






Proofreading experimentally assigned stereochemistry through Q2MM predictions in Pd-catalyzed allylic aminations

Jessica Wahlers¹, Jèssica Margalef ², Eric Hansen ¹, Armita Bayesteh ³, Paul Helquist ¹,
Montserrat Diéguez ², Oscar Pàmies ², Olaf Wiest ¹✉ & Per-Ola Norrby ^{4,5}✉

The palladium-catalyzed enantioselective allylic substitution by carbon or nitrogen nucleophiles is a key transformation that is particularly useful for the synthesis of bioactive compounds. Unfortunately, the selection of a suitable ligand/substrate combination often requires significant screening effort. Here, we show that a transition state force field (TSFF) derived by the quantum-guided molecular mechanics (Q2MM) method can be used to rapidly screen ligand/substrate combinations. Testing of this method on 77 literature reactions revealed several cases where the computationally predicted major enantiomer differed from the one reported. Interestingly, experimental follow-up led to a reassignment of the experimentally observed configuration. This result demonstrates the power of mechanistically based methods to predict and, where necessary, correct the stereochemical outcome.

¹Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556, USA. ²Departament de Química Física i Inorgànica, Universitat Rovira i Virgili, C/Marcel·li Domingo, 43007 Tarragona, Spain. ³Oral Product Development, Pharmaceutical Technology & Development, Operations, AstraZeneca, Gothenburg, Sweden. ⁴Data Science and Modelling, Pharmaceutical Sciences, R&D, AstraZeneca Gothenburg, Pepparedsleden 1, SE-431 83, Mölndal, Sweden. ⁵Department of Chemistry and Molecular Biology, University of Gothenburg, Gothenburg, Sweden. ✉email: owiest@nd.edu; Per-Ola.Norrby@astrazeneca.com

Computational chemistry has long promised the development of predictive methods in order to reduce the time needed to develop and optimize the conditions of reactions¹. This has become especially desirable for predicting stereoselectivity in asymmetric catalysis because the identification of a chiral catalyst that gives high enantiomeric excess (*ee*) for a given substrate requires significant effort. While high-throughput experimentation allows for many different reaction conditions to be tested at once, this method still remains costly, especially for testing many different ligands^{2,3}. Computational methods can not only predict which ligands would give the best results, reducing the time and cost needed to find the best catalyst⁴, but also give insight into the steric and electronic interactions that promote high stereoselectivity. Given the small energy differences involved, the computational methods need to be highly accurate while being fast enough to be useful for the synthetic chemist.

A reaction of wide use in the pharmaceutical industry is the palladium-catalyzed asymmetric allylic substitution due to its mild conditions and ability to stereoselectively form a bond to carbon with a wide range of nucleophiles (Fig. 1a)^{5–7}. Of particular interest is the allylic amination reaction, which forms a bond between a chiral carbon and an amine nitrogen. About 84% of pharmaceuticals contain at least one nitrogen atom, many of which are at a chirality center for which absolute configuration is important for desired therapeutic properties^{8,9}. While this substitution reaction has been widely studied to determine the scope and mechanism, new substrates or nucleophiles usually require a new ligand screen to find the optimal catalyst^{5,7,10–12}. The selectivity in this reaction depends on a complex interplay between steric interactions favoring a certain allyl geometry, dynamic interconversion through *exo-endo* isomerization of the allyl moiety, and electronic effects whereby the ligand can influence the regioselectivity of nucleophilic attack^{7,13}.

The catalytic cycle of this reaction proceeds^{7,14–16} through an oxidative addition to form the reactive η^3 -allyl palladium intermediate, which has been studied by X-ray crystallography. (Fig. 1b). The *exo* and *endo* isomers of the Pd-allyl species are

generally in rapid equilibrium with each other¹³. The nucleophile then attacks the allyl group in the stereoselectivity determining transition state. The most common chiral ligands to introduce stereoselectivity in this step are phosphorus and nitrogen based bidentate ligands^{7,17,18}. There has been interest in using P,N ligands because they can discriminate between the two terminal allylic carbons based on their electronic differentiation, directing the nucleophile towards the allylic carbons trans to the phosphorus atom. Some common ligands used for this reaction include the PHOX ligands, phosphite-oxazoline ligands, and aminoalkyl-phosphine ligands^{7,18–22}. These ligands can control *exo-endo* preference through the chiral oxazoline/amine moiety which, thanks to the trans phosphorus, is in close proximity to the reacting allyl terminus (Fig. 1c)¹⁷.

