

A Quantum-Guided Molecular Mechanics Force Field for the Ferrocene Scaffold

Jessica Wahlers, Anthony R. Rosales, Neil Berkel, Aaron Forbes, Paul Helquist, Per-Ola Norrby, and Olaf Wiest*



Cite This: *J. Org. Chem.* 2022, 87, 12334–12341



Read Online

ACCESS |



Metrics & More

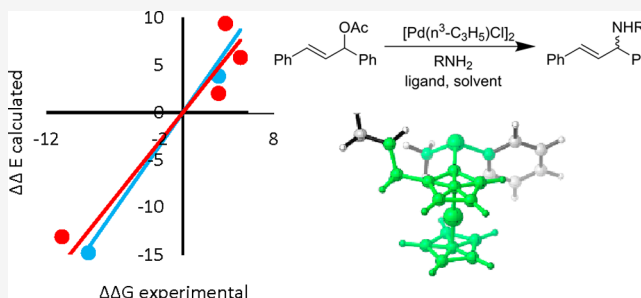


Article Recommendations



Supporting Information

ABSTRACT: Ferrocene derivatives have a wide range of applications, including as ligands in asymmetric catalysis, due to their chemical stability, rigid backbone, steric bulk, and ability to encode stereochemical information via planar chirality. Unfortunately, few of the available molecular mechanics force fields incorporate parameters for the accurate study of this important building block. Here, we present a MM3* force field for ferrocenyl ligands, which was generated using the quantum-guided molecular mechanics (Q2MM) method. Detailed validation by comparison to DFT calculations and crystal structures demonstrates the accuracy of the parameters and uncovers the physical origin of deviations through excess energy analysis. Combining the ferrocene force field with a force field for Pd–allyl complexes and comparing the crystal structures shows the compatibility with previously developed MM3* force fields. Finally, the ferrocene force field was combined with a previously published transition-state force field to predict the stereochemical outcomes of the aminations of Pd–allyl complexes with different amines and different chiral ferrocenyl ligands, with an R^2 of ~ 0.91 over 10 examples.



INTRODUCTION

Ferrocene and its many derivatives are many-faceted molecules with applications in supramolecular, medicinal, bioinorganic, and materials chemistry due to their unique combination of ease of synthesis, stability, redox activity, and structure, making them some of the most important organometallic molecules.^{1,2} Ferrocene's rigid backbone, steric bulk, and potential for planar chirality³ also make it an attractive scaffold for ligands in enantioselective catalysis. As a result, ferrocene ligands have been used in many reactions involving hydrogenation, addition to alkenes, allylic substitution, and cross-coupling reactions.^{3,4}

In recent years, the field of computational chemistry has moved from the rationalization of experimental results to the prediction of new reactions⁵ and, in some cases, the correction of reports of experimental observations in the literature.^{6,7} For example, computational approaches have been highly successful in the prediction of stereochemical outcomes of transition-metal-catalyzed reactions.^{8–10} The application of these powerful computational methods to ferrocene-containing systems is hindered by the lack of accurate parameters for the ferrocene moiety in most literature force fields (FFs). Although electronic structure methods such as density functional theory (DFT) can be used to get accurate structures of the intermediates and transition structures (TS), these methods are too slow for performing large ligand and substrate screens or sampling the conformational space^{11,12} of the species involved.

One possible approach to accelerate computational studies to a point that they are useful for guiding experimental studies is the use of surrogate methods such as ground-state or transition-state force fields (GSFF and TSFF, respectively) that are fitted toward electronic structure calculations at an appropriate level of theory. We have developed the quantum-guided molecular mechanics (Q2MM) method to automate the fitting of FF parameters to simplified model systems optimized by DFT.^{8,9} We were able to demonstrate that the resulting FFs combine the accuracy of electronic structure methods with the speed of molecular mechanics. It should be noted that because the FFs are fitted to small model systems rather than the full systems and do not use any information from experiments in the fitting process, the stereoselectivity calculations that use them are true predictions. Here, we present the application of the Q2MM method to fit a GSFF for ferrocenyl compounds using the functional form of the widely used MM3* FF.^{13–15} We also demonstrate the application of the resulting GSFF to the prediction of the structure and relative energy of a range of ferrocene derivatives.

Received: July 4, 2022

Published: September 6, 2022



Finally, we study the combination of this method with other previously published GSFFs and TSFFs for applications including stereoselectivity predictions. The physical origin of the deviations between experimental and computational results is also discussed.

RESULTS AND DISCUSSION

Topology of the GSFF. The unique bonding features of ferrocenes require a careful consideration of the topology of the FF parameters for the sandwich complex. Approaches previously used to represent ferrocene in FFs include rigid body,¹⁶ nonbonded,¹⁷ σ -bonding,¹⁸ and dummy-bonding^{19–22} topologies, as shown in Figure 1. The rigid body approach

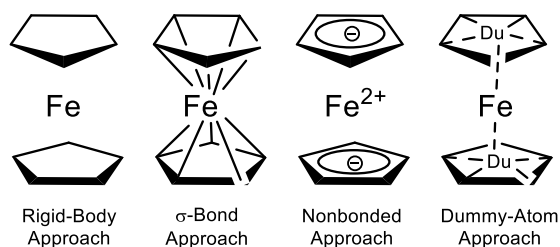


Figure 1. Different MM methods for modeling the ferrocene scaffold.

