A general reactivity map for predicting outcomes in palladium-catalyzed cross-coupling

Jingru Lu, Sofia Donnecke, Irina Paci,[*] David C. Leitch[*]

Department of Chemistry, University of Victoria, 3800 Finnerty Rd. Victoria BC, CANADA, V8P 5C2.

*ipaci@uvic.ca; dcleitch@uvic.ca.

ABSTRACT: Making accurate, quantitative predictions of chemical reactivity based on molecular structure is an unsolved problem in chemical synthesis, particularly for complex molecules. We report a generally applicable and mechanistically based structure-reactivity model for the oxidative addition of (hetero)aryl halides to palladium(0), which is a key step in myriad catalytic processes. This model links simple molecular descriptors – electrostatic potential, steric parameters, and bond strengths – to relative rates of oxidative addition for 71 electrophile substrates. Because oxidative addition often controls the rate and/or selectivity of palladium-catalyzed cross-coupling reactions, this model can be used to make predictions about catalytic reaction outcomes. Demonstrated applications include a quantitative model for the rate of 410 different Sonogashira coupling reactions, and successful site-selectivity predictions for a series of multihalogenated substrates relevant to the synthesis of pharmaceuticals and natural products.

The synthesis of structurally complex organic molecules relies on forging new chemical bonds between diverse molecular building blocks. Palladium-catalyzed cross-coupling is one of the most versatile and widely-used methods to link these molecular fragments,[1] with applications ranging from the manufacture of active pharmaceutical ingredients,[2] to the selective modification of biomolecules,[3] to the creation of new functional materials.[4] While much has been done to develop and understand new cross-coupling reactions and catalysts, less is known about how the specific molecular structures of the building blocks affect the likelihood of successful coupling. As a result, time- and resource-intensive reaction screening and optimization campaigns are often required for each new synthetic target. All too often these efforts fail to identify suitable conditions, impeding access to potentially promising new medicines and materials. Emerging approaches in reactivity prediction that combine high-throughput experimentation[5,6] with molecular descriptor sets[7,8] and multivariate statistical analysis including machine learning[9,10] can accelerate this process and increase success rates; however, the predictions generated by these approaches are rarely applicable beyond the specific reaction under investigation. Developing and refining the next generation of organic chemistry tools, including computer-aided synthesis design, automated reaction optimization, and predictive algorithms,[11] requires a general and quantitative framework that links molecular structure to reactivity for many different reactants and catalysts.

Here, we describe a general approach to predict catalytic reaction outcomes that focuses on mapping the reactivity of a key organometallic transformation common to myriad catalytic mechanisms; in this case, oxidative addition. Oxidative addition, the reaction between an electrophilic substrate and an electron-rich metal center,[12] is particularly relevant to palladium-catalyzed cross-coupling, where it often controls the reaction rate and/or selectivity[13] (Figure 1A).

We hypothesized that a quantitative structure-reactivity model for the oxidative addition of aryl halides to a catalytically-relevant low-valent palladium complex would enable predictions that are generally applicable to many palladium-catalyzed transformations under a variety of conditions. We have assembled this model by combining experimental relative rate data (Figure 1B) with easily-obtained molecular descriptors for a diverse set of substrates, with an emphasis on incorporating pharmaceutically-relevant heterocycles (Figure 1C). As hypothesized, the resulting model – a "reactivity map" – is general: it can predict catalytic rate constants as a function of substrate structure even when different catalysts/solvents are used, and can correctly identify the most reactive position(s) when multiple reaction sites are available. Importantly from both a fundamental and practical perspective, the predictive ability of this reactivity map extends well beyond the specific molecular structures and reaction conditions included in the initial data set. By revealing how subtle changes to the reacting molecules affect a key step in a catalytic mechanism, this map serves as a powerful predictive tool that can be successfully applied to multiple reaction classes, and will enable more sophisticated and accurate computer-aided retrosynthetic design.

To construct our reactivity map for cross-coupling catalysis, we conducted a series of oxidative addition competition experiments in THF and toluene solvent using a library of 71 (hetero)aryl halides (Figure 1B). By combining excess but equal amounts of each substrate with $Pd(PCy_3)_2$, we measured the oxidative addition product ratio by ^{31}P NMR spectroscopy, thus determining relative observed rate constants. As a check on these kinetic data, we constructed Hammett plots for four sets of substrates, obtaining reaction constants consistent with previous reports (Figures S32-37). Using the specific mechanistic features of oxidative addition to Pd(0) as a guide (*vide infra*), we considered potential molecular descriptors that would provide mechanistically meaningful correlations between substrate structures and oxidative addition reactivity (Figure 1C).[14] This led us to a combination of average molecular electrostatic potentials (*ESP*) as electronic descriptors for specific atoms in the substrate,[15] A-values as steric descriptors,[16] and the intrinsic bond strength index (*IBSI*) as a bond energy descriptor.[17] Importantly, these descriptors are easily calculated from electronic wavefunctions (using density functional theory), or are tabulated in the literature.

