

Data-Driven Prediction of Circular Dichroism-Based Calibration Curves for the Rapid Screening of Chiral Primary Amine *ee* Values

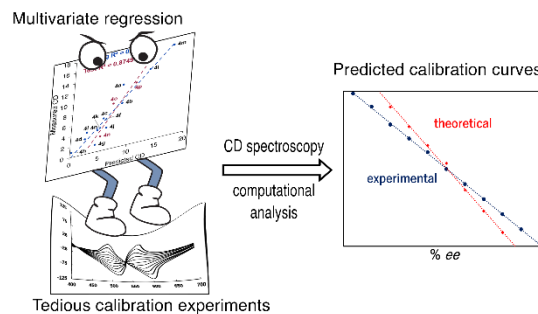
James R. Howard^a, Arya Bhakare^a, Zara Akhtar^a, Christian Wolf^{b,*}, and Eric V. Anslyn^{a,*}

^aDepartment of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States

^bDepartment of Chemistry, Georgetown University, Washington, DC 20057, United States

Supporting Information Placeholder

ABSTRACT: Here we describe the prediction of the circular dichroism (CD) response of a three-component chiroptical sensor for enantiomeric excess (*ee*) determination of chiral amines using a multivariate fit to electronic and steric parameters. These computationally derived parameters can be computed for nearly any amine and correlate well with the CD response of the 13 amines comprising the training set. The resulting model was used to accurately predict the CD response of a test set of chiral amines. Theoretical calibration curves were then created and used to determine the *ee* of solutions of unknown *ee*. Using this method, the error in *ee* determination differed by less than 10% compared to experimentally generated calibration curves.



Introduction

High-throughput experimentation (HTE) has become an integral part of reaction discovery and optimization.^{1–3} Asymmetric synthesis is an area in which HTE demonstrates prowess as it is used to screen thousands of reaction conditions exploring ligand, additive, solvent, and catalyst chemical space to afford rapid reaction optimization.⁴ Methods for conducting reactions on such scale are mature, and yet analysis of reaction outcome, such as yield determination, still relies heavily on serial chromatographic separation. Moreover, asymmetric synthesis requires chromatographic resolution of enantiomers to assess enantioselectivity, further slowing HTE analysis.⁵ Chiroptical methods address this bottleneck by enabling *ee* determination on timescales amenable to HTE.⁶ Chiroptical assays have been developed for numerous functional groups, and the techniques have proven useful in many reaction screening contexts.^{7–14} Asymmetric reaction analysis typically starts with the development and validation of a method that allows reliable *ee* determination of the anticipated product. If one chooses chiral HPLC for this purpose, one needs to obtain the reaction product in racemic form to search for a chiral stationary phase (CSP) and conditions that achieve sufficient enantioresolution. With that in hand, a chemist is able to determine *ee* values for that particular compound. If one resorts to chiroptical sensing, an enantiopure or enantioenriched reference is needed to construct a calibration curve with a suitable probe that interacts with the analyte to generate a quantifiable CD signal. In the end, both techniques require reference material before they can be developed and used, albeit a racemic sample is often more easily available than an enantiopure one. The workload increases further when the application scope of an asymmetric reaction protocol is under investigation – a very common scenario in synthetic meth-

odology development laboratories. For each new reaction product, a reference sample has to be prepared in a racemic or enantioenriched form to acquire a suitable HPLC or CD method, respectively. This is inefficient, time-consuming and decelerates the discovery pace. Chiral HPLC method development can be particularly arduous because it often is a trial-and-error approach at least when polysaccharide, cyclodextrin or protein CSPs are tested and there is no guarantee that satisfactory enantioseparation can eventually be achieved. In some cases, days to weeks are spent to no avail in search for a method that allows *ee* determination. We envisioned that one could remove this bottleneck and dependence with a rationally calculated, universally applicable calibration curve that correctly predicts induced CD intensities across a whole compound class, e.g. for all conceivable chiral amines. If possible, this would enable *ee* determination of any asymmetric reaction product without the need for a reference and individual method development.

Linear free energy relationships (LFERs) use physicochemical parameters that describe and predict chemical reactions and properties.¹⁵ Recently, the use of multi-parameter LFERs has found wide success in modeling reaction outcomes in asymmetric catalysis by guiding reaction discovery and optimization via screening.¹⁶ However, examples in which spectroscopic properties associated with analytical methods for screening are modeled and predicted are less common.¹⁷

For *ee* assays that rely on CD, the signal is largely governed by differences in the steric size of the groups on the analyte's point stereocenter.¹⁸ Our group has used this steric dependence to predict the calibration curves of a tripodal zinc (II) complex for the *ee* determination of chiral secondary alcohols, as well as a tripodal quinoline-Cu(II) sensor for chiral carboxylic acids.⁸ In the case of the alcohol sensor, the three-component assembly **1** adopts a preferred configurational twist (P or M)

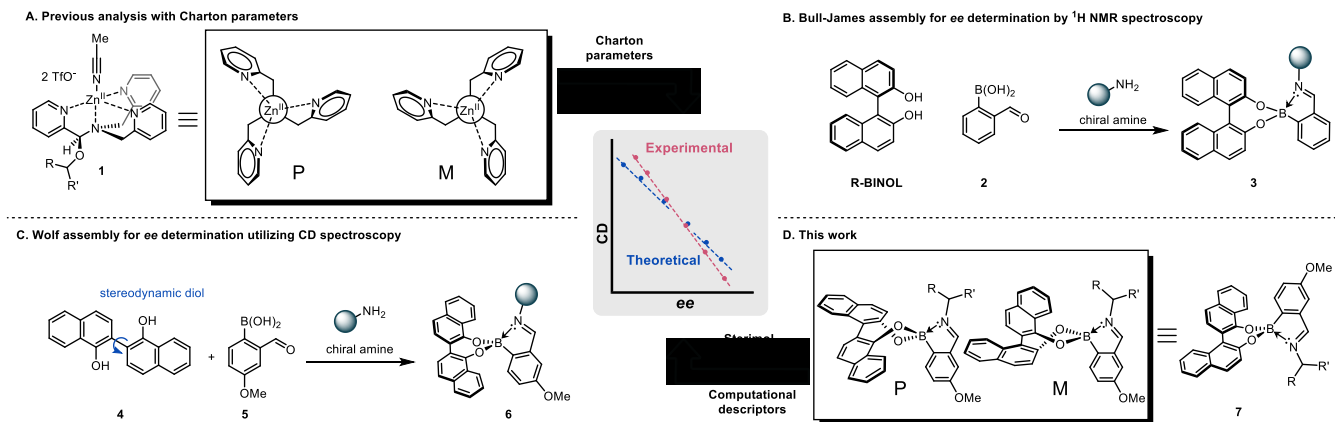


Figure 1. A) Three-component chiroptical sensor adopting *P* and *M* axial configurations, allowing *ee* determination of alcohols without prior calibration experiments. B) Bull-James assembly for *ee* determination of chiral amines via ^1H -NMR spectroscopy. C) Wolf assembly for *ee* determination of a variety of analytes via CD spectroscopy. D) Three-component chiroptical sensor for the *ee* determination of amines used in this study.