There have been a few methods developed to predict stereoselectivity in asymmetric catalysis. Calculation of the transition state structures and the energy difference between the structures leading to the R and S enantiomers by DFT^{14,16,23} is slow and typically does not sample a sufficiently large number of conformations. Another method is to predict stereoselectivity by fitting to various steric and electronic parameters²⁴. Recently, there has been a push to use machine learning methods, but these methods often need large data sets of high quality to train the model, and offer limited insight into details of the reaction mechanisms and which parameters contribute to high stereoselectivity²⁵.

Quantum Guided Molecular Mechanics (Q2MM) was developed to predict stereoselectivity, combining the speed of molecular mechanics (MM) with the accuracy of DFT^{26–29}. It uses transition state force fields (TSFFs) that are trained on electronic structure calculations of simplified models of the stereoselecting transition state. Because no empirical data are used to fit the force field, the results are true predictions. Once a force field has been developed, it can be used to perform a Monte-Carlo conformational search to determine the Boltzmann-averaged energy difference between the transition state structures that lead to the R and S enantiomers. CatVS is a program that automates the process of building full TS structures as well as adding conformational search parameters to the

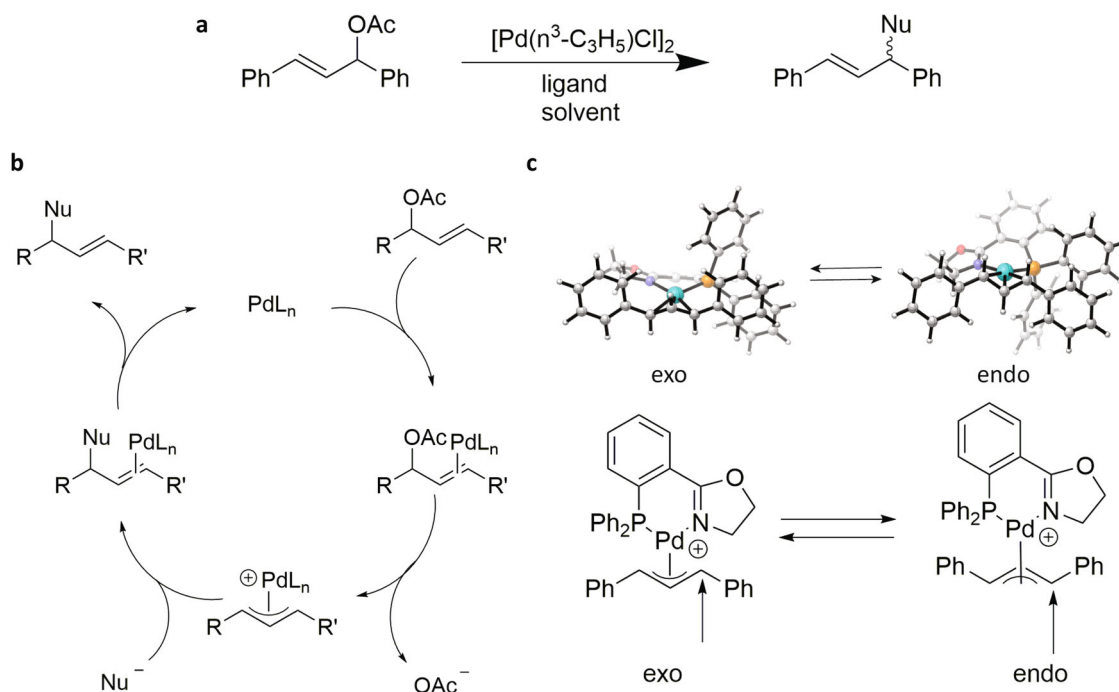


Fig. 1 Pd-catalyzed allylic amination reaction. **a** Reaction modeled for the TSFF being developed. **b** Simplified mechanism of the reaction. **c** Exo-endo isomerization of the allyl.