Multivariate linear regression analysis of the descriptor sets for each substrate versus the relative free energies of activation for oxidative addition ($\Delta G^{\ddagger}_{OA}$, kJ mol^{-1}, derived from the observed relative rate constants) reveals a strong correlation across the entire substrate library. Figure 1D shows the correlation from the THF data set, and the corresponding correlation for the toluene data set is in Figure S60. This model incorporates two *ESP* values: one for the carbon undergoing substitution (ESP_1), and a second for an atom adjacent to the reactive site (ESP_2). Similarly, two substituent A-values for groups R_1 and R_2 account for steric effects on the oxidative addition rate. While the steric effect of groups *ortho* to the reactive C–X bond is intuitive, the effect of R_2 for 2-halopyridine substrates is not initially obvious; however, our experimental results reveal the steric influence of R_2 is approximately equal to that of R_1 throughout the 2-halopyridine substrate set. As these two terms have very similar coefficients when treated separately during linear regression (Table S9), we opted to use the sum ($A_1 + A_2$) as a single descriptor. Finally, the *IBSI* enables a unified model that is suitable for both (hetero)aryl chlorides and bromides.
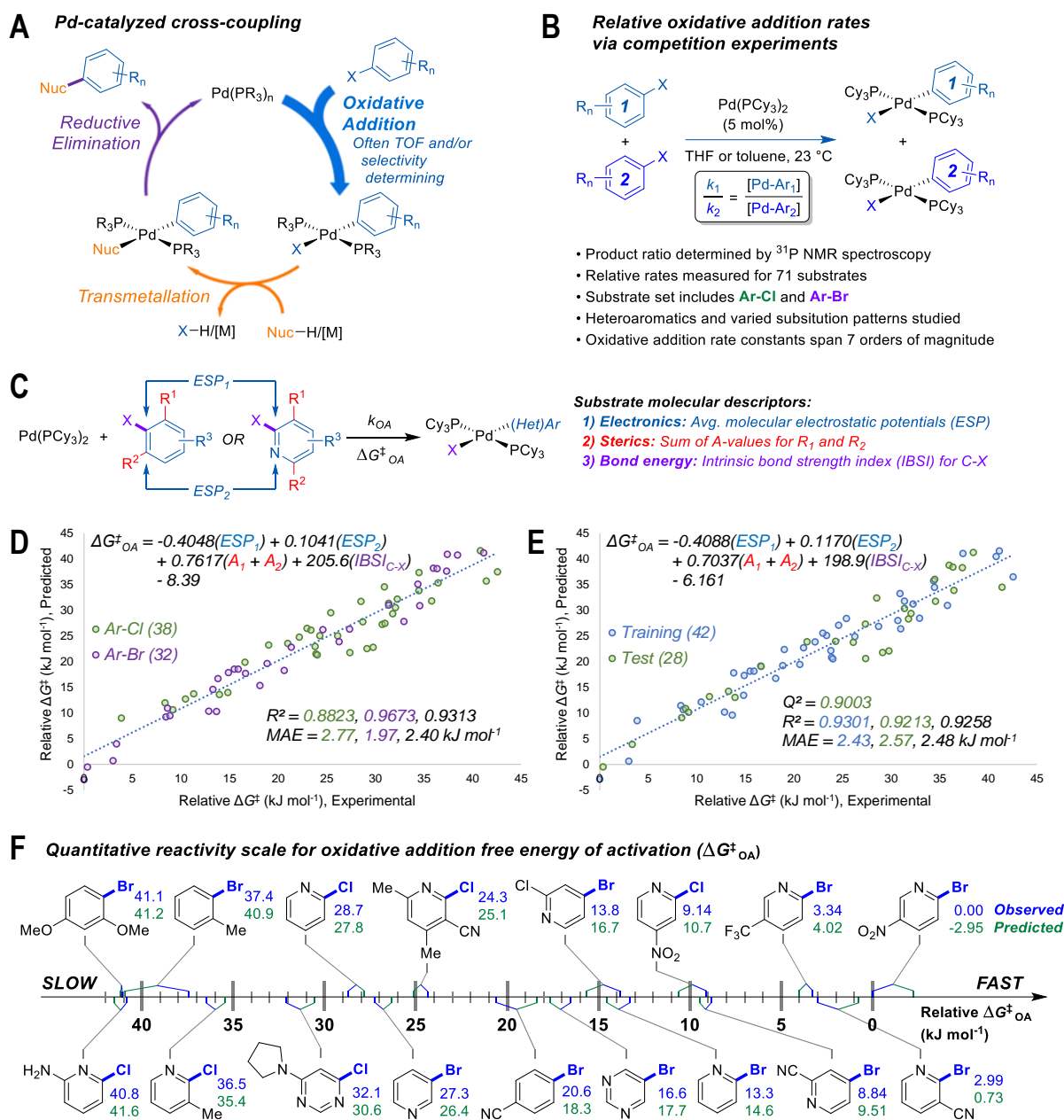
Figure 1. Design and performance of a general and quantitative reactivity map for oxidative addition to Pd(0). (**A**) Simplified mechanism for Pd-catalyzed cross-coupling, highlighting oxidative addition as the rate and/or selectivity determining step. (**B**) Competition experiment approach to map relative rates of oxidative addition. (**C**) Molecular descriptors used to model oxidative addition reactivity as a function of substrate structure. (**D**) Multivariate linear regression model of $\Delta G^{\ddagger}_{OA}$ for 70 substrates in THF, including all data points in regression analysis. (**E**) Representative multivariate linear regression model generated using a 60/40 training/test split. (**F**) Reactivity scale for oxidative addition to Pd(PCy$_3$)$_2$ in THF with selected substrates, giving experimental (blue) and predicted (green) $\Delta G^{\ddagger}_{OA}$ values; experimental $\Delta G^{\ddagger}_{OA}$ for 2-bromo-5-nitropyridine set to 0 kJ mol^{-1}.

We have evaluated the robustness of this linear model by regression analysis of five random training (60%) and test (40%) sets, and comparing the mean absolute errors (MAE) and predictive squared correlation coefficients (Q^2);[18] one example is shown in Figure 1E. All of these analyses give similar linear equations, and excellent agreement between predicted and experimental values in the test sets (Figures S61-64, Q^2 range: 0.8681-0.9178; test set mean absolute error range: 2.57-2.91 kJ mol^{-1}). We also partitioned the data into targeted training and test sets to evaluate out-of-sample prediction accuracy (Figures S65-68), again giving excellent agreement between predicted and experimental values (Q^2 range: 0.8480-0.9391; test set mean absolute error range: 2.30-2.92 kJ mol^{-1}). Furthermore, the linear regression model generated from oxidative addition reactions in toluene is very similar to the THF-based model, and the experimentally-determined $\Delta G^{\ddagger}_{OA}$ in both solvents have a strong linear correlation with one another (Figure S6). These observations indicate that the use of THF or toluene has minimal effect on relative oxidative addition rates within our data set. Our quantitative structure-reactivity map thus serves as an intrinsic oxidative addition reactivity scale for (hetero)aryl chlorides and bromides, accounting for sterics, electronics, and the strength of the C–X bond, with accurate predictions possible across a range of rates that spans more than 7 orders of magnitude (Figure 1F).

The generality and predictive power of this reactivity map is a direct result of its mechanistic foundation. Mechanisms for aryl halide oxidative addition to Pd(0) have been extensively studied both computationally[19–22] and experimentally,[23–27] and are generally proposed to proceed via initial coordination of the aromatic π-system to Pd (Figure 2A). Two bonding extremes can be envisioned for the π-complex intermediate, where the degree of polarization of the coordinated C=C or C=N bond in the substrate influences partial charge distribution in the π-complex. From this intermediate, two types of oxidative addition transition state have been proposed: a 3-centered, relatively non-polar transition state involving simultaneous Pd–C and Pd–X bond formation, and a polarized transition state with C–X heterolytic bond cleavage occurring alongside Pd–C bond formation; this latter pathway resembles the proposed mechanism for nucleophilic aromatic substitution (S$_N$Ar).[24,27]