upon binding a chiral analyte resulting in exciton coupling (Figure 1A). In this previous work, we demonstrated that the diastereomeric ratio (*dr*) of *P* and *M* isomers correlates to the CD signal intensity. This relationship can be used to generate theoretical calibration curves, thus bypassing the need for calibration experiments.^{18,19} Using only Charton steric parameters, the calibration curves enabled *ee* determination of unknown samples. However, the use of Charton parameters limits this approach as they are experimentally determined and have been measured for only selected functional groups, thus narrowing the range of chiral analytes accommodated by the model. Additionally, predictive models based solely on steric parameters would become unreliable for analyte CD responses that are also dependent on electronic effects. A more universal approach would be to use computationally generated parameters so that calibration curves for any chiral analyte could be known *a priori*.

Sterimol parameters are prime candidates for a computational approach, as they divide the steric description into dimensional components: width, length, etc.^{20,21} and have been widely impactful in asymmetric catalysis modeling.²² More recently, several groups have demonstrated the utility of distance resolved parameters, which describe the size of functional groups within a certain spatial region.²³ This dimensional specificity enables chemists to distinguish the most relevant parameters that describe the steric effects (e.g., proximal versus distal steric interactions). Further, if necessary, adding electronic parameters, such as Hammett σ values, IR-stretching frequencies, or computed HOMO/LUMO energies, would further deconvolute the CD response of the sensors, thereby revealing any non-covalent interactions (NCIs) responsible for the sensor function.

The predictive modeling of CD based *ee* assays has been reported. In a collaboration with our group, Sigman and coworkers parameterized chemical derivatives of the α - and β -chiral alcohol sensor **1** to guide the design to accommodate γ -chiral alcohols.²³ The goal was to optimize the sensor for this new analyte class, which required steric and electronic parameterization of sensor derivatives. Critically, such an unexplored approach should also lead to the prediction of *ee* calibration curves that reveal the relationship between analyte parameters and CD intensity.

To explore this prediction, we examined assemblies involving *ortho*-imino boronic acids. Such boronic acids are frequently used in sensing applications.^{24–26} For example, the Bull-James assembly using 2-formylbenzeneboronic acid (**2**, Figure 1B) involves both boronic acid/diol and amine/aldehyde condensations for ^1H -NMR *ee* analysis of chiral amines and chiral diols.^{27,28} The complex equilibria of these systems are well-studied by our group, making such assemblies attractive platforms for sensing applications.²⁹ Building upon this work, Wolf and coworkers introduced stereodynamic diol (**4**) which increased the sensitivity and enabled *ee* determination via CD analysis for a wide variety of analytes (Figure 1C). The high sensitivity of tropos biaryl sensors has been established^{30,31} making Wolf's assembly a suitable system to expand upon.

In the study reported herein, we combined steric and electronic modeling, as used in reaction discovery and optimization,³² to develop a predictive model of calibration curves for a three-component *ortho*-imino boronic acid-based sensor (Figure 1D). This removes the necessity to generate calibration curves for chiroptical sensing assays used in asymmetric reaction discovery and optimization.

Results and Discussion

Initial Sensor and Limitations

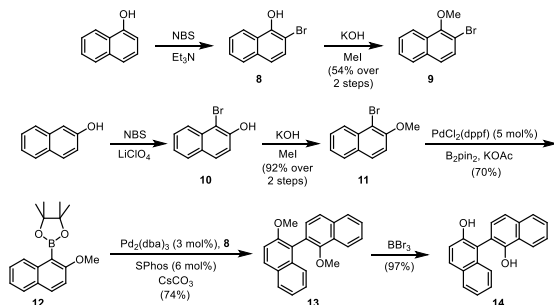
Before model construction, we first confirmed that the Wolf assembly behaved much like other *ortho*-imino boronic acid assemblies. To no surprise, the sensor speciated into numerous boronic acid-diol complexes which were observable by ^{11}B -NMR spectroscopy (Figure S1 in SI), but most importantly, we confirmed the dominance of species **6** which features an N-B dative bond in aprotic solvents. As we will discuss later, this influences our parameterization workflow.

Following the initial ^{11}B -NMR study, we sought to derive a relationship between the CD response of the Wolf assembly and several Sterimol parameters that describe the analyte. The stereodynamic binaphthol **4** is CD-silent until addition of a chiral amine and boronic acid **5**, after which it adopts either a *P* or *M* helical conformation. The energetic preference for one rotamer (i.e., *dr*) is what ultimately gives rise to a couplet in the CD spectrum at 232 nm.

To replicate our previous CD modeling approach (Figure 1A) in the context of the Wolf assembly, we first attempted to quantify the CD-active diastereomers which result from the twist in

the stereodynamic binaphthol chromophore.^{30,33} However, even at -80 °C, the diastereomeric conformers were not observable by ¹H-NMR spectroscopy (Figure S2 in SI). The fast rotation about the chiral axis in **4** is likely a result of the boron coordination which is expected to reduce the dihedral angle and the Gibbs free energy of activation to a value that does not allow NMR resolution of the P and M conformations at -80 °C.³⁴ To overcome this technical challenge, we redesigned sensor **6** to increase the barrier of rotation enough to observe the co-existing stereoisomers at a higher temperature.

Wolf assembly modification and characterization



Scheme 1. Synthesis of the asymmetric binaphthol ligand **14**.

Before synthesizing an analogue of **4**, we scoured literature values for similar biaryls with rotation barriers slightly higher than **4**. We ultimately arrived at binaphthol **14** which we calculated to possess a higher barrier to rotation (20.0 kcal mol⁻¹) than the initial binaphthol (Table 1 in SI for calculation details). The modified binaphthol ligand is easily synthesized by bromination of the 1-naphthol and 2-naphthol subunits with NBS and protection with methyl iodide. Miyaura borylation provided the boronic ester coupling partner, **12** which was subjected to Suzuki cross-coupling conditions with **9**. Demethylation with boron tribromide afforded the modified binaphthol **14** (Scheme 1).

