Quantitative Reactivity Models for Oxidative Addition to $L_2Pd(0)$: Additional Substrate Classes, Solvents, and Mechanistic Insights

Jingru Lu,[a] Holly Celuszak,[a] Irina Paci,*[a] and David C. Leitch*[a]

[a] Dr. J. Lu, H. Celuszak, Prof. Dr. I. Paci, Prof. Dr. D. C. Leitch
Department of Chemistry
University of Victoria
3800 Finnerty Rd., Victoria, BC V8P 5C2 (Canada)
E-mail: ipaci@uvic.ca; dcleitch@uvic.ca

Abstract: Quantitative molecular structure-reactivity models are useful for generating predictions to guide synthesis design, and in formulating and testing mechanistic hypotheses. We report an expanded multivariate linear regression (MLR) model for the rate of (hetero)aryl (*pseudo*)halide oxidative addition to $L_2Pd(0)$, here exemplified by $Pd(PCy_3)_2$. This builds on a prior model from our group, with additional substrate classes (aryl chlorides and iodides) and reaction solvents (THF, toluene, THF/DMF mixture). Overall solvent effects across the entire substrate set are minimal under these conditions, enabling a unified MLR model without introduction of new molecular descriptors beyond the original five. Examining the mechanistic origin of the two molecular electrostatic potential (*ESP*) descriptors led to generation of a simpler, four descriptor model that is suitable for aryl halides, but not for 2-halopyridines. Using this model we identified a mechanistic outlier, 2-pyridyl triflate, which undergoes a nucleophilic displacement oxidative addition that does not involve the adjacent nitrogen atom. Finally, we discuss the relationship between C–X bond strength and oxidative addition rates, and compare the intrinsic bond strength index (*IBSI*) to bond dissociation enthalpy (*BDE*) as a bond strength descriptor.

Introduction

Oxidative addition (OA) of carbon-heteroatom bonds to low-valent transition metal centers is one of the fundamental reactions in organometallic chemistry and catalysis.[1–3] The specific OA reaction of (hetero)aryl (*pseudo*)halides to group 10 metal centers – predominantly Pd(0) – is a key step in many catalytic cross-coupling reactions.[4–7] Crucial aspects of the reaction outcome (*e.g.* rate and selectivity) are often governed by the OA step;[8] as a result, many researchers have conducted mechanistic investigations of OA reactions to Pd(0) complexes computationally and/or experimentally.[9–38] Recent work indicates that OA to Pd(0) can proceed by different mechanisms depending on the specific reaction conditions (catalyst, substrate, solvent, *etc.*).[39–47] This rich mechanistic diversity makes it challenging to accurately prediction reaction outcomes for a given system.[48,49]

As one approach to this prediction problem, in 2022 we published a quantitative structure-reactivity model for predicting oxidative addition rates/selectivities for various substrate classes to $Pd(PCy_3)_2$.[50] This multivariate linear regression model correlates a small number of calculated substrate molecular descriptors to measured rates of oxidative addition collected *via* competition experiments (represented as $\Delta G^{\ddagger}$, free energies of activation) (Fig. 1). We further demonstrated how to use this model to make accurate predictions of both reaction rates and site-selectivity for various Pd-catalyzed reactions. In addition to their use in reactivity prediction/synthesis planning, quantitative structure-reactivity relationships are also useful tools in mechanistic elucidation.[51,52] We recently reported our mechanistic findings on substrates that are outliers to the initial model, revealing how molecular orbital symmetry can control which mechanism operates.[53]

Our initial multivariate model was built using 79 (Het)Ar–X substrates drawn from chloro/bromo pyridines, aryl bromides and aryl triflates, and used a single solvent (THF). According to several mechanistic studies,[10,11,33,34,43,47] there are two common mechanisms for (Het)Ar–X oxidative addition to $L_2Pd(0)$: the 'traditional' 3-centered type, and the nucleophilic displacement type (Fig. 1A). Using these mechanisms as a guide, we identified five molecular descriptors that correlate well with oxidative addition rates: molecular electrostatic potential at the reactive carbon and an adjacent atom (ESP_1 and ESP_2 respectively),[54–57] A-values to account for sterics of adjacent substituents,[58] the intrinsic bond strength index (*IBSI*),[59] and the pK_a value for the leaving group conjugate acid in DCE. This gives a quantitative multivariate linear regression (MLR) model of OA reactivity as a function of substrate structure (Fig. 1B).

Herein, we describe an expansion of this multivariate linear model, along with a discussion of specific mechanistic insights gleaned therefrom. We incorporate two new substrate classes into the model – aryl chlorides and (hetero)aryl iodides – without additional feature selection. We also collected rate data in two additional solvent systems – toluene and 1:1 THF/DMF – to assess solvent impacts on rates for different substrate types. Finally, we use the model to probe the interplay between molecular descriptors and mechanistic aspects of the oxidative addition reaction, including 2-pyridyl triflate as a mechanistic outlier, and correlations between C–X bond strength and OA rate. In a follow-up study, we further probe additional outlier substrates to elucidate specific explicit solvent effects on reactivity and selectivity.[60]

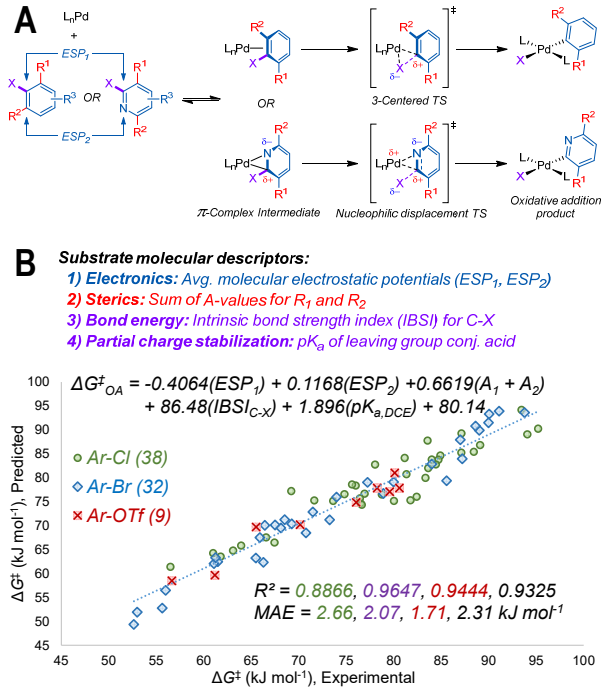

Figure 1. Overview of prior quantitative model for Pd oxidative addition reactivity.[50] **(A)** General mechanism for oxidative addition to $L_nPd(0)$, with π-complex intermediate preceding either Pd insertion into C–X bond, or a nucleophilic displacement of X^-. **(B)** Unified linear regression model of $\Delta G^{\ddagger}_{OA}$ for 79 Ar–Cl, Ar–Br, and Ar–OTf substrates in THF built using molecular descriptors. MAE = mean absolute error. Colour-coding on R^2, Q^2, and MAE values corresponds to the matching data subset, values in black are for all data.