The specific molecular descriptors in our model provide mechanistic insights into several aspects of oxidative addition reactivity. ESP_1 is related to the electrophilicity of the carbon undergoing oxidative addition, reflecting the degree of partial positive charge in the π-complex intermediate and transition state: a larger positive ESP_1 leads to a smaller $\Delta G^{\ddagger}_{OA}$ and thus faster oxidative addition. In contrast, ESP_2 reflects the degree of negative charge on the adjacent atom (C or N), where a larger negative ESP_2 leads to a faster oxidative addition. Considered together, these ESP terms indicate that a more polarized C=C or C=N bond in the substrate leads to faster oxidative addition. It follows from this analysis that the ESP at Pd in the intermediate and/or transition state should also affect the oxidative addition rate, with a larger positive ESP_{Pd} correlating with faster oxidative addition. We have confirmed this by determining ESP_{Pd} for a set of 11 calculated π-complex intermediates and 6 transition states (Figures S40-56). Remarkably, there is a linear correlation between ESP_{Pd} for the π-complexes and $\Delta G^{\ddagger}_{OA}$ (Figures 2B, S57), revealing the significant effect that substrate-catalyst bonding has on the electronic structure of Pd. While the influence of ancillary ligands is often invoked to explain organometallic reactivity, the substrate itself clearly has a profound impact on the reactivity of the metal center.
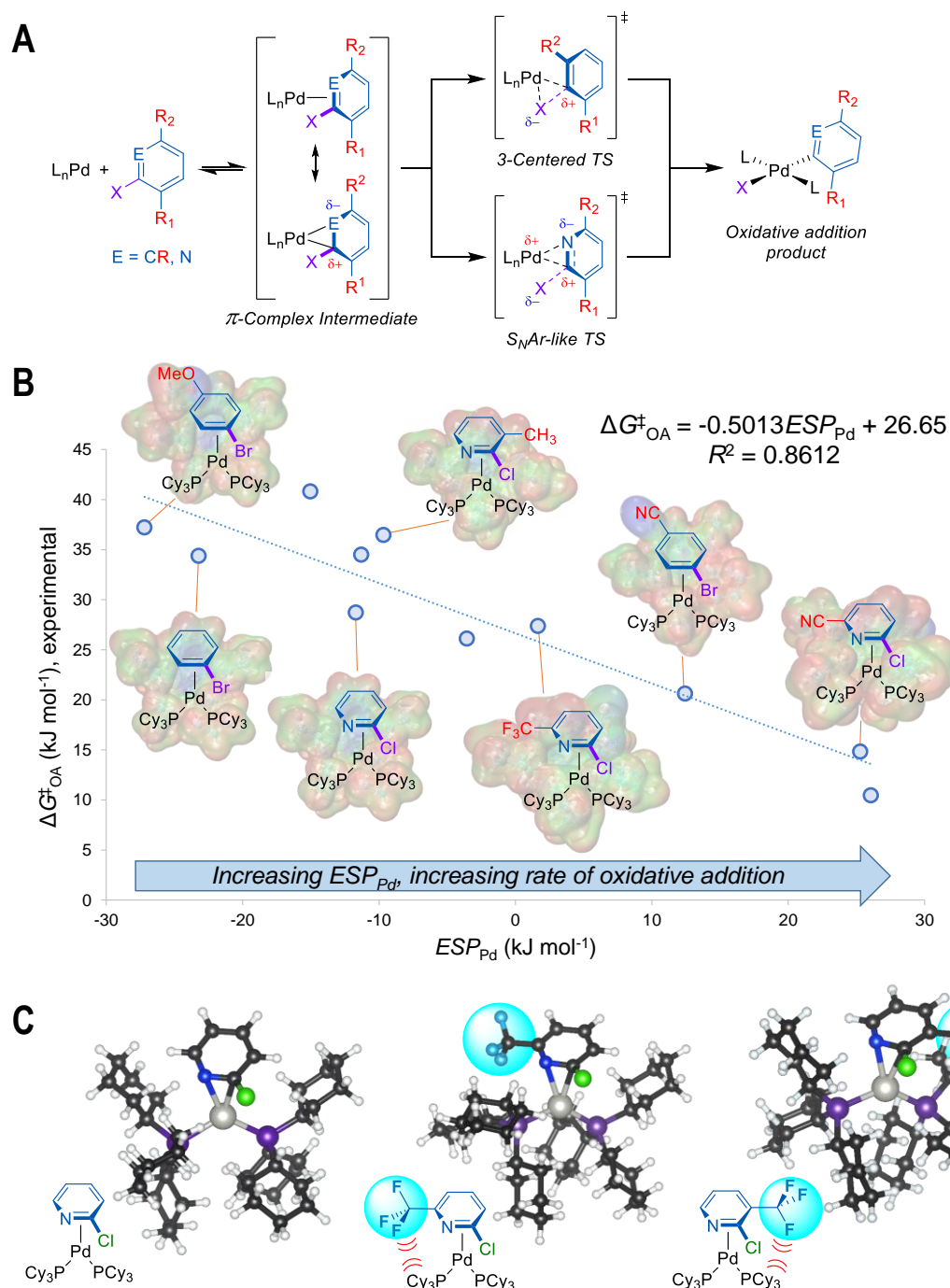
Figure 2. Mechanistic aspects of oxidative addition are revealed by reactivity mapping with molecular descriptors. (**A**) General mechanism for oxidative addition to $L_nPd(0)$, with π-complex intermediate preceding either Pd insertion into C–X bond, or an S_NAr-like displacement of X^-. (**B**) ESP_{Pd} for calculated π-complex intermediate structures correlates with oxidative addition rates; structures for 7 of 11 examples shown. (**C**) Calculated structures of π-complex intermediates reveal how steric strain induced by R^1 and R^2 (here, –CF_3 groups) in 2-halopyridines affect oxidative addition reactivity in equal proportions.

Analyzing how the substrate binds to the catalyst also sheds light on the observed equal weights of the steric A-values for R^1 and R^2 in affecting $\Delta G^{\ddagger}_{OA}$ for the 2-halopyridine substrate series. Larger substituents in these positions destabilize the π-complex intermediate and oxidative addition transition state through steric repulsion between the bound substrate and the ancillary phosphine ligands. Because the Pd center coordinates to the C=N bond, substituents adjacent to either C or N will occupy roughly equivalent positions with respect to the phosphines. Comparing the calculated structures of the π-complex intermediates for 2-chloropyridine to its 3- and 6-trifluoromethyl substituted analogues illustrates this feature of the substrate-catalyst interaction (Figure 2C). Importantly, this mechanistic insight into equal steric effects for R^1 and R^2 flows directly from our quantitative reactivity map.