involves holding the bonds and angles at fixed values.¹⁶ This method does not require parametrization but does not accurately capture the interactions between the metal and the cyclopentadienyl rings (Cp).²³ The nonbonded approach involves holding the molecule together through nonbonded electrostatic and van der Waals interactions.²⁴ This method requires only the parametrization of the charges but can lead to both high-energy contributions from these charges and accidental dissociation in the conformational exploration, again affecting the accuracy and transferability of the results. The σ -bonding approach involves placing a σ -bond between the Fe and the carbons of the Cp rings.^{18,25} This approach is widely applicable to many different metallocene complexes, but this method is difficult to parametrize due to the generation of many three-membered rings²⁶ and does not allow low-barrier rotation of the rings, which has been observed for many ferrocenes.²² The dummy-bonding approach tends to be more accurate when compared to other methods. It involves placing a dummy atom at the center of the Cp rings and connecting the dummy atom to each of the carbons in the Cp ring with a zero-order bond while the Fe atom is connected to the dummy atom with a σ -bond.^{20,27} This method is physically more realistic, as it keeps the Cp rings close to 180° from each other and allows the unhindered rotation of the Cp rings, but can be difficult to parametrize due to the required additional parameters for the dummy atom.²⁸ However, the Q2MM method is well-suited to the generation of such parameters. We therefore decided to use the dummy atom approach to develop a GSFF for the ferrocene scaffold that could be used for a diverse ligand library.

Using the Q2MM method,²⁹ parameters are automatically fit to reproduce the geometry and Hessian elements of a DFT-optimized reference data set for a given reaction.^{8,9} The goal is to develop GSFF parameters that describe ferrocene and can be paired with standard FF parameters to provide a highly accurate treatment of ferrocene-based ligands. The accuracy of the FF in comparison to DFT and experimental data is analyzed in detail. Finally, we will show how this GSFF can be

combined with an existing GSFF for palladium-catalyzed allylic substitutions.

Training Set for Ferrocenes. To develop a FF using the Q2MM method, a training set is needed to fit the added FF parameters. These structures are used to capture the steric and electronic interactions in the system. The geometries and eigenvalues of the Hessian Matrix of these structures are used to fit the FF parameters by minimizing a penalty function described previously.^{8,9}

A training set of seven structures consisting of common classes of bisphosphine ferrocene ligands, other substituted ferrocene derivatives, and an unsubstituted ferrocene (Figure 2) were calculated at the B3LYP-D3/LanL2DZ/6-31G** level

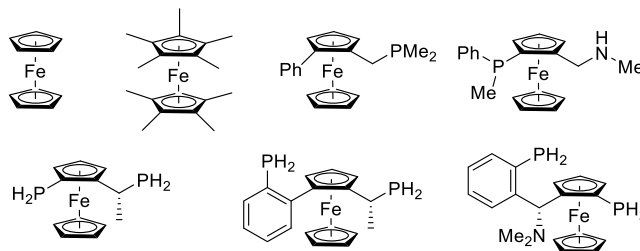


Figure 2. DFT-optimized training set used to fit the parameters for the ferrocene scaffold.

of theory. These structures were chosen to capture the geometries and energies of the Cp rings around the iron center. The full ligand structures were added to the training set to ensure that bonds and angles connecting the ferrocene moiety to the rest of the ligands were accurately depicted by the FF. The bonds, angles, and torsions directly involved with the Cp ring were parametrized in the FF.

Some of the MM parameters needed to be defined with higher accuracy to capture the rotations of the Cp rings. To obtain the needed training data, four different scans were added to the training set (Figure 3). These scans consisted of

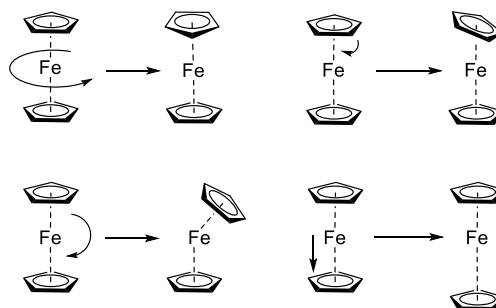


Figure 3. Additional DFT scans of the ferrocene moiety.

the dihedral rotation of the Cp rings, the angle bending of the Cp ring at the Fe center, the angle bending of the Cp rings at the dummy center, and the bond stretching between the dummy atom and Fe.

Parameterization of the Ferrocene Moiety. To develop predictive methods for the ferrocene-based chiral ligands, FF parameters for the ferrocene moiety within a ligand needed to be developed. Standard FFs do not contain such parameters, yet they should be compatible with available FFs. Only the interactions containing the atoms that comprise ferrocene and the atoms bonded to ferrocene (shown in green in Figure 4) are parametrized, while all other interactions are described by

the underlying force field, which in the case discussed here is the MM3* FF.^{13–15,26}

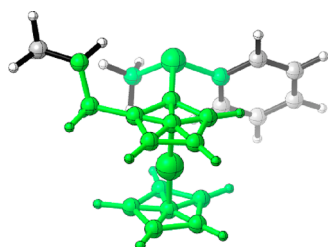


Figure 4. An example of a ferrocenyl ligand, where atoms fit in the FF are shown in green.