Results and Discussion

Oxidative Addition of Aryl Chlorides and Iodides

In our initially published model,[50] we focused mainly on halogenated heterocycle derivatives and aryl bromides/triflates due to their frequency of use in pharmaceutical synthesis. To assess this model's ability to extrapolate predictions for new substrate classes, and for reactions outside of the training data rate range, we collected relative rate data for a set of ten 4-R-PhCl substrates and seven (Het)Ar–I substrates (using the competition experiment approach described previously[50,61]) and scaled them onto an absolute $\Delta G^{\ddagger}$ scale.[53]

As expected, aryl chlorides are the least reactive species among the substrates in our training data set (Fig. 2B). This substrate set spans a Hammett series – from electron deficient (R = NO_2, CN) to electron rich (R = OCH_3, OH) – with the measured rates spanning 3 orders of magnitude. The Hammett plot for these 10 substituted aryl chlorides gives an excellent linear fit to both the σ_{para} substituent values (ρ = +2.7) and the σ- values (ρ = +1.9).[62] Both ρ values are much smaller than that obtained from prior work by Portnoy and Milstein,[11] who studied the oxidative addition mechanism of aryl chlorides to $Pd(dippp)_2$ (dippp = 1,3-*bis*(diisopropylphosphino)propane). They reported a reaction pathway where the bisligated Pd undergoes reversible dissociation of one dippp ligand, followed by oxidative addition of aryl chlorides to the (dippp)Pd(0) complex *via* a polarized nucleophilic displacement mechanism. Key evidence for this mechanism is the large positive Hammett ρ value (ρ = +5.2) they obtained from a similar set of 4-substituted chlorobenzenes. Using $Pd(PCy_3)_2$, the smaller ρ values we obtain for aryl chlorides and previously for aryl bromides (ρ = +2.3, also using σ-)[50] are consistent with oxidative addition occurring via a non-polar concerted 3-centered mechanism. Importantly, this is also consistent with computational work from Neufeldt and co-workers on the OA mechanistic dichotomy as a function of ligand and ligation state for Pd(0) species, where they observe the 3-centered mechanism is favoured for OA of PhCl and PhBr to $Pd(PCy_3)_n$ (n = 1, 2).[47]

Using these 10 aryl chlorides as an external validation set, we assessed our initial model for out-of-sample prediction accuracy. As shown in Fig. 2B, the model predicts the oxidative addition rates of aryl chlorides with high accuracy, with a mean absolute error (MAE) of 2.27 kJ mol^{-1} for the test set. This is despite the fact that the $\Delta G^{\ddagger}_{OA}$ values for the slowest-to-react Ar–Cl substrates fall beyond the range of those covered by the training set.

On the other end of the reactivity scale, aryl iodides are generally considered as the most reactive electrophile class, based on the generic reactivity order for oxidative addition to Pd(0): *ArI* > *ArOTf* ≈ *ArBr* >> *ArCl*.[23] A mechanistic study by Pflüger and coworkers[9,63] is consistent with a nonpolar 3-centered transition state for oxidative addition of substituted iodobenzenes to $Pd(PPh_3)_4$ based on a kinetic analysis. They obtained similar Hammett ρ values (ρ = +2 in THF, ρ = +2.3 ± 0.2 in toluene) to those we obtained for ArCl and ArBr substrates using $Pd(PCy_3)_2$. They also obtained similar activation energies for reactions in THF and toluene, two solvents which are significantly different in polarity. Later, Jutand and coworkers[34] investigated the oxidative addition mechanism of 2-iodopyridine derivatives to $Pd(PPh_3)_4$ using Hammett plots and DFT calculations of transition states, which also revealed that 2-iodopyridines undergo OA via a 3-centered transition state.

We measured the oxidative addition rates for 7 (hetero)aryl iodides to $Pd(PCy_3)_2$ in THF for a preliminary study of the structure-reactivity relationship of this electrophile class (Fig. 2A). We initially used these 7 (Het)Ar–I substrates as an external test to assess our initial model's performance. In stark contrast to the Ar–Cl set, the predicted rates are much slower than the observed rates, with a large MAE of 8.5 kJ mol^{-1} and a poor predictive squared correlation coefficient[64] (Q^2 = 0.3054) (Fig. 2B). Notably, the R^2 value for this external set is 0.9233, indicating good linear correlation between predicted and experimental values, but a systematic error in prediction accuracy.

Without additional feature selection, we carried out new linear regression analyses of a random 60% subset of the expanded 96 (Het)Ar–X dataset, which results in a very similar model with slightly adjusted coefficients (Fig. 2C). Inspection of the unnormalized coefficients indicates a greater weighting of the C–X bond strength (given here by the intrinsic bond strength index, *IBSI*[59]) and HX pK_a. Doing the regression analysis with normalized descriptors (min/max scaling) confirms a slight reweighting in favour of these two descriptors (Table 1).

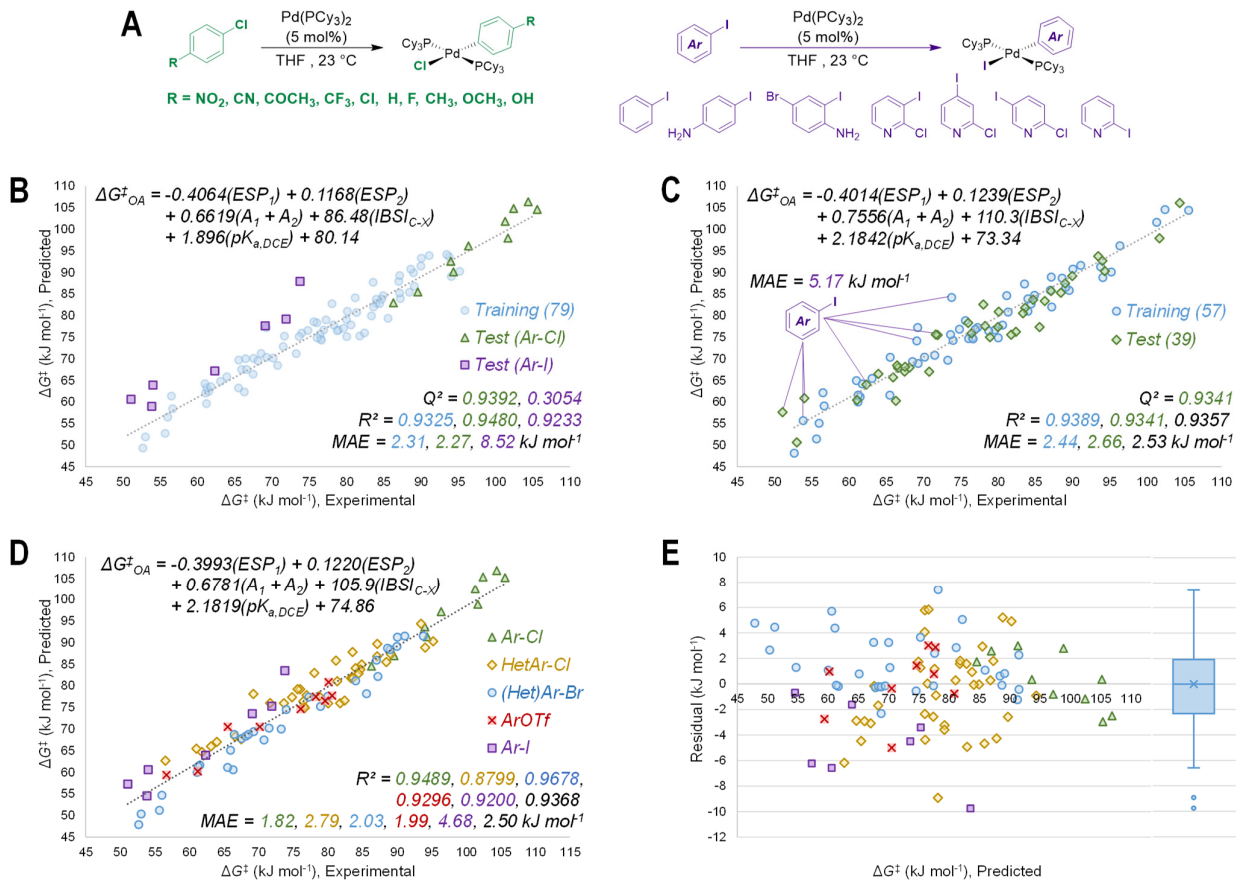

Figure 2. (A) Ar–Cl and Ar–I substrates added in this work. **(B)** Previously reported MLR model for (Het)Ar–X OA to $Pd(PCy_3)_2$ with Ar–Cl and Ar–I substrates used as external test sets, with excellent performance for Ar–Cl, but systematic overestimation of $\Delta G^{\ddagger}_{OA}$ for Ar–I. **(C)** Incorporating Ar–Cl and Ar–I substrates into training/test sets improves overall performance for the MLR model, without additional descriptors. **(D)** MLR model trained using all data with individual substrate classes indicated. **(E)** Residuals plot for the MLR model in D, with two outliers identified by box plot (4-iodoaniline and 4-chloro-2-methylpyrimidine). MAE = mean absolute error. Colour-coding on R^2, Q^2, and MAE values corresponds to the matching data subset, values in black are for all data.