To test our hypothesis that this oxidative addition reactivity map will be generally applicable in cross-coupling catalysis, we applied our $\Delta G^{\ddagger}_{OA}$ predictions to two case studies. The first test case involves modeling the initial rates of Sonogashira coupling reactions, which are drawn from published data sets containing 410 individual rates (29 substrates and 17 catalysts, Figure 3A).[28,29] We first predicted $\Delta G^{\ddagger}_{OA}$ for each of the 29 aryl bromides based on the equation from Figure 1D; while 9 of these substrates are included in our experimental oxidative addition data set, the other 20 are out-of-sample predictions. Remarkably, these predicted $\Delta G^{\ddagger}_{OA}$ values are linearly-correlated with the corresponding ln k values for all 17 investigated phosphine ligands (Figure S69-70); three of these – for P(nBu)$_3$, PCy$_3$, and P(tBu)$_3$ – are shown in Figure 3B. These correlations hold despite the fact that the Sonogashira reactions are conducted under different conditions (higher temperature, different solvent) than our oxidative addition experiments, and the fact that our predicted $\Delta G^{\ddagger}_{OA}$ is derived from a model using only PCy$_3$. Thus, $\Delta G^{\ddagger}_{OA}$ can be applied to quantitatively predict the outcome of catalytic reactions for out-of-sample substrates, out-of-sample reaction conditions, and even out-of-sample catalysts.

To further test the generality of $\Delta G^{\ddagger}_{OA}$ in predicting reaction rates, we assembled a single and unified linear model that accurately predicts ln k for the entire 410 Sonogashira reaction data set. Combining $\Delta G^{\ddagger}_{OA}$ with two calculated descriptors for the free phosphines – the average *ESP* at phosphorus, and the percent buried volume (%V$_{burr}$) at phosphorus[30] – we obtain excellent linear fit and predictive power with initial rates spanning 10 orders of magnitude (Figure 3C). These phosphine descriptors outperform the analogous descriptors calculated for the corresponding mono or bis(phosphine) Pd(0) complexes (Figure S71-72). The training and test sets used to build this model are from a random 60/40 split of substrate set #1, which focuses on electronic effects (Fig 3A). To challenge the model, we reserved substrate set #2, which focuses on steric effects, as an external validation set. Despite the fact that the training set contains no substrates with *ortho*-substituents, and therefore no information about steric effects on reaction rate, the model is still able to predict ln k for substrate set #2 with a mean absolute error of 0.729 ($Q^2 = 0.8009$). Only two reactions are identified as significant outliers (MAE > 2): 2,4,6-triisopropylphenylbromide with P(tBu)$_3$ and PAd$_2$(tBu), representing the most sterically-hindered substrate with the two largest ligands. The robustness of this model in the face of significantly different out-of-sample predictions stems directly from the generality of the reactivity map. The substrate molecular properties – electronics, sterics, and C–X bond strengths – are all encoded within the predicted $\Delta G^{\ddagger}_{OA}$ values by virtue of our diverse oxidative addition training set.

Improved prediction accuracy can be achieved by breaking the 410-member data set into two subsets based on phosphine ligand size. One set includes reactions using 13 phosphines with a %V$_{burr}$ < 75, and the other includes the 4 largest phosphines (%V$_{burr}$ > 75). For the small

phosphine set, a slight reweighting of the three descriptors leads to smaller MAEs for training, test, and external data sets (Figure 3D). For the large phosphines, we used a 70/30 training/test split due to the smaller number of data points. We also determined that a simpler, two-descriptor model is superior, with only $\Delta G^{\ddagger}_{OA}$ and $\%V_{burr}$ needed to make accurate predictions (Figure 3D). This partition into small and large phosphine sets is consistent with prior studies that describe a change in mechanism, where oxidative addition occurs from either bis(phosphine) or mono(phosphine) Pd(0) intermediates.[31,32] Notably, the predicted $\Delta G^{\ddagger}_{OA}$ is effective in both cases.
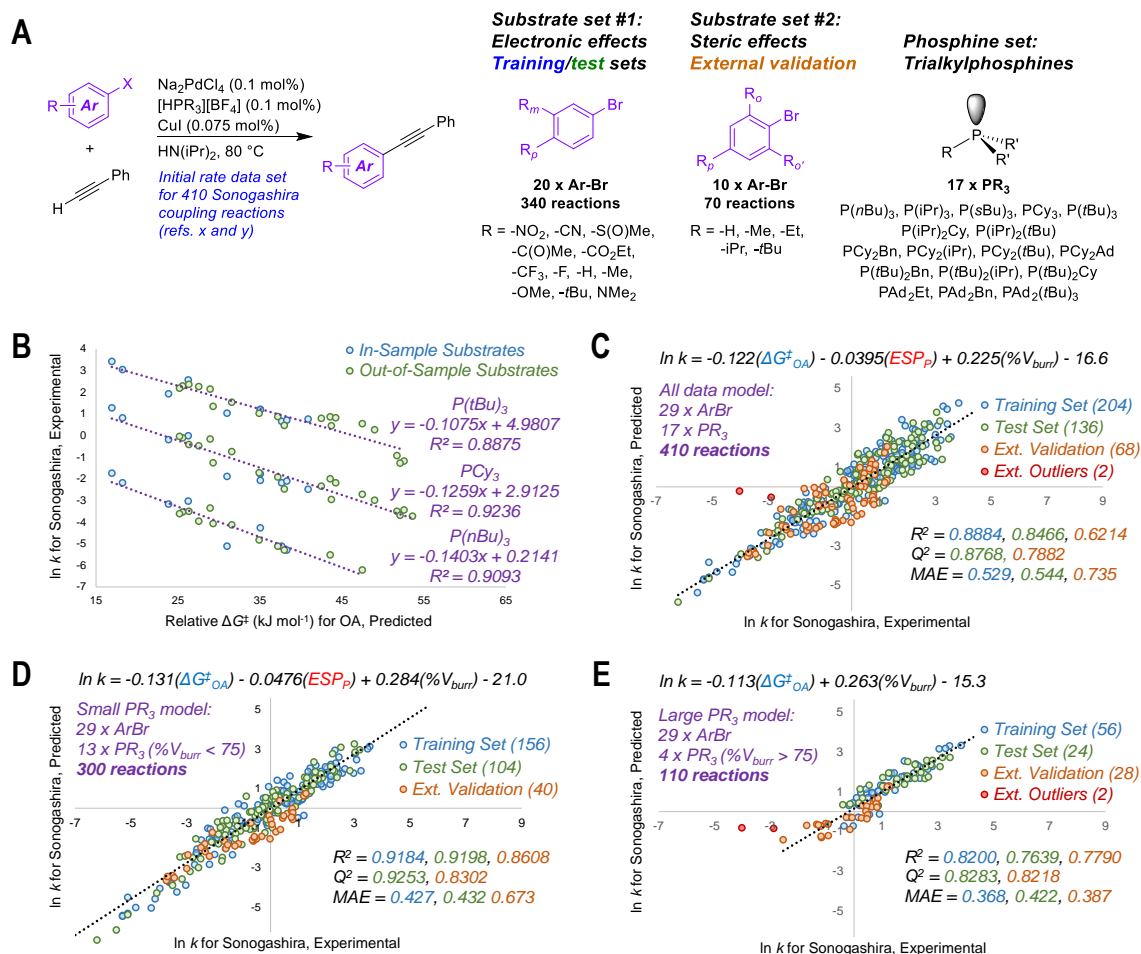