Replacing binaphthol **4** with **14** in the three-component assembly results in an additional stereocenter at boron due to the removal of the C₂ symmetry axis of the binaphthol ligand. Four diastereomers of assembly **7** therefore exist upon introduction of a chiral amine (Figure 2A). This arises due to the existence of three elements of chirality: the stereocenter of the chiral amine (R or S), the stereocenter at the boron atom (R or S), and the axial chirality of the binaphthol ligand (M or P). This produces 2³ = 8 stereoisomers (4 diastereomers and their enantiomers) which results in 4 resonances in the ¹H-NMR spectra (Figure 2B). Interconversion of two of these diastereomers occurs via ring puckering that interconverts the M and P conformers via rotation about the chiral axis, while a second set of diastereomers interconvert by dissociation of the N-B dative bond and rotation along the boron-benzene ring single bond. We confirmed the dynamic nature of this process by VT-NMR using (S)-1-cyclohexylethylamine and observed that the 4 signals of the diastereomers coalesce at approximately 95 °C (Figure 2C). The observation that two diastereomers do not coalesce before another two implies that the interconversion of all four diastereomers is coupled to the same rate determining step, which is most likely the breakage of the N-B dative bond, followed by a

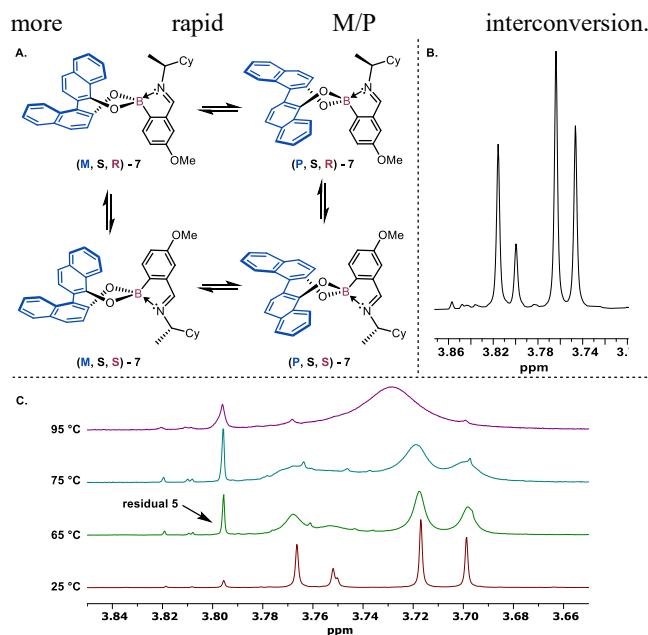


Figure 2. A) Four diastereomers of sensor **7** formed with amine (S)-1-cyclohexylethylamine. The stereochemical descriptors are in order: binaphthol axial chirality (M/P), amine stereocenter (S in this example), and boron stereocenter (R/S). B) ¹H-NMR spectrum featuring the four diastereomers of **7**. C) VT-NMR experiment showing coalescence of the methoxy resonances of the 4 diastereomers at 95 °C (DMSO-*d*₆, 600 MHz).

In general, the aromatic region of the spectra of **7** with different chiral amines was too crowded for meaningful interpretation. However, the methoxy group of boronic acid **5** served as a spectroscopic handle by which the diastereomers can be quantified. For most of the amines used in this study, four signals were observed in the δ4.2 – δ3.8 region corresponding to the methoxy groups of the interconverting diastereomers. The ¹H-NMR spectra of our initial training set (Figure 3B) were collected and integrated from δ4.2 – δ3.8 to provide the *dr* for our initial model. Although each resonance represents a single CD-active diastereomer, assigning each resonance to a particular diastereomer was difficult. The signals exhibit no spin-spin splitting or nuclear Overhauser effects (NOEs) to inform our diastereomer assignments. Nevertheless, we were able to calculate the *dr* using the following reasoning.

We recognized that two resonances were associated with the boron epimers and two were associated with the binaphthol axial epimers. It stands to reason that the sum of the integrations of two of resonances divided by the sum of the remaining two results in a *dr* value of either the boron or the axial stereogenic element. We found that the first and the third largest resonance by area represented the predominate axial twist because it correlated strongly to the CD signal (for a detailed discussion on *dr* measurements see SI). However, this experimental approach has substantial drawbacks as residual boronic acid **5** and its corresponding boronate ester were also present if the assembly had not fully formed or if **5** was in slight excess. Additionally, any analytes bearing an additional methoxy group crowded the δ4.2 – δ3.8 region making interpretation difficult. Despite these difficulties, we were able to measure this form of a *dr* for **15a-15f**.

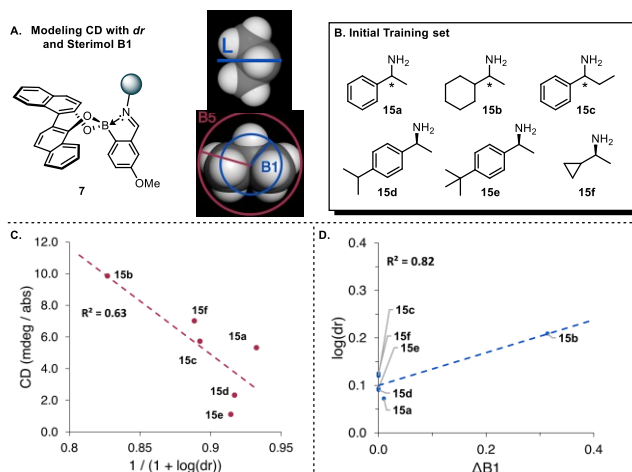


Figure 3. A) Sensor **7** and diagram of Sterimol parameters for an isopropyl group. B) Initial training set. C) Relationship between CD and dr . D) Relationship between $\Delta B1$ and $\log(dr)$.

These dr measurements were correlated to the CD signal associated with the binaphthol reporter chromophore at 232 nm using our previously established relationship between the two values ($CD = m * [1/(1 + \log(dr))]$, Figure 3C).¹⁹ While several amines (**15b**, **15c**, **15f**, and **15d**) produced a reasonably straight line, two of the lower dr amines (**15a** and **15e**) considerably deviated from the correlation. Nevertheless, we continued to explore how steric size differences of the two substituents of the amine stereocenter affected this relationship by generating three Sterimol values for each of the amine-incorporated assemblies, which represent the length (L), minimum width (B1), and maximum width (B5) of each of the substituents of the chiral amine (Figure 4B).