Table 1. Linear regression coefficients and % contribution to predicted $\Delta G^{\ddagger}_{OA}$ for min/max scaled descriptors.

Model	ESP_1	ESP_2	$(A_1 + A_2)$	*IBSI*	pK_a
Fig. 2B	-51.65	23.68	6.924	19.14	21.81
	42%	19%	6%	16%	18%
Fig. 2C	-55.79	23.36	6.357	28.51	23.72
	41%	17%	5%	21%	17%

This modified regression model does show improved prediction accuracy for the Ar–I substrates (3 of which are in the training set, 4 in the test set), with an MAE = 5.17 kJ mol^{-1}. Regression analysis using the entire set of 96 substrates as the training set gives a very similar model (Fig. 2D), with each of the subclasses showing excellent linear correlation between predicted and observed, as well as good MAE values (1.82–4.68 kJ mol^{-1}). Residuals analysis shows a random distribution, and a box plot of the residuals reveals only two significant outliers (4-iodoaniline and 4-chloro-2-methylpyrimidine). Importantly, including these two additional substrate sets has expanded the range of predicted rates to nearly 10 orders of magnitude, from 2-chloro-4-iodopyridine ($\Delta G^{\ddagger}_{OA}$ = 51.1 kJ mol^{-1}) to 4-chlorophenol ($\Delta G^{\ddagger}_{OA}$ = 105.6 kJ mol^{-1}).

Datasets in Other Solvents: Toluene and THF/DMF (1:1)

To study how solvent identity influences the overall oxidative addition reactivity and the corresponding quantitative model, we collected oxidative addition rates for targeted subsets of (Het)Ar–X to $Pd(PCy_3)_2$ in two additional solvent systems: toluene and a mixture of 1:1 THF/DMF. In the latter case, adding THF was necessary to ensure complete dissolution of $Pd(PCy_3)_2$ and the oxidative addition products, enabling reliable analysis by ^{31}P NMR spectroscopy. Along with the initial model data obtained using THF, these three solvents represent a range from non-polar to polar aprotic. Furthermore, they are often used in palladium-catalyzed reactions. Using the same competition experiment approach as for THF,[50] we collected relative rate data for (hetero)aryl chlorides, bromides, iodides and triflates in the two new solvent systems, along with four additional substrates in THF. Finally, we measured absolute rates for touchstone reactions in each solvent system, and used those values to convert the

relative rates into a scale of absolute rates for the entire array of reactions.[62] This approach produced absolute $\Delta G^{\ddagger}_{OA}$ scales for 100 substrates in THF, 49 in 1:1 THF/DMF, and 50 in toluene.

Using these $\Delta G^{\ddagger}_{OA}$ data, we built solvent-specific MLR models by correlating the measured energies in each solvent to the 5 molecular descriptors (*ESP*$_1$, *ESP*$_2$, *A*$_1$+*A*$_2$, *IBSI* and p*K*$_a$). Two substrates – 2-pyridyl triflate and 2-pyridyl tosylate – were omitted from these MLR analyses (*vide infra*). Fig. 3 shows the linear models generated from the entire dataset in each solvent. We also performed five random training/test splits (60/40 for THF, 70/30 for others) for each dataset, which give statistically identical results for the MLR coefficients and corresponding predictions.[62]

As is clear from simple inspection of the data and the resulting models (Fig. 3B-D), there appears to be little/no difference in OA rates for these substrates across the three solvents investigated. The MLR coefficients are very similar for the three models, indicating that there is no significant, systematic difference in how the five descriptors correlate to $\Delta G^{\ddagger}_{OA}$ values. The relative weighting of the five descriptors is also very similar across the three models. This result is not what we had hypothesized; instead, we anticipated that OA with substrates undergoing the more polar nucleophilic displacement mechanism would be significantly accelerated in the THF/DMF solvent system, in accord with prior kinetic results from Maes, Jutand, and co-workers.[34] They measured a 6-fold increase in the observed rate constant for OA of 2-chloropyridine to $Pd(PPh_3)_4$ when changing from THF to DMF solvent, and a corresponding 2.5-fold increase for 2-bromopyridine, whereas our experimental data show essentially equal rates between THF and THF/DMF for these (and many other) substrates. In our case, we are using a different supporting ligand (PCy_3 versus PPh_3), different Pd speciation (PdL_2 versus PdL_4), and a less polar solvent medium (THF/DMF mixture versus only DMF). Importantly, we *do* measure reaction constants (ρ) from Hammett plots that are in accord with those obtained by Maes and Jutand. For 2-Cl-5-*Z*-pyridines, we obtain $\rho_{THF} = 4.8$[50] and $\rho_{THF/DMF} = 4.6$, similar to the reported values for $Pd(PPh_3)_4$ of $\rho_{THF} = 4.3$ and $\rho_{DMF} = 3.9$.[34]

The very similar observed $\Delta G^{\ddagger}_{OA}$ values and corresponding MLR outcomes in each solvent led us to generate a single, unified model using a random training set (60%) of the aggregated data, with no additional descriptors (Fig. 3E). The statistics for this model show it possesses excellent prediction accuracy for the test set, despite the absence of any solvent-based descriptors. We also do not observe any significant outliers that we attribute to solvent effects. Remarkably, we can generate a similarly-performing model using only 10% of the data (20 points) as a random training set ($Q^2 = 0.897$, $MAE_{test} = 2.82$ kJ mol^{-1}).[62]