Figure 3. Translating oxidative addition predictions to successful quantitative models of catalytic reactivity. (**A**) General reaction scheme and chemical space explored for 410 Sonogashira reactions, with two distinct substrate sets; initial rates determined previously [28,29]. (**B**) Univariate linear correlations between predicted $\Delta G^{\ddagger}_{OA}$ for oxidative addition to Pd(PCy$_3$)$_2$ and ln k for Sonogashira coupling with three phosphine ligands; out-of-model substrates represent ArBr molecules not included in oxidative addition training set. (**C**) Unified three-descriptor model for predicting Sonogashira ln k for the entire set of 410 reactions (29 substrates, 17 ligands), with data partitioned into training (60% of substrate set #1), test (40% of substrate set #1), and external validation (substrate set #2). (**D**) Subset of the model focusing on 13 "small" phosphines with $\%V_{burr} < 75$, with three descriptors. (**E**) Subset of the model focusing on 4 "large" phosphines with $\%V_{burr} > 75$, with only two descriptors required.
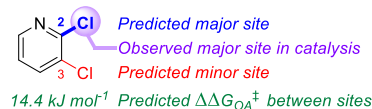
7

Our second case study tests the use of $\Delta G^{\ddagger}_{OA}$ in predicting site selectivity for cross-coupling when multiple reactive positions are present (Figure 4). Previous approaches to this problem include NMR spectroscopic analysis,[33] the distortion-interaction transition state model,[34] and qualitative arguments based on empirical observations;[13,35] however, none of these methods is able to rapidly and accurately predict site selectivity for a new substrate based only on its structure. In our case, each of the descriptors in the $\Delta G^{\ddagger}_{OA}$ prediction model are easy to obtain for each halogenated site within a molecule; therefore, one can determine the most reactive site in a substrate (or combination of substrates) simply by finding the lowest predicted $\Delta G^{\ddagger}_{OA}$. Assessing a series of multihalogenated heterocycles with reported experimental selectivities[35] reveals that these $\Delta G^{\ddagger}_{OA}$ predictions are correct across a diverse range of substrates, including many heterocycle classes (isoquinolines, diazines, 5-membered rings) not included in our initial training set (Figure 4A). These predictions also correctly identify when a C–Cl position is more reactive than a C–Br position in the same molecule, and vice versa.

Because these predictions are based on relative $\Delta G^{\ddagger}_{OA}$, they are also quantitative. A large $\Delta\Delta G^{\ddagger}_{OA}$ indicates very high selectivity for one site over another, while a small $\Delta\Delta G^{\ddagger}_{OA}$ indicates poor or potentially tunable selectivity. Three specific examples of this are given in Figure 4B. For Suzuki-Miyaura coupling of methyl 2,6-dichloronicotinate, our model predicts $\Delta\Delta G^{\ddagger}_{OA}$ = 5.6 kJ mol^{-1}, favoring C6 by ~10:1 at room temperature. The observed selectivity using Pd(PPh$_3$)$_4$ as a catalyst at 65 °C is 5:1 C6 to C3, corresponding to $\Delta\Delta G^{\ddagger}$ = 4.5 kJ mol^{-1}.[36] For 3,6-dichloro-4-methoxypyridazine, our model predicts $\Delta\Delta G^{\ddagger}_{OA}$ = 1.0 kJ mol^{-1}, favoring C6 by only ~3:1 (though this difference is smaller than the model MAE). The observed selectivity using Pd(PPh$_3$)$_4$ as a catalyst at 100 °C is 3:1 C6 to C3, corresponding to $\Delta\Delta G^{\ddagger}$ = 3.4 kJ mol^{-1}.[37] Finally, the predicted $\Delta\Delta G^{\ddagger}_{OA}$ for 2-chloro-5-bromopyridine is 0.2 kJ mol^{-1}, indicating effectively no intrinsic selectivity. We confirmed this by carrying out an oxidative addition reaction between this substrate and Pd(PCy$_3$)$_2$, which gives a 1:1 mixture of C2 and C5-derived Pd(II) products. Notably, previous studies of Buchwald-Hartwig amination site selectivity with this substrate reveal ligand-controlled reactivity at either C2 or C5.[38,39] The realization of tunable selectivity in this case is likely aided by the similar reactivity of these two positions, as revealed by our reactivity map.
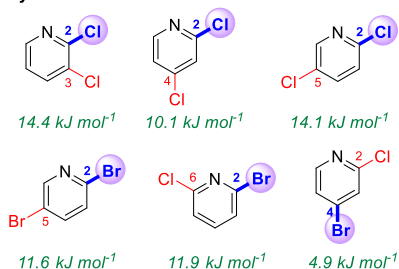
As an additional demonstration of how $\Delta G^{\ddagger}_{OA}$ could be used in synthetic planning, we have applied these predictions to the reported synthesis of an advanced intermediate toward Dragmacidin D (Figure 4C).[40] This sequence relies on two regioselective cross-coupling reactions when multiple Ar–Br positions are present. Predicted $\Delta G^{\ddagger}_{OA}$ values for each of these sites is consistent with the observed selectivity. Thus, applying $\Delta G^{\ddagger}_{OA}$ predictions to hypothetical synthetic sequences could be used to plan a selective route and/or prioritize different potential routes prior to commencing experimental investigations.