Four substructure sections in the MM3* format were developed to describe ferrocene. The first section to be fitted was the ferrocene core comprised of the iron, dummy atoms, and atoms in the Cp ring. The initial parameters for the equilibrium bond lengths and angle values were set by taking the averages of the values in the training set structures. Bond and angle force constants were set to an estimated value between 0.5 and 1.0 based on the strength or weakness of the bonds and angles. The bond dipoles and torsions were initially set to zero at the beginning of the parametrization. Many of the angles that contained the dummy atom were set to the ideal values, and the force constant was set to zero. These parameters were not optimized in the FF.

According to the general Q2MM procedure,⁹ bond dipoles were the first parameters to be optimized in the FF. The next set of parameters to be optimized was the bond and angle force constants. To parametrize the ideal bond lengths and angle values, they need to be tethered and parametrized with the bond and angle force constants to ensure that they did not deviate too much from the average values. Once the equilibrium values were fitted, the force constants were reoptimized. The results of the scans shown in Figure 3 were then incorporated into the optimization to fit the bonds and angles around the dummy atom and iron. The torsions were the last set of parameters optimized. A skipped torsion was added to define the rotation of the Cp rings. This parameter was added for the $C_{cp}-D-D-C_{cp}$ interaction to capture the energies of the staggered and eclipsed conformations of the Cp rings. A V5 torsional term, which is typically defined to describe the rotation of five-membered rings, was added to describe this rotation, and the scans described above were used to fit this parameter.

The next three substructure sections of the FF were used to fit the different types of ligands. The ligand sections differ in the type of atom connected to the Cp ring of the ferrocene moiety. The connecting atoms included an sp^3 -hybridized carbon atom, an sp^2 -hybridized carbon atom, and a phosphorus atom. These sections were optimized in the same manner as described above.

Once the FF was fully optimized, it was tested to see how well it is able to reproduce the reference structures. The bond lengths, angles, torsions, and Hessian eigenvalues were calculated by the MM FF and compared to data for the DFT reference structures. In all these plots, the slope and R^2 value were close to one (see Figure S1 in the Supporting Information). Significantly, there are no systematic major deviations in the DFT and MM structures. These results match those obtained for our previous FFs, where the interval

validation graphs had slopes of close to one for the bond lengths, bond angles, torsions, and Hessian eigenvalues.

External Validation. With the FF able to reproduce the training set structures well, a validation set of six chiral ligands (Figure 5) was used to test how well the FF reproduced

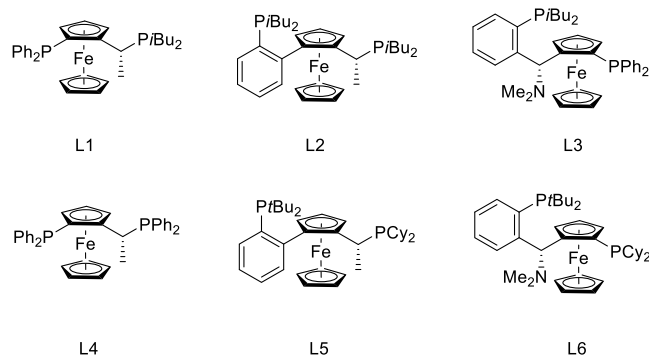


Figure 5. Validation set structures used to compare the relative energies of DFT- and MM-optimized structures.

structures different from the training set. The chiral carbon centers of these six structures were also inverted, which led to the generation of a diastereomer due to planar chirality. To test the performance of the FF, the energies of the diastereomeric structures were calculated and compared to the energies calculated by DFT methods. A Monte Carlo search in MacroModel was performed to find the lowest-energy conformer of each of the structures. The five lowest-energy conformations were then reoptimized at the B3LYP-D3/LanL2DZ/6-31G** level of theory. A vibrational analysis was then performed to confirm the structures as minima on the potential energy surface.

To compare the energy difference of the diastereomers of the six structures to the energy difference calculated with the FF, the lowest-energy conformation of the DFT-optimized structures was subjected to a MM optimization. This allowed the heavy atoms to slightly relax and avoided large energy contributions from small bond deviations while remaining structurally close to the DFT-calculated structure. The energy difference was then calculated for the diastereomers of the six ligands and compared between the DFT and MM methods (Table 1).

Table 1. Relative Energies of Ligand Diastereomers Calculated with QM and MM Validation Set Methods

ligand	$\Delta E^{MM,opt}$ (kJ/mol)	$\Delta E^{QM,opt}$ (kJ/mol)
L1	6.5	6.8
L2	7.3	9.2
L3	−3.3	−4.5
L4	5.7	4.0
L5	6.8	0.8
L6	1.9	−13.1

The relative energies of the diastereomers for L1–L4 calculated by DFT and MM were similar, with a mean unsigned error (MUE) of 1.3 kJ/mol. When comparing the structures calculated with the two methods, there is little deviation around the ferrocene core. There was a significant energy difference between the DFT- and MM-optimized diastereomers of L5 and L6, with MUEs of 6.6 and 14.9 kJ/

mol, respectively. Analyzing the largest errors between DFT and MM can give insights into the limits of the MM3 FF as well as any systematic errors that are present in the literature FF.^{6,7} The two ligands with the largest deviations, L5 and L6, contain cyclohexyl substituents on the phosphorus atom. There was a significant deviation in the $C_{cp}-C_{sp^3}-P$ angle in (S)-L5 between the DFT- and MM-optimized structures, which had values of 114.2° and 121.8°, respectively (Figure 6).