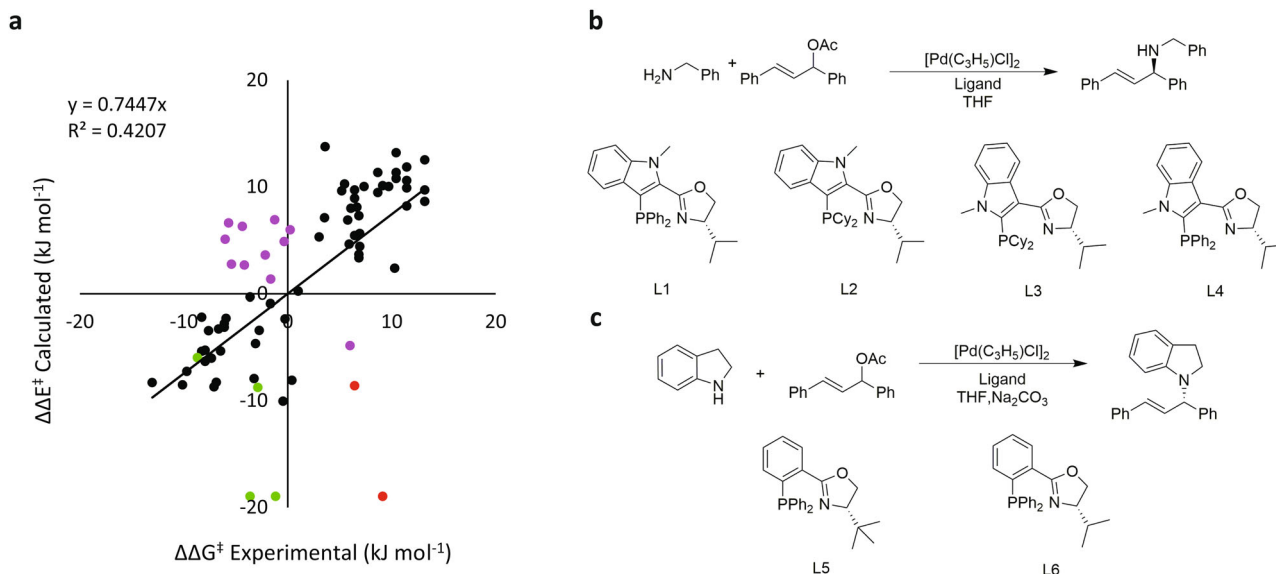


Fig. 2 Comparison of relative energies of the experimental values to the calculated MM values. **a** The largest systematic errors in the TSFF are for ligands containing an indole backbone (green), examples of predicting opposite absolute configuration with a PHOX ligand (red), and examples of predicting opposite absolute configuration with a phosphite-oxazoline ligand (purple). **b** Reactions that are catalyzed by ligands with an indole backbone (green data points). **c** Reaction of the two examples that give the opposite absolute configuration when catalyzed by the PHOX ligands (red data points).

full system³⁰. These energy differences are then compared and validated by the experimental results.

A ground state force field of the reactive intermediate for this reaction was previously developed to study steric interactions that contribute most to the stereoselectivity of the reaction^{31–33}. However, predictions using the ground state force field requires manual inspection of geometries and assumptions about preferred nucleophilic attack vectors. For the rapid screening of new ligands, substrates, and nucleophiles, a TSFF is better suited to predict stereoselectivity, since it is the difference in transition state energies rather than ground states that govern preference for formation of a particular stereoisomer of the major product. Computational insight could also elucidate which interactions influence selectivity to find the optimal ligand for a given substrate and nucleophile.

Here, we describe the development of a TSFF for the palladium-catalyzed allylic amination reaction to predict stereoselectivity as well as understand the interactions in the transition state that lead to higher selectivity. Numerous discrepancies between the predicted and literature values are re-investigated experimentally, and found to be due to misassigned absolute stereochemistry in the literature.

Results and discussion

A training set consisting of 21 simplified TS structures (see Supplementary Fig. 1 and Supplementary Table 1) that capture the steric and electronic information around the reaction coordinate and metal center was used to parameterize the TSFF. In addition, one structure representing a full ligand (achiral) and a full allyl structure was included to ensure that the interactions being parameterized accurately describe the steric and electronic interactions as well as capture the geometry of a full system. The reference structures were optimized using M06-D3/LANL2DZ/6-31 + G* (for details see Methods section), and the TSFF was parameterized by Q2MM as described earlier^{26,27}. Internal validation of the optimized parameters such as structural data and Hessian eigenvalues between the QM and MM optimized transition structures is shown in the Supplementary Information. Minor deviations in the bond length of the forming bond between the allylic carbon and the amine are observed for cases with

sterically bulky ligands where the forming bond is usually shorter. No significant deviations between QM and MM in the angles and torsions of the training set are observed. Overall, the R^2 values for the internal validation ranges from 0.988 to 0.998 for geometric and Hessian eigenvalues, respectively, and 0.822 for charges, which are typical values for internal validations of TSFFs^{28,34,35}.

The next step is the external validation by prediction of selectivities for ligand-substrate combinations from the literature that are not part of the training set. Using CatVS³⁰, the libraries of TS structures can rapidly and automatically be prepared for conformational searches by merging substrate, ligand, and nucleophile sub-libraries onto a template. The calculation of each pair of diastereomeric transition states takes between 15 and 60 min on a single core, making this method suitable for high-throughput calculations on even a modest cluster. The output is given as differences in TS energies for forming the two enantiomeric products, and also as enantiomeric ratio and excess, calculated from Eq. 1. For cases with more than two competing transition states, the ratio is obtained by a Boltzmann summation over diastereomeric pathways where positive values indicate a preference of the S enantiomer:

$$\begin{aligned} \text{enantiomeric ratio} : er &= e^{\Delta\Delta G^\ddagger/RT} \\ \text{enantiomeric excess} : ee &= 100\% \frac{er-1}{er+1} \end{aligned} \quad (1)$$