As we expand this quantitative reactivity map to incorporate additional sets of reaction conditions and catalysts, we anticipate it will not only continue to shed new light on the mechanistic aspects of cross-coupling, but also find widespread use in refining and augmenting computer-aided synthesis design and automated reaction discovery/optimization. We also anticipate that this approach of building predictive reactivity models based on individual steps in catalytic mechanisms will lead to more accurate and more general predictions that can be applied to many catalytic transformations.
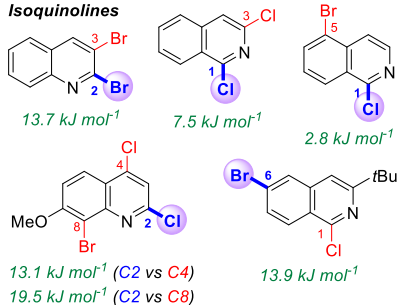
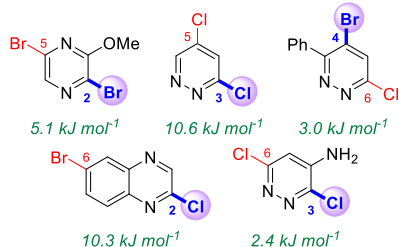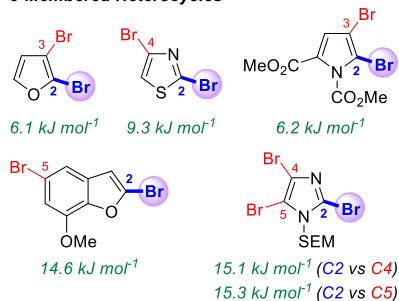**A** — *Site selectivity in cross-coupling of multihalogenated heterocycles*

Predicted major site
Observed major site in catalysis
Predicted minor site

14.4 kJ mol⁻¹ *Predicted ΔΔG$_{OA}^{‡}$ between sites*

Pyridines

14.4 kJ mol⁻¹   10.1 kJ mol⁻¹   14.1 kJ mol⁻¹

11.6 kJ mol⁻¹   11.9 kJ mol⁻¹   4.9 kJ mol⁻¹

Quinolines and Isoquinolines

13.7 kJ mol⁻¹   7.5 kJ mol⁻¹   2.8 kJ mol⁻¹

13.1 kJ mol⁻¹ (C2 vs C4)
19.5 kJ mol⁻¹ (C2 vs C8)

13.9 kJ mol⁻¹

Diazines

5.1 kJ mol⁻¹   10.6 kJ mol⁻¹   3.0 kJ mol⁻¹

10.3 kJ mol⁻¹   2.4 kJ mol⁻¹

5-Membered Heterocycles

6.1 kJ mol⁻¹   9.3 kJ mol⁻¹   6.2 kJ mol⁻¹

14.6 kJ mol⁻¹   15.1 kJ mol⁻¹ (C2 vs C4)
15.3 kJ mol⁻¹ (C2 vs C5)

B

Predicted: Moderate selectivity
5.6 kJ mol⁻¹

PhB(OH)$_2$ (1.05 equiv)
Pd(PPh$_3$)$_4$ (5 mol%)
K$_2$CO$_3$ (3 equiv)
MeOH, reflux
ref. 36

Observed:
C6:C2 = 5:1 (4.5 kJ mol⁻¹ at 65 °C)

Predicted: Poor selectivity
1.0 kJ mol⁻¹
(C6 favored, within error)

ArB(OH)$_2$ (1.05 equiv)
Pd(PPh$_3$)$_4$ (5 mol%)
Na$_2$CO$_3$ (2 equiv)
toluene/EtOH/H$_2$O, 100 °C
Ar = 4-MeOPh
ref. 37

Observed:
C6:C3 ~3:1 (3.4 kJ mol⁻¹ at 100 °C)

60%   20%
(+ 14% diarylated)

Predicted: Poor selectivity
0.2 kJ mol⁻¹

Note: observe ~1:1 C2:C5 oxidative addition with Pd(PCy$_3$)$_2$

N-Boc-piperazine
Pd$_2$(dba)$_3$ (2 mol%)
ligand (6 mol%)
NaOtBu (1.5 equiv)
toluene, 100 °C
ref. 38

OR
PhNH$_2$ (1 equiv)
Pd(dba)$_2$ (5 mol%)
ligand (5-10 mol%)
NaOtBu (1.25 equiv)
dioxane, 80 °C
ref. 39

Observed:
Selectivity is tunable based on ligand

Ligand:	C2 : C5
IPr•HCl	1 : 10
IMes•HCl	3.3 : 1
dppf	1 : 1
dtbpf	8.5 : 1
dmapf	>99 : 1

C

8.0 kJ mol⁻¹
2.9 kJ mol⁻¹
45.7 kJ mol⁻¹

Pd(PPh$_3$)$_4$ (10 mol%)
Na$_2$CO$_3$ (2 M in water)
benzene/MeOH, reflux
ref. 40, 52% yield

21.0 kJ mol⁻¹
40.0 kJ mol⁻¹

TBAF
95% yield

19.3 kJ mol⁻¹
37.9 kJ mol⁻¹
46.0 kJ mol⁻¹

Core structure of Dragmacidin D

Pd(PPh$_3$)$_4$ (13 mol%)
CuI (20 mol%)
DMF, 80 °C
ref. 40, 92% yield

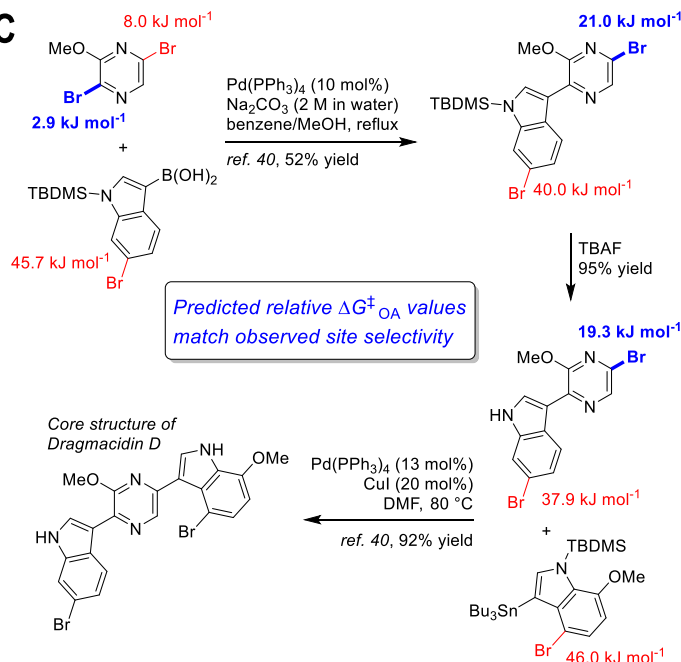*Predicted relative ΔG$^{‡}_{OA}$ values match observed site selectivity*

Figure 4. Site selectivity predictions for cross-coupling reactions. (**A**) Predicted and observed selectivities for multihalogenated heterocycles undergoing Suzuki-Miyaura cross-coupling reactions. (**B**) Selectivity predictions for dihalogenated heterocycles with small ΔΔ$G^{‡}_{OA}$ between the two sites, and observed product ratios consistent with predictions. (**C**) Application of selectivity predictions to reported synthesis of the core structure of Dragmacidin D.[40]

References:

(1)   Campeau, L.-C.; Hazari, N. Cross-Coupling and Related Reactions: Connecting Past Success to the Development of New Reactions for the Future. *Organometallics* **2019**, *38*, 3–35. https://doi.org/10.1021/acs.organomet.8b00720.