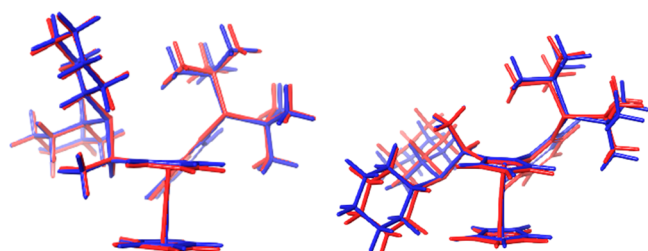


Figure 6. Lowest-energy conformers of (R)-L5 (left) and (S)-L5 (right) optimized by MM (blue) and DFT (red) methods.

In the MM-optimized structure, that angle is larger than the reference angle (108.7°), which causes a significant change in the calculated energy. In contrast, (R)-5 exhibits little deviation between the MM- and QM-optimized structures. We hypothesize that in the ligand with the (S)-configuration, the distortion of the angle is caused by a problem in the nonbonded substructure section in which the steric repulsion between the cyclohexyl group and the Cp moiety is overestimated, a known error in the nonbonded substructure section of the MM3 FF.³⁰ This suggests that the origin of the deviation in the calculated geometries is the underlying MM3 GSFF parameters and not the Q2MM-fitted parameters describing the ferrocene moiety.

Validation with Crystal Structures. Another method of validating the structures produced by the FF is to compare them to the crystal structures, which allows for an assessment of how well the force field can reproduce structures obtained experimentally. To allow for a validation relevant to catalysis, the ferrocene FF was combined with a previously developed FF that describes a Pd–allyl complex.^{31,32} A search of the Cambridge Structural Database (CSD) provided 19 crystal structures containing both Pd–allyl and ferrocene substructures, which were used as a validation set to compare structures calculated by the FF to experimental data. It should be noted that the computed structures do not account for crystal packing forces or crystals containing different conformations.

The 19 crystal structures were minimized by both DFT and MM methods and compared to a frozen or constrained structure of the same method. The excess energy, which is defined as the energy difference between an experimentally observed structure and a computationally optimized structure,³³ was calculated. Since all crystal structures contain small distortions due to, for example, crystal packing forces, the force field will give a positive excess energy. If the computational optimized structure and the crystal structure are close in energy, the excess energy is low. If the excess energy is high, there is a deviation between the crystal structure and the computationally optimized structure. The excess energy approach has been used for validation purposes in previous studies of fitted MM3* parameters for Rh–polypyridyl complexes³³ and Pd–allyl complexes.³¹

The energy differences for the 19 crystal structures were calculated between the constrained or frozen structures and the fully optimized structures (Table 2). The excess energies were

Table 2. Excess Energies of Crystal Structures Calculated with QM and MM Methods^a

CSD code	R factor	excess energy (QM)	excess energy (MM, frozen atoms)	excess energy (MM, constrained atoms)
ZOMYOP	2.60%	97.8	147.4	53.9
BAGLEB	5.41%	265.2	332.4	84.1
ECOYUR	4.48%	53	65.2	44.9
EREZOQ	5.61%	59.2	99.7	56.4
HOHRUS	4.66%	136.5	166.1	61.7
HOHSAZ	4.64%	39.2	69.7	29.6
HOHSED	6.47%	428.4	492.7	170.9
HOHSIH	5.47%	56.9	72.5	32.8
NIBVUP	5.23%	120	186.2	83.1
OGOPOP	4.00%	142.6	152.7	51.5
OJJJUN	2.69%	57.6	128.4	54.6
PUZTEJ	4.30%	146	162.6	55.9
PUZTIN	4.90%	165.1	137.4	39.0
UBOVUC	3.20%	74.5	105.3	54.9
UBOWAJ	3.10%	96.5	113.9	45.9
UDACAE	2.71%	25.7	68.7	46.2
UDIFIW	3.57%	46.4	87.9	46.2
YEPZAV	3.66%	103.5	141.0	48.5
YEPZEZ	5.90%	193.6	195.9	60.4

^aAll energies are in kilojoules per mole (kJ/mol).

calculated when the structures were constrained for both QM and MM, i.e., only the hydrogen atoms were allowed to optimize. In these cases, the QM structures had lower excess energies than the MM structures but were consistent in the relative energy ordering. For the MM values, the calculations were also modified to allow the heavy atoms to optimize with a constraint of a 500 kJ/mol force constant while still allowing the hydrogen atoms to optimize. The excess energies were greatly reduced and closer in energy to the excess energies calculated with QM.

A small number of structures exhibited high excess energies in both the QM and MM calculations, which is often indicative of errors in the crystal structure; this is in line with the observation that the structures with the highest R factors also tend to have the highest excess energies, as shown in Table 2. Two common types of errors are seen in these crystal structures. The first error is caused by two conformations of the allyl group overlapping in the same crystal, causing errors in the bond lengths and angles of the allyl moiety. The two structures with the largest errors were in the crystal structures were BAGLEB and HOHSED, which contained errors in the allyl substructure. This includes an allylic bond length of 1.19 Å in the BAGLEB crystal structure and an allylic angle of 144.5° in the HOHSED structure (Figure 7 left), both of which deviate significantly from the expected values for allylic systems. Another common error observed in these crystal structures is large alternations in the bond lengths of aromatic groups due to libration. The atomic positions in crystal structures are averages over several cocrystallizing rotameric structures. When the averaged positions lie on a curve, the apparent distances between such averaged positions will be different from the actual bond lengths, generally shorter. As