A validation dataset containing 77 structures (Supplementary Figs. 3 and 4, and Supplementary Table 3) assembled from the literature^{19,21,36–43} was used to test the performance of the TSFF for systems different than the training set (Fig. 2a). 1,3-Diphenyl propenyl was used as the allyl component reacting with 16 different amines, catalyzed by the Pd-complexes of 53 different P,N ligands. Most ligands, including PHOX and norbornyl ligands as well as ligands with different substituents on the nitrogen are well described by the force field. The experimental free energy differences between ensembles leading to the enantiomeric product, $\Delta\Delta G^\ddagger$, was derived from Eq. 2 where positive $\Delta\Delta G^\ddagger$ indicate that the S enantiomer is the preferred product

$$\Delta\Delta G^\ddagger = RT \ln(er) \quad er = \frac{100\% + ee}{100\% - ee} \quad (2)$$

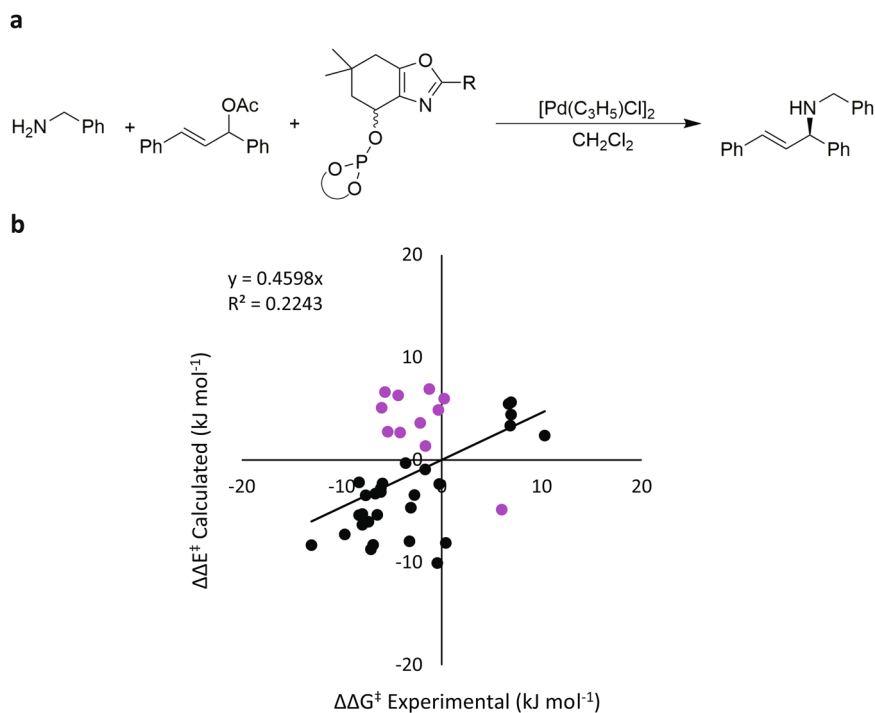


Fig. 3 Comparison of relative energies of the experimental values to the calculated MM values for 39 phosphite-oxazole ligands. a Reaction corresponding to the 11 mismatched data points. **b** Calculated vs. experimental stereoselectivity with mismatched cases in purple.

The final test showed larger deviations than are usually seen with Q2MM. The mean unsigned error (MUE) over the 77 cases was 4.4 kJ/mol and the R^2 value only 0.41 (Fig. 2). Although these values are not as good as those of several published TSFFs^{26,27}, it is clear from Fig. 2a that the vast majority of cases in the validation set is reproduced well and that the deviations are due to a small number (<20%) of cases with significant differences between the computed and experimental results.

Historically, the path to systematic improvements of force fields is through the detailed analysis of the outliers⁴⁴. Such an analysis for the results in Supplementary Table 3 indicates that the high MUE originates from a few systematic deviations that are color-coded in Fig. 2a. The first set of ligands where the predictions deviate from the experimental results are IndPHOX ligands, shown in green in Fig. 2a. Experimentally, L1 and L4 give very different selectivities of 52 % ee and 94 % ee, respectively⁴³. Sterically, the ligands are very similar, and thus the force field predicts that these two ligands should give similar selectivity results with L1 giving 93.5 % ee and L4 giving 95.3 % ee. Similar results are obtained for the related ligands L2 and L3, where the selectivities are predicted to be too high. In L1 and L2, the phosphorus is connected to the very electron-rich 3-position of the indole. It is plausible that the resulting catalytic activity is so high that the nucleophilic attack is faster than the exo-endo isomerization. The Q2MM model depends on a Curtin-Hammett situation where the exo and endo isomers are in rapid equilibrium. If this effect is negated by a too fast nucleophilic attack, the reaction becomes stereospecific, and a racemic allylic acetate will be formed in such a situation yielding low selectivity. Thus, this seems to be a case of a change in mechanism for which the Q2MM-derived TSFF is therefore not applicable.