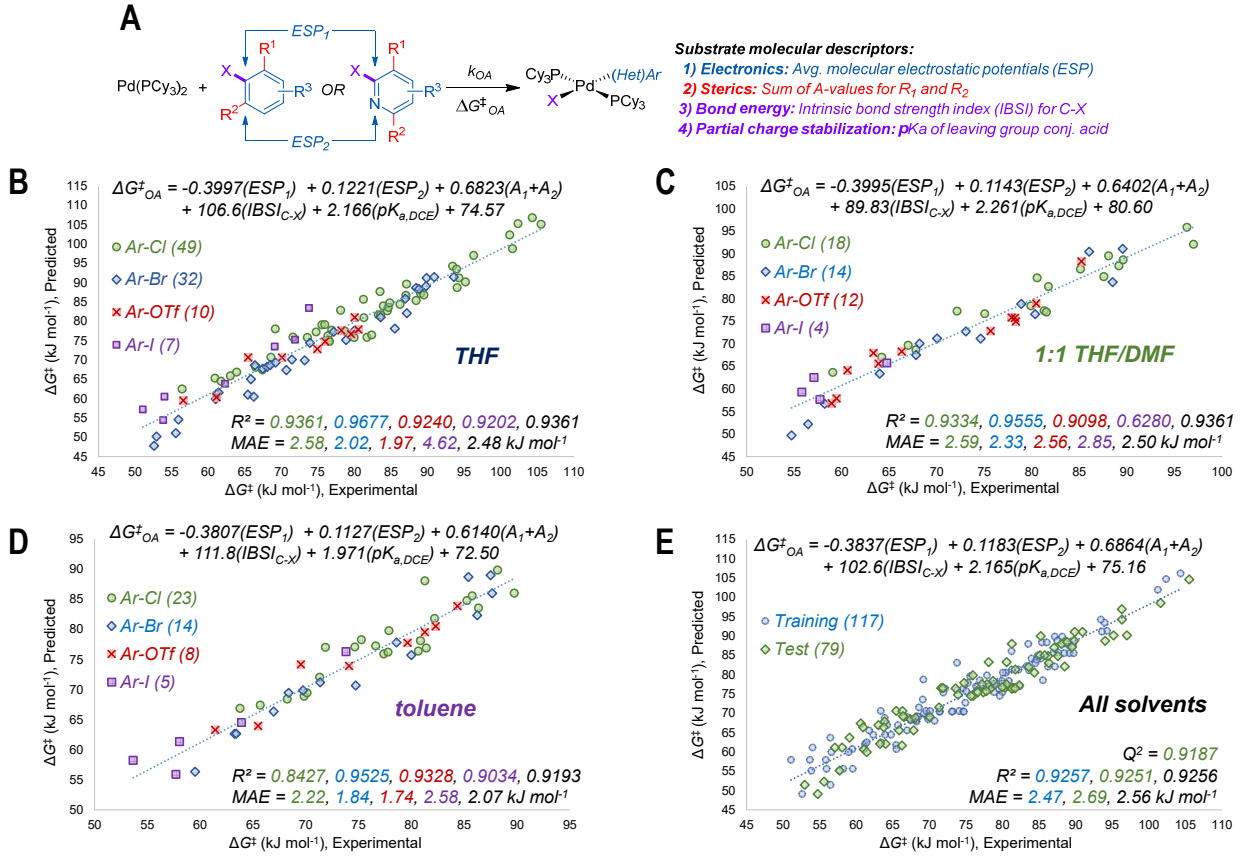

Figure 3. (A) Molecular descriptors used to model oxidative addition reactivity as a function of substrate structure. **(B)** MLR model of $\Delta G^{\ddagger}_{OA}$ for 98 substrates with data obtained in THF. **(C)** MLR model of $\Delta G^{\ddagger}_{OA}$ for 48 substrates with data obtained in 1:1 THF/DMF. **(D)** MLR model of $\Delta G^{\ddagger}_{OA}$ for 50 substrates with data obtained in toluene. **(E)** Unified model for 196 data points in all three solvents, with random selection of training (60%) and test (40%) sets. MAE = mean absolute error. Colour-coding on R^2, Q^2, and MAE values corresponds to the matching data subset, values in black are for all data.

Given the apparent lack of a pronounced solvent effect from the results in Fig. 3, we do have an important caveat to note. Our data do *not* indicate solvent has no effect on OA reactivity. There are several instances where specific solvent effects *are* observed, and these are the subject of a contemporaneous follow-up study.[60] An unfortunate consequence of data aggregation is that phenomena affecting a small subset of the population can be obscured during a holistic analysis. As well, while the MAE values we obtain are generally very good (2-3 kJ mol^{-1}), these do represent rate differences of approximately 2–3 fold at room temperature. Our overall interpretation of the results in Fig. 3 is that *in general*, the choice of toluene, THF, or a THF/DMF mixture has an energetically small effect on the rate of OA to $Pd(PCy_3)_2$.

A Four-Descriptor Model for Halobenzene Oxidative Addition

While there are a number of ways to subdivide our substrate training set, one mechanistically-relevant division is between: 2-X-pyridines, which have a nitrogen atom adjacent to the reactive site; and X-arenes (aryl (*pseudo*)halides and 3-, 4-, or 5-halopyridines, which have only carbon atoms adjacent to the reactive site. The difference in the putative transition state structures (Fig. 1A) between the two limiting OA mechanisms is partly due to the neighbouring atom involvement in bonding to Pd. In a nucleophilic displacement mechanism, palladium-nitrogen coordination occurs at the formation of both the π-complex intermediate and transition state. The partial negative charge in the pyridyl ring is stabilized by the electronegative nitrogen, which helps to lower the energy of this polarized transition state. In contrast, in a 3-centered mechanism, the adjacent carbon is only involved in forming the π-complex, and is not involved in bonding to Pd in the transition state.

A somewhat surprising aspect of all the models discussed thus far is that they appear to function well regardless of substrate type. Both the nucleophilic displacement (generally favoured for 2-X-pyridines) and 3-centered mechanisms are represented in the dataset. On one hand, this is a practical advantage, where a single MLR model can function across a wide range of (Het)Ar–X structures. On the other hand, our ability to use this MLR model to differentiate between the two mechanisms is diminished.

We considered that a simpler, four-descriptor (*ESP*$_1$, *A*$_1$, *IBSI* and pK_a) model should be applicable to electrophiles undergoing OA via the 3-centered mechanism (Fig. 4). Computational analysis of myriad OA transition states by Neufeldt and co-workers indicates a major differentiator between the 3-centered and nucleophilic displacement mechanisms is the involvement of the adjacent atom (Y_{ortho}); this is quantified in their work using the Pd–Y_{ortho} distance in the transition structure.[47] In our MLR models, the neighbouring atom properties are incorporated as the *ESP*$_2$ descriptor. It stands to reason that omitting this descriptor should have lead to good prediction accuracy involving substrates undergoing 3-centered OA mechanisms, but poor predictions for those undergoing the nucleophilic displacement mechanism.

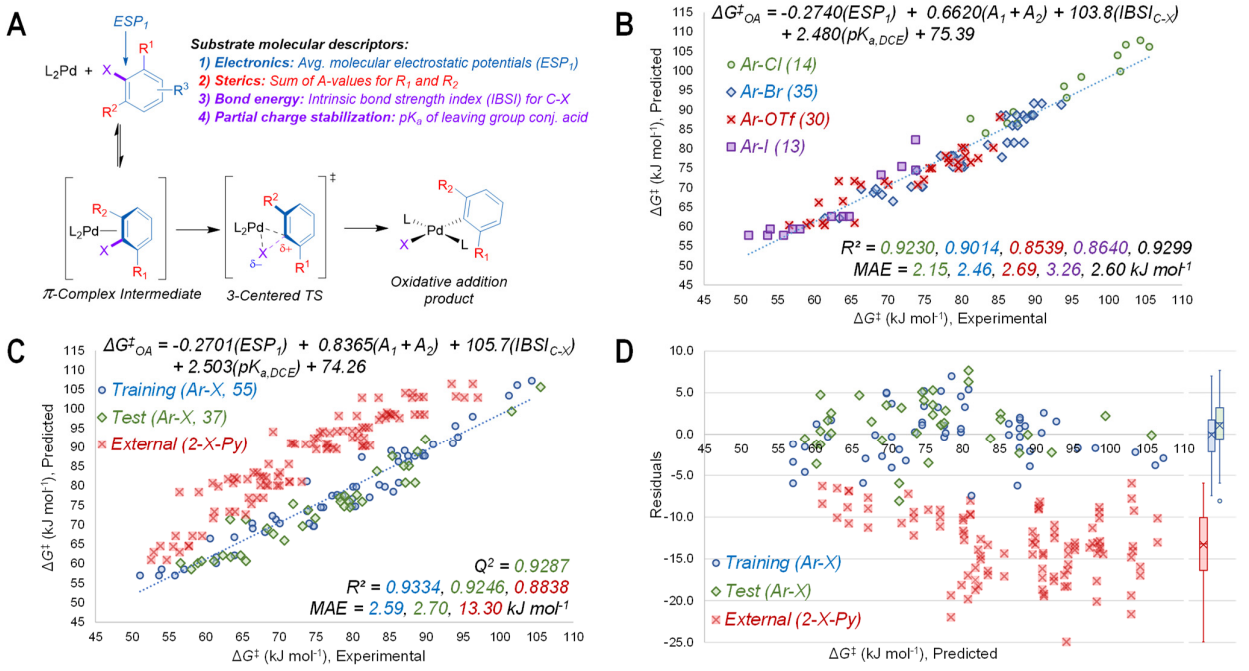

Figure 4. (A) General 3-centered OA mechanism for oxidative addition to $L_2Pd(0)$, with no/minimal Pd–Y_{ortho} bonding in the transition state. **(B)** Four descriptor MLR model for $\Delta G^{\ddagger}_{OA}$ generated using all Ar–X (X = Cl, Br, OTf, I) substrates in all three solvents. **(C)** Four descriptor MLR model generated from random training (60%) and test (40%) sets from Ar–X substrates, with 2-halopyridines used as an external test set. Poor predictions are consistent with this substrate class generally undergoing a nucleophilic displacement OA mechanism. **(D)** Residual plot for the predicted $\Delta G^{\ddagger}_{OA}$ values from Fig. 3C. Box plots for the training, test, and external datasets show systematic divergence between the Ar–X and 2–X–Py sets. MAE = mean absolute error. Colour-coding on R^2, Q^2, and MAE values corresponds to the matching data subset, values in black are for all data.