Sterimol parameters are highly sensitive to conformation. For instance, the length of a butyl group varies greatly depending on the number of methylene units in *anti* or *gauche* conformations.²⁰ Additionally, the local environment of a substituent influences its accessible conformations, such as that of the boronic acid assembly, and thereby the steric parameters. Thus, we opted to model the entire assembly as opposed to just the isolated substituents of interest. The resulting Sterimol values are then Boltzmann averaged as they would be for an experimentally determined steric parameter (e.g., Charton or A-values).

To gain a realistic conformer ensemble, each assembly was subjected to a Conformer-Rotamer Ensemble Search Tool (CREST) conformer search (Figure 4A).³⁶ The conformer-rotamer ensembles (CREs) were further optimized at the B3LYP/6-31G(d,p) level and then filtered for similarity based on root-mean squared deviation (RMSD) of atomic positions. After duplicate structures were removed, the remaining structures of the CRE were subjected to a single-point energy calculation at the M06-2X/Def2TZVP level. Only conformers within 5 kcal/mol were considered for the calculation of descriptors. The Sterimol parameters were collected using a modified Python script based on Paton's wSterimol program.²⁰

Using a stepwise multivariate linear regression, we identified the difference between the B1 values ($\Delta B1$) of the two substituents on the amine stereocenter to be the only significant predictor. However, the LFER between $\log(dr)$ and $\Delta B1$ was poor (Figure 3D). Almost all of the amines in the initial training set had $\Delta B1$ values close to 0, resulting in a large distribution of

points along the y-axis. This was a strong indication that additional parameters must be included for a more nuanced picture of the relevant steric, and possibly electronic, effects necessary to predict CD intensity. Despite the limitations of this steric-only approach, we were able to gain insights into the degree to which steric interactions influence CD intensity.

Initially, we found that **15d** and **15e** were not well tolerated as they possessed larger $\Delta B1$ values (~ 0.7 Å), yet their measured dr values were similar to the dr values of **15a**. The large $\Delta B1$ was attributed to distal steric bulk of **15d** and **15e**, which bear an isopropyl and *tert*-butyl group, respectively, at the *para*-position of the benzene ring. We concluded that at some distance along the L Sterimol axis of the substituents, steric effects became inconsequential towards the CD response of sensor **7**. To address this, we moved to distance-resolved Sterimol parameters which are calculated on substructures of the parent molecule truncated a predefined distance. We found Sterimol parameters calculated within 3.8 Å to give the best fit with dr (see SI for details on distance-resolved Sterimol parameters).

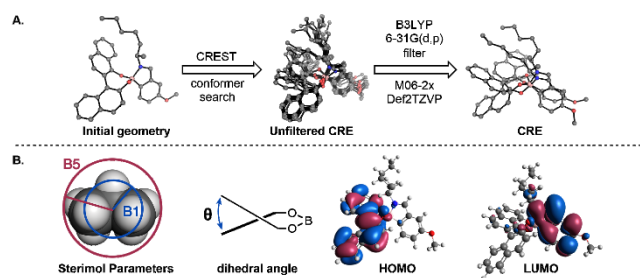


Figure 4. A) Conformer-rotamer ensemble generation workflow using CREST. B) Select features generated for each analyte.

We sought to improve our model by calculating a variety of electronic parameters. Taking lessons from the multivariate analysis of catalytic reactions, we began by calculating HOMO/LUMO energies, natural bond orbital populations (NBOs), vibrational frequencies, polarizability, and dipole moments (Figure 4B). Select structural features such as bond lengths and the dihedral angle between the 1-naphthol and 2-naphthol subunits were also calculated.

The increased number of predictive parameters would prove challenging with such a small initial training set. The primary bottleneck for increasing the training set size was determining dr values, as many amines possessed coincidental resonances in the ¹H-NMR spectrum which prevented us from measuring an accurate dr . Therefore, although we have historically correlated substituent effects to dr to gain insight into how structural parameters influence the thermodynamics (i.e., LFERS), followed by then correlating the dr values to CD intensities (e.g., Figure 3),^{8,18,19} we opted to directly predict CD instead of using dr as an intermediate predictor. While we lose the ability to calculate thermodynamic differences between diastereomers, we are still able to gain insight into which structural parameters most influence the CD intensities, and in turn postulate about steric and electronic interactions within the three-component assembly.

Using the scikit-learn package of Python, we attempted several multivariate linear regression (MLR) techniques to relate our descriptors to CD. Both ridge regression and stepwise MLR showed promise, however, the relationship between the selected features and CD response was ambiguous. Due to the large number of descriptors and limited size of the training set, we arrived at LASSO regression, a popular MLR technique for

problems with few observables and many descriptors.³⁷ LASSO implements a cost function during the regression process which penalizes models for including additional parameters. During the training process, irrelevant variable coefficients ultimately reduce to zero and are removed allowing for a more stable and interpretable model.³⁸

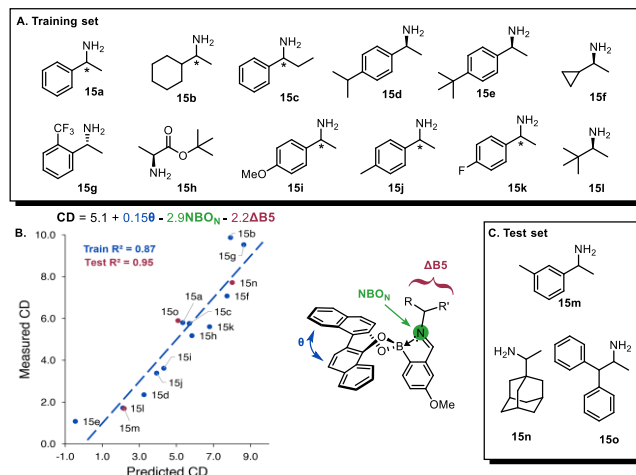


Figure 5. A) Training set for second generation model. An asterisk indicates both enantiomers were analyzed. B) Multivariate regression equation for predicting CD (mdeg / abs) and graphical explanation of selected parameters. C) Test set for second generation model.