More interesting are cases where the predicted stereoselectivity is high but opposite to the one reported in the literature. These include two examples of PHOX ligands (L5 and L6 in Fig. 2c) shown in red in Fig. 2a³⁹ and a series of reactions using a phosphite-oxazole ligand shown in purple in Fig. 2a and discussed below. The force field predicts that the absolute product

configuration should be *R* for the two PHOX ligands while the experimental results have *S* as the absolute stereochemistry. While the absolute configuration was determined by a crystal structure, the reported Flack parameter was 0.1 with an uncertainty of 0.6. This large degree of uncertainty could suggest that the absolute stereochemistry was reported incorrectly. L6 has previously been used by another group with similar reaction conditions, but using benzylamine rather than indoline as the nucleophile²¹. In that case, the absolute configuration predicted by the force field matches the absolute configuration described in the literature which was reported as *R*. A substrate screen performed by Liu and coworkers using a *D*-camphor-based chiral P,N ligand did not show a change in the absolute stereochemistry between the benzylamine and indoline nucleophiles³⁷. This suggests that there should not have been a change in the absolute stereochemistry when using L6 for the benzylamine and indoline nucleophiles that was seen experimentally. To study this, the stereochemistry assignment was reexplored experimentally (see Supplementary Information). Comparison of the chromatographic eluting order and the polarimetric analysis of the aminated product using ligand L5 with the literature indicated that the major enantiomer formed is the (*R*)-(-)-1-(1,3-diphenylallyl)indoline as predicted by the calculations.

The possibility for the mismatch between computed and reported absolute stereochemistry was also explored for the phosphite-oxazole ligands (Fig. 3b) for which a larger dataset is available. In total, 39 different ligand-substrate combinations for this reaction were studied^{36,37}, 11 of which showed the mismatch (Fig. 3a). Specifically, the TSFF predicts that the absolute configuration to be *S* while the literature reports an absolute configuration of *R* for the products. An analysis of the 28 cases where the predicted and reported stereochemistry match (black in Fig. 3a) did not show any significant differences to the 11 cases that did.

We therefore initiated experimental studies to check the original stereochemical assignment. For that purpose, we reexamined several of the mismatched phosphite-oxazole ligands in allylic amination of (*rac*)-1,3-diphenyl allyl acetate with

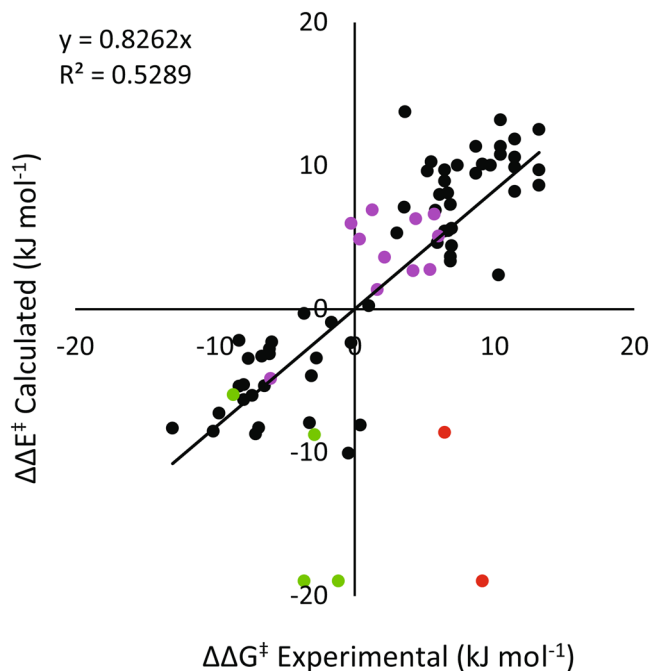


Fig. 4 Comparison of relative energies of the corrected experimental values to the calculated MM values. The absolute configurations for the 11 data points in purple have been corrected.

benzylamine (see Supplementary Table 5). In all cases, chromatographic comparison of the aminated product to known samples revealed that the original assignment in the literature was incorrect, and that the dominant stereoisomer was the one predicted by the Q2MM force field. This shows that the predictions of the model in this case are qualitatively and quantitatively correct even when they contradict assignments of the absolute stereochemistry in the literature.