Using either the entire set of 92 data points (52 Ar–X substrates in three solvent systems, Fig. 4B), or a random 60% training set (Fig. 4C), to generate a four-descriptor MLR model results in strong linear correlation and good MAE values for each leaving group subset (2.15–3.26 kJ mol^{-1}), including better performance for Ar–I prediction (compare to Fig. 2D and 3B). As expected, this model performs poorly with an external test set of 2–X–pyridine substrates (Fig. 4C and 4D, red points). The four-descriptor model systematically overestimates the $\Delta G^{\ddagger}_{OA}$ values for these substrates, with a mean residual of -13.30 kJ mol^{-1}. We do observe a large residual range, from -5 to -25 kJ mol^{-1}, where the smallest residuals are within the residual range of the training/test sets of Ar–X substrates (Fig. 4D). Nevertheless, for the vast majority of 2–X–pyridines examined, there is a clear separation between predicted and experimental $\Delta G^{\ddagger}_{OA}$ values, which we attribute to the two different mechanisms.

Examining the 2–X–pyridines with |residual| < 8 kJ mol^{-1} (*i.e.* within the training/test set range) does reveal counter examples. Two of these low |r| substrates are reported to undergo a 3-centered OA mechanism instead of the nucleophilic displacement mechanism: 2-chloro-5-aminopyridine (|r| = 5.95 [THF/DMF], 7.77 [THF] kJ mol^{-1}),[53] and 2-iodopyridine (|r| = 6.86 [THF/DMF], 6.88 [toluene] kJ mol^{-1}).[34] Other |r| < 8 kJ mol^{-1} substrates include a set of 2-bromo-5-EWG-pyridines (EWG = NO_2, CN, F, CF_3); however, all of these substrates have smaller |r| values in the 5-descriptor model, rendering ambiguous any mechanistic-sorting based on residual values.

Mechanistic Contrast between 2-Pyridyl Tosylate and 2-Pyridyl Triflate Oxidative Addition to $L_2Pd(0)$

As two special cases, we retained 2-pyridyl tosylate and 2-pyridyl triflate as external test substrates. For the former, we sought to examine whether our expanded model(s) could predict reactivity for a new leaving group without additional feature selection (and, unfortunately, we did not observe oxidative addition between simple aryl tosylates and $Pd(PCy_3)_2$ under our reaction conditions). For the latter, we wanted to assess what we predicted would be an extremely reactive species using our original five-descriptor model (Fig. 1B).

Incorporating all data points except for the 2-pyridyl sulfonates gives the five-descriptor and four-descriptor MLR models shown in Fig. 5A and 5B respectively (compare to Fig. 3E and 4C respectively for 60/40 training/test models). Using the five-descriptor model, the predicted $\Delta G^{\ddagger}_{OA}$ value for 2-pyridyl tosylate (92.8 kJ mol^{-1}) is within 5 kJ mol^{-1} of the experimental values obtained in THF (88.1 kJ mol^{-1}) and THF/DMF (87.9 kJ mol^{-1}), even though –OTs is a "new" leaving group. In stark contrast, 2-pyridyl triflate is predicted to be much faster (56.1 kJ mol^{-1}) than is observed (65.6 kJ mol^{-1}), making it a clear outlier based on a box plot of the residuals. We confirmed this experimental value with three independent measurements using three different competition experiment partners.[62]

The situation is reversed when using the four-descriptor model (Fig. 5B). The $\Delta G^{\ddagger}_{OA}$ for 2-pyridyl tosylate is predicted poorly, with |r| > 15 kJ mol^{-1}. This is consistent with the 'average' 2–X–pyridine substrate, the set of which has an MAE of 13.3 kJ mol^{-1} in the four-descriptor model (Fig. 4C); thus, this substrate likely undergoes the expected nucleophilic displacement mechanism. The $\Delta G^{\ddagger}_{OA}$ for 2-pyridyl triflate, on the other hand, is well predicted, with |r| of only 2.5 kJ mol^{-1}. This implies 2-pyridyl triflate undergoes a 3-centered mechanism, or at the very least, is mechanistically distinct from the other 2–X–pyridines. This prompted us to investigate the OA of 2-pyridyl triflate computationally.

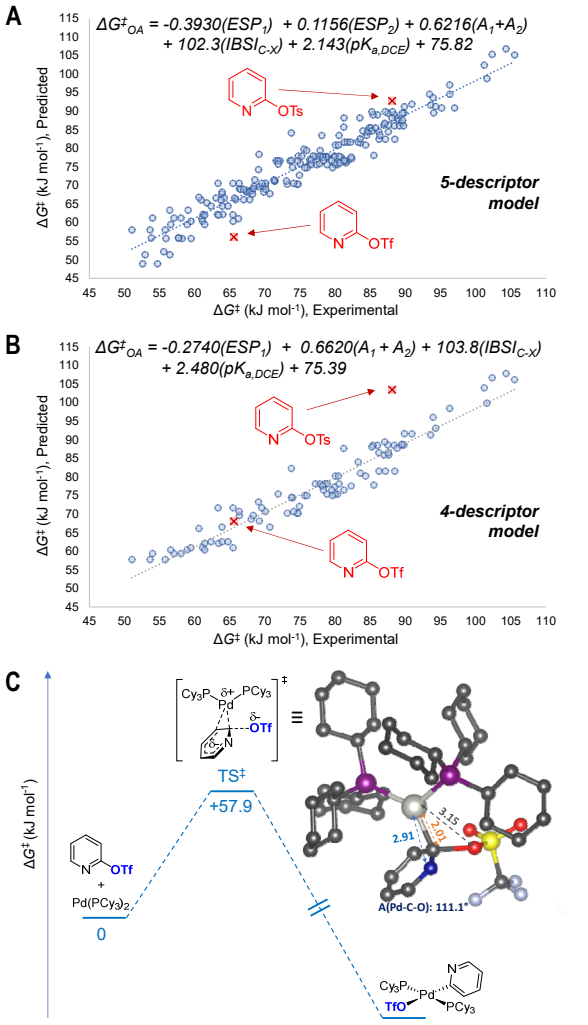

Figure 5. Comparing MLR models for predicting $\Delta G^{\ddagger}_{OA}$ of 2-pyridyl sulfonates. **(A)** $\Delta G^{\ddagger}_{OA}$ prediction using the five descriptor model, which works well for 2-pyridyl tosylate, but 2-pyridyl triflate is a significant outlier. **(B)** $\Delta G^{\ddagger}_{OA}$ prediction using the four descriptor model (only Ar–X data points), which has 2-pyridyl tosylate as a significant outlier (consistent with a nucleophilic displacement mechanism), but which works well for 2-pyridyl triflate, suggesting a mechanistic divergence. **(C)** DFT calculated transition state for the oxidative addition of 2-pyridyl triflate to $Pd(PCy_3)_2$ in THF, which is consistent with a nucleophilic displacement mechanism involving C3 as Y_{ortho} rather than N. TS and INT search, Gibbs energy correction at RI BP86/def2-SVP, def2/J, D3BJ/def2-TZVP(Pd)/CPCM; Single point energy at RI-B2PLYP/def2-TZVP, def2-TZVP/C, D3/CPCM.

