Title

A Combined DFT, Energy Decomposition, and Data Analysis Approach to Investigate the Relationship Between Noncovalent Interactions and Selectivity in a Flexible DABCOnium/Chiral Anion Catalyst System

Journal

ACS Catalysis, 12(19)

ISSN

2155-5435

Authors

Miller, Edward

Khanh, Binh

Read, Jacquelyne A

et al.

Publication Date

2022-10-07

DOI

10.1021/acscatal.2c03077

Peer reviewed

HHS Public Access

Author manuscript

ACS Catal. Author manuscript; available in PMC 2023 October 07.

Published in final edited form as:

ACS Catal. 2022 October 07; 12(19): 12369–12385. doi:10.1021/acscatal.2c03077.

A Combined DFT, Energy Decomposition, and Data Analysis Approach to Investigate the Relationship Between Noncovalent Interactions and Selectivity in a Flexible DABCOnium/Chiral Anion Catalyst System

Edward Miller,
Department of Chemistry, University of California, Berkeley, California 94720, United States

Binh Khanh Mai,
Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States

Jacquelyne A. Read,
Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, United States

William C. Bell,
Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States

Jeffrey S. Derrick,
Department of Chemistry, University of California, Berkeley, California 94720, United States

Peng Liu,
Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States

F. Dean Toste
Department of Chemistry, University of California, Berkeley, California 94720, United States

Abstract

Developing strategies to study reactivity and selectivity in flexible catalyst systems has become an important topic of research. Herein, we report a combined experimental and computational study aimed at understanding the mechanistic role of an achiral DABCOnium cofactor in a regio- and enantiodivergent bromocyclization reaction. It was found that electron-deficient aryl substituents enable rigidified transition states via an anion–π interaction with the catalyst, which drives the selectivity of the reaction. In contrast, electron-rich aryl groups on the DABCOnium result in

Corresponding Authors: **Peng Liu** – *Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States;* pengliu@pitt.edu; **F. Dean Toste** – *Department of Chemistry, University of California, Berkeley, California 94720, United States;* fdtoste@berkeley.edu.

Author Contributions

E.M. and B.K.M. contributed equally.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.2c03077.

Experimental data, additional computational results, Cartesian coordinates of lowest-energy conformers of optimized structures, energies of all conformers of each intermediate and TS, energy terms from EDA calculations, and X-ray crystal structure of **2f'** (PDF)

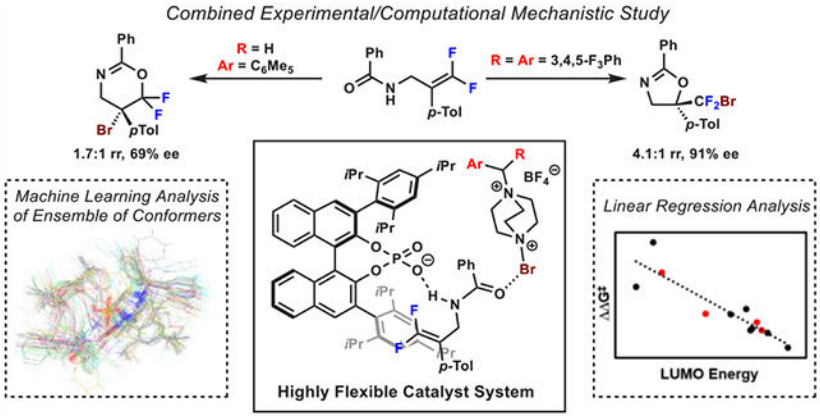
Cartesian coordinates of all conformers from computational study (ZIP)

The authors declare no competing financial interest.

Complete contact information is available at: https://pubs.acs.org/10.1021/acscatal.2c03077

significantly more flexible transition states, where interactions between the catalyst and substrate are more important. An analysis of not only the lowest-energy transition state structures but also an ensemble of low-energy transition state conformers via energy decomposition analysis and machine learning was crucial to revealing the dominant noncovalent interactions responsible for observed changes in selectivity in this flexible system.

Graphical Abstract



Keywords

computational analysis; linear free energy relationship; energy decomposition analysis; machine learning; asymmetric catalysis; conformers

INTRODUCTION

High selectivity, rate enhancement, and chemospecificity are features of enzyme-catalyzed reactions that chemists strive to emulate with synthetic catalysts.[1] Unlike nature's process of evolution, the rational and deliberate design of small-molecule catalysts necessitates precise structural changes that ideally provide predictable and rationalizable effects on reactivity and selectivity. In the field of asymmetric catalysis, the desire to reliably tune the steric and electronic profile of the chiral environment to affect the selectivity of a reaction has resulted in the extensive use of rigid, C_2-symmetric ligands and organocatalysts,[2] where flexibility has traditionally been considered an undesirable feature. In these systems, classical physical organic techniques, in conjunction with locating transition state (TS) structures by density functional theory (DFT), have become commonplace for understanding the interactions underpinning selectivity.[3] In the case of traditional chiral catalysts, their relative inflexibility allows for computational studies, where often only the lowest-energy structure of key intermediates and TSs are used to ascertain interactions affecting selectivity.

Recently, catalyst architectures have departed from tradition by embracing flexibility as a design element, leading to novel reactivity and selectivity, as well as new mechanistic questions.[4] This is exemplified by a report where a phosphite–oxazoline ligand, featuring a flexible biaryl motif, was employed for Pd-catalyzed allylic substitution (Scheme 1A).[5,6] This flexible ligand outperformed established representative ligands for overall

substrate tolerance, providing higher enantioselectivities across a series of model substrates. Unique mechanistic questions about selectivity in a flexible organocatalyst system have been demonstrated by a guanidine-/bisthiourea-promoted enantiodivergent Mannich-type reaction, where switching the polarity of the solvent was able to invert the enantioselectivity of the reaction (Scheme 1B)[7] An Eyring analysis demonstrated that the solvent-dependent enantioselectivity of the conformationally flexible catalyst was a result of switching between an enthalpically and entropically controlled mechanism. Within the realm of organocatalysis, some of the most prolifically studied flexible systems have been small-molecule peptide catalysts.[8,9] Recent work by the Miller and Baker groups demonstrated that, in a peptide-catalyzed, atroposelective bromination,[10,11] a more flexible peptide enhanced enantioselectivity, as compared to a less flexible analogue (Scheme 1C).[12] This outcome resulted from the catalyst's ability to more easily access selective and reactive catalyst conformations in transition states. Despite these advances, understanding the origins of selectivity in conformationally flexible catalyst systems remains an ongoing challenge.

In most of the previous mechanistic studies of flexible catalyst systems, flexibility arises from freely rotatable σ bonds in the chiral catalyst. As a result, the primary selectivity-determining interactions considered are those between the catalyst and substrate. The interplay of reaction components becomes more challenging to deconvolute in the field of ionpairing and phase-transfer catalysis,[13,14] where flexibility can result from intermolecular interactions between multiple catalyst components. In these reactions, interactions among the substrate, charged catalyst, and an oppositely charged reagent must be considered, where changing the structure and relative position of one component can have a multifaceted effect on reactivity and selectivity. Another practical challenge in understanding the interactions between the multiple flexible reaction components is the existence of several different types of noncovalent interactions between each pair of components. Recently, energy decomposition analysis (EDA) calculations have been used to quantitatively analyze the different types of catalyst–substrate interactions, such as London dispersion, electrostatics, polarization, charge transfer, and steric (Pauli) repulsions.[15] Although EDA studies have been successfully applied to relatively rigid catalysts, analyzing EDA results in flexible systems is uniquely challenging due to the high dimensionality produced by aggregating data from multiple conformers. Inspired by the use of machine learning (ML) algorithms to process complex data sets and predict chemical reactivity/selectivity,[16] we surmised that ML algorithms may also be employed to facilitate the interpretation of EDA results.

Recently, the Toste group reported a regio- and enantioselective bromocyclization of difluoroalkenes to access CF_2Br-containing stereocenters under chiral anion phase transfer[17] (CAPT) conditions (Scheme 1D).[18] During the optimization of this reaction, it was observed that the identity of the brominating reagent effected the regio- and enantioselectivity of the reaction. This effect was so pronounced that the favored regioisomer, as well as the enantiomer of both *exo*- and *endo*-cyclization products, could be inverted ($-25\% < ee_{exo} < 91\%$) by the choice of the arene on the 1,4-diazabicyclo[2.2.2]octane (DABCO)nium reagent. With a brominating reagent derived from **DAB^2**, the *endo* cyclization product was the major regioisomer (1.7:1 rr) and was produced in 69% ee, while the *exo* cyclization was obtained in −25% ee. These selectivities were inverted with a brominating reagent

derived from **DAB^{12}**, where the *exo* cyclization product was the major regioisomer (4.1:1 rr) and was produced in 91% ee, while the *endo* cyclization was obtained in −55% ee. A better understanding of how this achiral cofactor interacts with the chiral phosphate and substrate to affect the regio- and enantioselectivity would provide insight on how to leverage noncovalent interactions in a flexible catalyst system to achieve selectivity; however, the lack of "well-defined" ground states and TSs presents a significant challenge. Each cyclization pathway involves several conformers with similar energies and distinct structures, featuring different stabilizing and destabilizing interactions. As a result, information derived from lowest-energy structures alone leads to an incomplete understanding of the interactions that control the selectivity.

Herein, a combined experimental and computational approach was developed to study the observed switching of selectivity by the identity of the DABCOnium reagent. The flexibility of the system makes it such that a significant number of TSs, which contain varied conformations of each component, are energetically accessible. These TSs must be accounted for when predicting the reactivity/selectivity and analyzing the effects of the DABCOnium structure. Therefore, in our calculations, noncovalent interactions in not only lowest-energy TS structures but also other low-energy conformers were analyzed using a combined DFT, EDA, and data analysis approach. To provide a complete picture on the origin of selectivity, the dominant interactions responsible for regio- and enantioselectivity were revealed by evaluating the feature importance in a random forest model developed using different EDA energy terms in the ensemble of low-energy TS conformers.

COMPUTATIONAL DETAILS

General Workflow for Computational Analysis.

In order to investigate the origin of regio- and enantioselectivity in this flexible catalyst system, we employed a combined energy decomposition analysis/machine learning (EDA/ML) approach to reveal the dominant noncovalent interactions in the bromocyclization transition states (Scheme 2). After conformational sampling (step 1) and DFT geometry optimizations (step 2), EDA calculations were performed to dissect noncovalent interactions in low-energy TS conformers into chemically meaningful energy terms (step 3). Next, a feature importance analysis was carried out using ML algorithms to identify the most important factors that control the selectivity for this reaction (steps 4–7). The computational details of each step in the workflow are provided below.

Conformational Sampling and Density Functional Theory Calculations.

Due to the conformational flexibility of this system, an extensive conformational search was performed to identify the low-energy conformers of each stationary point. Conformational sampling was carried out for each intermediate and TS using metadynamics simulations and a genetic *z*-matrix crossing (iMTD-GC) approach in the gas phase with the conformer–rotamer ensemble sampling (CREST) algorithm.[19] The CREST metadynamics simulations were performed with the semiempirical tight-binding method GFN2-xTB[20] using the default settings for the iMTD-GC workflow with a time step of 1.0 ps and total run times of 432.0 and 468.0 ps for **DAB^{12}** and **DAB^{2}** (Scheme 1D), respectively.

During TS sampling, the forming/breaking bond distances, including the C–Br bond of the Br–DABCOnium and the forming N–Br and C–O bonds with the difluoroalkene, were fixed to values obtained from DFT-optimized TS geometries (2.25, 2.39, and 2.55 Å, respectively). Conformers within 8 kcal/mol of the lowest-energy conformer were included in a subsequent analysis. To validate whether the 8 kcal/mol energy threshold is adequate, energy values from the GFN2-xTB method were plotted against the Gibbs free energies from DFT calculations for all conformers of the bromocyclization TSs (Figure S2). It was found that all low-energy conformers at the DFT level are within 6 kcal/mol relative to the lowest-energy conformers at the GFN2-xTB method. The average error between the GFN2-xTB- and DFT-predicted relative energies of the different TS conformers is 2.1 ± 1.6 kcal/mol, which indicates that the energy threshold for conformational sampling (8 kcal/mol) is adequate. Next, the structural similarity of low-energy conformers was analyzed using root-mean-square deviation (RMSD), where conformers that have RMSD less than 0.5 Å were considered as duplicates and discarded. Each of the unique conformers was then reoptimized without constraint with DFT, using the Gaussian 16 program[21] on Pitt CRC, XSEDE,[22] and Frontera supercomputers. The DFT geometry optimizations of intermediate and TS conformers were performed in the gas phase at the B3LYP-D3/6-31G(d) level of theory. Vibrational frequency calculations at the same level of theory were performed for all structures to confirm the nature of each stationary point as a minimum or TS structure. Single-point energy calculations were carried out on the DFT-optimized geometries using the M06-2X functional[23] and 6-311+G(d,p) basis set.[24] Solvation energy corrections using toluene as the solvent were calculated using the SMD solvation model[25] in the single-point energy calculations. Intrinsic reaction coordinate (IRC) calculations were carried out for the lowest-energy conformer of all TSs to confirm the connections between TSs and ground state structures. Tetrafluoroborate (BF_4^-) was included as a counterion in all conformational sampling and DFT calculations.

The conformational sampling generated many relatively low-energy conformers for each of the four possible bromocyclization TSs. Boltzmann weighting of the conformers from DFT calculations was used to compute predicted enantiomeric excess (ee) values using eq 1:

$$\text{ee}\ (\%) = \frac{e^{-\Delta G_{\text{eff}}(R)/RT} - e^{-\Delta G_{\text{eff}}(S)/RT}}{e^{-\Delta G_{\text{eff}}(R)/RT} + e^{-\Delta G_{\text{eff}}(S)/RT}} \quad (1)$$

Here, $\Delta G_{\text{eff}}(R)$ and $\Delta G_{\text{eff}}(S)$ are the effective activation free energies in pathways leading to the R and S enantiomers of the product, respectively

$$\Delta G_{\text{eff}}(R) = G_{\text{eff}}(\mathbf{TS} - R) - G_{\text{eff}}(\mathbf{Int{-}2}) \quad (2)$$

$$\Delta G_{\text{eff}}(S) = G_{\text{eff}}(\mathbf{TS} - \mathbf{S}) - G_{\text{eff}}(\mathbf{Int{-}2}) \quad (3)$$

where **TS-*R*** and **TS-*S*** are the bromocyclization transition states leading to the R and S enantiomers of the product, respectively, **Int-2** is the trimolecular complex of (*R*)-TRIP, Br-DABCOnium, and the substrate, and G_{eff} is the Boltzmann-weighted effective Gibbs free energy of each stationary point:

$$G_{\text{eff}} = -RT \ln\left(\sum_{i}^{n} e^{-G_i/RT}\right) \quad (4)$$

Energy Decomposition Analysis to Investigate Non-covalent Interactions between the Reaction Components.