Having experimentally confirmed that the computationally predicted absolute stereochemistry is correct, the overall MUE over 77 cases decreased to 3.2 kJ/mol (Fig. 4). This value is still affected by the a small number of data points where we believe a mechanistic shift has invalidated the Q2MM model as discussed earlier. Excluding the IndPHOX results (green dots) as being out of scope due to change in mechanism the remaining 95% of the 77 cases are predicted by the TSFF with an MUE of 2.8 kJ/mol and an R^2 of 0.72, which is typical Q2MM-derived force fields^{26,27}.

To conclude, mechanism-based prediction of using Q2MM-derived TSFF has shown an ability not only to predict reaction outcome in advance of experimental work but also to correct stereochemical assignments of sets of reported data. We note that other methods that are based on machine learning (ML), which are becoming rapidly popular in the field, are highly sensitive to such errors in input data and are unable to correct for such systematic errors for more than a very small percentage of the training data. Instead, the highly complex correlation models obtained from many ML methods will attempt to fit the spurious data, resulting in models that give erroneous assignments for sets within the applicability domain. Although it is obvious that stringent quality control of the training for ML models is essential, this is in practice difficult to do for large datasets from the literature. We thus believe that fast TSFF calculations, which describe the underlying causes for the observed selectivities through physically meaningful parameters, provide a new, high-throughput tool to “proofread” stereochemical assignments that could be highly useful for researchers engaged in studies of asymmetric synthesis.

Methods

DFT calculations of the training set were performed in the gas phase using Gaussian⁴⁵. The M06⁴⁶ functional form was used with a D3 empirical dispersion correction⁴⁷. The basis sets used were LANL2DZ for palladium and 6-31 + G* for all other atoms. CHELPG⁴⁸ with a vdW radius of 2.4 Å for palladium was used to calculate the partial charges. Frequency analysis confirmed that the transition state structures contained one negative vibration corresponding to the formation of the carbon–nitrogen bond.

The TSFF parameters for the atoms involved in bond formation (see Supplementary Information) were fit and optimized using the Q2MM method. The MM3* force field⁴⁹ was used as the functional form of the TSFF and for any parameter that were not being fit. The full TS systems were automatically generated by CatVS and subjected to 40,000 steps of Monte Carlo conformational search using the mixed torsional/low-mode sampling in MacroModel⁵⁰ with a constant dielectric of 1.0. The resulting conformations of the diastereomeric transition states were, after Boltzmann averaging, used for prediction of selectivity²⁷.

Code availability

An open-source version of the Q2MM/CatVS code, together with a library of the currently available TSFFs, reaction templates and ligand libraries, is available to the scientific community free of charge as part of the Q2MM package for the generation of TSFFs in the GitHub repository (<https://github.com/Q2MM/q2mm>).

Data availability

All other data are available from the authors upon reasonable request.

Received: 1 June 2021; Accepted: 3 November 2021;

Published online: 18 November 2021

References

- Houk, K. N. & Liu, F. Holy grails for computational organic chemistry and biochemistry. *Acc. Chem. Res.* **50**, 539–543 (2017).
- Krska, S. W., DiRocco, D. A., Dreher, S. D. & Shevlin, M. The evolution of chemical high-throughput experimentation to address challenging problems in pharmaceutical synthesis. *Acc. Chem. Res.* **50**, 2976–2985 (2017).
- McDougal, N. T., Virgil, S. C. & Stoltz, B. M. High-throughput screening of the asymmetric decarboxylative alkylation reaction of enolate-stabilized enol carbonates. *Synlett* **2010**, 1712–1716 (2010).
- Patrascu, M. B. et al. From desktop to benchtop with automated computational workflows for computer-aided design in asymmetric catalysis. *Nat. Catal.* **3**, 574–584 (2020).
- Trost, B. M. & Van Vranken, D. L. Asymmetric transition metal-catalyzed allylic alkylations. *Chem. Rev.* **96**, 395–422 (1996).
- Diéguez, M. & Pàmies, O. Biaryl phosphites: new efficient adaptative ligands for Pd-catalyzed asymmetric allylic substitution reactions. *Acc. Chem. Res.* **43**, 312–322 (2010).
- Pàmies, O. et al. Recent advances in enantioselective Pd-catalyzed allylic substitution: from design to applications. *Chem. Rev.* **121**, 4373–4505 (2021).
- Vitaku, E., Smith, D. T. & Njardarson, J. T. Analysis of the structural diversity, substitution patterns, and frequency of nitrogen heterocycles among U.S. FDA approved pharmaceuticals: miniperspective. *J. Med. Chem.* **57**, 10257–10274 (2014).
- Hu, L., Cai, A., Wu, Z., Kleij, A. W. & Huang, G. A mechanistic analysis of the palladium-catalyzed formation of branched allylic amines reveals the origin of the regio- and enantioselectivity through a unique inner-sphere pathway. *Angew. Chem. Int. Ed.* **58**, 14694–14702 (2019).
- Johannsen, M. & Jørgensen, K. A. Allylic amination. *Chem. Rev.* **98**, 1689–1708 (1998).
- Trost, B. M. Asymmetric allylic alkylation, an enabling methodology. *J. Org. Chem.* **69**, 5813–5837 (2004).
- Lu, Z. & Ma, S. Metal-catalyzed enantioselective allylation in asymmetric synthesis. *Angew. Chem. Int. Ed. Engl.* **47**, 258–297 (2008).
- Hansson, S. et al. Effects of phenanthroline type ligands on the dynamic processes of (η^3 -allyl) palladium complexes. Molecular structure of (2,9-dimethyl-1,10-phenanthroline)[(1, 2, 3- η)-3-methyl-2-butenyl]-chloropalladium. *Organometallics* **12**, 4940–4948 (1993).
- Cusumano, A. Q., Stoltz, B. M. & Goddard, W. A. III Reaction mechanism, origins of enantioselectivity, and reactivity trends in asymmetric allylic alkylation: a comprehensive quantum mechanics investigation of a C(sp³)-C(sp³) cross-coupling. *J. Am. Chem. Soc.* **142**, 13917–13933 (2020).
- Trost, B. M., Machacek, M. R. & Aponick, A. Predicting the stereochemistry of diphenylphosphino benzoic acid (DPPBA)-based palladium-catalyzed asymmetric allylic alkylation reactions: a working model. *Acc. Chem. Res.* **39**, 747–760 (2006).