DFT calculations on the OA of 2-pyridyl triflate to $Pd(PCy_3)_2$ give a $\Delta G^{\ddagger}_{OA}$ of 57.9 kJ mol^{-1}, which is reasonably consistent with the experimental $\Delta G^{\ddagger}_{OA}$ value. While we were not able to locate a pre-coordination intermediate between Pd and the π-system, we did locate a transition state structure (Fig. 5C). The geometry of

this transition state reveals a pathway distinct from either the 3-centered mechanism, or the typical nucleophilic displacement mechanism for 2–X–pyridines. The long Pd···N distance of 2.91 Å is well outside of the Pd···Y_{ortho} distances observed by Neufeldt and co-workers in a series of calculated nucleophilic displacement transition states (2.18–2.56 Å). Instead, there is a short Pd···C3 distance of 2.37 Å, and a long Pd···OTf distance of 3.15 Å, both of which are consistent with a nucleophilic displacement transition state as per the Neufeldt group's analysis.[47]

The oddity of this particular OA mechanism is that, contrary to other 2–X–pyridine OA reactions studied computationally, the Pd center is partially bonded to the *carbon* adjacent to the C–X site rather than the nitrogen. This is consistent with the two prediction outcomes from Fig 5A and B. For the five-descriptor model, our assumption of a Pd···N TS interaction leads to incorrectly including the ESP_2 as the value for the N atom, giving the large underestimation of the $\Delta G^{\ddagger}_{OA}$ value. Importantly, including the *ESP* for C3 rather than N as ESP_2 leads to a predicted $\Delta G^{\ddagger}_{OA}$ value of 62.8 kJ mol^{-1}, much closer to the experimental value of 65.6 kJ mol^{-1}. Furthermore, this example shows how quantitative structure-reactivity models of this type can be used as mechanistic probes, where discrepancies between predicted and actual outcomes signal the potential for mechanistic divergence.[53] The solvent effects on site-selectivity due to such mechanistic shifts are discussed in a follow-up paper.[60]

The Relationship Between C–X Bond Strength and Oxidative Addition Rates

In our 2022 report on the initial OA model, we included a brief univariate analysis of the *IBSI* descriptor and $\Delta G^{\ddagger}_{OA}$ values for (hetero)aryl chlorides, bromides, and triflates, stating that "bond strength on its own is a poor predictor of oxidative addition reactivity when comparing two substrates with either the same or different leaving groups."[50] The additional data generated in the present study, including a fourth substrate class in (hetero)aryl iodides, enables a more extensive evaluation of how the *IBSI* relates to experimental oxidative addition reactivity (Fig. 6).

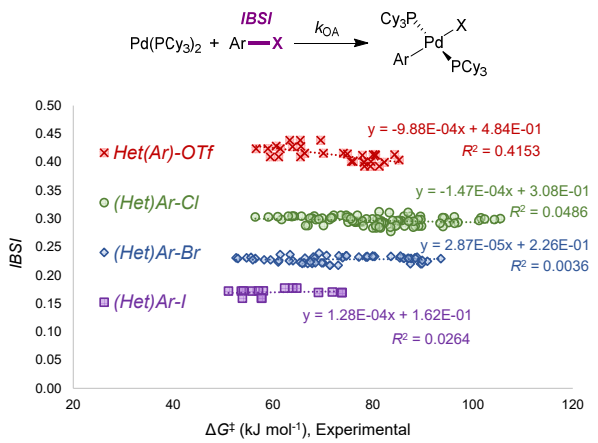

Figure 6. Univariate correlation of $\Delta G^{\ddagger}_{OA}$ against the *IBSI* descriptor to assess relationships between bond strength and oxidative addition rate for four (Het)Ar–X substrate classes (X = Cl, Br, I, OTf); no significant correlation is observed, except for a slight inverse correlation for ArOTf substrates (R^2 = 0.42).

A univariate plot of *IBSI* versus experimental $\Delta G^{\ddagger}_{OA}$ values for all substrates in all solvents obtained thus far adopts a tiered structure, with bond strength index values falling within a narrow range for a given leaving group. In contrast, $\Delta G^{\ddagger}_{OA}$ values for a given leaving group range widely, with the (Het)Ar–Cl series going from 55 – 105 kJ mol^{-1}. Consistent with our initial analysis, there is poor correlation between the *IBSI* and experimental $\Delta G^{\ddagger}_{OA}$ values generally and also within each substrate class; however, we do note a slight *inverse* correlation for the (Het)Ar–OTf substrates, where a higher *IBSI* (i.e. stronger bond) corresponds to a *lower* $\Delta G^{\ddagger}_{OA}$.

Further examination of the relationship between *IBSI* and $\Delta G^{\ddagger}_{OA}$ within four substrate subclasses reveals additional inverse correlations (Fig. 7). Breaking the set of (Het)Ar–Cl substrates into Ar–Cl and 2–Py–Cl, *IBSI* is clearly inversely correlated with $\Delta G^{\ddagger}_{OA}$ in both cases, especially so for the aryl chloride substrates. The same trends are observed for Ar–Br and 2–Py–Br, where stronger C–X bonds corresponds to smaller $\Delta G^{\ddagger}_{OA}$ values.

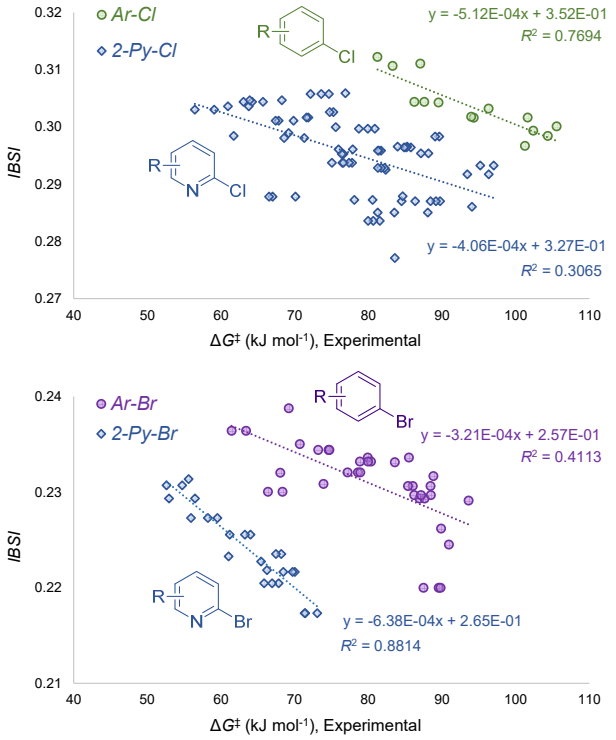

Figure 7. Univariate correlation of $\Delta G^{\ddagger}_{OA}$ against the *IBSI* descriptor within four substrate classes, differentiating 2-halopyridines from simple aryl halides; an *inverse* correlation between bond strength and activation energy is observed.

This counterintuitive result led us to consider the nature of the *IBSI* as a descriptor, and it use in our MLR models. Clearly, *IBSI* is suitable for differentiating between different leaving group types, with values for a given (*pseudo*)halide falling within a tight range (Fig. 6). Unlike other bond strength descriptors, the *IBSI* is closely correlated to the bond stretching force constant. We reasoned that the correlations in Fig. 7 may be simply representing the fact that more polar C–X bonds are stronger, and that more polar C–X bonds also lead to faster OA. Accordingly, we examined univariate correlations between the ESP_1 descriptor

and the *IBSI* for all substrates, and within each subclass, with R^2 values shown in Table 2. Overall, ESP_1 and *IBSI* are not correlated when considering the entire substrate set (which initially led us to include it as an independent descriptor in MLR models); however, within each subclass there is a moderate to strong positive linear correlation, where larger positive ESP_1 values are correlated to larger *IBSI* values (and hence stronger bonds). This is entirely consistent with bond strength increasing as the electronegativity difference between bonded atoms increases.

Table 2. R^2 values for correlation between ESP_1 and two different C–X bond strength descriptors (*IBSI* or *BDE*) for (Het)Ar–X substrates.

ESP_1 vs.	All	ArCl	PyCl	ArBr	PyBr	ArOTf	ArI
IBSI	0.1015	0.6824	0.3038	0.5944	0.8412	0.3384	0.3743
BDE	0.0526	0.4520	0.1549	0.1763	0.0677	0.0142	0.0612

As an alternative bond strength descriptor for MLR analysis, we also evaluated C–X *BDE* values obtained using the ALFABET *BDE* estimator (Fig. 8).[65] As shown in Table 2, these values are not correlated to ESP_1, either in general or within each subclass. These *BDE* values are also poorly (though slightly positively) correlated with $\Delta G^{\ddagger}_{OA}$ values within each leaving group set (Fig. 8A) and also within the Ar–X and 2–Py–X subsets.[62] Using *BDE* in place of *IBSI* for model generation leads to nearly identical results, with both five (Fig. 8B) and four (Fig. 8C) descriptor models closely resembling those presented previously (compare to Fig. 3E and 4C respectively). Thus, *BDE* is a suitable alternative for representing bond strength in these MLR models that does not have any correlation with other descriptors (*i.e.* ESP_1).

