeus*. *Tetrahedron* **1997**, *53*, 4309-4316.
25. Thansandote, P.; Hulcoop, D. G.; Langer, M.; Lautens, M. Palladium-Catalyzed Annulation of Haloanilines and Halobenzamides Using Norbornadiene as an Acetylene Synthon: A Route to Functionalized Indolines, Isoquinolinones, and Indoles. *The Journal of Organic Chemistry* **2009**, *74*, 1673-1678.

## Methods

### General procedure for ligand-controlled heteroannulative synthesis of 3-substituted indolines

*N*-tosyl-bromoaniline **1a** (163 mg, 0.500 mmol, 1.0 equiv.), myrcene **2a** (103  $\mu$ L, 0.600 mmol, 1.2 equiv.), di(1-adamantyl)-*n*-butylphosphine (7.2 mg, 0.020 mmol, 0.04 equiv.), potassium carbonate (83 mg, 0.600 mmol, 1.2 equiv.), and tris(dibenzylideneacetone)dipalladium(0) (4.6 mg, 0.005 mmol, 0.010 equiv.) were weighed out in the above-mentioned order into a 1-dram vial equipped with a stir bar and a cap with a silicone septum. The vial was then placed under a nitrogen atmosphere and charged with 1 mL of freshly degassed anisole/dimethylformamide (95:5) solvent mixture. The reaction was stirred at 100 °C for 24 hours. After cooling to room temperature, regioisomeric ratios were determined by HPLC analysis of the crude reaction mixture and the reaction mixture was filtered with ethyl acetate through cotton plug. The solvents were removed under reduced pressure and the crude mixture was purified by flash column chromatography to obtain products **3**.

### Data availability

Crystallographic data are available free of charge from the Cambridge Crystallographic Data Centre under references CCDC 2263163. Additional experimental details, experimental procedures, crystal structures, compound characterization, NMR spectra for all new compounds, ligand data sets, computational details, and Python scripts are available in the Supplementary Information.

### Acknowledgements

The authors thank the University of Rochester and the National Science Foundation (CHE 22-38081) for financial support. D.R. is supported by the NIH T32 CBI Training Program (T32-GM118283), and J.V. is supported by Bristol-Myers-Squibb (Graduate Fellowship in Synthetic Organic Chemistry). We thank Amanda Canfield for checking the experimental procedure for preparation of compound **3j** in Table 1a, Dr. William Brennessel for conducting X-ray crystallography, and Natalie Seeger and Prof. Matthew Sigman for their assistance with the Kraken database.

### Author contributions

D.R. and J.V. conducted all the experimental work, analyzed the data, and completed the DFT calculations. J.V. developed the Python script and performed the regression analysis. J.V. conceived of the project, and D.R., J.V., and S.M.P. designed the project and wrote the manuscript. S.M.P. directed the research.

### Competing interests

The authors declare no competing interests.