Distortion and interaction energy terms[26] were computed to analyze factors that affect TS stability. The distortion energies (ΔE_{dist}) of Br-DABCOnium, chiral phosphoric acid (CPA), and the difluoroalkene substrate are the energies required to distort each component from its ground state geometry to the geometry in the bromocyclization TS. The pairwise noncovalent interaction energies (ΔE_{int}) between CPA and DABCOnium, between CPA and the difluoroalkene substrate, and between DABCOnium and the substrate in all TS conformers were computed from the interaction energies between each pair of the reaction components at their TS geometries. To exclude the bonding interactions in the analysis, the bromine atom is excluded in the noncovalent interaction energy and EDA calculations. The interaction energies (ΔE_{int}) were then further dissected into several chemically meaningful energy terms, including ΔE_{Pauli} (Pauli repulsion), ΔE_{elstat} (electrostatic interaction), ΔE_{disp} (London dispersion), ΔE_{polar} (polarization), ΔE_{ct} (charge transfer), and ΔE_{solv} (solvation), by using a second-generation energy decomposition analysis based on absolutely localized molecular orbitals (ALMO-EDA),[27] implemented in the Q-Chem 5.3 package[28] (eq 5):

$$\Delta E_{\text{int}} = \Delta E_{\text{Pauli}} + \Delta E_{\text{elstat}} + \Delta E_{\text{disp}} + \Delta E_{\text{polar}} + \Delta E_{\text{ct}} + \Delta E_{\text{solv}} \quad (5)$$

The EDA calculations were computed in toluene at the M06-2X/6-311G(d,p) level of theory. Here, the EDA calculations were performed to analyze the different pairwise noncovalent interactions between two reaction components at their TS geometry in the absence of the third component. It should be noted that the sum of the pairwise interaction energy terms does not equal the overall interaction energy of the three-component system because covalent interactions involving bromine and three-body noncovalent interaction terms were not included. Previous studies have shown that three-body terms mostly include polarization and charge transfer and thus are expected to be small in the present system due to the minimal overlap of frozen occupied orbitals from all three fragments.[29]

B-Factor Calculations.

The flexibility of intermediates and TSs was quantified using *B*-factor values[30] (B_i, Å^2) obtained from root-mean-square fluctuation (ρ_i^{RMSF}) calculations of the conformer ensembles using the *gmx rmsf* command implemented in Gromacs 2020.3 software.[31]

$$B_i = \frac{8\pi^2}{3}(\rho_i^{\text{RMSF}})^2 \quad (6)$$

The *B*-factor is widely used as an indicator for interpreting the flexibility and dynamics of proteins from X-ray crystallography data[32a,b] and molecular dynamics simulations.[32c–g]

Here, we use a similar approach to compute *B*-factor values from intermediate and TS conformer ensembles to analyze the flexibility of different reaction components along the reaction pathway. Small *B*-factor values indicate a conformationally rigid component and/or tight binding that restricts conformational change, whereas large *B*-factor values suggest greater flexibility. The *B*-factors computed in this study are unweighted, and therefore their magnitude should not be directly compared with *B*-factors from X-ray crystallography data or MD trajectories.

All conformers from conformational sampling were used to calculate the *B*-factor values, where the *B*-factor value of (*R*)-TRIP was calculated by averaging the *B*-factor values of the 12 C(sp^2) atoms in the aryl rings, excluding those in the binaphthyl backbone. The flexibility of the DABCO core was quantified by averaging the *B*-factor values of the eight heavy atoms (carbon and nitrogen) of bicyclo[2.2.2]octane, whereas the *B*-factors of the substituents on DABCOnium were obtained by averaging the *B*-factor values of the C(sp^2) atoms in the aryl groups, as well as the methylene/methine carbon (13 atoms for $CHAr_2$ in **DAB^{12}** and seven atoms for CH_2Ar in **DAB^{2}**). The flexibility of the substrate was quantified by the average *B*-factors of its 21 heavy atoms (carbon, nitrogen, oxygen, and fluorine).

Machine Learning Model Development and Validation and Feature Importance Score Calculations.

The training and validation of machine learning models for data analysis and feature importance score calculations were performed using the scikit-learn 0.24.0 package.[33] Eight regression algorithms were used, including random forest, decision tree, LASSO, elastic-net, Gaussian process regression, ridge regression, and support vector regression using linear and radial basis function kernels. To avoid issues of overfitting, which are common for small data sets, multifold cross-validation was performed after training each ML model.[34] Random samplings were carried out 10 times with 20-fold cross-validation at each iteration. Learning curves were plotted to validate that the data set is sufficient in size to train the ML model (Figure S3).

Importance scores of the different features, i.e. EDA energy terms, were computed to evaluate the significance of each energy term on regio- and enantioselectivity. The feature importance scores for the enantioselectivity of the *exo* and *endo* products were calculated using all low-energy conformers of **TS**-*siExo* and **TS**-*reExo* in the *exo*-selective pathway and **TS**-*reEndo* and **TS**-*siEndo* in the *endo*-selective pathway, respectively. For regioselectivity, feature importance scores were calculated using low-energy TS conformers leading to the preferred enantiomer of each regioisomeric product. To validate the computed feature importance scores in identifying the dominant factors that influence the selectivities, correlations between the most important features and the activation barriers of the selectivity-determining TS conformers were examined (Figure S12). The feature importance scores reported in the main text were obtained from the random forest algorithm. A feature importance analysis was also performed using the decision tree algorithm, which yielded the same conclusions regarding the dominant factors on selectivities (Figure S13). In a previous study,[35] a similar approach has been applied to identify important features to design new catalyst generations for selective Cr(P,N)-catalyzed ethylene oligomerization.

RESULTS AND DISCUSSION

Selectivity Studies.

In order to better understand how the structure of the achiral brominating reagent affects the regio- and enantioselectivity of the bromocyclization, we designed and synthesized a set of DABCOnium-based brominating reagents with varying steric and electronic profiles about the DABCOnium substituent (Figure 1A). Seven DABCOnium cations (**DAB^{3}**, **DAB^{5}**, **DAB^{6}**, **DAB^{7}**, **DAB^{8}**, **DAB^{10}**, and **DAB^{11}**) were designed and synthesized to augment the initial five that were explored in our previous study.[18] Subjecting amide **1a** to the bromination reagents derived from these novel reagents resulted in *exo*- and *endo*-cyclization products (**2a** and **2a'**, respectively) in varying regioselectivity (−30 to 49% re, $\Delta\Delta G^{\ddagger}$ = −0.37 to 0.63 kcal/mol), as well as enantioselectivities for the *exo* (5–65% ee, $\Delta\Delta G^{\ddagger}$ = 0.06–0.92 kcal/mol) and *endo* (4–46% ee, $\Delta\Delta G^{\ddagger}$ = 0.05–0.59 kcal/mol) products. When these data are combined with the five data points from our previous study, the overall ranges in observed regioselectivity (−30 to 61% re, $\Delta\Delta G^{\ddagger}$ = −0.37 to 0.84 kcal/mol) as well as enantioselectivities for the *exo* (−25 to 91% ee, $\Delta\Delta G^{\ddagger}$ = −0.30 to 1.8 kcal/mol) and *endo* (−55 to 69% ee, $\Delta\Delta G^{\ddagger}$ = −0.72 to 1.0 kcal/mol) products increase. Under these conditions, the reagents [**(DAB^{2})$_2$Br**](**BF$_4$**)$_3$ and [**(DAB^{12})$_2$Br**](**BF$_4$**)$_3$ displayed disparate selectivities, where [**(DAB^{2})$_2$Br**](**BF$_4$**)$_3$ provided *R*-**2a'** as the major regioisomer in moderate enantioselectivity (69% ee). This outcome is contrasted by the results with the reagent [**(DAB^{12})$_2$Br**](**BF$_4$**)$_3$, which favored *R*-**2a** as the major regioisomer in high enantioselectivity (91% ee).

The range in the observed enantioselectivity values inspired a statistical fitting analysis,[17l,36] which has been successfully utilized in chiral anion catalyzed reactions with configurationally flexible, doubly axial chiral phosphoric acids.[37] Following this workflow, the goal was to relate the observed regio- and enantioselectivities to parameters derived from computed ground-state geometries of the DABCOnium cations. A conformational search on the **DAB$^{1–12}$** structures was performed with Schrödinger,[38] using the MacroModel suite with an OPLS_2005 force field. Conformers within 1.30 kcal/mol (5 kJ/mol) of the lowest energy conformer were submitted to a geometry optimization using the def2-TZVP basis set and M06-2x functional. Using a cutoff limit of 2.5 kcal/mol, the parameters of each low-energy conformer were weighted using a Boltzmann distribution using relative energies based on the energy conformation. A set of parameters was generated that reflect the steric and electronic profile of the reagents and condensed utilizing a Boltzmann distribution for conformers within 2.5 kcal/mol (pages S19–S22 in the Supporting Information). Our primary intention was to understand what factors affected the enantioselectivity of the *exo* cyclization, as this product was the focus of the previous study and displayed the largest variation in $\Delta\Delta G^{\ddagger}$. The data set was partitioned into a training set and validation set (eight training points and four validation points) to probe the reagent parameters relevant to enantioselectivity.

Performing a linear regression analysis revealed a statistically robust model which showed a strong correlation between the LUMO energy of the DABCOnium and the enantioselectivity of the *exo* product (Figure 1B). As the LUMO of the DABCOnium

is localized on an arene π^* orbital, we hypothesized that this might be indicative of a potential noncovalent interaction between the catalyst and the reagent, leading to the range in observed enantioselectivities. This process was repeated for the enantioselectivity of the *endo* cyclization product, which revealed a two-parameter model featuring the LUMO energy of the DABCOnium as the major contributor (Figure S1). In order to investigate the significance of the LUMO energy on these enantioselectivities, as well as to ascertain which interactions were responsible for the observed selectivities, TS calculations were initiated utilizing the reagents that provided disparate senses of selectivity under these reaction conditions ([$\mathbf{(DAB^2)_2Br}$]-$\mathbf{(BF_4)_3}$ and [$\mathbf{(DAB^{12})_2Br}$]$\mathbf{(BF_4)_3}$).

Mechanism and Selectivity-Determining Step of Bromocyclization.

Based on the experimental results, as well as previously reported CAPT-catalyzed reactions, a proposed catalytic cycle is shown in Figure 2A. The reaction begins with an anion exchange between the insoluble, dimeric DABCOnium reagent [$\mathbf{(DAB^n)_2Br}$]$\mathbf{(BF_4)_3}$ and the ionized phase-transfer catalyst. This anion metathesis generates a soluble ion pair of the monomeric brominated DABCOnium with (*R*)-TRIP as the counteranion (**Int-1**). Upon binding of substrate **1a**, the trimolecular complex **Int-2** is formed, featuring a halogen bond between the amide carbonyl and the electrophilic bromine. A concerted bromocyclization of **Int-2** occurs, where transfer of the bromonium cation and cyclization occur synchronously. DFT calculations suggest that a stepwise mechanism involving bromonium transfer to form a benzylic carbocation requires a higher activation energy than the concerted pathways (see Figure S4 in the Supporting Information). The four possible bromocyclization TSs leading to both enantiomers of the *exo* and *endo* cyclization products (**2a** and **2a**', respectively) are shown in Figure 2B.

Hammett Studies.

The computed TSs for the bromination step for both *exo* and *endo* cyclizations revealed that the transfer of the electrophilic bromine to the difluoroalkene is concerted with cyclization of the amide moiety. Based on this observation, it was hypothesized that there should be a relatively mild electronic effect of the arene in conjugation with the difluoroalkene on the relative rates of *exo* vs *endo* cyclization. In order to acquire experimental evidence to test this hypothesis, a series of substrates bearing neutral and electron-deficient difluorostyrenes were synthesized to probe the effect of electronic variation on the regioselectivity of the reaction (Figure 3A). After subjecting each substrate to the bromocyclization conditions with the reagents [$\mathbf{(DAB^2)_2Br}$]-$\mathbf{(BF_4)_3}$ and [$\mathbf{(DAB^{12})_2Br}$]$\mathbf{(BF_4)_3}$, it was observed that more electron deficient arenes disfavored the *exo*-cyclization pathway. By plotting the logarithms of the regiomeric ratios for both reagents against the σ^+ Hammett parameters,[39,40] trend lines with relatively shallow negative slopes (−2.0 and −1.5) were obtained.

Previous reports of reactions involving benzylic carbocation intermediates from bromination of olefins feature ρ values between −4 and −5.[41] While this difference could be evidence for a relatively concerted bromine transfer step, one mechanistic possibility is that both *exo*- and *endo*-cyclization rates are similarly sensitive to the electronics of the arene. As a consequence, these effects would negate one another, resulting in relatively little change in regioselectivity across the series.

To deconvolute the individual effects of *exo*- and *endo*-cyclization rates, additional competition experiments were conducted to obtain relative rates of both cyclization modes (Figure 3B). Plotting the logarithm of the relative *exo*-cyclization selectivities (exo_X/exo_H) against the respective ρ^+ values reproduced trend lines similar to the regioselectivity. The regioselectivities and *exo*-competition selectivities were then used to computationally derive the relative rates of *endo* cyclization, which revealed relatively little effect of arene electronics. The relatively shallow slopes of the regioselectivity and *exo*-cyclization rates are in agreement with a concerted bromine transfer with relatively little carbocation buildup at the benzylic position. Additionally, the insensitivity of arene electronics on the *endo* cyclization corroborates the mechanistic hypothesis that the stabilization of α-carbocations by fluorine is the main factor leading to *endo* cyclization.[42]

Activation Energies of Low-Energy TS Conformers and Computed Regio- and Enantioselectivities.

Transition state calculations were performed with **DAB^{2}** and **DAB^{12}** reagents as the achiral cofactors, as these provided products with opposite senses of stereochemistry via activation of opposite faces of the difluoroalkene. Under the reaction conditions, the *exo* product (*R*-**2a**) is favored in the reaction with **DAB^{12}**, whereas the *endo* product (*R*-**2a**') is favored with **DAB^{2}**. Therefore, we surmised that a detailed computational analysis comparing these two reagents will provide insight into the origins of the regio- and enantioselectivities.

The xTB-DFT conformational sampling of the four possible bromocyclization TSs (see Computational Details) led to many relatively low energy conformers. In the bromocyclization with **DAB^{12}**, there are 14, 25, 13, and 23 unique conformers within 3 kcal/mol relative to the lowest-energy conformer of **TS1-***siExo*, **TS1-***reExo*, **TS1-***reEndo*, and **TS1-***siEndo*, respectively (Figure 4A), where TS1-*siExo* has the lowest activation barrier. These calculations agree with the experimental regio- and enantioselectivity results, showing that the *R-exo* product is favored in the **DAB^{12}**-mediated reaction. Based on Boltzmann-weighted activation energies and considering all low-energy conformers of **TS1-***siExo* and **TS1-***reExo*, the computed ee is >99%, which is higher than the experimentally observed 91% ee for the *exo* product. The *exo*:*endo* selectivity derived from Boltzmann-weighted activation energies is ~10^3:1, which correctly predicts the *exo* product to be favored, although the calculations overestimate the experimental regioselectivity for this reaction. Moreover, our results indicate that the *exo* and *endo* products favor activation of the same face of the olefin (forming *R*-**2a** and *S*-**2a**', respectively), and the enantioselectivity for the *endo* products is lower than that for the *exo* products (−95% vs >99% ee, respectively), which also agrees with experimental results.