As a final observation, we again note that bond strength – whether described by *IBSI* or *BDE* values – is alone not a reliable predictor of OA reactivity. *BDE* values are often used to rationalize OA site-selectivity for multiply halogenated systems, with (for example) the higher reactivity of 2- versus 4-chloropyridine reflected in the corresponding C–Cl *BDE*s (392 kJ mol^{-1} vs. 404 kJ mol^{-1} respectively, values from ALFABET). In many cases comparing *BDE* values does give accurate qualitative predictions; however, substrate steric and electronic factors (not to mention ligand effects and catalyst speciation effects) often override the propensity for the weaker bond to be activated. This is clear from the entire-dataset comparison in Fig. 8A, where there are many cases of substrates with higher *BDE* values undergoing faster OA than others with lower *BDE*s. This is not to say bond strength is unimportant in determining OA reactivity, but rather that using it as the sole predictor or rationale for relative reactivity or site-selectivity should be done with caution.

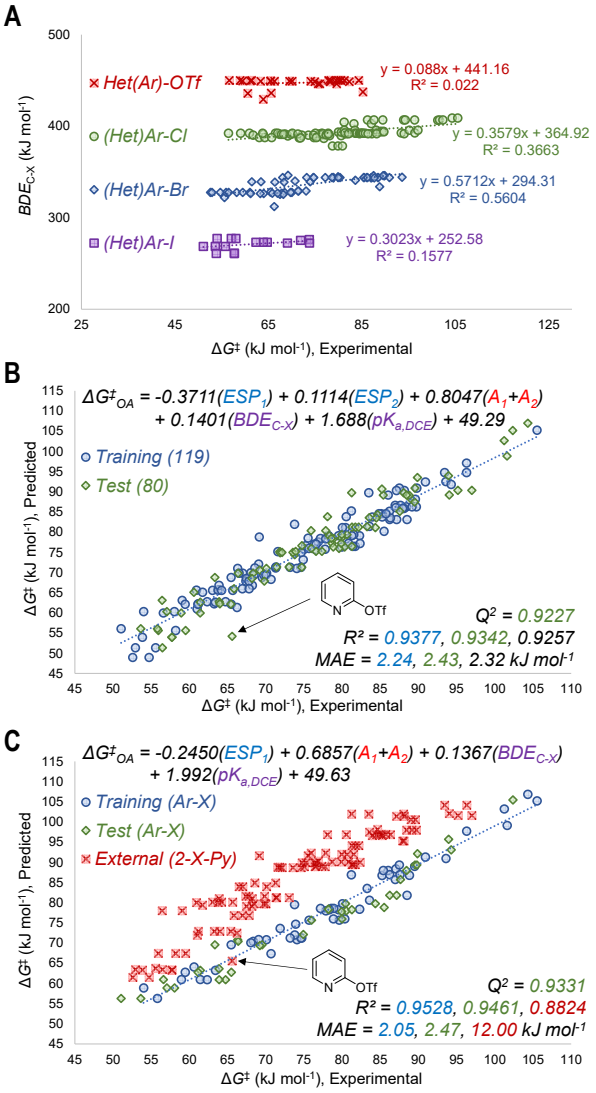

Figure 8. Use of $BDE_{C–X}$ values as a bond strength descriptor. **(A)** Univariate correlation of $\Delta G^{\ddagger}_{OA}$ against *BDE* to assess relationships between bond strength and oxidative addition rate for four (Het)Ar–X substrate classes (X = Cl, Br, I, OTf). **(B)** Five descriptor MLR model incorporating *BDE in lieu* of *IBSI* (compare to Fig. 3E). **(C)** Four descriptor MLR model incorporating *BDE in lieu* of *IBSI* (compare to Fig. 4C); 2–X–Py substrates (external test set) are still near-universal outliers, except 2-pyridyl triflate ($\Delta G^{\ddagger}_{OA}$ = 65.6 kJ mol^{-1}).

Conclusion

Quantitative structure-reactivity relationships are an effective tool for predicting how molecules will react based on their molecular features, and for interrogating reaction mechanisms across a wide range of structures. In the realm of catalysis, there is significant value in creating QSRR models for individual steps in catalytic mechanisms, as these models may be more general than those built using a single catalytic reaction. In the realm of organopalladium catalysis, understanding and predicting oxidative addition reactivity is crucial, as this step is common to myriad mechanisms, and is often rate and/or selectivity determining.

In this work, we have expanded our initially-reported multivariate linear regression model for the reactivity of (hetero)aryl (*pseudo*)halide electrophiles toward oxidative addition to $L_2Pd(0)$. Additional substrate classes and reaction solvents further validate our choice of molecular descriptors, and also enable us to interrogate the mechanistic importance of these descriptors. By contrasting the two mechanistic types for (Het)Ar–X oxidative addition – 3-centered and nucleophilic displacement – we have generated two models to aid in mechanism differentiation. This led to the identification of 2-pyridyl triflate as a nucleophilic displacement outlier, where DFT calculations indicate participation of the neighbouring carbon atom rather than the usual nitrogen atom.

We also assessed the relationship between oxidative addition rates and two different descriptors of bond strength – *IBSI* and *BDE*. While both of these are useful in generating MLR models, *IBSI* has a hidden correlation with the ESP_1 descriptor within subclasses of (Het)Ar–X substrates, leading to inverse linear correlations between *IBSI* and $\Delta G^{\ddagger}_{OA}$ within these same subclasses. In contrast, *BDE* is poorly correlated with both ESP_1 and $\Delta G^{\ddagger}_{OA}$.

Finally, we note that solvent identity appears to have little general effect on $\Delta G^{\ddagger}_{OA}$ values across our dataset. A more comprehensive study of specific solvent effects and their origins is the subject of a contemporaneous follow-up paper.[60]

Supporting Information

Detailed experimental procedures, characterization data, and data tables, computational methods, tables of molecular descriptors, and additional statistical plots (PDF). Additional supporting files contain a complete machine-readable table of experimental data and molecular descriptors (xlsx), and coordinate files for all calculated structures (xyz in zip folder). The authors have cited additional references within the Supporting Information.[66-83]

Acknowledgements

We acknowledge and respect the Lekwungen peoples on whose traditional territory the University of Victoria (UVic) stands, and the Songhees, Esquimalt and WSÁNEĆ peoples whose historical relationships with the land continue to this day. We also acknowledge funding from the New Frontiers in Research Fund – Exploration and the NSERC Discovery Grant programs. D.C.L. thanks the Research Corporation for Science Advancement for a Cottrell Scholar Award.

Keywords: structure-reactivity relationships • organopalladium • oxidative addition • solvent effects • cross-coupling