During model development, LASSO regression found 3 variables to be significant: binaphthol dihedral angle (θ), $\Delta B5$, and NBO population of the imine nitrogen atom (NBO_N) (Eq. given in Figure 5B). Visual inspection of the regression line shows excellent correlation with the predictor variables, while the slight scatter results in an $R^2 = 0.87$. Of the three variables, the dihedral angle is unsurprisingly positively correlated with CD, as this angle between coupling chromophores partially governs CD intensity via exciton coupling, which makes its inclusion into the regression equation imperative.^{33,39} The twist of the reporter chromophore arises from a combination of steric effects imposed by the substituents of the amine, as well as attractive NCIs, each of which varies between analytes. It is possible that this angle parameter also serves as a “catch-all” for minute steric and electronic effects not accounted for by other descriptors.

In contrast to θ , the other two variables inclusion is not immediately obvious, showing the power of using a blind multi-parameter fit. The NBO_N values were negatively correlated with CD (Eq. given in Figure 5B). NBO analysis gives insight into the distribution of charge in a molecule.⁴⁰ In sensor **7**, the NBO_N value reflects the electron density of the imine nitrogen atom and thus the degree of N-B dative bond formation. Amines with electron donating *p*-substituted benzene rings (e.g., **15d**, **15e**, **15i**, and others) had the largest NBO_N values and the lowest CD signals of the training set. This suggests that electron rich substituents hinder point-to-axial chirality transfer to the reporter binaphthol ligand. Additionally, each amine featuring electron withdrawing substituents, such as **15g**, **15h**, and **15k**, exhibited diminished NBO_N values. Our computational investigation revealed that the electron deficient moieties of each substituent often resided directly above the naphthyl groups for the lowest lying conformers of the CRE. We suspect this donor-acceptor interaction brings the stereocenter of the analyte closer to the reporter chromophore, thus enhancing point-to-axial chirality transfer.

While our initial model (Figure 3) was falsely predicated on steric effects being the only factor of the CD response of **7**, the multiparameter fit finds that steric interactions are still influential in the sensor’s function. However, to our surprise, $\Delta B5$ (differences between the B5 values of the two substituents on the amine stereocenter) was negatively, not positively, correlated with CD despite our previous findings with sensor **1**. This negatively correlated steric parameter, $\Delta B5$, was calculated in the same manner as $\Delta B1$ our initial model. Because the training set is mostly comprised of amines with methyl groups at the stereocenter, which have the same B5 value, this term is effectively a measure of the steric size of the amines along the N-C* axis (where C* is the stereocenter of the amine). The largest $\Delta B5$ values are found for **15g** and **15h**, which feature an *o*-trifluoromethylbenzene ring and a *tert*-butoxycarbonyl group, respectively. These substituents add steric bulk near the reporter chromophore and reduce CD intensity. We further explored this assertion by calculating Sterimol parameters for along the N-C* axis. When B5 is measured along this bond, a strong negative correlation between CD intensity and B5 was observed for all but two amines we tested (see SI for details). Amines **15d** and **15e** also had a substantial contribution from this term due to the large *p*-substituted benzene ring substituents.

Although the relationship between the $\Delta B5$ parameter and CD induction can be rationalized by a repulsive steric argument, certain attractive NCIs appear to be enhancing point-to-axial chirality transfer which may counteract this effect. With bulky amines (e.g., **15b** and **15l**), the C-H bond at the amine stereocenter was found in the CRE to reside directly above the aryl-aryl bond of the binaphthol ligand. This is best rationalized as the minimum energy conformer adopting a geometry which places the medium sized substituent (methyl) and the large substituent (cyclohexyl or *tert*-butyl) away from the binaphthol chromophore. However, amines bearing aromatic rings (**15a**, **15c**, **15d**, **15e**, and others) adopted conformations which placed their rings much closer or directly over the binaphthol ligand, apparently maximizing attractive NCIs. This positioning maximizes donor-acceptor interactions between the binaphthol and amine substituents, which enhances point-to-axial chirality transfer. Such interactions are clearly visible in the computed NCI plots (Figure 6) which are used to visualize van der Waals interactions, hydrogen bonds, and other NCIs.⁴¹ The computed NCI plots revealed the presence of large dispersion interactions between aromatic amine substituents and the binaphthol chromophore, while much smaller NCI surfaces were observed for bulkier aliphatic amines. This finding is in agreement with the fact that electron withdrawing groups greatly enhance the CD signal of sensor **7**.

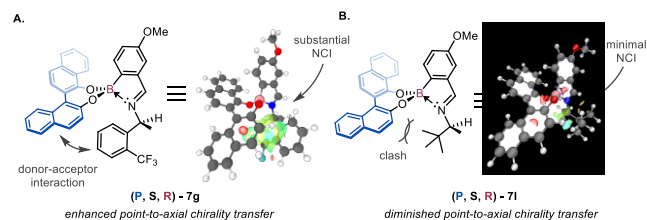


Figure 6. A) NCI plot of **7g** showing π - π stacking geometry and a large NCI surface. B) NCI plot of **7l** exhibiting diminished NCI surface as the *tert*-butyl group is placed away from the naphthalene ring.

The combination of NBO_N , which likely embodies the NCIs (Figure 6), and $\Delta B5$ in the regression equation demonstrates the

importance of point-to-axial chirality transfer in the sensor's function. The CD intensity of **7** is maximized when NBO_N is small indicating a strong N-B dative bond and the point stereocenter is "pulled" closer to the reporter chromophore by NCIs. Conversely, once substituents at the stereocenter are too large, steric interference "pushes" the stereocenter away and hinders point-to-axial chirality transfer.

To validate our model with a test set, we made predictions for three additional amines with qualitatively similar features to those of the training set (Figure 5). The largest NBO_N of the test set was found with **15m** which unsurprisingly gives CD responses similar to other electron rich aromatic amines (e.g., **15e**, and **15d**). This combined with an intermediate ΔB5 value of 2.3 explained why **15m** gave one of the lowest CD values of the amines we explored. Conversely, **15n** and **15o** had similarly low NBO_N values and thus larger CD signals. However, the ΔB5 of **15o** substantially attenuated the predicted CD signal owing to the high steric congestion between the benzhydryl group and the reporter chromophore.

Single-blind *ee* determination study

To demonstrate the utility of these predictions, theoretical calibration curves were generated by our model (Eq. given in Figure 5B). This was done by calculating the expected CD values of mixtures of enantiomers knowing that the enantiomers will have equal and opposite CD values, with mixtures thereof proportionally contributing to the CD (where racemic mixtures have $\text{CD} = 0$). These calculated calibration curves were overlaid with experimentally generated calibration curves (Figure 7) for two amines of the test set (**15n** and **15o**) because the larger CD intensities of these amines give a larger dynamic range. One researcher (ZA) prepared two solutions each of **15n** and **15o**. Each sample's *ee* was only known by researcher ZA. Another researcher (JH) analyzed these solutions with sensor **7** and used both the predicted and experimental calibration curves to calculate *ee* values.