(2)   Magano, J.; Dunetz, J. R. Large-Scale Applications of Transition Metal-Catalyzed Couplings for the Synthesis of Pharmaceuticals. *Chem. Rev.* **2011**, *111*, 2177–2250. https://doi.org/10.1021/cr100346g.

(3)   Vinogradova, E. V.; Zhang, C.; Spokoyny, A. M.; Pentelute, B. L.; Buchwald, S. L. Organometallic Palladium Reagents for Cysteine Bioconjugation. *Nature* **2015**, *526*, 687–691. https://doi.org/10.1038/nature15739.

(4)   Wang, K.; Osuka, A.; Song, J. Pd-Catalyzed Cross Coupling Strategy for Functional Porphyrin Arrays. *ACS Cent. Sci.* **2020**, *6*, 2159–2178. https://doi.org/10.1021/acscentsci.0c01300.

(5)   Mennen, S. M.; Alhambra, C.; Allen, C. L.; Barberis, M.; Berritt, S.; Brandt, T. A.; Campbell, A. D.; Castañón, J.; Cherney, A. H.; Christensen, M.; Damon, D. B.; Eugenio de Diego, J.; García-Cerrada, S.; García-Losada, P.; Haro, R.; Janey, J.; Leitch, D. C.; Li, L.; Liu, F.; Lobben, P. C.; MacMillan, D. W. C.; Magano, J.; McInturff, E.; Monfette, S.; Post, R. J.; Schultz, D.; Sitter, B. J.; Stevens, J. M.; Strambeanu, I. I.; Twilton, J.; Wang, K.; Zajac, M. A. The Evolution of High-Throughput Experimentation in Pharmaceutical Development and Perspectives on the Future. *Org. Process Res. Dev.* **2019**, *23*, 1213–1242. https://doi.org/10.1021/acs.oprd.9b00140.

(6)   Allen, C. L.; Leitch, D. C.; Anson, M. S.; Zajac, M. A. The Power and Accessibility of High-Throughput Methods for Catalysis Research. *Nature Catal.* **2019**, *2*, 2–4. https://doi.org/10.1038/s41929-018-0220-4.

(7)   Durand, D. J.; Fey, N. Computational Ligand Descriptors for Catalyst Design. *Chem. Rev.* **2019**, *119*, 6561–6594. https://doi.org/10.1021/acs.chemrev.8b00588.

(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*, 827–836. https://doi.org/10.1021/acs.accounts.0c00745.

(9)   Ahneman, D. T.; Estrada, J. G.; Lin, S.; Dreher, S. D.; Doyle, A. G. Predicting Reaction Performance in C–N Cross-Coupling Using Machine Learning. *Science* **2018**, *360*, 186–190. https://doi.org/10.1126/science.aar5169.

(10)   Santiago, C. B.; Guo, J.-Y.; Sigman, M. S. Predictive and Mechanistic Multivariate Linear Regression Models for Reaction Development. *Chem. Sci.* **2018**, *9*, 2398–2412. https://doi.org/10.1039/C7SC04679K.

(11) Shen, Y.; Borowski, J. E.; Hardy, M. A.; Sarpong, R.; Doyle, A. G.; Cernak, T. Automation and Computer-Assisted Planning for Chemical Synthesis. *Nature Rev. Methods Primers* **2021**, *1*, 1–23. https://doi.org/10.1038/s43586-021-00022-5.

(12) Labinger, J. A. Tutorial on Oxidative Addition. *Organometallics* **2015**, *34*, 4784–4795. https://doi.org/10.1021/acs.organomet.5b00565.

(13) Reeves, E. K.; Entz, E. D.; Neufeldt, S. R. Chemodivergence between Electrophiles in Cross-Coupling Reactions. *Chem. Eur. J.* **2021**, *27*, 6161–6177. https://doi.org/10.1002/chem.202004437.

(14) Sandford, C.; Fries, L. R.; Ball, T. E.; Minteer, S. D.; Sigman, M. S. Mechanistic Studies into the Oxidative Addition of Co(I) Complexes: Combining Electroanalytical Techniques with Parameterization. *J. Am. Chem. Soc.* **2019**, *141*, 18877–18889. https://doi.org/10.1021/jacs.9b10771.

(15) A. Grimmel, S.; Reiher, M. The Electrostatic Potential as a Descriptor for the Protonation Propensity in Automated Exploration of Reaction Mechanisms. *Faraday Discuss.* **2019**, *220*, 443–463. https://doi.org/10.1039/C9FD00061E.

(16) Hirsch, J. A. Table of Conformational Energies—1967. In *Topics in Stereochemistry*; John Wiley & Sons, Ltd, 1967; pp 199–222. https://doi.org/10.1002/9780470147108.ch4.

(17) Klein, J.; Khartabil, H.; Boisson, J.-C.; Contreras-García, J.; Piquemal, J.-P.; Hénon, E. New Way for Probing Bond Strength. *J. Phys. Chem. A* **2020**, *124*, 1850–1860. https://doi.org/10.1021/acs.jpca.9b09845.

(18) Consonni, V.; Ballabio, D.; Todeschini, R. Comments on the Definition of the Q^2 Parameter for QSAR Validation. *J. Chem. Inf. Model.* **2009**, *49*, 1669–1678. https://doi.org/10.1021/ci900115y.

(19) Senn, H. M.; Ziegler, T. Oxidative Addition of Aryl Halides to Palladium(0) Complexes: A Density-Functional Study Including Solvation. *Organometallics* **2004**, *23*, 2980–2988. https://doi.org/10.1021/om049963n.

(20) Ahlquist, M.; Norrby, P.-O. Oxidative Addition of Aryl Chlorides to Monoligated Palladium(0): A DFT-SCRF Study. *Organometallics* **2007**, *26*, 550–553. https://doi.org/10.1021/om0604932.

(21) McMullin, C. L.; Jover, J.; Harvey, J. N.; Fey, N. Accurate Modelling of Pd(0) + PhX Oxidative Addition Kinetics. *Dalton Trans.* **2010**, *39*, 10833–10836. https://doi.org/10.1039/C0DT00778A.

(22) Kashihara, M.; Gordon, C. P.; Copéret, C. Reactivity of Substituted Benzenes toward Oxidative Addition Relates to NMR Chemical Shift of the Ipso-Carbon. *Org. Lett.* **2020**, *22*, 8910–8915. https://doi.org/10.1021/acs.orglett.0c03300.