In the bromocyclization with **DAB^{2}**, there are 38, 29, 21, and 22 unique conformers within 3 kcal/mol relative to the lowest-energy conformer of **TS2-***siExo*, **TS2-***reExo*, **TS2-***reEndo*, and **TS2-***siEndo*, respectively (Figure 4B). Based on the computed activation energies, the regio- and enantioselectivities are diminished for this reaction, which agrees with the experimental results. Additionally, the DFT results predict that the lowest energy barrier is now associated with **TS2-***reEndo*, which corresponds to *R*-**2a**' as the major product. This was confirmed experimentally by isolation of the enantioenriched *endo*-product from the

reaction with [**(DAB^{2})$_{2}$Br**][**(BF$_{4}$)$_{3}$**] and determining the absolute stereochemistry by single-crystal X-ray diffraction (see the Supporting Information).

From an analysis of the Boltzmann-weighted activation free energies, the predicted ee values for *endo* and *exo* products derived from [**(DAB^{2})$_{2}$Br**][**(BF$_{4}$)$_{3}$**] are 44% and −17%, respectively, which are in agreement with the experimental ee values of 69% and −25% for *endo* and *exo* products, respectively. Moreover, the *exo*:*endo* selectivity derived from Boltzmann-weighted activation energies is 1:2.2, agreeing with the experimentally observed regioselectivity of 1:1.7. Overall, the DFT results are consistent with the experimental findings, reproducing all of the key trends in the regio- and enantioselectivity with the reagents [**(DAB^{2})$_{2}$Br**][**(BF$_{4}$)$_{3}$**] and [**(DAB^{12})$_{2}$Br**][**(BF$_{4}$)$_{3}$**].

Structures of the Lowest-Energy TS Isomers.

Optimized structures of the lowest-energy conformers of the four competing TSs with **DAB^{12}** and **DAB^{2}** are shown in Figure 5. The relative positions of CPA, DABCOnium, and the substrate differ significantly in these TS structures, which are expected to affect several different types of interactions between these components. For example, in **TS1-***siExo*, the aryl substituents on **DAB^{12}** are placed into the unoccupied quadrant **I**, leading to strong interactions between **DAB^{12}** and the CPA (see Figure S5), including an anion–π interaction between the phosphate center and the electron-deficient arene on the DABCOnium. A similar situation is found for **TS1-***siEndo*, where the aryl substituents on **DAB^{12}** are placed in the unoccupied quadrant **I**, leading to an anion–π interaction. In contrast, in **TS1-***reExo* and **TS1-***reEndo*, **DAB^{12}** is placed into the occupied quadrant **II**, leading to weaker interactions with the CPA. The stronger interaction of the electron-deficient 3,4,5-trifluoroarene on **DAB^{12}** with the chiral phosphate catalyst in **TS1-***siExo* in comparison to **TS1-***reExo* is consistent with the parametrization results implicating the DABCOnium LUMO energy as an important parameter for predicting *exo* enantioselectivity. Additionally, strong hydrogen bonds between the NH moiety of the substrate and oxyanion of CPA were found in both **TS1-***siExo* and **TS1-***siEndo* but were absent in **TS1-***reExo* and **TS1-***reEndo*. Between **TS1-***siExo* and TS1-*siEndo*, several different structural features exist, all of which may affect their relative stability. In **TS1-***siEndo*, several CH–π interactions between the *i*Pr group of the CPA and substrate are observed, which cannot be found in **TS1-***reEndo*. This result indicates that London dispersion between CPA and substrate is stronger in **TS1-***siEndo* than in **TS1-***reEndo*. On the other hand, **TS1-***siEndo* appears to be more distorted than **TS1-***reEndo*—although the C–Br distances in these two TSs are similar, the forming C–O distance in **TS1-***siEndo* (2.57 Å) is much shorter than that in **TS1-***siExo* (2.91 Å).

A comparison between these TSs and those containing **DAB^{2}** shows that the aryl group on **DAB^{2}** is positioned far from the phosphate in all TSs, which is in stark contrast to the TSs with **DAB^{12}**. Instead of the anion–π and CH–π interactions that likely affect the stability of TSs with **DAB^{12}**, the less stable TS isomers with **DAB^{2}** (**TS2-***siEndo* and **TS2-***reExo*) appear to be destabilized by steric repulsions between CPA and the substrate. While an examination of these lowest energy conformers is valuable for understanding what types of interactions may be controlling selectivity, the immense variability in the relative position of

the reaction components in the TS conformers precludes using this analysis as the sole basis for establishing selectivity-determining interactions.

This issue can be seen in the overlay of all low-energy conformers of **TS-***siExo*, **TS-***reExo*, **TS-***reEndo*, and **TS-***siEndo* with both DABCOnium reagents (Figure 6), where great variability in the relative position is observed. In the case of **DAB^{12}** (Figure 6A), we found that, while (*R*)-TRIP is relatively rigid and blocks quadrants **II** and **IV**, both **DAB^{12}** and the substrate are relatively flexible and can be placed on several positions relative to the CPA. In most of the low-energy conformers of the favored *exo* TS and *endo* TS (**TS1-***siExo* and **TS1-***siEndo*), both aryl groups of **DAB**$^{12\ 2}$ are placed in the unoccupied quadrant **I**. Similar to the lowest energy conformer, this conformation leads to favorable noncovalent interactions with the CPA, including London dispersion and an anion—$\boldsymbol{\pi}$ interaction between the CPA and the electron-deficient aryl groups on **DAB^{12}**. Moreover, in most of the conformers of **TS1-***siExo* and **TS1-***siEndo*, a strong hydrogen bond between the –NH moiety of the substrate and oxyanion of the CPA is present, which does not appear in most of the conformers of **TS1-***reExo*. These results indicate that DABCOnium–CPA and CPA–substrate interactions both contribute to the ee of *exo* and *endo* products.

The computed TS structures with **DAB^{2}** (Figure 6B) differ greatly from those of **DAB**$^{12\ 2}$. More specifically, **DAB^{2}** cannot fit into quadrant **I** because of the steric profile of the pentamethylbenzyl group. The distance from the chiral pocket and the phosphate center implies that electrostatic attraction and London dispersion between the CPA and DABCOnium are not crucial interactions that affect the selectivity of the reaction with **DAB^{2}**. Due to the weaker noncovalent interactions between CPA and **DAB^{2}**, as compared to **DAB^{12}**, this system is significantly more flexible, where the DABCOnium reagent can be placed in several locations around the CPA. Due to the flexibility of this system, it is challenging to identify the main factors that affect the regio- and enantioselectivity by an examination of individual TSs.

Flexibility of Intermediates and Bromocyclization TSs.

To quantitatively analyze the flexibility of the DABCOnium, CPA, and difluoroalkene substrate in the three-component TSs, we computed the average unweighted *B*-factor values for atoms in these fragments in **Int-1**, **Int-2**, and the four bromocyclization TSs with **DAB^{12}** and **DAB^{2}** (Table 1). The *B*-factor is a widely used displacement parameter in protein crystallography.[32] Here, we use *B*-factor values derived from the intermediate and TS conformers to establish a flexibility profile of the different reaction components (see Table S1 for *B*-factor calculations with a different fragment selection).

From the data presented in Table 1, in the absence of the difluoroalkene substrate, the CPA and DABCOnium in **Int-1a** and **Int-1b** are relatively flexible, as indicated by the large *B*-factor values of (*R*)-TRIP and **DAB^{n}**. For both reagents, the binding of the difluoroalkene substrate (**Int-1** to **Int-2**) leads to relatively little change in the flexibility of (*R*)-TRIP. In the case of the ion pair with **DAB^{2}**, binding of the substrate increases the flexibility of the pentamethylbenzyl group[43] in **Int-2b**, as demonstrated by an increase of its average *B*-factor from 194.7 to 346.9 Å^2; with a moderate rigidification of the DABCO core, the

B-factor decreases from 131.1 to 77.2 Å^2. On the other hand, for **DAB^{12}**, both the DABCO aryl substituent and the DABCO core in **Int-2a** are significantly more rigid than in **Int-1a**, indicating tighter binding of **DAB^{12}** compared to **DAB^{2}** in the (*R*)-TRIP/**DAB^{n}**/**1a** ternary complex.

Generally, the flexibility of the DABCO substituent decreased in the bromocyclization TS, compared to reactant complexes **Int-2a** and **Int-2b**. The decrease of flexibility is especially prominent in **TS1-***siExo*, which leads to the major regio- and enantiomeric product in the **DAB^{12}**-mediated reaction. On the other hand, in all four TS isomers with **DAB^{2}** (**TS2**), the pentamethylbenzyl group remains relatively flexible, as indicated by the large *B*-factors (85.5–353.7 Å^2), which suggests that **DAB^{2}** does not bind as tightly as **DAB^{12}** in the transition states. The substrate becomes less flexible in the TSs than in the reactant complex, where it binds to the CPA/DABCOnium complex via a relatively weak halogen bond. Although the decrease in substrate flexibility in the cyclization TS is expected, the substrate is less flexible in **TS1-***siExo* and **TS2-***siExo* than in other TS isomers, where it is hypothesized that the tighter binding of the DABCOnium leads to this effect. These results are in agreement with the experimental data, implying that the aryl groups play a crucial role in the selectivity-determining TSs. Additionally, they reflect the hypothesis that **DAB^{12}** interacts strongly with the phosphate center via noncovalent interactions, as compared to **DAB^{2}**, which exhibits weaker interactions with the catalyst.

Energy Decomposition Analysis and Machine Learning to Reveal the Dominant Interactions Affecting Regio- and Enantioselectivity.

As previously stated, relying on an analysis of the lowest-energy conformers of each TS to arrive at mechanistic conclusions would paint an incomplete picture of this highly flexible system. The variability in geometry, as demonstrated by the overlay of all low-energy TS structures and the *B*-factor values (Figure 6 and Table 1), can lead to a situation where each conformer is stabilized and destabilized by different types of noncovalent interactions, including steric repulsions, London-dispersion interactions, and through-space electrostatic interactions between the CPA and the DABCOnium. EDA calculations were performed for each of the low-energy TS conformers to quantitatively describe distortion energies of the three reaction components and the different types of noncovalent interaction energies between these components. These include the distortion energies of each of the three reaction components (ΔE_{dist}(CPA), ΔE_{dist}(DAB), and ΔE_{dist}(sub)) and the interaction energies between each pair of the components: ΔE_{int}(1), between CPA and the substrate; ΔE_{int}(2), between CPA and DAB; ΔE_{int}(3), between DAB and the substrate. Then each pairwise interaction energy is further dissected into Pauli repulsion (ΔE_{Pauli}), electrostatic attraction (ΔE_{elstat}), London dispersion (ΔE_{disp}), polarization (ΔE_{polar}), charge transfer (ΔE_{ct}), and solvation (ΔE_{solv}) energy terms (eq 5). This created a large data set that is information-rich but too complex for manual interpretation.

To analyze these data, modern data analysis tools and machine learning (ML) algorithms were used to investigate which types of noncovalent interactions are important for the regio- and enantioselectivities. After the TS conformational sampling using GFN2-xTB/DFT and EDA calculations of the low-energy conformers (all conformers within a 3 kcal/mol

free energy window; (Scheme 2, steps 1–3), this approach utilizes ML algorithms to analyze the data set of EDA energy terms of all low-energy TS conformers (21 energy terms, i.e. features, for each conformer). Three different data sets were prepared for each reaction to develop models for the three types of selectivities. For the model predicting enantioselectivity for the *exo* product, all low-energy conformers of the two enantioselectivity-determining transition states **TS-***siExo* and **TS-***reExo* in the *exo*-selective pathways were used as the data set. For the model predicting enantioselectivity for the *endo* product, low-energy conformers of **TS-***reEndo* and **TS-***siEndo* were used. The regioselectivity model was established using all low-energy conformers of **TS1-***siExo*/**TS1-***siEndo* for **DAB^{12}** and **TS2-***reExo*/**TS2-***reEndo* for **DAB^{2}**, all of which are TSs leading to the major enantiomer of each regioisomeric product.

Each data set was split into training and validation sets with a 75%:25% ratio. Because some of the EDA energy terms may strongly correlate with each other and using these correlated features may cause multicollinearity and produce less reliable regression models,[44] for any two strongly correlated features (| Pearson correlation coefficient, r| > 0.75),[45] the feature with poorer correlation with the activation Gibbs free energies ($\Delta G^{\ddagger}$) was removed (step 4; see Figures S5–S8 for details). Next, the remaining noncorrelating features were used to develop ML models for predicting the $\Delta G^{\ddagger}$ values of conformers of regio- and enantioselectivity-determining TSs with eight different regression algorithms (step 5). The root-mean-square error (RMSE) of the eight regression algorithms for predicting $\Delta G^{\ddagger}$ was calculated to identify the best-performing ML model.[16a,b,35,46] In our analysis, the random forest algorithm provided the smallest RMSE values in all cases (Figures S10 and S11) and was used in subsequent feature importance score calculations.

Next, feature importance scores of the different descriptors in the random forest model were calculated to rank the relative significance of each descriptor (i.e., EDA energy term) for regio- and enantioselectivity (step 6). The computed feature importance scores for the bromocyclization using **DAB^{12}** (Figure 7A) indicate that the London dispersion interaction between (*R*)-TRIP and **DAB^{12}** (ΔE_{disp}(2)) is the most important factor that determines the enantioselectivity of the *exo* product.[47] These results, in combination with the strong correlation between ΔE_{disp}(2) and $\Delta E_{\mathrm{elstat}}$) (electrostatic energies between (*R*)-TRIP and **DAB^{12}**) (r = 0.94, Figure S6), are consistent with the interpretation of the parametrization results—the *exo* enantioselectivity correlating with the DABCOnium LUMO energy is indicative of a noncovalent interaction between the DABCOnium and CPA (Figure 1B). These results are also consistent with the short distance between the DABCOnium aryl substituents and CPA in most of the low-energy conformers of **TS1-***siExo* (Figures 5A and 6A). Additionally, based on the EDA/ML analysis, it was found that the CPA–substrate London dispersion (ΔE_{disp}(1)) is the most important factor for the enantioselectivity of the *endo* product (Figure 7B), and substrate distortion (ΔE_{dist}(sub)) is the most important factor for regioselectivity (Figure 7C) in the bromocyclization with **DAB^{12}**. Interactions between the substrate and CPA or DABCOnium are likely not important for selectivity, which accounts for the observation that the modification of substituent groups on the substrate has a relatively minor effect on the enantioselectivity of the *exo* product, as seen in a previous

report[18] and the enantioselectivity of the Hammett series (see page S16 in the Supporting Information).