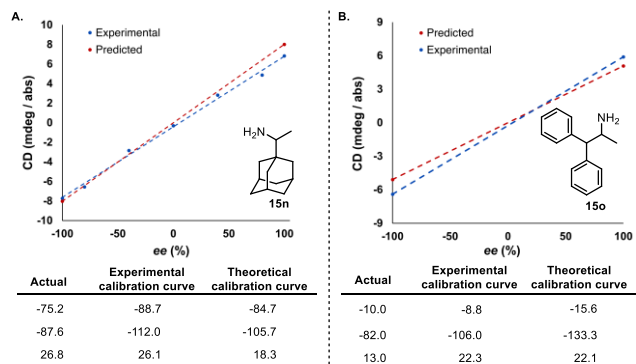


Figure 7. A) Predicted and experimental calibration curves for **15n** and results from single-blind *ee* determination study. B) Predicted and experimental calibration curves for **15o** and results from single-blind *ee* determination study.

When applied to **15n**, the experimental and theoretical calibration curves match and provide similar errors when used to determine the *ee* of unknown samples. The difference in errors between the experimental and theoretical calibration curves is approximately 6% for the three randomized samples. As the table accentuates, our experience is that the error associated with chiroptical *ee* assays increases with the *ee* of the sample.⁴²

The multivariate equation (Figure 5B) predicted a slightly lower CD intensity at 100% *ee* for **15o**, which is reflected in the

diminished slope of the theoretical calibration curve. As a result, the *ee* values of the unknown samples were overestimated using the theoretical calibration curve. At higher *ee* values, the disparity between the predicted and experimental calibration curves is exaggerated, which in turn makes the *ee* measurements less accurate in these ranges. Irrespective of the increased disparity, the differences in errors found for the *ee* values between the experimental and theoretical curves was around 10% (i.e., 11.4%). Moreover, the role of a theoretical calibration curve is to serve as a coarse analytical tool when optimizing asymmetric transformations. After all, in a screening experiment, the goal is to quickly measure *ee* values and the greatest enantiomeric excess values could be confirmed separately by chiral HPLC. Most importantly, the multiparameter fit introduced here serves as a useful tool for predicting whether or not chiroptical *ee* determination would be a worthwhile endeavor for a desired analyte based upon its structural features. For instance, a high-throughput reaction optimization campaign would be less effective if the predicted CD signal, based upon θ , ΔB5 , and NBO_N , is too low to accurately differentiate *ee* values during screening.

Conclusion

Determination of *ee* using chiroptical assays has demonstrated great utility in recent years. Its speed relative to chromatography is superior, however, this rate increase is only possible once calibration curves have been generated. In this study, we demonstrated the prediction of calibration curves for *ee* assays using a data-driven approach that analyzes a three-component assembly based on point-to-axial chirality transfer. After parameterizing a set of α -chiral primary amines in both electronic and steric domains, LASSO regression predicted the CD signal at 100% *ee*, allowing for the generation of calibration curves. Solutions of unknown enantiomeric composition were subject to the three-component assembly for *ee* determination. The average difference between the predicted and experimental calibration curves was 8.7%.

We believe that our findings will have far-reaching implications and, if adapted, will streamline future asymmetric reaction discovery projects by eliminating the need for reference materials and traditional analytical method development including trial-and-error efforts that can become very time-consuming and after all may prove unsuccessful. The understanding of the physicochemical underpinnings that determine the CD signal intensities obtained with a broadly applicable sensing assay and the prediction of a universal calibration curve thus remove current bottlenecks and roadblocks.

This data-driven approach also gives insight into factors which influence point-to-axial chirality transfer. Within the context of assembly **7**, while steric interactions between the groups on the stereocenter of the amine with the binaphthol were important, attractive NCIs and the extent of N-B dative bond formation were invaluable in predicting sensor function. Thus, chirality transfer for *ee* assays can be significantly more complex than a simple steric analysis may predict.

Important to the future development of chiroptical assays for *ee* determination, we have shown that one can circumvent calibration experiments via a multi-parameter fit that encompasses a series of steric and electronic factors. Just as the synthetic methodology community is increasing using such analyses to understand and optimize catalyst performance, the same approach can be applied to analytical chemistry pursuits within the organic chemistry community.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Experimental procedures and characterization data (PDF)

Spreadsheet containing computational parameters (XLSX)

AUTHOR INFORMATION

Corresponding Author

Eric V. Anslyn – Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States; orcid.org/0000-0002-5137-8797; Email: anslyn@austin.texas.edu

Christian Wolf – Department of Chemistry, Georgetown University, Washington, DC 20057, United States; orcid.org/0000-0002-4447-3753; Email: cw27@georgetown.edu

Authors

James R. Howard – Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States; orcid.org/0000-0003-1097-5031

Arya A. Bhakare – Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States; orcid.org/0000-0002-1841-4995

Zara Akhtar – Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, United States; orcid.org/0000-0003-1097-5031

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

We thank Max Moskowitz for the synthesis of binaphthol **4** and Dean Tantillo at UC Davis for suggesting to use binaphthol **14**. The authors thank the National Institute of Health for financial support for this work (5R01GM077437-12) as well as the National Science Foundation (CHE-1764135). The authors also acknowledge NIH Grant 1 S10 OD021508-1 (2015) for funding the 500 MHz NMR spectrometer upon which the NMR spectra were collected. The authors acknowledge the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing high performance computing resources that have contributed to the research results reported within this paper. EVA gratefully acknowledges support from the Welch Regents Chair (F-0046).