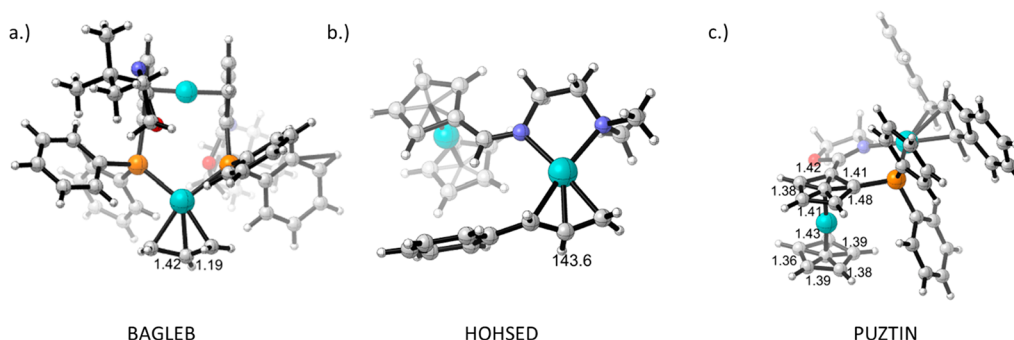


Figure 7. Crystal structures with large excess energies.

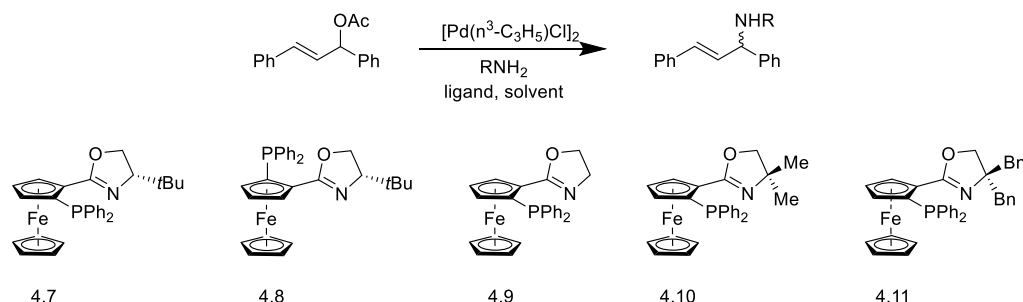


Figure 8. Validation set used for the stereoselectivity predictions of the Pd-catalyzed allylic amination reaction catalyzed by ligands containing the ferrocene moiety.

examples, the bond lengths in the two Cp rings of the ferrocene group in PUZTIN ranged from 1.36 to 1.48 Å (Figure 7 right), whereas the bond lengths in both DFT and MM3* structures are much more equal. This suggests that the large excess energy values observed for the outliers in Table 2 are due to errors in the crystal structure and not errors in the MM3* parameters derived to describe the ferrocene scaffold.

Stereoselectivity Predictions. As a first test of the applicability of the ferrocene FF for stereoselectivity predictions for asymmetric reactions catalyzed by ferrocenyl ligands, it was combined with the Pd-allyl GSFF, which was shown previously to provide qualitative predictions of the stereochemical outcome of these reactions.³⁴ A set of five Pd-allyl ground-state structures catalyzed by ligands containing the ferrocene moiety was constructed (Figure 8).³⁵ These ligands are similar to the ligands with which the FF was trained in that they contain a similar P–N ligand backbone but differ in that they contain a pyridine ring, which was not present in the training set for these ferrocenyl ligands. Therefore, these studies are a case study for the robustness of the parameters. A Monte Carlo multiple-minimum search was performed and minimized on the combined Pd-allyl and ferrocene GSFF. A benzyl amine nucleophile was then added to the lowest-energy conformer to form the diastereomeric transition states that lead to the (*R*)- and (*S*)- enantiomers. These structures were then optimized to a TS using the same DFT method described above.

The energy difference between the (*R*)- and (*S*)-enantiomers was calculated and compared to the experimental selectivity (Figure 9). The overall MUE for the five structures was 5.6 kJ/mol. There is a significant error for L11 which contains two benzyl groups on the oxazoline. This causes a significant steric interaction with the phenyl group on the allyl moiety. Because a GSFF was used instead of a TSFF, the selectivity predictions are not expected to be quantitatively

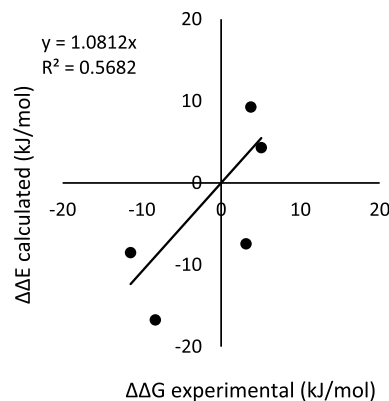


Figure 9. Comparison of the energy differences of the TS structures that lead to (*R*) and (*S*)-enantiomers between the DFT-optimized structures and the experimental results.

accurate; however, they do allow the full conformational search of the transition state needed to make accurate stereoselectivity predictions¹¹ and provide insights into the steric and electronic interactions of the transition state.