We then plotted the CPA–**DAB^{12}** and CPA–substrate London dispersion interaction and substrate distortion energies ($\Delta E_{disp}(2)$, $\Delta E_{disp}(1)$, and ΔE_{dist}(sub), respectively) for all low-energy conformers of the selectivity-determining TSs (Figure 7, bottom). In the low-energy conformers of **TS1**-*siExo*, London dispersion between (*R*)-TRIP and **DAB^{12}** is stronger than in **TS1**-*reExo* (Figure 7A), whereas, in the low-energy conformers of **TS1**-*siEndo*, London dispersion between CPA and the substrate is stronger than that in **TS1**-*reEndo*. This corroborates the feature importance calculations, where the CPA–DABCOnium and CPA–substrate dispersion interactions were found to be the most important for the enantioselectivity of the *exo* and *endo* products, respectively. The more favorable dispersion energies in **TS1**-*siExo* and **TS1**-*siEndo* stabilize these transition states leading to the experimentally observed major enantiomers. Additionally, substrate distortion in **TS1**-*siExo* is smaller than that in **TS1**-*siEndo*, which is consistent with the EDA/ML analysis revealing substrate distortion as the most important factor that determines the *exo*-favored regioselectivity for the reaction with **DAB^{12}**.

This analysis was then performed with the EDA data derived from the TS conformers with **DAB^{2}** (Figure 8). While no single feature was found to have a high importance score for the enantioselectivity of the *exo* product, Pauli repulsion between CPA and substrate ($\Delta E_{Pauli}(1)$) is the most important factor that affects both the enantioselectivity of the *endo* product and the regioselectivity. The lack of important features controlling the enantioselectivity of the *exo* product is consistent with the low ee observed experimentally. Here, the lack of sufficient steric control that differentiates the two enantioselectivity-determining TSs, **TS2**-*siExo* and **TS2**-*reExo*, is evidenced by the comparable Pauli repulsions between CPA and **DAB^{2}** ($\Delta E_{Pauli}(2)$) (Figure 8A, bottom). On the other hand, the CPA–substrate Pauli repulsions ($\Delta E_{Pauli}(1)$) in **TS2**-*reEndo* are smaller than those in **TS2**-*siEndo* and **TS2**-*reExo* (Figure 8B,C, bottom). This diminished steric repulsion in **TS2**-*reEndo* leads to the (*R*)-*endo* product as the major regio- and enantiomer observed experimentally.

CONCLUSION

In summary, a combined experimental and computational study was undertaken to examine the relationship between noncovalent interactions and reaction selectivities in a highly flexible catalyst system. A set of DABCOnium-based brominating reagents was designed and synthesized to probe their effects on the regio- and enantioselectivity of a bromocyclization reaction. Linear regression analysis on this data set revealed a strong dependence of the LUMO energy of the DABCOnium on the observed enantioselectivities of the reaction, prompting the hypothesis that this was indicative of a noncovalent interaction between the DABCOnium reagent and the chiral phosphate. This hypothesis was interrogated by DFT calculations of almost 200 low-energy TS conformers, which was required to account for the flexibility of the three-component reaction system. Selectivity-determining TSs were calculated for various concerted bromocyclization pathways with the reagents displaying opposite senses of selectivity ([**(DAB^{2})$_2$Br**][**BF$_4$)$_3$** and [**(DAB^{12})$_2$Br**]

$(BF_4)_3$). The concerted nature of these TSs was supported with linear free energy relationships, revealing relatively little carbocation buildup at the benzylic position.

We employed a combined energy decomposition analysis (EDA)/machine learning (ML) workflow to identify the dominant interactions that control enantio- and regioselectivities. This workflow uses EDA to calculate distortion energies and different types of interaction energy terms between each pair of the three reaction components for all low-energy TS conformers. Then, the most significant EDA energy term (feature) is identified using ML feature importance score calculations. This EDA/ML workflow revealed that non-covalent interactions between the DABCOnium and CPA are the dominant factors for the enantioselectivity of the *exo* product in the reaction with **$[(DAB^{12})_2Br](BF_4)_3$**. Additionally, a *B*-factor analysis suggests a rigidifying interaction in the most favorable TS isomer, which is significantly more rigid than other higher energy TSs. The significance of the DABCOnium/CPA anion–π interaction is corroborated by the parametrization results and the rigidifying TSs. In the case of **$[(DAB^{2})_2Br](BF_4)_3$**, it was found that the *endo* enantioselectivity was primarily driven by CPA–substrate noncovalent interactions, which was corroborated by the significant flexibility of the DABCOnium in this system, as well as its position far from the chiral pocket/phosphate center.

Thus, it appears that DABCOnium flexibility and selectivity are inversely correlated when **$[(DAB^{12})_2Br](BF_4)_3$** is employed, as a strong interaction with the CPA rigidifies the lowest energy TS, localizing the DABCOnium to a small set of locations in relation to the catalyst. When **$[(DAB^{2})_2Br](BF_4)_3$** is used, the DABCOnium has minimal interactions with the other reaction components, leading to high flexibility and variability in its position. Taken together, changing the achiral cofactor from the electron-rich and bulky **DAB^2** to electron-deficient **DAB^{12}** changes the mechanism from one in which interactions between the DABCOnium and CPA are not selectivity-determining to a mechanism where these interactions are the primary, selectivity-determining factors. More broadly, this work demonstrates a combined experimental and computational approach that may be applied to interrogate mechanistic questions, such as the effects of noncovalent interactions on regio- and stereoselectivity, in a flexible catalyst system.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

E.M. and F.D.T. gratefully acknowledge the National Institutes of Health (R35 GM118190) for financial support, the College of Chemistry CheXray (NIH Shared Instrumentation Grant S10-RR027172) for X-ray crystallographic data, and the College of Chemistry Molecular Graphics and Computation Facility (NIH S10OD023532). We thank Dr. Hasan Celik and UC Berkeley's NMR facility in the College of Chemistry (CoC-NMR) for spectroscopic assistance. Instruments at the CoC-NMR are supported in part by NIH S10OD024998. J.A.R. thanks the NIH for a Ruth L. Kirschstein NRSA postdoctoral fellowship (F32GM128351), which supported this work. P.L. and B.K.M. acknowledge the University of Pittsburgh and the NIH (R35GM128779) for financial support. DFT calculations were carried out at the Center for Research Computing at the University of Pittsburgh, the Extreme Science and Engineering Discovery Environment (XSEDE), and the TACC Frontera Supercomputer supported by the National Science Foundation grant number ACI-1548562. We also thank Prof. Matthew S. Sigman, Dr. Suhong Kim, Dr. Jacob Tracy, Dr. Patrick Moon, Dr. Alec Christian, Dr. Banruo Huang, Dr. Rene Rahimoff, and Dr. Danny Thach for helpful discussions.

Data Availability Statement

Jupyter notebook python scripts for the feature importance score calculations are freely available on GitHub: https://github.com/pengliugroup/feature-importance.

REFERENCES

(1). (a) Knowles JR Enzyme Catalysis: Not Different, Just Better. Nature 1991, 350, 121–124. [PubMed: 2005961] (b) Garcia-Viloca M; Gao J; Karplus M; Truhlar DG How Enzymes Work: Analysis by Modern Rate Theory and Computer Simulations. Science 2004, 303, 186–195. [PubMed: 14716003] (c) Warshel A; Sharma PK; Kato M; Xiang Y; Liu H; Olsson MHM Electrostatic Basis for Enzyme Catalysis. Chem. Rev 2006, 106, 3210–3235. [PubMed: 16895325] (d) Knowles RR; Jacobsen EN Attractive Noncovalent Interactions in Asymmetric Catalysis: Links Between Enzymes and Small Molecule Catalysts. Proc. Natl. Acad. Sci. U. S. A 2010, 107, 20678–20685. [PubMed: 20956302]

(2). Whitesell JK C_2 Symmetry and Asymmetric Induction. Chem. Rev 1989, 89, 1581–1590.

(3). (a) For select examples of this approach with chiral organocatalysts, see: Jindal G; Kisan HK; Sunoj RB Mechanistic Insights on Cooperative Catalysis through Computational Quantum Chemical Methods. ACS Catal. 2015, 5, 480–503.(b) Reddi Y; Tsai C-C; Avila CM; Toste FD; Sunoj RB Harnessing Noncovalent Interactions in Dual-Catalytic Enantiose-lective Heck-Matsuda Arylation. J. Am. Chem. Soc 2019, 141, 998–1009. [PubMed: 30562010] (c) Yepes D; Neese F; List B; Bistoni G Unveiling the Delicate Balance of Steric and Dispersion Interactions in Organo-catalysis Using High-Level Computational Methods. J. Am. Chem. Soc 2020, 142, 3613–3625. [PubMed: 31984734]

(4). (a) Crawford JM; Sigman MS Conformational Dynamics in Asymmetric Catalysis: Is Catalyst Flexibility a Design Element? Synthesis 2019, 51, 1021–1036. [PubMed: 31235980] (b) Shao H; Chakrabarty S; Qi X; Takacs JM; Liu P Ligand Conformational Flexibility Enables Enantioselective Tertiary C-B Bond Formation in the Phosphonate-Directed Catalytic Asymmetric Alkene Hydroboration. J. Am. Chem. Soc 2021, 143, 4801–4808. [PubMed: 33750118]

(5). Pàmies O; Diéguez M; Claver C New Phosphite-Oxazoline Ligands for Efficient Pd-Catalyzed Substitution Reactions. J. Am. Chem. Soc 2005, 127, 3646–3647. [PubMed: 15771469]

(6). Bellini R; Magre M; Biosca M; Norrby P-O; Pàmies O; Diéguez M; Moberg C Conformational Preferences of a Tropos Biphenyl Phosphinooxazoline-a Ligand with Wide Substrate Scope. ACS Catal. 2016, 6, 1701–1712.

(7). Sohtome Y; Tanaka S; Takada K; Yamaguchi T; Nagasawa K Solvent-Dependent Enantiodivergent Mannich-Type Reaction: Utilizing a Conformationally Flexible Guanidine/Bisthiourea Organo-catalyst. Angew. Chem., Int. Ed 2010, 49, 9254–9257.

(8). Jarvo ER; Copeland GT; Papaioannou N; Bonitatebus PJ; Miller SJ A Biomimetic Approach to Asymmetric Acyl Transfer Catalysis. J. Am. Chem. Soc 1999, 121, 11638–11643.

(9). Fierman MB; O'Leary DJ; Steinmetz WE; Miller SJ Structure-Selectivity Relationships and Structure for a Peptide-Based Enantioselective Acylation Catalyst. J. Am. Chem. Soc 2004, 126, 6967–6971. [PubMed: 15174866]

(10). Diener ME; Metrano AJ; Kusano S; Miller SJ Enantioselective Synthesis of 3-Arylquinazolin-4(3H)-ones via Peptide-Catalyzed Atroposelective Bromination. J. Am. Chem. Soc 2015, 137, 12369–12377. [PubMed: 26343278]

(11). Metrano AJ; Abascal NC; Mercado BQ; Paulson EK; Miller SJ Structural Studies of β-Turn-Containing Peptide Catalysts for Atroposelective Quinazolinone Bromination. Chem. Commun 2016, 52, 4816–4819.

(12). Stone EA; Hosseinzadeh P; Craven TW; Robertson MJ; Han Y; Hsieh S-Y; Metrano AJ; Baker D; Miller SJ Isolating Conformers to Assess Dynamics of Peptidic Catalysts Using Computationally Designed Macrocyclic Peptides. ACS Catal. 2021, 11, 4395–4400. [PubMed: 34659874]

(13). Mahlau M; List B Asymmetric Counteranion-Directed Catalysis: Concept, Definition, and Applications. Angew. Chem., Int. Ed 2013, 52, 518–533.

(14). Brak K; Jacobsen EN Asymmetric Ion-Pairing Catalysis. Angew. Chem., Int. Ed 2013, 52, 534–561.

(15). (a) Horn PR; Mao Y; Head-Gordon M Probing non-covalent interactions with a second generation energy decomposition analysis using absolutely localized molecular orbitals. Phys. Chem. Chem. Phys 2016, 18, 23067–23079. [PubMed: 27492057] (c) Lu G; Liu RY; Yang Y; Fang C; Lambrecht DS; Buchwald SL; Liu P Ligand-Substrate Dispersion Facilitates the Copper-Catalyzed Hydroamination of Unactivated Olefins. J. Am. Chem. Soc 2017, 139, 16548–16555. [PubMed: 29064694] (d) Zhao L; von Hopffgarten M; Andrada DM; Frenking G Energy decomposition analysis. WIREs Comput. Mol. Sci 2018, 8, 1345.(e) Xi Y; Su B; Qi X; Pedram S; Liu P; Hartwig JF Application of Trimethylgermanyl-Substituted Bisphosphine Ligands with Enhanced Dispersion Interactions to Copper-Catalyzed Hydroboration of Disubstituted Alkenes. J. Am. Chem. Soc 2020, 142, 18213–18222. [PubMed: 32962336] (f) Saper NI; Ohgi A; Small DW; Semba K; Nakao Y; Hartwig JF Nickel-catalysed anti-Markovnikov hydroarylation of unactivated alkenes with unactivated arenes facilitated by non-covalent interactions. Nat. Chem 2020, 12, 276–283. [PubMed: 32042137] (g) Hu L; Gao H; Hu Y; Lv X; Wu Y-B; Lu G Origin of Ligand Effects on Stereoinversion in Pd-Catalyzed Synthesis of Tetrasubstituted Olefins. J. Org. Chem 2021, 86, 18128–18138. [PubMed: 34878798] (h) Hao H; Qi X; Tang W; Liu P Energy Decomposition Analysis Reveals the Nature of Lone Pair-π Interactions with Cationic π Systems in Catalytic Acyl Transfer Reactions. Org. Lett 2021, 23, 4411–4414. [PubMed: 34010010] (i) Hu L; Gao H; Hu Y; Lv X; Wu Y-B; Lu G Origins of regio- and stereoselectivity in Cu-catalyzed alkyne difunctionalization with CO2 and organoboranes. Org. Chem. Front 2022, 9, 1033–1039. (j) Gao H; Hu L; Hu Y; Lv X; Wu Y-B; Lu G How the electron-deficient Cp ligand facilitates Rh-catalyzed annulations with alkynes. Org. Chem. Front 2022, 9, 979–988.(k) Hu Y; Hu L; Gao H; Lv X; Wu Y-B; Lu G Computational study of Cu-catalyzed 1,2-hydrocarboxylation of 1,3-dienes with CO_2: Pauli repulsion-controlled regioselectivity of Cu-Bpin additions. Org. Chem. Front 2022, 9, 2240–2248.(l) Bickelhaupt FM; Guerra CF; Mitoraj M; Sagan F; Michalak A; Pancd S; Frenking G Clarifying notes on the bonding analysis adopted by the energy decomposition analysis. Phys. Chem. Chem. Phys 2022, 24, 15726–15735. [PubMed: 35730200] (m) Liu S; Hirao H Energy Decomposition Analysis of the Nature of Coordination Bonding at the Heme Iron Center in Cytochrome P450 Inhibition. Chem. Asian J 2022, 13, No. e202200360.