(23) Amatore, C.; Jutand, A.; Khalil, F.; M'Barki, M. A.; Mottier, L. Rates and Mechanisms of Oxidative Addition to Zerovalent Palladium Complexes Generated *in Situ* from Mixtures of Pd^0(dba)$_2$ and Triphenylphosphine. *Organometallics* **1993**, *12*, 3168–3178. https://doi.org/10.1021/om00032a045.

(24) Portnoy, M.; Milstein, D. Mechanism of Aryl Chloride Oxidative Addition to Chelated Palladium(0) Complexes. *Organometallics* **1993**, *12*, 1665–1673. https://doi.org/10.1021/om00029a026.

(25) Barrios-Landeros, F.; Carrow, B. P.; Hartwig, J. F. Effect of Ligand Steric Properties and Halide Identity on the Mechanism for Oxidative Addition of Haloarenes to Trialkylphosphine Pd(0) Complexes. *J. Am. Chem. Soc.* **2009**, *131*, 8141–8154. https://doi.org/10.1021/ja900798s.

(26) Mitchell, E. A.; Jessop, P. G.; Baird, M. C. A Kinetics Study of the Oxidative Addition of Bromobenzene to Pd(PCy$_3$)$_2$ (Cy = Cyclohexyl) in a Nonpolar Medium: The Influence on Rates of Added PCy$_3$ and Bromide Ion. *Organometallics* **2009**, *28*, 6732–6738. https://doi.org/10.1021/om900679w.

(27) Maes, B. U. W.; Verbeeck, S.; Verhelst, T.; Ekomié, A.; von Wolff, N.; Lefèvre, G.; Mitchell, E. A.; Jutand, A. Oxidative Addition of Haloheteroarenes to Palladium(0): Concerted versus S$_N$Ar-Type Mechanism. *Chem. Eur. J.* **2015**, *21*, 7858–7865. https://doi.org/10.1002/chem.201406210.

(28) an der Heiden, M. R.; Plenio, H. The Effect of Steric Bulk in Sonogashira Coupling Reactions. *Chem. Commun.* **2007**, 972–974. https://doi.org/10.1039/B616608C.

(29) an der Heiden, M. R.; Plenio, H.; Immel, S.; Burello, E.; Rothenberg, G.; Hoefsloot, H. C. J. Insights into Sonogashira Cross-Coupling by High-Throughput Kinetics and Descriptor Modeling. *Chem. Eur. J.* **2008**, *14*, 2857–2866. https://doi.org/10.1002/chem.200701418.

(30) Clavier, H.; Nolan, S. P. Percent Buried Volume for Phosphine and N-Heterocyclic Carbene Ligands: Steric Properties in Organometallic Chemistry. *Chem. Commun.* **2010**, *46*, 841–861. https://doi.org/10.1039/B922984A.

(31) Schoenebeck, F.; Houk, K. N. Ligand-Controlled Regioselectivity in Palladium-Catalyzed Cross Coupling Reactions. *J. Am. Chem. Soc.* **2010**, *132*, 2496–2497. https://doi.org/10.1021/ja9077528.

(32) Niemeyer, Z. L.; Milo, A.; Hickey, D. P.; Sigman, M. S. Parameterization of Phosphine Ligands Reveals Mechanistic Pathways and Predicts Reaction Outcomes. *Nature Chem.* **2016**, *8*, 610–617. https://doi.org/10.1038/nchem.2501.

(33) Handy, S. T.; Zhang, Y. A Simple Guide for Predicting Regioselectivity in the Coupling of Polyhaloheteroaromatics. *Chem. Commun.* **2006**, 299–301. https://doi.org/10.1039/B512948F.

(34) Legault, C. Y.; Garcia, Y.; Merlic, C. A.; Houk, K. N. Origin of Regioselectivity in Palladium-Catalyzed Cross-Coupling Reactions of Polyhalogenated Heterocycles. *J. Am. Chem. Soc.* **2007**, *129*, 12664–12665. https://doi.org/10.1021/ja075785o.

(35) Almond-Thynne, J.; C. Blakemore, D.; C. Pryde, D.; C. Spivey, A. Site-Selective Suzuki–Miyaura Coupling of Heteroaryl Halides – Understanding the Trends for Pharmaceutically Important Classes. *Chem. Sci.* **2017**, *8*, 40–62. https://doi.org/10.1039/C6SC02118B.

(36) Yang, W.; Wang, Y.; Corte, J. R. Efficient Synthesis of 2-Aryl-6-Chloronicotinamides via PXPd2-Catalyzed Regioselective Suzuki Coupling. *Org. Lett.* **2003**, *5*, 3131–3134. https://doi.org/10.1021/ol035188g.

(37) Blaise, E.; Kümmerle, A. E.; Hammoud, H.; de Araújo-Júnior, J. X.; Bihel, F.; Bourguignon, J.-J.; Schmitt, M. Access to 4-Alkylaminopyridazine Derivatives via Nitrogen-Assisted Regioselective Pd-Catalyzed Reactions. *J. Org. Chem.* **2014**, *79*, 10311–10322. https://doi.org/10.1021/jo501930s.

(38) Ji, J.; Li, T.; Bunnelle, W. H. Selective Amination of Polyhalopyridines Catalyzed by a Palladium−Xantphos Complex. *Org. Lett.* **2003**, *5*, 4611–4614. https://doi.org/10.1021/ol0357696.

(39) Keylor, M. H.; Niemeyer, Z. L.; Sigman, M. S.; Tan, K. L. Inverting Conventional Chemoselectivity in Pd-Catalyzed Amine Arylations with Multiply Halogenated Pyridines. *J. Am. Chem. Soc.* **2017**, *139*, 10613–10616. https://doi.org/10.1021/jacs.7b05409.

(40) Yang, C.-G.; Liu, G.; Jiang, B. Preparing Functional Bis(Indole) Pyrazine by Stepwise Cross-Coupling Reactions: An Efficient Method to Construct the Skeleton of Dragmacidin D. *J. Org. Chem.* **2002**, *67*, 9392–9396. https://doi.org/10.1021/jo026450m.

Author contributions: *Jingru Lu*: Conceptualization, Methodology, Investigation, Validation, Formal analysis, Writing. *Sofia Donnecke*: Investigation, Formal analysis, Writing. *Irina Paci*: Conceptualization, Methodology, Supervision, Writing. *David C. Leitch*: Conceptualization, Methodology, Formal analysis, Supervision, Writing.

Competing interests: Authors declare no competing interests.

Data and materials availability: All data is available in the main text or the supplementary materials.