As a more quantitative application of the ground-state ferrocene FF for stereoselectivity predictions, it was paired with a TSFF for the attack of nitrogen nucleophiles on a Pd-allyl complex for which we had previously reported successful predictions of enantioselectivity for other classes of chiral ligands.⁷ A mixed Monte Carlo multiple-minimum/low-mode search was performed for 40 000 steps for benzyl amine and NH₃ nucleophiles. The selectivities were calculated and compared to the experimental values, with an MUE of 5.6 kJ/mol for the benzyl amine nucleophiles and an MUE of 3.7 kJ/mol for the NH₃ nucleophile (Figure 10). While the MUE values between the two methods used to predict the selectivity are similar, the slope and R^2 values of the respective selectivity

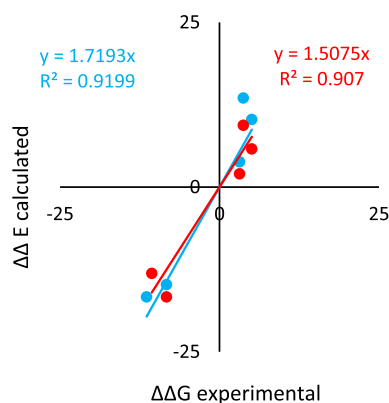


Figure 10. Comparison of the energy differences of the TS structures that lead to the (R)- and (S)-enantiomers between the MM-optimized structures and the experimental results for the benzylamine (blue) and NH_3 (red) nucleophiles.

comparisons are different. When DFT methods were used to optimize the ground-state structures to transition-state structures to calculate the selectivity, the slope was 1.08 and $R^2 = 0.57$. (Figure 9). While the slope was close to one, the low R^2 value shows a modest correlation. When the TSFF was used to calculate the selectivity, the slopes were 1.72 and 1.51 and $R^2 = 0.92$ and 0.91 for the benzyl amine and the NH_3 nucleophiles, respectively (Figure 10). Although the slope is too high, i.e., selectivities will be overestimated, the data have a correlation coefficient similar to those of previous TSFFs.^{7–9} This suggests that the FF systematically overpredicts the selectivity but correctly rank-orders them to identify the most promising ligand for a given substrate.

The transition structure calculated for ligand 4.10 and the benzylamine nucleophile (Figure 11) had the largest deviation

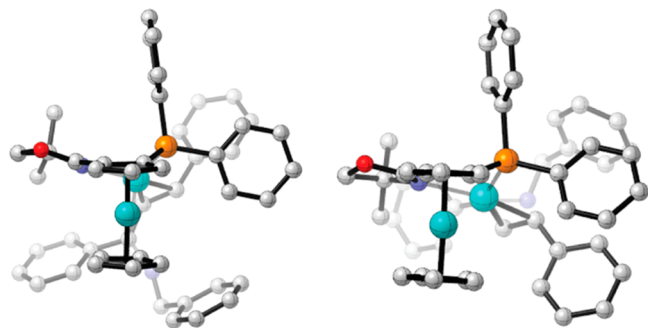


Figure 11. Lowest-energy conformations of the transition states leading to the (R)-product (left) and the (S)-product (right) for the amination of Pd-allyl with benzylamine using 4.10 as the ligand.

between the calculated and experimental selectivities and was analyzed further. The reported experimental selectivity is 60.3% ee,³⁵ while the calculated selectivity is 98.7% ee. When the lowest-energy conformers of the transition structures leading to the major (S)-enantiomer and the minor (R)-enantiomer were analyzed, the significant difference was in the aromatic interactions with the ferrocene. In the transition structure leading to the (S)-product, there were three aromatic groups surrounding the ferrocene moiety, while in the TS leading to the (R)-product there were two aromatic systems (Figure 11). This would suggest that the nonbonded energy for the transition state leading to the (S)-product should be

higher in energy. However, a comparison of the contributions to the MM energy between the two lowest energy conformers shows that the van der Waals energy is 7 kJ/mol lower in the TS leading to the (S)-product. This is most likely due to the underlying error in the MM3 GSFF used to describe the nonbonded interactions in aromatic systems. Despite this outlier, calculations using the ferrocene FF combined with the TSFF for the Pd-allyl reaction will allow the rapid identification of promising ligand–substrate combinations.

CONCLUSIONS

Force field parameters were developed to describe the ferrocenyl moiety that could be paired with other FF parameters. This FF was able to accurately reproduce the crystal structures of 19 Pd-allyl complexes. While there were some deviations from the crystal structure and the MM-optimized structures, most were due to errors in the crystal structures. Stereoselectivity predictions were made by combining the ferrocene FF with both a Pd-allyl GSFF for the reactive intermediate and a Pd-allyl TSFF. The MUE for both methods was 5.6 kJ/mol. This shows the FF systematically but consistently overestimates the selectivity but is still useful for the rapid selection of ligands for use in enantioselective reactions.

COMPUTATIONAL METHODS

All DFT calculations were performed in the gas phase using Gaussian 09.³⁶ The B3LYP functional with the empirical dispersion correction³⁷ was used with a LanL2DZ basis set to describe iron and palladium atoms, and 6-31G** was used to describe all other atoms. The CHELPG³⁸ scheme was used to calculate partial charges. Frequency analysis confirmed ground-state or transition-state structures. Dummy atoms were placed at the centers of the Cp rings after the QM optimizations were completed.

The ferrocene parameters were fit and optimized using the automated fitting procedure of the Q2MM method. The MM3* FF^{13–15} was used as the functional form to fit to and for default parameters that were not fit by the Q2MM method. All molecular mechanics (MM) calculations were completed using Macro Model.³⁹

Energy excess calculations^{31,33} were performed by having both frozen and constrained heavy atoms to allow hydrogen positions to optimize. These calculations were performed using both DFT and MM methods. The reported MM excess energies were calculated by comparing a single point calculation of the constrained and frozen structures. The constrained approach implemented a 500 kJ/mol force constant on heavy atoms.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.joc.2c01553>.