(16). (a) Ahneman DT; Estrada JG; Lin S; Dreher SD; Doyle AG Predicting reaction performance in C-N cross-coupling using machine learning. Science 2018, 360, 186–190. [PubMed: 29449509] (b) Zahrt AF; Henle JJ; Rose BT; Wang Y; Darrow WT; Denmark SE Prediction of Higher-Selectivity Catalysts by Computer-Driven Workflow and Machine Learning. Science 2019, 363, 247.(c) Sandfort F; Strieth-Kalthoff F; Kühnemund M; Beecks C; Glorius F A Structure-Based Platform for Predicting Chemical Reactivity. Chem. 2020, 6, 1379–1390.(d) Friederich P; dos Passos Gomes G; Bin RD; Aspuru-Guzik A; Balcells D Machine learning dihydrogen activation in the chemical space surrounding Vaska's complex. Chem. Sci 2020, 11, 4584–4601. [PubMed: 33224459] (e) Li X; Zhang S-Q; Xu L-C; Hong X Predicting Regioselectivity in Radical C-H Functionalization of Heterocycles through Machine Learning. Angew. Chem., Int. Ed 2020, 59, 13253–13259.(f) Janet JP; Duan C; Nandy A; Liu F; Kulik HJ Navigating Transition-Metal Chemical Space: Artificial Intelligence for First-Principles Design. Acc. Chem. Res 2021, 54, 532–545. [PubMed: 33480674]

(17). (a) Rauniyar V; Lackner AD; Hamilton GL; Toste FD Asymmetric Electrophilic Fluorination Using and Anionic Chiral Phase-Transfer Catalyst. Science 2011, 334, 1681–1684. [PubMed: 22194571] (b) Phipps RJ; Hiramatsu K; Toste FD Asymmetric Fluorination of Enamides: Access to α-Fluoroimines using an Anionic Chiral Phase Transfer Catalyst. J. Am. Chem. Soc 2012, 134, 8376–8379. [PubMed: 22574822] (c) Wang Y-M; Wu J; Hoong C; Rauniyar V; Toste FD Enantioselective Halocyclization Using Reagents Tailored for Chiral Anion Phase-Transfer Catalysis. J. Am. Chem. Soc 2012, 134, 12928–12931. [PubMed: 22830953] (d) Shunatona HP; Früh N; Wang Y-M; Rauniyar V; Toste FD Enantioselective Fluoroamination: 1,4-Addition to Conjugated Dienes Using Anionic Phase-Transfer Catalysis. Angew. Chem., Int. Ed 2013, 52, 7724–7727.(e) Phipps RJ; Toste FD Chiral Anion Phase-Transfer Catalysis Applied to the Direct Enantioselective Fluorinative Dearomatization of Phenols. J. Am. Chem. Soc 2013, 135, 1268–1271. [PubMed: 23330962] (f) Romanov-Michailidis F; Guené e L; Alexakis A Enantioselective

Organocatalytic Fluorination' Induced Wagner-Meerwein Rearrangement. Angew. Chem., Int. Ed 2013, 52, 9266–9270.(g) Xie W; Jiang G; Liu H; Hu J; Pan X; Zhang H; Wan X; Lai Y; Ma D Highly Enantioselective Bromocyclization of Tryptamines and its Application in the Synthesis of (−)-Chimonanthine. Angew. Chem., Int. Ed 2013, 52, 12924–12927.(h) Wu J; Wang Y-M; Drljevic A; Rauniyar V; Phipps RJ; Toste FD A Combination of Directing Groups and Chiral Anion Phase-Transfer Catalysis for Enantioselective Fluorination of Alkenes. Proc. Natl. Acad. Sci. U. S. A 2013, 110, 13729–13733. [PubMed: 23922394] (i) Yang X; Phipps RJ; Toste FD Asymmetric Fluorination of α-Branched Cyclohexanones Enabled by a Combination of Chiral Anion Phase Transfer Catalysis and Enamine Catalysis using Protected Amino Acids. J. Am. Chem. Soc 2014, 136, 5225–5228. [PubMed: 24684209] (j) Zi W; Wang Y-M; Toste FD An in situ Directing Group Strategy for Chiral Anion Phase-Transfer Fluorination of Allylic Alcohols. J. Am. Chem. Soc 2014, 136, 12864–12867. [PubMed: 25203796] (k) Shen Z; Pan X; Lai Y; Hu J; Wan X; Li X; Zhang H; Xie W Chiral Ion-Pair Organocatalyst Promotes Highly Enantioselective 2-exo Iodocycloetherification of Allyl Alcohols. Chem. Sci 2015, 6, 6986–6990. [PubMed: 29861937] (l) Neel AJ; Milo A; Sigman MS; Toste FD Enantiodivergent Fluorination of Allylic Alcohols: Data Set Design Reveals Structural Interplay between Achiral Directing Group and Chiral Anion. J. Am. Chem. Soc 2016, 138, 3863–3875. [PubMed: 26967114]

(18). Miller E; Kim S; Gibson K; Derrick JS; Toste FD Regio-and Enantioselective Bromocyclization of Difluoroalkenes as a Strategy to Access Tetrasubstituted Difluoromethylene-Containing Stereocenters. J. Am. Chem. Soc 2020, 142, 8946–8952. [PubMed: 32352775]

(19). Pracht P; Bohle F; Grimme S Automated Exploration of the Low-Energy Chemical Space with Fast Quantum Chemical Methods. Phys. Chem. Chem. Phys 2020, 22, 7169–7192. [PubMed: 32073075]

(20). Bannwarth C; Ehlert S; Grimme S GFN2-xTB - An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. J. Chem. Theory Comput 2019, 15, 1652–1671. [PubMed: 30741547]

(21). Frisch MJ; Trucks GW; Schlegel HB; Scuseria GE; Robb MA; Cheeseman JR; Scalmani G; Barone V; Petersson GA; Nakatsuji H; Li X; Caricato M; Marenich AV; Bloino J; Janesko BG; Gomperts R; Mennucci B; Hratchian HP; Ortiz JV; Izmaylov AF; Sonnenberg JL; Williams; Ding F; Lipparini F; Egidi F; Goings J; Peng B; Petrone A; Henderson T; Ranasinghe D; Zakrzewski VG; Gao J; Rega N; Zheng G; Liang W; Hada M; Ehara M; Toyota K; Fukuda R; Hasegawa J; Ishida M; Nakajima T; Honda Y; Kitao O; Nakai H; Vreven T; Throssell K; Montgomery JA Jr.; Peralta JE; Ogliaro F; Bearpark MJ; Heyd JJ; Brothers EN; Kudin KN; Staroverov VN; Keith TA; Kobayashi R; Normand J; Raghavachari K; Rendell AP; Burant JC; Iyengar SS; Tomasi J; Cossi M; Millam JM; Klene M; Adamo C; Cammi R; Ochterski JW; Martin RL; Morokuma K; Farkas O; Foresman JB; Fox D J. Gaussian 16 Rev. C.01; Gaussian Inc. 2016.

(22). Towns J; Cockerill T; Dahan M; Foster I; Gaither K; Grimshaw A; Hazlewood V; Lathrop S; Lifka D; Peterson GD; Roskies R; Scott JR; Wilkins-Diehr N XSEDE: Accelerating Scientific Discovery. Comput. Sci. Eng 2014, 16, 62–74.

(23). Zhao Y; Truhlar DG The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. Theor. Chem. Acc 2008, 120, 215–241.

(24). Single-point calculations were also performed using the ωB97X-D functional. Both M06-2X and ωB97X-D give the same lowest-energy conformer for each transition state. The conclusion from our calculations will not change when a different DFT functional is employed.

(25). Marenich AV; Cramer CJ; Truhlar DG Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. J. Phys. Chem. B 2009, 113, 6378–6396. [PubMed: 19366259]

(26). (a) Ess DH; Houk KN Distortion/Interaction Energy Control of 1,3-Dipolar Cycloaddition Reactivity. J. Am. Chem. Soc 2007, 129, 10646–10647. [PubMed: 17685614] (b) Bickelhaupt FM; Houk KN Analyzing Reaction Rates with the Distortion/Interaction-Activation Strain Model. Angew. Chem., Int. Ed 2017, 56, 10070–10086.

(27). (a) Horn PR; Mao Y; Head-Gordon M Defining the contributions of permanent electrostatics, Pauli repulsion, and dispersion in density functional theory calculations of intermolecular interaction energies. J. Chem. Phys 2016, 144, 114107. [PubMed: 27004862] (b) Horn PR; Mao Y; Head-Gordon M Probing non-covalent interactions with a second generation energy decomposition analysis using absolutely localized molecular orbitals. Phys. Chem. Chem. Phys 2016, 18, 23067–23079. [PubMed: 27492057] (c) Mao Y; Demerdash O; Head-Gordon M; Head-Gordon T Assessing Ion-Water Interactions in the AMOEBA Force Field Using Energy Decomposition Analysis of Electronic Structure Calculations. J. Chem. Theory Comput 2016, 12, 5422–5437. [PubMed: 27709939] (d) Mao Y; Loipersberger M; Kron KJ; Derrick JS; Chang CJ; Sharada SM; Head-Gordon M Consistent Inclusion of Continuum Solvation in Energy Decomposition Analysis: Theory and Application to Molecular CO_2 Reduction Catalysts. Chem. Sci 2021, 12, 1398–1414.

(28). Shao Y; Gan Z; Epifanovsky E; Gilbert ATB; Wormit M; Kussmann J; Lange AW; Behn A; Deng J; Feng X; Ghosh D; Goldey M; Horn PR; Jacobson LD; Kaliman I; Khaliullin RZ; Kuś T; Landau A; Liu J; Proynov EI; Rhee YM; Richard RM; Rohrdanz MA; Steele RP; Sundstrom EJ; Woodcock HL; Zimmerman PM; Zuev D; Albrecht B; Alguire E; Austin B; Beran GJO; Bernard YA; Berquist E; Brandhorst K; Bravaya KB; Brown ST; Casanova D; Chang C-M; Chen Y; Chien SH; Closser KD; Crittenden DL; Diedenhofen M; DiStasio RA; Do H; Dutoi AD; Edgar RG; Fatehi S; Fusti-Molnar L; Ghysels A; Golubeva-Zadorozhnaya A; Gomes J; Hanson-Heine MWD; Harbach PHP; Hauser AW; Hohenstein EG; Holden ZC; Jagau T-C; Ji H; Kaduk B; Khistyaev K; Kim J; Kim J; King RA; Klunzinger P; Kosenkov D; Kowalczyk T; Krauter CM; Lao KU; Laurent AD; Lawler KV; Levchenko SV; Lin CY; Liu F; Livshits E; Lochan RC; Luenser A; Manohar P; Manzer SF; Mao S-P; Mardirossian N; Marenich AV; Maurer SA; Mayhall NJ; Neuscamman E; Oana CM; Olivares-Amaya R; O'Neill DP; Parkhill JA; Perrine TM; Peverati R; Prociuk A; Rehn DR; Rosta E; Russ NJ; Sharada SM; Sharma S; Small DW; Sodt A; Stein T; Stück D; Su Y-C; Thom AJW; Tsuchimochi T; Vanovschi V; Vogt L; Vydrov O; Wang T; Watson MA; Wenzel J; White A; Williams CF; Yang J; Yeganeh S; Yost SR; You Z-Q; Zhang IY; Zhang X; Zhao Y; Brooks BR; Chan GKL; Chipman DM; Cramer CJ; Goddard WA; Gordon MS; Hehre WJ; Klamt A; Schaefer HF; Schmidt MW; Sherrill CD; Truhlar DG; Warshel A; Xu X; Aspuru-Guzik A; Baer R; Bell AT; Besley NA; Chai J-D; Dreuw A; Dunietz BD; Furlani TR; Gwaltney SR; Hsu C-P; Jung Y; Kong J; Lambrecht DS; Liang W; Ochsenfeld C; Rassolov VA; Slipchenko LV; Subotnik JE; Van Voorhis T; Herbert JM; Krylov AI; Gill PMW; Head-Gordon M Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. Mol. Phys 2015, 113, 184–215.

(29). (a) Horn PR; Mao Y; Head-Gordon M Probing non-covalent interactions with a second generation energy decomposition analysis using absolutely localized molecular orbitals. Phys. Chem. Chem. Phys 2016, 18, 23067–23079. [PubMed: 27492057] (b) Bell F; Ruan QN; Golan A; Horn PR; Ahmed M; Leone SR; Head-Gordon M Dissociative Photoionization of Glycerol and its Dimer Occurs Predominantly via a Ternary Hydrogen-Bridged Ion–Molecule Complex. J. Am. Chem. Soc 2013, 135, 14229–14239. [PubMed: 23924376]

(30). Willis BTM; Pryor AW Thermal vibrations in crystallography. Acta Crystallogr. 1975, A31, 879.

(31). Abraham MJ; Murtola T; Schulz R; Páll S; Smith JC; Hess B; Lindahl E GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015, 1–2, 19–25.

(32). (a) Trueblood KN; Bürgi H-B; Burzlaff H; Dunitz JD; Gramaccioli CM; Schulz HH; Shmueli U; Abrahams SC Atomic Displacement Parameter Nomenclature. Report of a Subcommittee on Atomic Displacement Parameter Nomenclature. Acta Cryst. A 1996, 52, 770–781.(b) Sun Z; Liu Q; Qu G; Feng Y; Reetz MT Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. Chem. Rev 2019, 119, 1626–1665. [PubMed: 30698416] (c) Lu WC; Wang CZ; Yu EW; Ho KM Dynamics of the Trimeric AcrB Transporter Protein Inferred From a B-Factor Analysis of the Crystal Structure. Proteins: Struct., Funct., Bioinf 2006, 62, 152–158.(d) Huang J; Xie D-F; Feng Y Engineering thermostable (R)-selective amine transaminase from Aspergillus terreus through in silico design employing B-factor and folding free energy calculations. Biochem. Biophys. Res. Commun 2017, 483, 397–402. [PubMed: 28017723] (e) Nowalk JA; Fang C; Short AL; Weiss RM; Swisher JH; Liu P; Meyer TY Sequence-Controlled Polymers Through Entropy-Driven Ring-

Opening Metathesis Polymerization: Theory, Molecular Weight Control, and Monomer Design. J. Am. Chem. Soc 2019, 141, 5741–5752. [PubMed: 30714723] (f) Land H; Campillo-Brocal JC; Humble MS; Berglund P B-factor Guided Proline Substitutions in Chromobacterium violaceum Amine Transaminase: Evaluation of the Proline Rule as a Method for Enzyme Stabilization. ChemBioChem. 2019, 20, 1297–1304. [PubMed: 30637901] (g) Tang H; Shi K; Shi C; Aihara XH; Zhang J; Du G Enhancing subtilisin thermostability through a modified normalized B-factor analysis and loop-grafting strategy. J. Biol. Chem 2019, 294, 18398–18407. [PubMed: 31615894]

(33). Pedregosa F; Varoquaux G; Gramfort A; Michel V; Thirion B; Grisel O; Blondel M; Prettenhofer P; Weiss R; Dubourg V; Vanderplas J; Passos A; Cournapeau D; Brucher M; Perrot M; Duchesnay E Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res 2011, 12, 2825–2830.