16. McPherson, K. E., Croatt, M. P., Morehead, A. T. Jr & Sargent, A. L. DFT mechanistic investigation of an enantioselective tsuji–trost allylation reaction. *Organometallics* **37**, 3791–3802 (2018).
17. Helmchen, G. & Pfaltz, A. Phosphinoxazolines—a new class of versatile, modular P, N-ligands for asymmetric catalysis. *Acc. Chem. Res.* **33**, 336–345 (2000).
18. Holtzman, B. S. et al. Ligand and base additive effects on the reversibility of nucleophilic addition in palladium-catalyzed allylic aminations monitored by nucleophile crossover. *Tetrahedron Lett.* **58**, 432–436 (2017).
19. Császár, Z. et al. Aminoalkyl-phosphine (P, N) ligands with pentane-2, 4-diyol backbone in asymmetric allylic substitution reactions. *Monatsh. Chem.* **148**, 2069–2077 (2017).
20. Biosca, M. et al. An improved class of phosphite-oxazoline ligands for Pd-catalyzed allylic substitution reactions. *ACS Catal.* **9**, 6033–6048 (2019).
21. Caminiti, N. S. et al. Reversible nucleophilic addition can lower the observed enantioselectivity in palladium-catalyzed allylic amination reactions with a variety of chiral ligands. *Tetrahedron Lett.* **56**, 5445–5448 (2015).
22. Jailet, A. et al. Design of P-chirogenic aminophosphine–phosphinite ligands at both phosphorus centers: origin of enantioselectivities in Pd-catalyzed allylic reactions. *J. Org. Chem.* **85**, 14391–14410 (2020).
23. Lam, Y.-H., Grayson, M. N., Holland, M. C., Simon, A. & Houk, K. N. Theory and modeling of asymmetric catalytic reactions. *Acc. Chem. Res.* **49**, 750–762 (2016).
24. Reid, J. P. & Sigman, M. S. Comparing quantitative prediction methods for the discovery of small-molecule chiral catalysts. *Nat. Rev. Chem.* **2**, 290–305 (2018).
25. Zahrt, A. F. et al. Prediction of higher-selectivity catalysts by computer-driven workflow and machine learning. *Science* **363**, eaau5631 (2019).
26. Hansen, E., Rosales, A. R., Tutkowski, B., Norrby, P.-O. & Wiest, O. Prediction of Stereochemistry using Q2MM. *Acc. Chem. Res.* **49**, 996–1005 (2016).
27. Rosales, A. R. et al. Application of Q2MM to predictions in stereoselective synthesis. *Chem. Commun.* **54**, 8294–8311 (2018).
28. Rosales, A. R. et al. Transition state force field for the asymmetric redox-relay heck reaction. *J. Am. Chem. Soc.* **142**, 9700–9707 (2020).
29. Le, D. N. et al. Hydrogenation catalyst generates cyclic peptide stereocentres in sequence. *Nat. Chem.* **10**, 968–973 (2018).
30. Rosales, A. R. et al. Rapid virtual screening of enantioselective catalysts using CatVS. *Nat. Catal.* **2**, 41–45 (2019).
31. Hagelin, H., Åkermark, B. & Norrby, P.-O. New molecular mechanics (MM3*) force field parameters for calculations on (η^3 -Allyl) palladium complexes with nitrogen and phosphorus ligands. *Organometallics* **18**, 2884–2895 (1999).
32. Hagelin, H., Svensson, M., Åkermark, B. & Norrby, P.-O. Molecular mechanics (MM3*) force field parameters for calculations on palladium olefin complexes with phosphorus ligands. *Organometallics* **18**, 4574–4583 (1999).
33. Norrby, P.-O., Åkermark, B., Haeffner, F., Hansson, S. & Blomberg, M. Molecular mechanics (MM2) parameters for the (η^3 -allyl) palladium moiety. *J. Am. Chem. Soc.* **115**, 4859–4867 (1993).
34. Donoghue, P. J., Helquist, P., Norrby, P.-O. & Wiest, O. Development of a Q2MM force field for the asymmetric rhodium catalyzed hydrogenation of enamides. *J. Chem. Theory Comput.* **4**, 1313–1323 (2008).