Coordinates, point charges, and energies for DFT- and MM-optimized structures; optimized TSFF parameters in the MM3 format; and validation plots (PDF)

AUTHOR INFORMATION

Corresponding Author

Olaf Wiest – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States; orcid.org/0000-0001-9316-7720; Email: owiest@nd.edu

Authors

Jessica Wahlers – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States

Anthony R. Rosales – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States; orcid.org/0000-0002-9262-4325

Neil Berkel – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States

Aaron Forbes – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States

Paul Helquist – Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, Indiana 46556, United States; orcid.org/0000-0003-4380-9566

Per-Ola Norrby – Data Science and Modelling, Pharmaceutical Sciences, R&D, AstraZeneca Gothenburg, Mölndal SE-431 83, Sweden; orcid.org/0000-0002-2419-0705

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acs.joc.2c01553>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of this research from the National Science Foundation (CHE-1855908) and AstraZeneca. A.R.R. received support through a CBBI Fellowship (NIH, T32 GM075762).

REFERENCES

- (1) Larik, F. A.; Saeed, A.; Fattah, T. A.; Muqadar, U.; Channar, P. A. Recent advances in the synthesis, biological activities and various applications of ferrocene derivatives. *Appl. Organomet. Chem.* **2017**, *31*, e3664.
- (2) Wilkinson, G.; Rosenblum, M.; Whiting, M. C.; Woodward, R. B. The structure of iron bis-cyclopentadienyl. *J. Am. Chem. Soc.* **1952**, *74*, 2125–2126.
- (3) Dai, L. X.; Tu, T.; You, S. L.; Deng, W. P.; Hou, X. L. Asymmetric catalysis with chiral ferrocene ligands. *Acc. Chem. Res.* **2003**, *36*, 659–667.
- (4) Gómez Arrayás, R.; Adrio, J.; Carretero, J. C. Recent applications of chiral ferrocene ligands in asymmetric catalysis. *Angew. Chem. Int. Ed.* **2006**, *45*, 7674–7715.
- (5) Houk, K. N.; Liu, F. Holy grails for computational organic chemistry and biochemistry. *Acc. Chem. Res.* **2017**, *50*, 539–543.
- (6) Rosales, A. R.; Ross, S. P.; Helquist, P.; Norrby, P.-O.; Sigman, M. S.; Wiest, O. Transition state force field for the asymmetric redox-relay heck reaction. *J. Am. Chem. Soc.* **2020**, *142* (21), 9700–9707.
- (7) Wahlers, J.; Margalef, J.; Hansen, E.; Bayesteh, A.; Helquist, P.; Diéguez, M.; Pàmies, O.; Wiest, O.; Norrby, P.-O. Proofreading experimentally assigned stereochemistry through Q2MM predictions in Pd-catalyzed allylic aminations. *Nature Comm.* **2021**, *12*, 6719.
- (8) Hansen, E.; Rosales, A. R.; Tutkowski, B.; Norrby, P. O.; Wiest, O. Prediction of Stereochemistry using Q2MM. *Acc. Chem. Res.* **2016**, *49*, 996–1005.
- (9) Rosales, A. R.; Quinn, T. R.; Wahlers, J.; Tomberg, A.; Zhang, X.; Helquist, P.; Wiest, O.; Norrby, P. O. Application of Q2MM to predictions in stereoselective synthesis. *Chem. Commun.* **2018**, *54*, 8294–8311.
- (10) Reid, J. P.; Sigman, M. S. Comparing quantitative prediction methods for the discovery of small-molecule chiral catalysts. *Nature Rev. Chem.* **2018**, *2*, 290–305.
- (11) Lee, J. M.; Zhang, X.; Norrby, P. O.; Helquist, P.; Wiest, O. Stereoselectivity in (acyloxy)borane-catalyzed mukaiyama aldol reactions. *J. Org. Chem.* **2016**, *81*, 5314–5321.
- (12) Xu, L.; Zhang, X.; McCamant, M. S.; Sigman, M. S.; Wu, Y. D.; Wiest, O. Mechanism and Selectivity in the Pd-Catalyzed Difunctionalization of Isoprene. *J. Org. Chem.* **2016**, *81*, 7604–7611.
- (13) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular mechanics. The MM3 force field for hydrocarbons. 1. *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
- (14) Lii, J. H.; Allinger, N. L. Molecular mechanics. The MM3 force field for hydrocarbons. 2. Vibrational frequencies and thermodynamics. *J. Am. Chem. Soc.* **1989**, *111*, 8566–8575.
- (15) Lii, J. H.; Allinger, N. L. Molecular mechanics. The MM3 force field for hydrocarbons. 3. The van der Waals' potentials and crystal data for aliphatic and aromatic hydrocarbons. *J. Am. Chem. Soc.* **1989**, *111*, 8576–8582.
- (16) Shubina, E. S.; Esptein, L. M.; Timofeeva, T. V.; Struchkov, Y. T.; Kreindlin, A. Z.; Fadeeva, S. S.; Rybinskaya, M. I. Intramolecular hydrogen bonds and conformations of ferrocenyl- and non-amethylferrocenylcarbinols. *J. Organomet. Chem.* **1988**, *346*, 59–66.
- (17) De Hatten, X.; Cournia, Z.; Huc, I.; Smith, J. C.; Metzler-Nolte, N. Force-field development and molecular dynamics simulations of ferrocene-peptide conjugates as a scaffold for hydrogenase mimics. *Chem.-Eur. J.* **2007**, *13*, 8139–8152.
- (18) Rudziński, J. M.; Ōsawa, E. Molecular mechanics of bridged ferrocene derivatives: Conformational energy surfaces of [3]-, [4]- and [45] ferrocenophanes. *J. Phys. Org. Chem.* **1992**, *5*, 382–394.
- (19) Mackie, S. C.; Baird, M. C. Investigation of the Applicability of Molecular Mechanics (MMX) Calculations to a Conformational Investigation of Organometallic Compounds of the Types (η^5 -C₅H₅)Fe(CO)2CH₂R, (η^5 -C₅H₅)Fe(CO)2C(=O)R, and (η^5 -C₅H₅)Fe(CO)(PPh₃)R (R = H, Alkyl, Acyl). *Organometallics* **1992**, *11*, 3712–3724.
- (20) Bosnich, B. Molecular mechanics force fields for cyclopentadienyl complexes. *Chem. Soc. Rev.* **1994**, *23*, 387–395.
- (21) Johnston, B. F.; Rankin, D. W. H.; Robertson, H. E.; Hughes, R. P.; Lompfrey, J. R. Molecular structure of Ru(η^5 -C₅Me₅)(η^5 -C₅F₅) by gas-phase electron diffraction and density functional theory. *Organometallics* **2002**, *21*, 4840–4846.
- (22) Morrison, C. A.; Bone, S. F.; Rankin, D. W. H.; Robertson, H. E.; Parsons, S.; Coxall, R. A.; Fraser, S.; Howell, J. A. S.; Yates, P. C.; Fey, N. Conformational properties of substituted ferrocenes: Experimental and theoretical studies of the molecular structures of 1,1'-di-tert-butylferrocene and isopropylferrocene. *Organometallics* **2001**, *20*, 2309–2320.
- (23) Blom, R.; Hammel, A.; Haaland, A.; Weidlein, J.; Timofeeva, T. V.; Struchkov, Y. T. Molecular structures of tris-(methylcyclopentadienyl)-scandium and -ytterbium as studied by gas phase electron diffraction and molecular mechanics calculations: the scandium atom is too small to accommodate three pentahapto cyclopentadienyl rings. *J. Organomet. Chem.* **1993**, *462*, 131–139.
- (24) Menger, F. M.; Sherrod, M. J. Docking Calculations on Ferrocene Complexation with Cyclodextrins. *J. Am. Chem. Soc.* **1988**, *110*, 8606–8611.
- (25) Rudziński, J. M.; Ōsawa, E. Conformational interconversion of [1.1] ferrocenophane: A molecular mechanics study. *J. Phys. Org. Chem.* **1993**, *6*, 107–112.
- (26) Timofeeva, T. V.; Lii, J. H.; Allinger, N. L. Molecular Mechanics Explanation of the Metallocene Bent Sandwich Structure. *J. Am. Chem. Soc.* **1995**, *117*, 7452–7459.
- (27) Doman, T. N.; Bosnich, B.; Landis, C. R. Molecular Mechanics Force Fields for Linear Metallocenes. *J. Am. Chem. Soc.* **1992**, *114*, 7264–7272.
- (28) Fey, N. Organometallic molecular modelling - The computational chemistry of metallocenes: A review. *J. Chem. Technol. Biotechnol.* **1999**, *74*, 852–862.