(34). (a) Hawkins DM The Problem of Overfitting. J. Chem. Inf. Comput. Sci 2004, 44, 1–12. [PubMed: 14741005] (b) Guo S; Bocklitz T; Neugebauer U; Popp J Common mistakes in cross-validating classification models. Anal. Methods 2017, 9, 4410–4417.

(35). Maley SM; Kwon D-H; Rollins N; Stanley JC; Sydora OL; Bischof SM; Ess DH Quantum-mechanical transition-state model combined with machine learning provides catalyst design features for selective Cr olefin oligomerization. Chem. Sci 2020, 11, 9665–9674. [PubMed: 34094231]

(36). (a) Harper KC; Sigman MS Three-Dimensional Correlation of Steric and Electronic Free Energy Relationships Guides Asymmetric Propargylation. Science 2011, 333, 1875–1878. [PubMed: 21960632] (b) Harper KC; Bess EN; Sigman MS Multidimensional Steric Parameters in the Analysis of Asymmetric Catalytic Reactions. Nat. Chem 2012, 4, 366–374. [PubMed: 22522256] (c) Milo A; Bess EN; Sigman MS Interrogating Selectivity in Catalysis Using Molecular Vibrations. Nature 2014, 507, 210–214. [PubMed: 24622199] (d) Mougel V; Santiago CB; Zhizhko PA; Bess EN; Varga J; Frater G; Sigman MS; Copéret C Quantitatively Analyzing Metathesis Catalyst Activity and Structural Features in Silica-Supported Tungsten Imido-Alkylidene Complexes. J. Am. Chem. Soc 2015, 137, 6699–6704. [PubMed: 25938259] (e) Niemeyer ZL; Milo A; Hickey DP; Sigman MS Parameterization of Phosphine Ligands Reveals Mechanistic Pathways and Predicts Reaction Outcomes. Nat. Chem 2016, 8, 610–617. [PubMed: 27219707] (f) Sigman MS; Harper KC; Bess EN; Milo A The Development of Multidimensional Analysis Tools for Asymmetric Catalysis and Beyond. Acc. Chem. Res 2016, 49, 1292–1301. [PubMed: 27220055] (g) Yamamoto E; Hilton MJ; Orlandi M; Saini V; Toste FD; Sigman MS Development and Analysis of a Pd(0)-Catalyzed Enantioselective 1,1-Diarylation of Acrylates Enabled by Chiral Anion Phase Transfer. J. Am. Chem. Soc 2016, 138, 15877–15880. [PubMed: 27960315] (h) Christian AH; Jia S; Cao W; Zhang P; Meza AT; Sigman MS; Chang CJ; Toste FD A Physical Organic Approach to Tuning Reagents for Selective and Stable Methionine Bioconjugation. J. Am. Chem. Soc 2019, 141, 12657–12662. [PubMed: 31361488]

(37). (a) Tsai C-C; Sandford C; Wu T; Chen B; Sigman MS; Toste FD Enantioselective Intramolecular Allylic Substitution via Synergistic Palladium/Chiral Phosphoric Acid Catalysis: Insight into Stereoinduction through Statistical Modeling. Angew. Chem., Int. Ed 2020, 132, 14755–14763. (b) Miró J; Gensch T; Ellwart M; Han S-J; Lin H-H; Sigman MS; Toste FD Enantioselective Allenoate-Claisen Rearrangement Using Chiral Phosphate Catalysts. J. Am. Chem. Soc 2020, 142, 6390–6399. [PubMed: 32182422]

(38). Schrödinger Release 2014-3, MacroModel; Schrödinger LLC: 2014.

(39). Brown HC; Okamoto Y Electrophilic Substituent Constants. J. Am. Chem. Soc 1958, 80, 4979–4987.

(40). The regiomeric ratio (exo/endo) was used as a proxy for the ratio of rates (k_{exo}/k_{endo}), as this is likely representative of the difference in relative rates in this intramolecular competition reaction.

(41). (a) Rolston JH; Yates K Polar additions to the styrene and 2-butene systems. III. Kinetics and linear free energy relations. J. Am. Chem. Soc 1969, 91, 1483–1491.(b) Dubois JE; Schwarcz A Essai de correlation reactivite-structure relative aux vitesses d'addition du brome sur les derives substitues du styrene. Tetrahedron Lett. 1964, 5, 2167–2173.

(42). (a) Suda M Reactions of 1,1-Difluoro-1-Olefins with Electrophilic Reagents. Tetrahedron Lett. 1980, 21, 2555–2556.(b) Morikawa T; Kumadaki I; Shiro M Electrophilic Cyclization Reaction of gem-Difluoroolefin Derivatives: Syntheses of 6,6-Difluorotetrahydro-2-pyrones and 2,2-Difluorotetrahydropyran via Halogen Induced Cyclization. Chem. Pharm. Bull 1985,

33, 5144–5146.(c) Fujita T; Kinoshita R; Takanohashi T; Suzuki N; Ichikawa J Ring-Size-Selective Construction of Fluorine-Containing Carbocycles via Intramolecular Iodoaryltion of 1,1-Difluoro-1-alkenes. Beilstein J. Org. Chem 2017, 13, 2682–2689. [PubMed: 29564005]

(43). Only six atoms on the phenyl ring are used to calculate *B*-factor values. The larger *B*-factor of the pentamethylphenyl group is not due to the rotamers of the methyl groups. The tetrafluoroborate anion and all hydrogen atoms are excluded in flexibility calculations.

(44). Santiago CB; Guo J-Y; Sigman MS Predictive and mechanistic multivariate linear regression models for reaction development. Chem. Sci 2018, 9, 2398–2412. [PubMed: 29719711]

(45). Kim Y; Mittal A; Robichaud DJ; Pilath HM; Etz BD; John PC St.; Johnson DK; Kim S Prediction of Hydroxymethylfurfural Yield in Glucose Conversion through Investigation of Lewis Acid and Organic Solvent Effects. ACS Catal. 2020, 10, 14707–14721.

(46). (a) Janet JP; Kulik HJ Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure-Property Relationships. J. Phys. Chem. A 2017, 121, 8939–8954. [PubMed: 29095620] (b) Nielsen MK; Ahneman DT; Riera O; Doyle AG Deoxyfluorination with Sulfonyl Fluorides: Navigating Reaction Space with Machine Learning. J. Am. Chem. Soc 2018, 140, 5004–5008. [PubMed: 29584953]

(47). The following energy terms were found to strongly correlate with $\Delta E_{disp}(2)$ and thus also contribute to the enantioselectivity: $\Delta E_{elstat}(2)$, $\Delta E_{Pauli}(2)$, $\Delta E_{Polar}(2)$, $\Delta E_{ct}(2)$, and $\Delta E_{solv}(2)$.

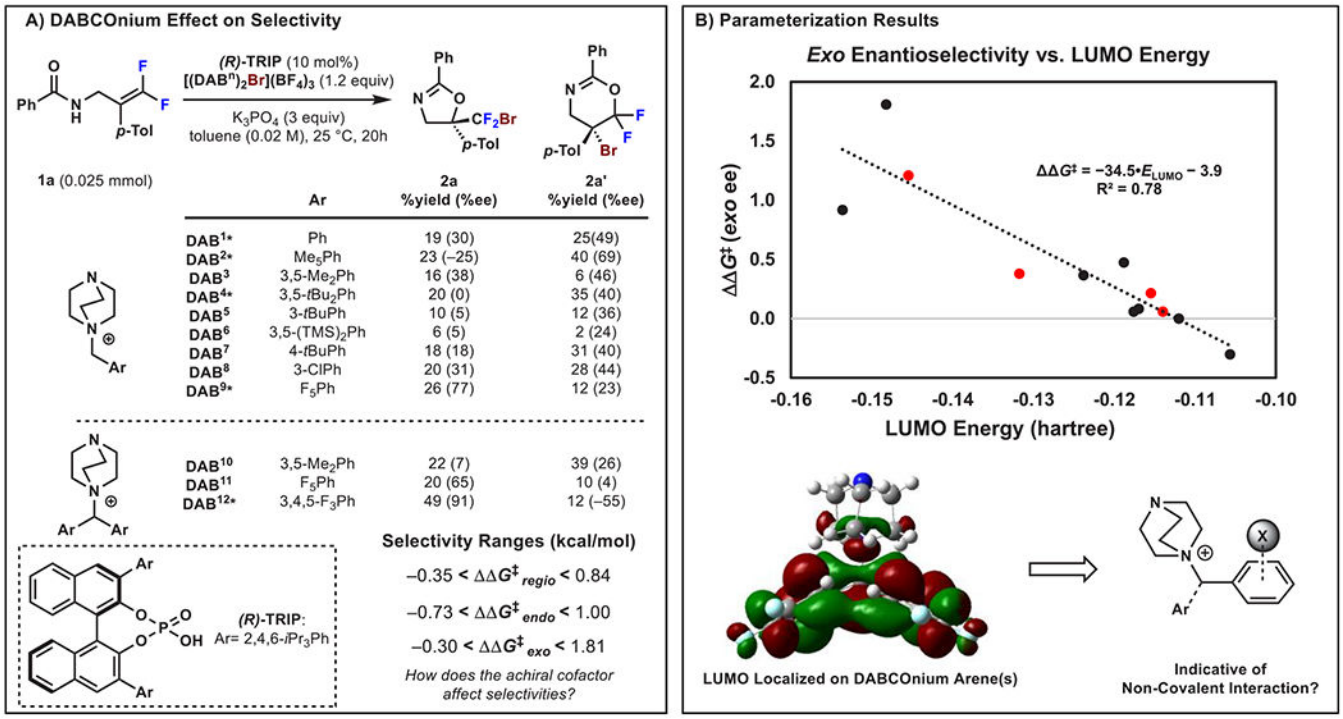


Figure 1.

(A) Effect of DABCOnium reagent on the regio- and enantioselectivity of CAPT-catalyzed bromocyclization of difluoroalkenes (* denotes previously reported data[18]). (B) Linear regression model relating the LUMO energy of the DABCOnium library to *exo* and *endo* enantioselectivities. Data points in black represent the training set; data points in red represent the validation set. $\Delta\Delta G^{\ddagger}$ (*exo* ee) was calculated from the equation $\Delta\Delta G^{\ddagger}$(*exo* ee) = $RT\ln$(*R*-2a/*S*-2a), where T = 298.15 K.

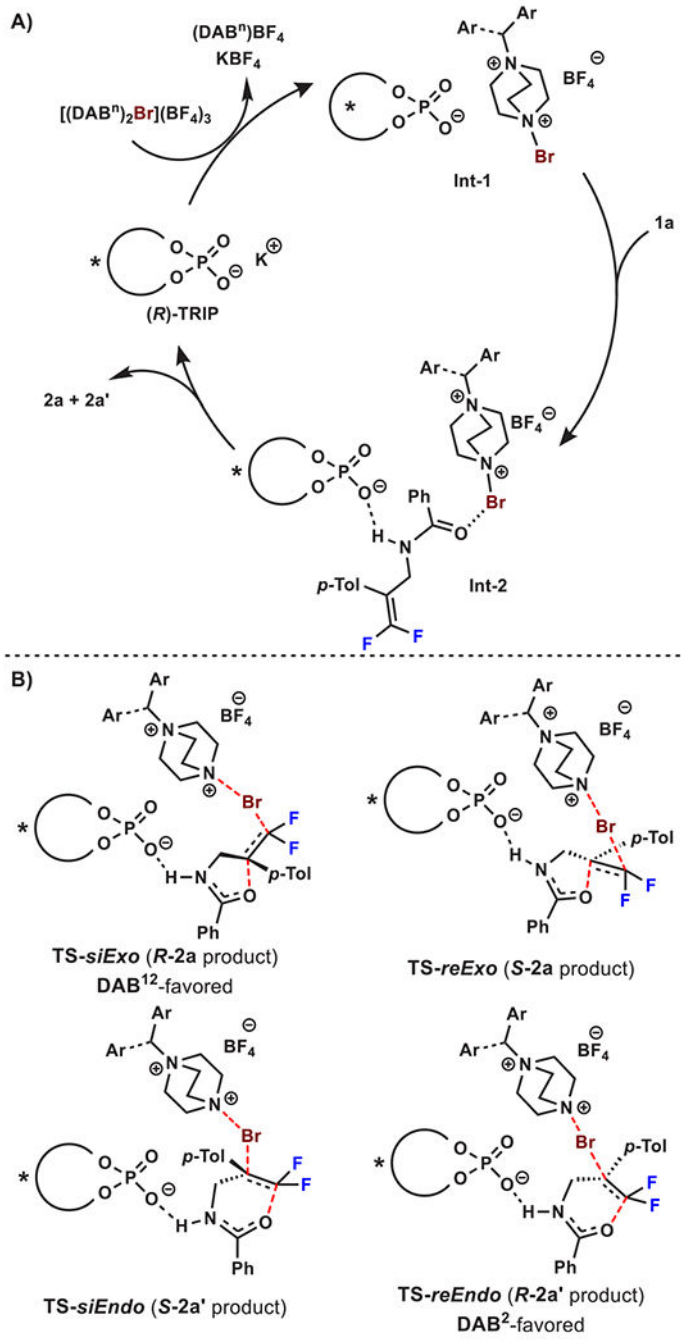


Figure 2.

(A) Proposed catalytic cycle and (B) TSs leading to the four possible products of the bromocyclization of difluoroalkene **1a**.