REFERENCES

- (1) Shevlin, M. Practical High-Throughput Experimentation for Chemists. *ACS Med. Chem. Lett.* **2017**, *8* (6), 601–607. <https://doi.org/10.1021/acsmchemlett.7b00165>.
- (2) 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.* **2017**, *50* (12), 2976–2985. <https://doi.org/10.1021/acs.accounts.7b00428>.
- (3) *High-Throughput Analysis in the Pharmaceutical Industry*; Wang, P. G., Ed.; Critical reviews in combinatorial chemistry; CRC Press: Boca Raton, FL, 2009.
- (4) Mennen, S. M.; Alhambra, C.; Allen, C. L.; Barberis, M.; Bertritt, S.; Brandt, T. A.; Campbell, A. D.; Castañón, J.; Cherney, A. H.; Christensen, M.; Damon, D. B.; Eugenio de Diego, J.; García-Cerrada, S.; García-Losada, P.; Haro, R.; Janey, J.; Leitch, D. C.; Li, L.; Liu, F.; Lobben, P. C.; MacMillan, D. W. C.; Magano, J.; McInturff, E.; Monfette, S.; Post, R. J.; Schultz, D.; Sitter, B. J.; Stevens, J. M.; Strambeanu, I. I.; Twilton, J.; Wang, K.; Zajac, M. A. The Evolution of High-Throughput Experimentation in Pharmaceutical Development and Perspectives on the Future. *Org. Process Res. Dev.* **2019**, *23* (6), 1213–1242. <https://doi.org/10.1021/acs.oprd.9b00140>.
- (5) Herrera, B. T.; Pilicer, S. L.; Anslyn, E. V.; Joyce, L. A.; Wolf, C. Optical Analysis of Reaction Yield and Enantiomeric Excess: A New Paradigm Ready for Prime Time. *J. Am. Chem. Soc.* **2018**, *140* (33), 10385–10401. <https://doi.org/10.1021/jacs.8b06607>.
- (6) Pilicer, S. L.; Dragna, J. M.; Garland, A.; Welch, C. J.; Anslyn, E. V.; Wolf, C. High-Throughput Determination of Enantiopurity by Microplate Circular Dichroism. *J. Org. Chem.* **2020**, *85* (16), 10858–10864. <https://doi.org/10.1021/acs.joc.0c01395>.
- (7) Nan, J.; Yan, X.-P. A Circular Dichroism Probe for L-Cysteine Based on the Self-Assembly of Chiral Complex Nanoparticles. *Chem. - Eur. J.* **2010**, *16* (2), 423–427. <https://doi.org/10.1002/chem.200902479>.
- (8) Joyce, L. A.; Maynor, M. S.; Dragna, J. M.; da Cruz, G. M.; Lynch, V. M.; Canary, J. W.; Anslyn, E. V. A Simple Method for the Determination of Enantiomeric Excess and Identity of Chiral Carboxylic Acids. *J. Am. Chem. Soc.* **2011**, *133* (34), 13746–13752. <https://doi.org/10.1021/ja205775g>.
- (9) Gholami, H.; Zhang, J.; Anyika, M.; Borhan, B. Absolute Stereochemical Determination of Asymmetric Sulfoxides via Central to Axial Induction of Chirality. *Org. Lett.* **2017**, *19* (7), 1722–1725. <https://doi.org/10.1021/acs.orglett.7b00495>.
- (10) Bentley, K. W.; Nam, Y. G.; Murphy, J. M.; Wolf, C. Chirality Sensing of Amines, Diamines, Amino Acids, Amino Alcohols, and α -Hydroxy Acids with a Single Probe. *J. Am. Chem. Soc.* **2013**, *135* (48), 18052–18055. <https://doi.org/10.1021/ja410428b>.
- (11) Zardi, P.; Wurst, K.; Licini, G.; Zonta, C. Concentration-Independent Stereodynamic g-Probe for Chiroptical Enantiomeric Excess Determination. *J. Am. Chem. Soc.* **2017**, *139* (44), 15616–15619. <https://doi.org/10.1021/jacs.7b09469>.
- (12) Herrera, B. T.; Moor, S. R.; McVeigh, M.; Roesner, E. K.; Marini, F.; Anslyn, E. V. Rapid Optical Determination of Enantiomeric Excess, Diastereomeric Excess, and Total Concentration Using Dynamic-Covalent Assemblies: A Demonstration Using 2-Aminocyclohexanol and Chemometrics. *J. Am. Chem. Soc.* **2019**, *141* (28), 11151–11160. <https://doi.org/10.1021/jacs.9b03844>.
- (13) Hassan, D. S.; Kariapper, F. S.; Lynch, C. C.; Wolf, C. Accelerated Asymmetric Reaction Screening with Optical Assays. *Synthesis* **2022**, *54* (11), 2527–2538. <https://doi.org/10.1055/a-1754-2271>.
- (14) Sripada, A.; Thanzeel, F. Y.; Wolf, C. Unified Sensing of the Concentration and Enantiomeric Composition of Chiral Compounds with an Achiral Probe. *Chem* **2022**, *8* (6), 1734–1749. <https://doi.org/10.1016/j.chempr.2022.03.008>.
- (15) Williams, W. L.; Zeng, L.; Gensch, T.; Sigman, M. S.; Doyle, A. G.; Anslyn, E. V. The Evolution of Data-Driven Modeling in Organic Chemistry. *ACS Cent. Sci.* **2021**, *7* (10), 1622–1637. <https://doi.org/10.1021/acscentsci.1c00535>.
- (16) Granda, J. M.; Donina, L.; Dragone, V.; Long, D.-L.; Cronin, L. Controlling an Organic Synthesis Robot with Machine Learning to Search for New Reactivity. *Nature* **2018**, *559* (7714), 377–381. <https://doi.org/10.1038/s41586-018-0307-8>.
- (17) Mamede, R.; Pereira, F.; Aires-de-Sousa, J. Machine Learning Prediction of UV–Vis Spectra Features of Organic Compounds Related to Photoreactive Potential. *Sci. Rep.* **2021**, *11* (1), 23720. <https://doi.org/10.1038/s41598-021-03070-9>.