- (29) Hansen, E.Q2MM. Github, 2016. <https://github.com/Q2MM>.
- (30) Paton, R. S.; Goodman, J. M. Hydrogen bonding and π -stacking: How reliable are force fields? A critical evaluation of force field descriptions of nonbonded interactions. *J. Chem. Inf. Mod.* **2009**, *49*, 944–955.
- (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* **1999**, *18*, 2884–2895.
- (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* **1999**, *18*, 4574–4583.
- (33) Brandt, P.; Norrby, T.; Åkermark, B.; Norrby, P. O. Molecular Mechanics (MM3*) Parameters for Ruthenium(II)-Polypyridyl Complexes. *Inorg. Chem.* **1998**, *37*, 4120–4127.
- (34) Peña-Cabrera, E.; Norrby, P.-O.; Sjögren, M.; Vitagliano, A.; De Felice, V.; Oslob, J.; Ishii, S.; O'Neill, D.; Åkermark, B.; Helquist, P. Molecular mechanics predictions and experimental testing of asymmetric palladium-catalyzed allylation reactions using new chiral phenanthroline ligands. *J. Am. Chem. Soc.* **1996**, *118*, 4299–4313.
- (35) You, S.-L.; Hou, X.-L.; Dai, L.-X.; Yu, Y.-H.; Xia, W. Role of planar chirality of S, N-and P, N-ferrocene ligands in palladium-catalyzed allylic substitutions. *J. Org. Chem.* **2002**, *67*, 4684–4695.
- (36) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, rev. A1; Gaussian, Inc.: Wallingford, CT, 2009.
- (37) 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.* **2010**, *132*, 154104.
- (38) 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.* **1990**, *11*, 361–373.
- (39) *Macromodel*, Schrödinger release 2018-3; Schrödinger, LLC: New York, NY, 2018.