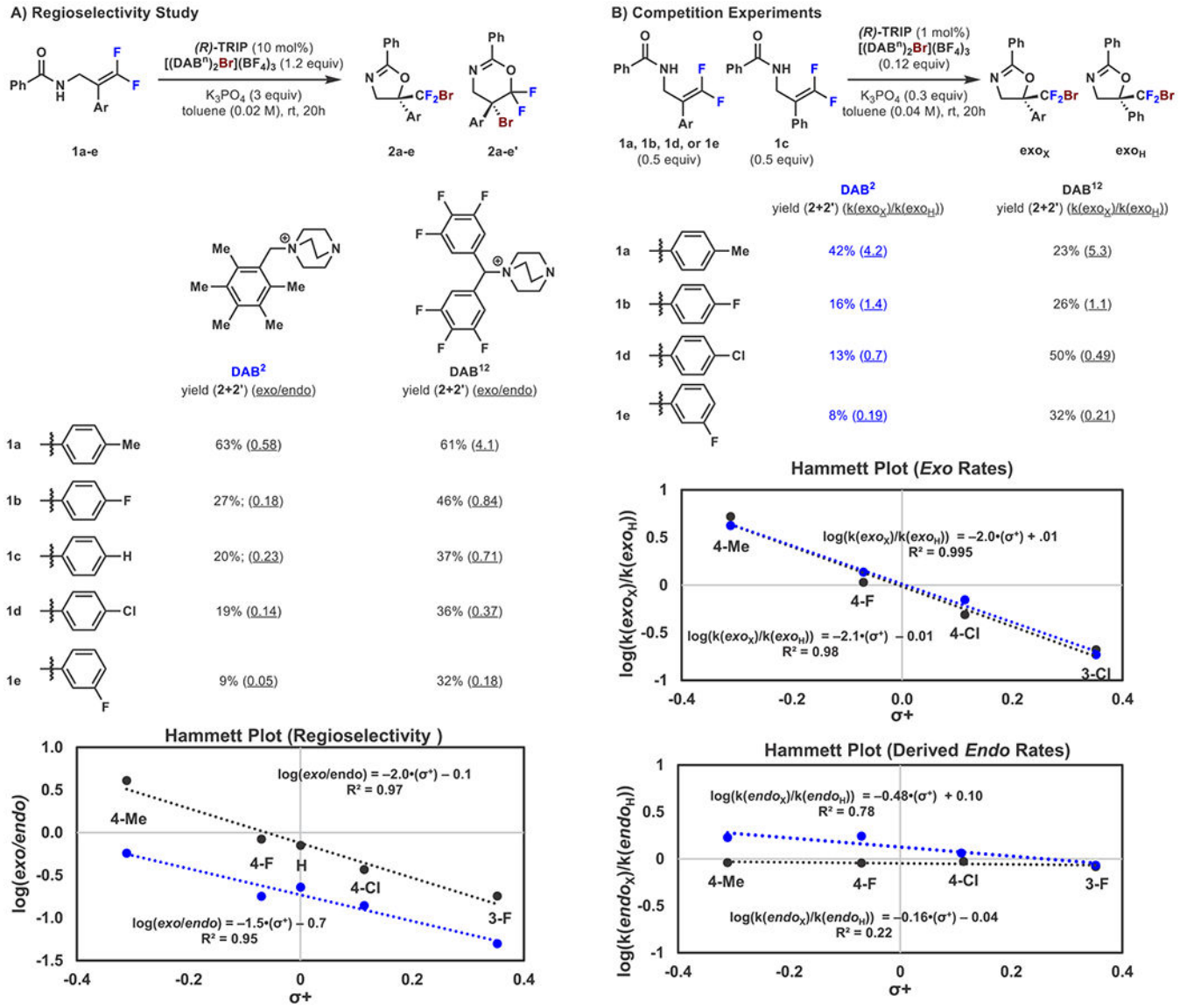


Figure 3.

Hammett plot relating arene electronics to (A) regioselectivity and (B) relative rates of *exo* and *endo* cyclization. *endo* competition rates were derived from the equation $\frac{\text{endo}_X}{\text{endo}_H} = \frac{R_H \times \text{Comp}_{X/H}}{R_X}$, where $R_H = \frac{exo_H}{endo_H}$, $R_X = \frac{exo_X}{endo_X}$, and $Comp_{X/H} = \frac{exo_X}{exo_H}$. Data points in black represent **DAB^{12}**; data points in blue represent **DAB^{2}**.

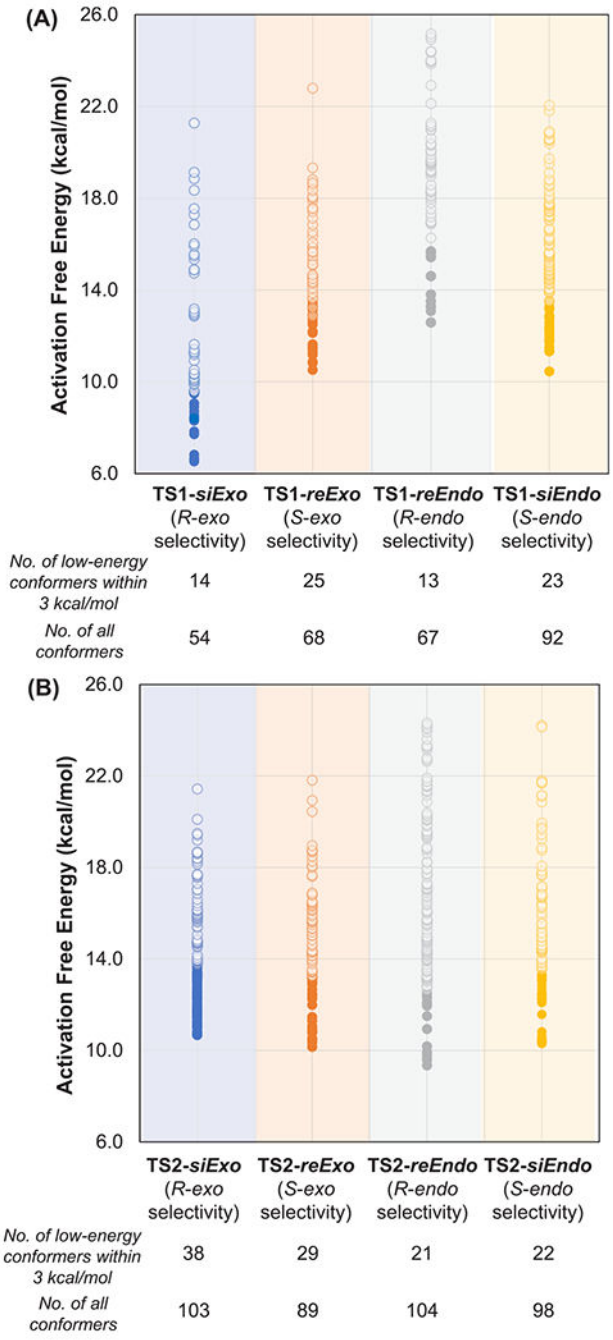


Figure 4.

Activation free energies (kcal/mol) of all low-energy conformers of each TS in the (A) **DAB^{12}**- and (B) **DAB^{2}**-mediated bromocyclization of **1a**. Energy values of all conformers are relative to the lowest-energy conformer of **Int-2a** and **Int-2b** for **DAB^{12}** and **DAB^{2}**, respectively. Solid and empty circles represent conformers within and beyond 3 kcal/mol relative to the lowest-energy conformer of each TS, respectively.

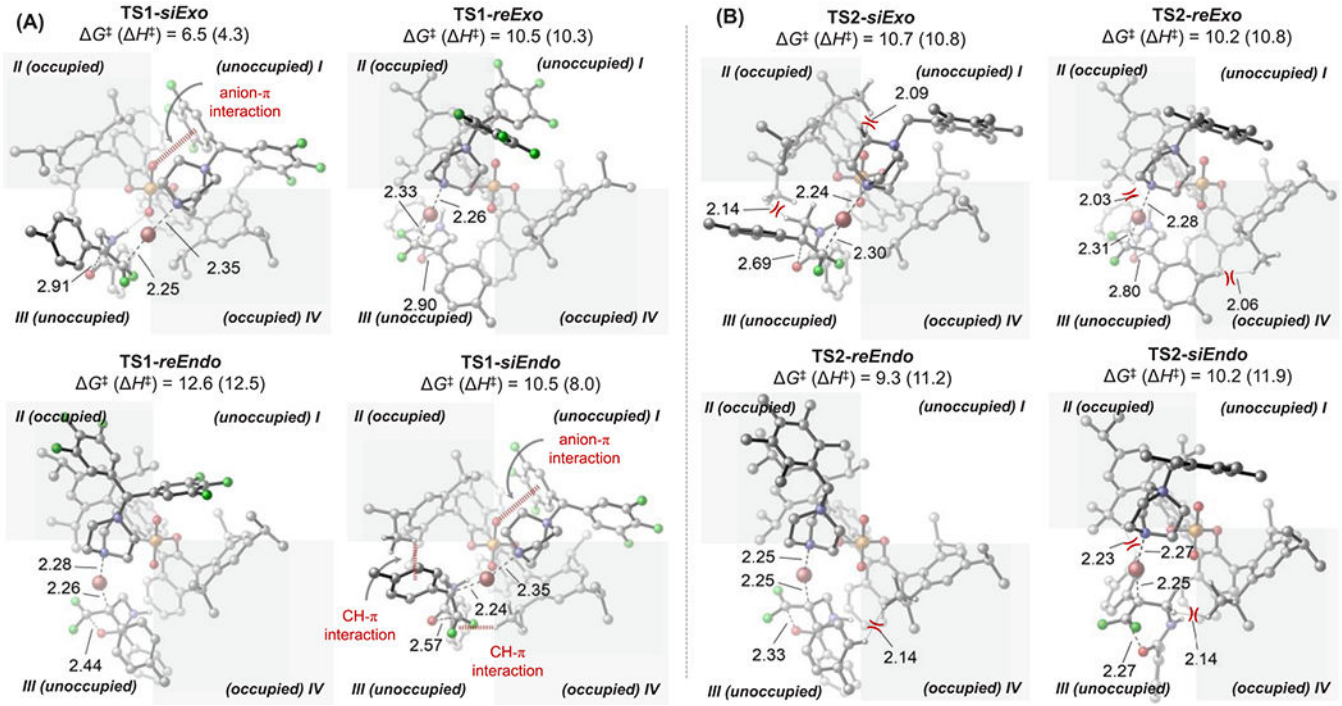


Figure 5.

Optimized structures of the lowest-energy conformers of the four bromocyclization TSs with (A) **DAB^{12}** and (B) **DAB^{2}**. The BF_4^- anion and most hydrogen atoms are omitted for clarity. Gibbs free energies and enthalpies (in kcal/mol) are relative to the lowest-energy conformer of **Int-2a** and **Int-2b** for **DAB^{12}** and **DAB^{2}**, respectively.

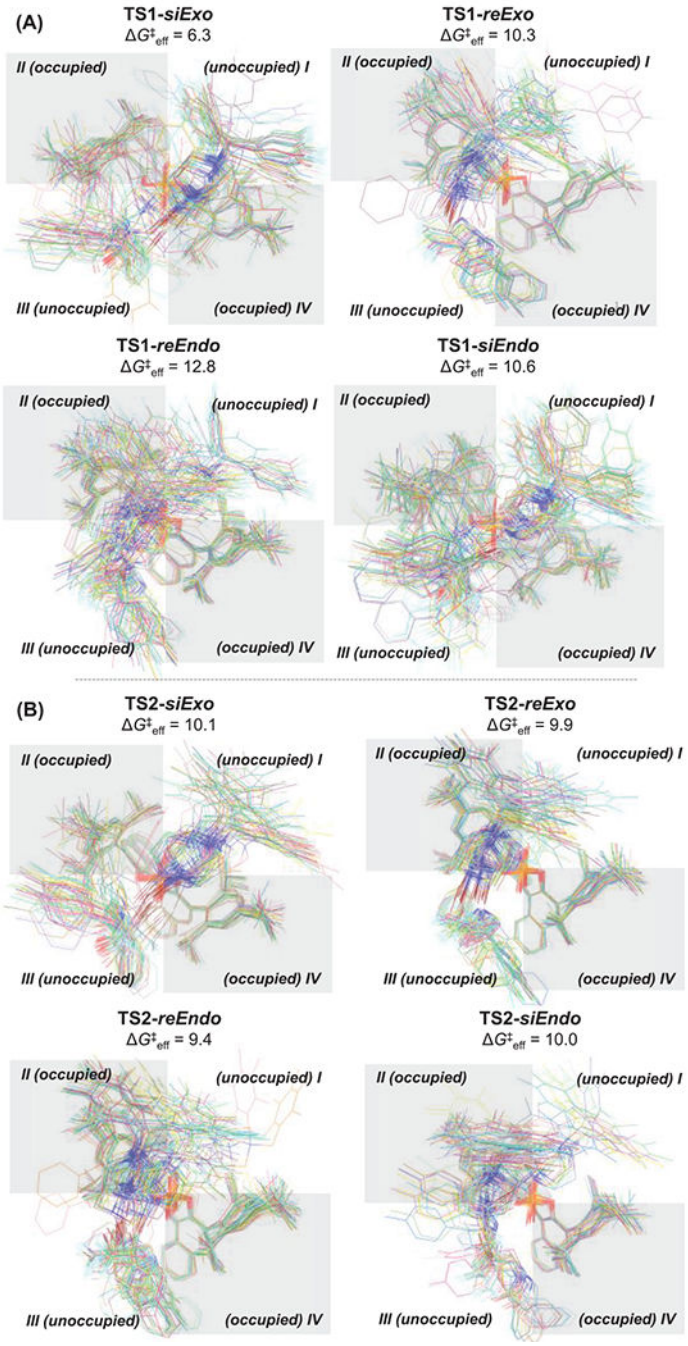


Figure 6.

Overlay of all low-energy conformers within 3 kcal/mol relative to the most stable conformer of the TSs with (A) **DAB^{12}** and (B) **DAB^{2}**. The BF_4^- and all hydrogen atoms are omitted for clarity. Effective activation free energies of TSs ($\Delta G^{\ddagger}_{eff}$, kcal/mol) are relative to the effective free energies of **Int-2a** for **DAB^{12}** and **Int-2b** for **DAB^{2}**.

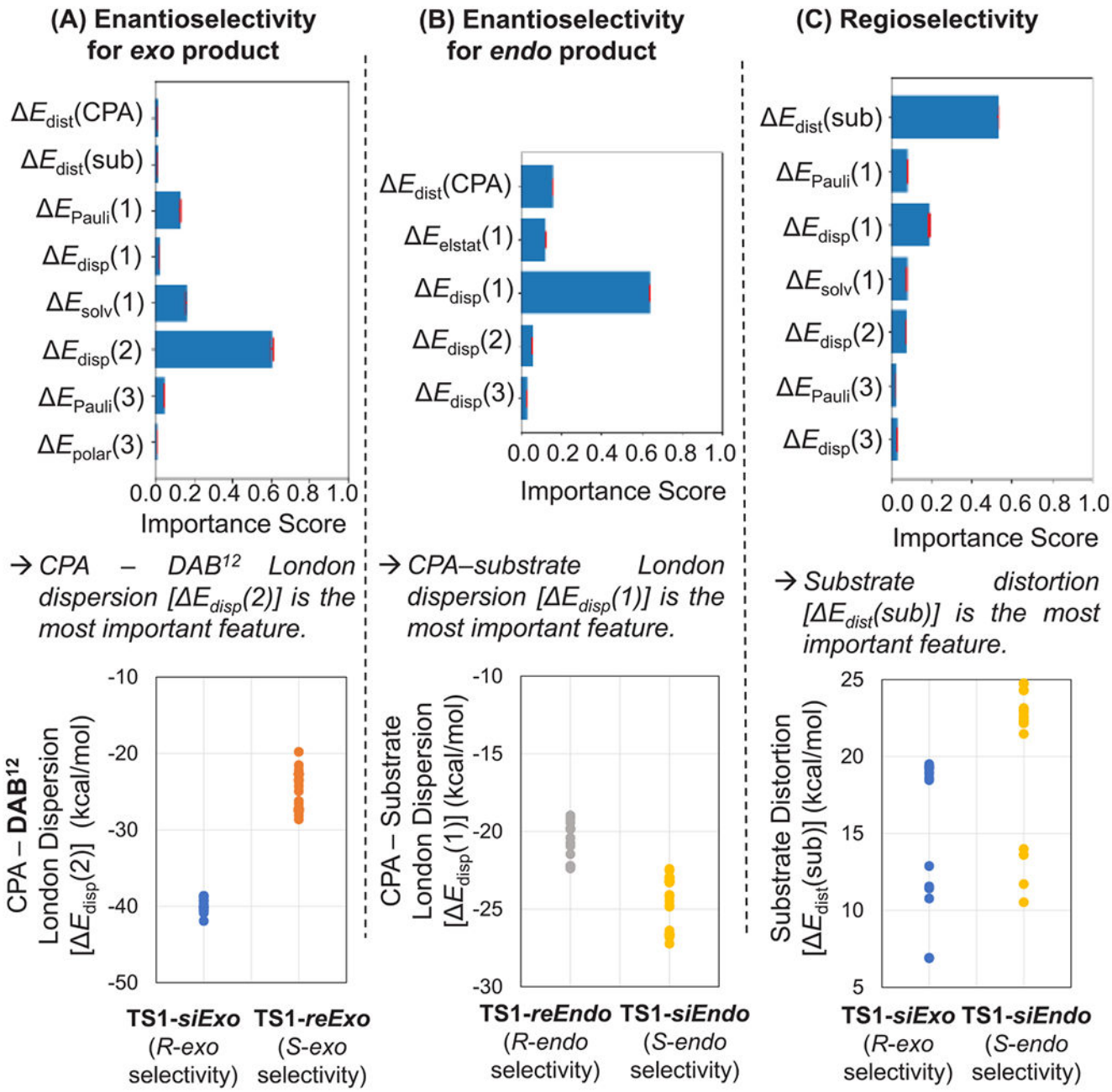


Figure 7.