[1] J. K. Stille, K. S. Y. Lau, *Acc. Chem. Res.* **1977**, *10*, 434–442.
[2] J. Hartwig, *Organotransition Metal Chemistry: From Bonding to Catalysis*, University Science Books, Sausalito, Calif, **2009**.
[3] J. A. Labinger, *Organometallics* **2015**, *34*, 4784–4795.
[4] Norio. Miyaura, Akira. Suzuki, *Chem. Rev.* **1995**, *95*, 2457–2483.
[5] B. T. Ingoglia, S. L. Buchwald, *Org. Lett.* **2017**, *19*, 2853–2856.
[6] L.-C. Campeau, N. Hazari, *Organometallics* **2019**, *38*, 3–35.
[7] D. J. Jones, M. Lautens, G. P. McGlacken, *Nat. Catal.* **2019**, *2*, 843–851.
[8] V. Palani, M. A. Perea, R. Sarpong, *Chem. Rev.* **2022**, *122*, 10126–10169.
[9] C. Amatore, F. Pfluger, *Organometallics* **1990**, *9*, 2276–2282.
[10] C. Amatore, A. Jutand, F. Khalil, M. A. M'Barki, L. Mottier, *Organometallics* **1993**, *12*, 3168–3178.
[11] M. Portnoy, D. Milstein, *Organometallics* **1993**, *12*, 1665–1673.
[12] A. Jutand, A. Mosleh, *Organometallics* **1995**, *14*, 1810–1817.
[13] F. Paul, J. Patt, J. F. Hartwig, *Organometallics* **1995**, *14*, 3030–3039.
[14] J. F. Hartwig, F. Paul, *J. Am. Chem. Soc.* **1995**, *117*, 5373–5374.
[15] C. Amatore, E. Carre, A. Jutand, M. A. M'Barki, *Organometallics* **1995**, *14*, 1818–1826.
[16] C. Amatore, G. Broeker, A. Jutand, F. Khalil, *J. Am. Chem. Soc.* **1997**, *119*, 5176–5185.
[17] C. Amatore, A. Jutand, *Coord. Chem. Rev.* **1998**, *178–180*, 511–528.
[18] A. Jutand, K. K. (Mimi) Hii, M. Thornton-Pett, J. M. Brown, *Organometallics* **1999**, *18*, 5367–5374.
[19] L. M. Alcazar-Roman, J. F. Hartwig, A. L. Rheingold, L. M. Liable-Sands, I. A. Guzei, *J. Am. Chem. Soc.* **2000**, *122*, 4618–4630.
[20] L. M. Alcazar-Roman, J. F. Hartwig, *J. Am. Chem. Soc.* **2001**, *123*, 12905–12906.
[21] L. M. Alcazar-Roman, J. F. Hartwig, *Organometallics* **2002**, *21*, 491–502.
[22] A. H. Roy, J. F. Hartwig, *J. Am. Chem. Soc.* **2003**, *125*, 8704–8705.
[23] A. H. Roy, J. F. Hartwig, *Organometallics* **2004**, *23*, 194–202.
[24] H. M. Senn, T. Ziegler, *Organometallics* **2004**, *23*, 2980–2988.
[25] L. J. Gooßen, D. Koley, H. Hermann, W. Thiel, *Chem. Commun.* **2004**, 2141–2143.
[26] Y. Macé, A. R. Kapdi, I. J. S. Fairlamb, A. Jutand, *Organometallics* **2006**, *25*, 1795–1800.
[27] S. Shekhar, P. Ryberg, J. F. Hartwig, *Org. Lett.* **2006**, *8*, 851–854.
[28] M. Ahlquist, P.-O. Norrby, *Organometallics* **2007**, *26*, 550–553.
[29] G. Paladino, D. Madec, G. Prestat, G. Maitro, G. Poli, A. Jutand, *Organometallics* **2007**, *26*, 455–458.
[30] F. Barrios-Landeros, B. P. Carrow, J. F. Hartwig, *J. Am. Chem. Soc.* **2008**, *130*, 5842–5843.
[31] F. Barrios-Landeros, B. P. Carrow, J. F. Hartwig, *J. Am. Chem. Soc.* **2009**, *131*, 8141–8154.
[32] E. A. Mitchell, P. G. Jessop, M. C. Baird, *Organometallics* **2009**, *28*, 6732–6738.
[33] C. L. McMullin, J. Jover, J. N. Harvey, N. Fey, *Dalton Trans.* **2010**, *39*, 10833–10836.
[34] B. U. W. Maes, S. Verbeeck, T. Verhelst, A. Ekomié, N. von Wolff, G. Lefèvre, E. A. Mitchell, A. Jutand, *Chem. Eur. J.* **2015**, *21*, 7858–7865.
[35] B. A. Anjali, C. H. Suresh, *ACS Omega* **2017**, *2*, 4196–4206.
[36] S. Wagschal, L. A. Perego, A. Simon, A. Franco-Espejo, C. Tocqueville, J. Albaneze-Walker, A. Jutand, L. Grimaud, *Chem. Eur. J.* **2019**, *25*, 6980–6987.
[37] J. Becica, G. Gaube, W. A. Sabbers, D. C. Leitch, *Dalton Trans.* **2020**, *49*, 16067–16071.
[38] N. Pipaón Fernández, G. Gaube, K. J. Woelk, M. Burns, D. P. Hruszkewycz, D. C. Leitch, *ACS Catal.* **2022**, *12*, 6997–7003.
[39] F. Barrios-Landeros, J. F. Hartwig, *J. Am. Chem. Soc.* **2005**, *127*, 6944–6945.
[40] M. Besora, F. Maseras, *Dalton Trans.* **2019**, *48*, 16242–16248.
[41] E. K. Reeves, O. R. Bauman, G. B. Mitchem, S. R. Neufeldt, *Isr. J. Chem.* **2020**, *60*, 406–409.
[42] E. K. Reeves, E. D. Entz, S. R. Neufeldt, *Chem. Eur. J.* **2021**, *27*, 6161–6177.
[43] J. P. Norman, N. G. Larson, S. R. Neufeldt, *ACS Catal.* **2022**, *12*, 8822–8828.
[44] E. K. Elias, S. M. Rehbein, S. R. Neufeldt, *Chem. Sci.* **2022**, *13*, 1618–1628.
[45] C. Joshi, J. M. Macharia, J. A. Izzo, V. Wambua, S. Kim, J. S. Hirschi, M. J. Vetticatt, *ACS Catal.* **2022**, *12*, 2959–2966.
[46] G. M. Ibsen, V. H. Menezes da Silva, J. C. Pettigrew, S. R. Neufeldt, *Chem. Asian J.* **2023**, *18*, e202300036.
[47] M. J. Kania, A. Reyes, S. R. Neufeldt, *ChemRxiv* **2024**, DOI: 10.26434/chemrxiv-2024-h2s9s-v2.
[48] M. Kashihara, C. P. Gordon, C. Copéret, *Org. Lett.* **2020**, *22*, 8910–8915.

[49] J. Rio, H. Liang, M.-E. L. Perrin, L. A. Perego, L. Grimaud, P.-A. Payard, *ACS Catal.* **2023**, *13*, 11399–11421.
[50] J. Lu, S. Donnecke, I. Paci, D. C. Leitch, *Chem. Sci.* **2022**, *13*, 3477–3488.
[51] P. R. Wells, *Chem. Rev.* **1963**, *63*, 171–219.
[52] Corwin. Hansch, A. Leo, R. W. Taft, *Chem. Rev.* **1991**, *91*, 165–195.
[53] J. Lu, N. D. Schley, I. Paci, D. C. Leitch, *ChemRxiv* **2024**, DOI 10.26434/chemrxiv-2024-lrdz4-v2.
[54] C. H. Suresh, P. Alexander, K. P. Vijayalakshmi, P. K. Sajith, S. R. Gadre, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6492–6499.
[55] F. B. Sayyed, C. H. Suresh, *New J. Chem.* **2009**, *33*, 2465–2471.
[56] G. S. Remya, C. H. Suresh, *Phys. Chem. Chem. Phys.* **2016**, *18*, 20615–20626.
[57] S. R. Gadre, C. H. Suresh, N. Mohan, *Molecules* **2021**, *26*, 3289.
[58] J. A. Hirsch, in *Topics in Stereochemistry*, John Wiley & Sons, Ltd, **1967**, pp. 199–222.
[59] J. Klein, H. Khartabil, J.-C. Boisson, J. Contreras-García, J.-P. Piquemal, E. Hénon, *J. Phys. Chem. A* **2020**, *124*, 1850–1860.
[60] J. Lu, H. Celuszak, I. Paci, D. C. Leitch, *ChemRxiv* **2024**, DOI:
[61] J. Lu, I. Paci, D. C. Leitch, *Chem. Sci.* **2022**, *13*, 12681–12695.
[62] See the Supporting Information for more details.
[63] J.-F. Fauvarque, F. Pflüger, M. Troupel, *J. Organomet. Chem.* **1981**, *208*, 419–427.
[64] V. Consonni, D. Ballabio, R. Todeschini, *J. Chem. Inf. Model.* **2009**, *49*, 1669–1678.
[65] P. C. St. John, Y. Guan, Y. Kim, S. Kim, R. S. Paton, *Nat. Commun.* **2020**, *11*, 2328.