35. Lime, E. et al. Stereoselectivity in asymmetric catalysis: the case of ruthenium-catalyzed ketone hydrogenation. *J. Chem. Theory Comput.* **10**, 2427–2435 (2014).
36. Mazuela, J., Pàmies, O. & Diéguez, M. A new modular phosphite-pyridine ligand library for asymmetric Pd-catalyzed allylic substitution reactions: a study of the key Pd- π -allyl intermediates. *Chem. Eur. J.* **19**, 2416–2432 (2013).
37. Diéguez, M. & Pàmies, O. Modular phosphite–oxazoline/oxazine ligand library for asymmetric Pd-catalyzed allylic substitution reactions: scope and limitations—origin of enantioselectivity. *Chem. Eur. J.* **14**, 3653–3669 (2008).
38. Liu, Q. L. et al. A D-camphor-based schiff base as a highly efficient N,P ligand for enantioselective palladium-catalyzed allylic substitutions. *ChemCatChem* **8**, 1495–1499 (2016).
39. Zhao, Q., Zhuo, C.-X. & You, S.-L. Enantioselective synthesis of N-allylindoles via palladium-catalyzed allylic amination/oxidation of indolines. *RSC Adv.* **4**, 10875–10878 (2014).
40. Magre, M., Biosca, M., Norrby, P. O., Pàmies, O. & Diéguez, M. Theoretical and experimental optimization of a new amino phosphite ligand library for asymmetric palladium-catalyzed allylic substitution. *ChemCatChem* **7**, 4091–4107 (2015).
41. Popa, D., Marcos, R., Sayalero, S., Vidal-Ferran, A. & Pericas, M. A. Towards continuous flow, highly enantioselective allylic amination: ligand design, optimization and supporting. *Adv. Synth. Catal.* **351**, 1539–1556 (2009).
42. Borràs, C. et al. Amino-P ligands from iminosugars: new readily available and modular ligands for enantioselective Pd-catalyzed allylic substitutions. *Organometallics* **37**, 1682–1694 (2018).
43. Wang, Y., Vaismaa, M. J. P., Hämäläinen, A. M., Tois, J. E. & Franzén, R. Utilization of IndPHOX-ligands in palladium-catalysed asymmetric allylic aminations. *Tetrahedron Asymmetry* **22**, 524–529 (2011).
44. Dauber-Osguthorpe, P. & Hagler, A. T. Biomolecular force fields: where have we been, where are we now, where do we need to go and how do we get there? *J. Comput. Aided. Mol. Des.* **33**, 133–203 (2019).
45. Frisch, M. J. et al. *Gaussian 16 Rev. B.01* (2016).
46. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120**, 215–241 (2008).
47. Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **132**, 154104 (2010).
48. Breneman, C. M. & Wiberg, K. B. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comput. Chem.* **11**, 361–373 (1990).
49. Allinger, N. L., Yuh, Y. H. & Li, J. H. Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.* **111**, 8551–8566 (1989).
50. Schrödinger, Inc. MacroModel Release 2018-3 (Schrödinger, LLC, 2018).

Acknowledgements

This work was supported financially by NSF (CHE-1855900) and AstraZeneca (O.W.). M.D. and O.P. thank the Spanish Ministry of Science and Innovation (PID2019-104904GB-I00) and the Catalan Government (2017SGR1472). We would like to thank Prof. Björn Åkermark who, together with P.H., gave P.-O.N. the challenge to computationally predict selectivity in Pd-catalyzed allylation reactions in 1986.

Author contributions

E.H. and P.-O.N. wrote the code, J.W. and A.B. performed calculations, J.M., M.D., and O.P. performed experiments. P.H., P.-O.N. and O.W. designed the study, and all authors analyzed the data and contributed to the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-021-27065-2>.

Correspondence and requests for materials should be addressed to Olaf Wiest or Per-Ola Norrby.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons

Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021