Using feature importance to identify dominant factors controlling the selectivities in the bromocyclization with **DAB[12]**: (top) feature importance scores of EDA energy terms for predicting enantioselectivities of the *exo* (A) and *endo* (B) products, as well as regioselectivity (C); (bottom) values of the most important EDA energy terms for all low-energy TS conformers.

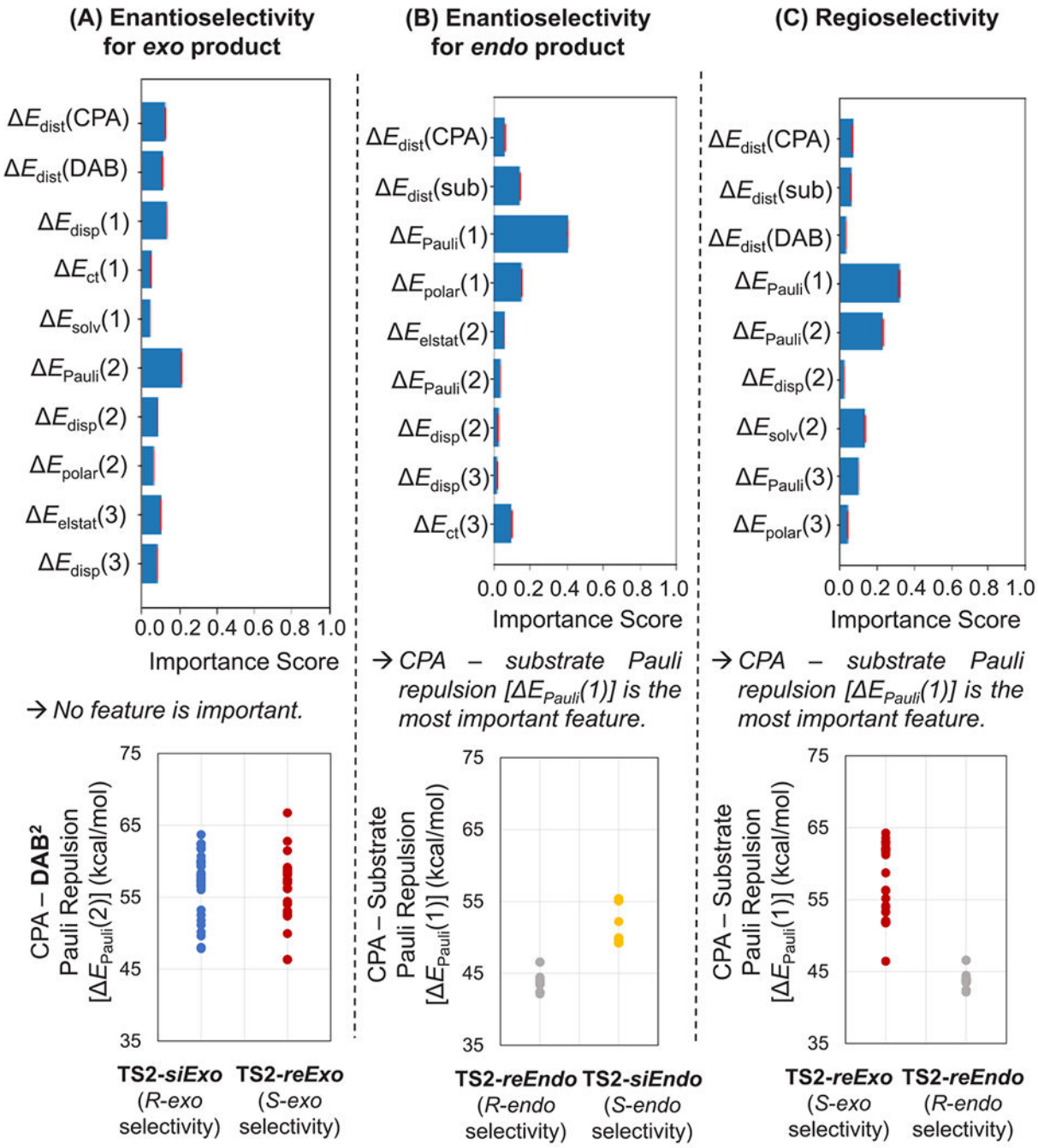
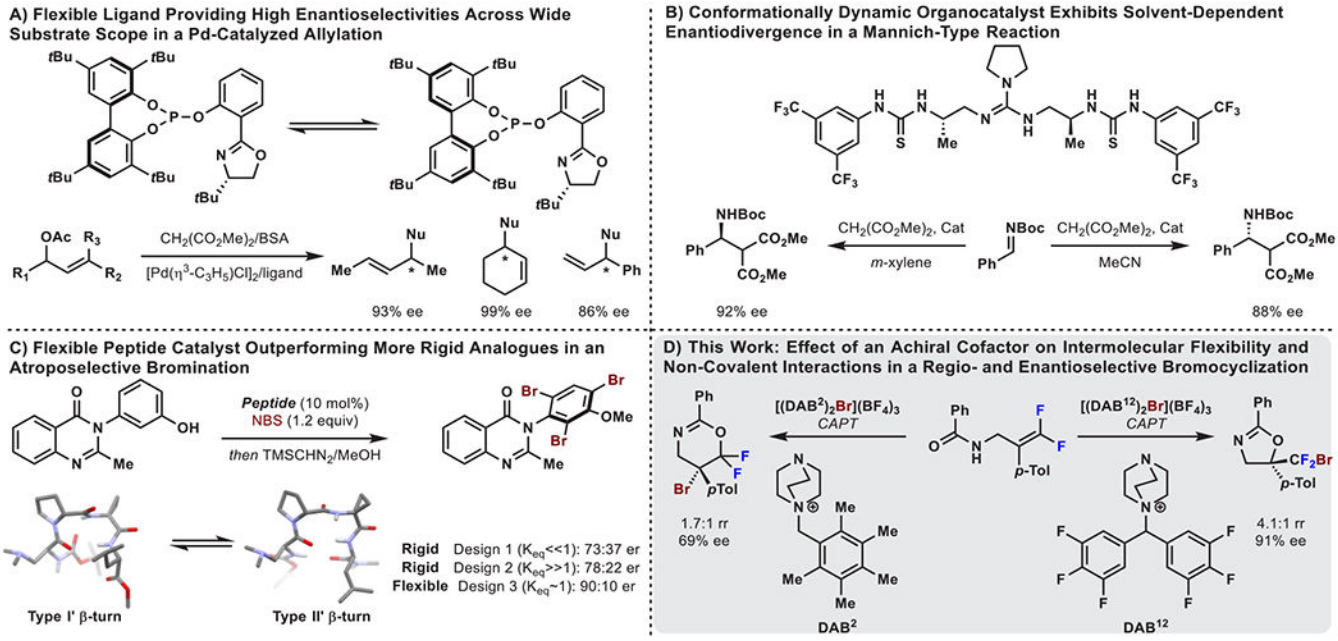


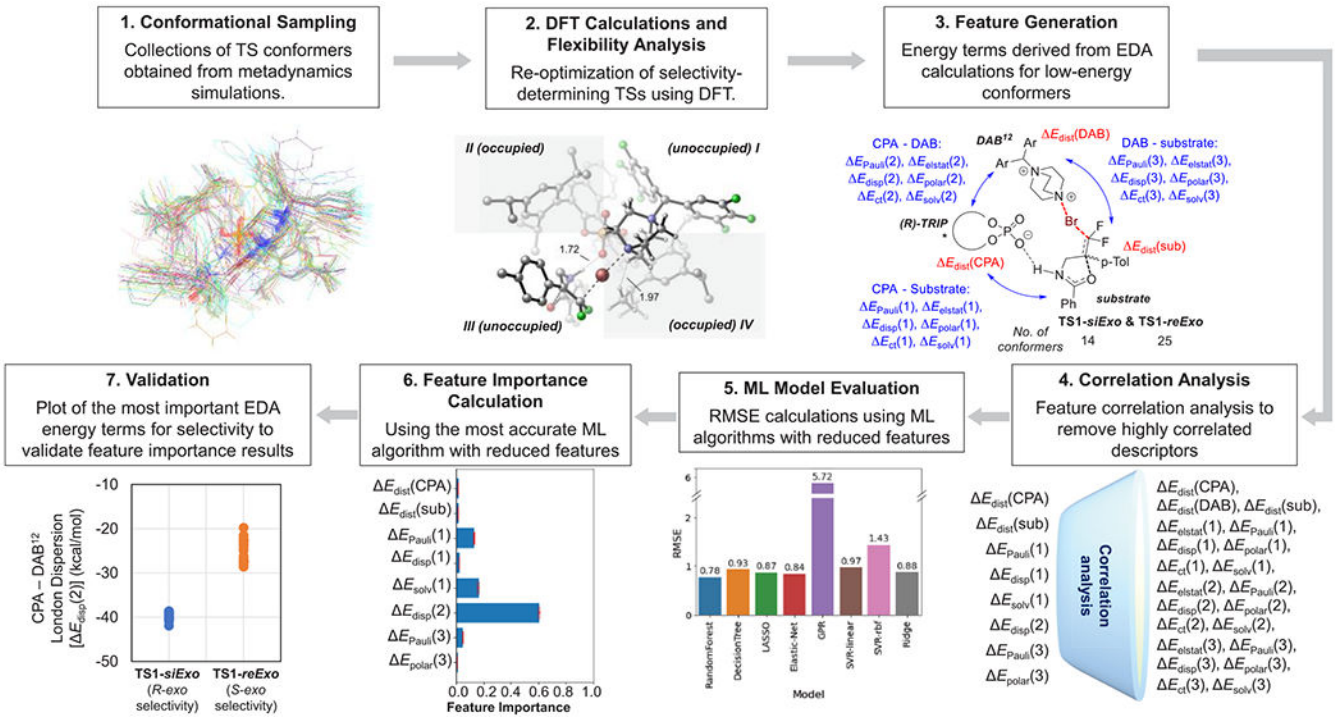
Figure 8.

Using feature importance to identify dominant factors controlling the selectivities in the bromocyclization with **DAB²**: (top) feature importance scores of EDA energy terms for predicting enantioselectivities of *exo* (A) and *endo* (B) products, as well as regioselectivity (C); (bottom) values of the most important EDA energy terms for all low-energy TS conformers.



Scheme 1.

Overview of Flexible Catalyst Systems Presenting Novel Enantioselectivity Trends

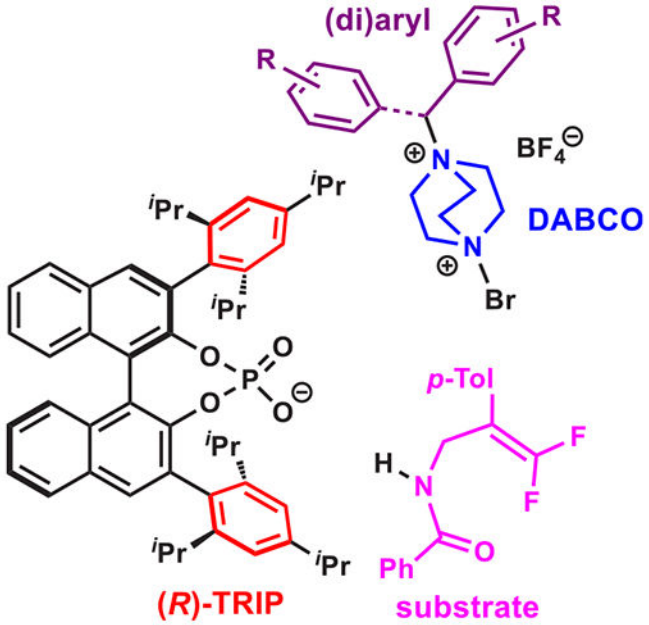


Scheme 2.

General Workflow of the Energy Decomposition Analysis/Machine Learning (EDA/ML) Approach to Reveal Dominant Noncovalent Interactions That Control Regio- and Enantioselectivity

Table 1.

Average Unweighted *B*-factors (Å^2) of the CPA, DABCOnium, and Substrate in Intermediates and TSs[a]



DAB	Fragment	Int-1a	Int-2a	TS1-*siExo*	TS1-*reExo*	TS1-*reEndo*	TS1-*siEndo*
DAB[12]	(*R*)-TRIP	99.6	76.5	31.2	32.0	44.7	37.6
	DABCO	169.3	17.7	14.1	31.0	76.0	24.7
	$CH(3,4,5\text{-}F_3Ph)_2$	280.0	96.5	44.3	138.0	112.5	71.9
	substrate **(2a)**		364.9	74.4	172.3	135.8	115.0
		Int-1b	**Int-2b**	**TS2-*siExo***	**TS2-*reExo***	**TS2-*reEndo***	**TS2-*siEndo***
DAB[2]	(*R*)-TRIP	85.3	116.1	38.2	53.1	157.4	101.4
	DABCO	131.1	77.2	37.5	57.0	198.0	178.9
	$CH_2(Me_5Ph)$	194.7	346.9	85.5	107.2	233.3	353.7
	substrate **(2a)**		428.0	47.7	228.5	118.1	174.9
	Less Flexible					More Flexible	

[a]Average *B*-factor values of highlighted atoms in each fragment (all H atoms were excluded). Definitions: **Int-1**, CPA-DABCOnium complex; **Int-2**, CPA-DABCOnium-substrate complex; **TS1/TS2**, bromocyclization transition states.