- (18) You, L.; Berman, J. S.; Anslyn, E. V. Dynamic Multi-Component Covalent Assembly for the Reversible Binding of Secondary Alcohols and Chirality Sensing. *Nat. Chem.* **2011**, *3* (12), 943–948. <https://doi.org/10.1038/nchem.1198>.
- (19) Lin, C.-Y.; Lim, S.; Anslyn, E. V. Model Building Using Linear Free Energy Relationship Parameters—Eliminating Calibration Curves for Optical Analysis of Enantiomeric Excess. *J. Am. Chem. Soc.* **2016**, *138* (26), 8045–8047. <https://doi.org/10.1021/jacs.6b03928>.
- (20) Brethomé, A. V.; Fletcher, S. P.; Paton, R. S. Conformational Effects on Physical–Organic Descriptors – the Case of Sterimol Steric Parameters. 14.
- (21) Ariëns, E. J. *Drug Design; Medicinal chemistry*; Academic press: New York, 1971; Vol. 3.
- (22) Wu, K.; Doyle, A. G. Parameterization of Phosphine Ligands Demonstrates Enhancement of Nickel Catalysis via Remote Steric Effects. *Nat. Chem.* **2017**, *9* (8), 779–784. <https://doi.org/10.1038/nchem.2741>.
- (23) Dotson, J. J.; Anslyn, E. V.; Sigman, M. S. A Data-Driven Approach to the Development and Understanding of Chiroptical Sensors for Alcohols with Remote γ -Stereocenters. *J. Am. Chem. Soc.* **2021**, *143* (45), 19187–19198. <https://doi.org/10.1021/jacs.1c09443>.
- (24) Shcherbakova, E. G.; James, T. D.; Anzenbacher, P. High-Throughput Assay for Determining Enantiomeric Excess of Chiral Diols, Amino Alcohols, and Amines and for Direct Asymmetric Reaction Screening. *Nat. Protoc.* **2020**, *15* (7), 2203–2229. <https://doi.org/10.1038/s41596-020-0329-1>.
- (25) Bull, S. D.; Davidson, M. G.; Van den Elsen, J. M.; Fossey, J. S.; Jenkins, A. T. A.; Jiang, Y.-B.; Kubo, Y.; Marken, F.; Sakurai, K.; Zhao, J. Exploiting the Reversible Covalent Bonding of Boronic Acids: Recognition, Sensing, and Assembly. *Acc. Chem. Res.* **2013**, *46* (2), 312–326.
- (26) Chatterjee, S.; Anslyn, E. V.; Bandyopadhyay, A. Boronic Acid Based Dynamic Click Chemistry: Recent Advances and Emergent Applications. *Chem. Sci.* **2021**, *12* (5), 1585–1599. <https://doi.org/10.1039/D0SC05009A>.
- (27) Fossey, J. S.; Anslyn, E. V.; Brittain, W. D. G.; Bull, S. D.; Chapin, B. M.; Le Duff, C. S.; James, T. D.; Lees, G.; Lim, S.; Lloyd, J. A. C.; Manville, C. V.; Payne, D. T.; Roper, K. A. Rapid Determination of Enantiomeric Excess via NMR Spectroscopy: A Research-Informed Experiment. *J. Chem. Educ.* **2017**, *94* (1), 79–84. <https://doi.org/10.1021/acs.jchemed.6b00355>.
- (28) Metola, P.; Anslyn, E. V.; James, T. D.; Bull, S. D. Circular Dichroism of Multi-Component Assemblies for Chiral Amine Recognition and Rapid Ee Determination. *Chem Sci* **2012**, *3* (1), 156–161. <https://doi.org/10.1039/C1SC00496D>.
- (29) Chapin, B. M.; Metola, P.; Lynch, V. M.; Stanton, J. F.; James, T. D.; Anslyn, E. V. Structural and Thermodynamic Analysis of a Three-Component Assembly Forming *Ortho* - Iminophenylboronate Esters. *J. Org. Chem.* **2016**, *81* (18), 8319–8330. <https://doi.org/10.1021/acs.joc.6b01495>.
- (30) Bentley, K. W.; Nam, Y. G.; Murphy, J. M.; Wolf, C. Chirality Sensing of Amines, Diamines, Amino Acids, Amino Alcohols, and α -Hydroxy Acids with a Single Probe. *J. Am. Chem. Soc.* **2013**, *135* (48), 18052–18055. <https://doi.org/10.1021/ja410428b>.
- (31) Ni, C.; Zha, D.; Ye, H.; Hai, Y.; Zhou, Y.; Anslyn, E. V.; You, L. Dynamic Covalent Chemistry within Biphenyl Scaffolds: Reversible Covalent Bonding, Control of Selectivity, and Chirality Sensing with a Single System. *Angew. Chem. Int. Ed.* **2018**, *57* (5), 1300–1305. <https://doi.org/10.1002/anie.201711602>.
- (32) Santiago, C. B.; Guo, J.-Y.; Sigman, M. S. Predictive and Mechanistic Multivariate Linear Regression Models for Reaction Development. *Chem. Sci.* **2018**, *9* (9), 2398–2412. <https://doi.org/10.1039/C7SC04679K>.
- (33) Person, V.; Monde, K.; Berova, N.; Nakanishi, K. A New Approach in Exciton-Coupled Circular Dichroism (ECCD)-Insertion of an Auxiliary Stereogenic Center. 8.
- (34) Mazzanti, A.; Chiarucci, M.; Bentley, K. W.; Wolf, C. Computational and DNMR Investigation of the Isomerism and Stereodynamics of the 2,2'-Binaphthalene-1,1'-Diol Scaffold. *J. Org. Chem.* **2014**, *79* (8), 3725–3730. <https://doi.org/10.1021/jo5005229>.
- (35) Superchi, S.; Bisaccia, R.; Casarini, D.; Laurita, A.; Rosini, C. Flexible Biphenyl Chromophore as a Circular Dichroism Probe for Assignment of the Absolute Configuration of Carboxylic Acids. *J. Am. Chem. Soc.* **2006**, *128* (21), 6893–6902. <https://doi.org/10.1021/ja058552a>.
- (36) Grimme, S. Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. *J. Chem. Theory Comput.* **2019**, *15* (5), 2847–2862. <https://doi.org/10.1021/acs.jctc.9b00143>.
- (37) Eklund, M.; Norinder, U.; Boyer, S.; Carlsson, L. Choosing Feature Selection and Learning Algorithms in QSAR. *J. Chem. Inf. Model.* **2014**, *54* (3), 837–843. <https://doi.org/10.1021/ci400573c>.
- (38) Tibshirani, R. Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **1996**, *58* (1), 267–288.
- (39) Telfer, S. G.; McLean, T. M.; Waterland, M. R. Exciton Coupling in Coordination Compounds. *Dalton Trans.* **2011**, *40* (13), 3097. <https://doi.org/10.1039/c0dt01226b>.
- (40) Weinhold, F.; Landis, C. R. *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective*, 1st ed.; Cambridge University Press, 2003. <https://doi.org/10.1017/CBO9780511614569>.
- (41) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* **2010**, *132* (18), 6498–6506. <https://doi.org/10.1021/ja100936w>.
- (42) Moor, S. R.; Howard, J. R.; Herrera, B. T.; Anslyn, E. V. High-Throughput Screening of α -Chiral-Primary Amines to Determine Yield and Enantiomeric Excess. *Tetrahedron* **2021**, *94*, 132315. <https://doi.org/10.1016/j.tet.2021.132315>.