

The First Frustrated Lewis Pairs Database: Machine Learning and Cheminformatics-Aided Prediction of Small Molecule Activation

Published as part of *The Journal of Physical Chemistry C* special issue "J. Karl Johnson Festschrift".

Jingyun Ye,* Chathura Wijethunga, and Megan McEwen



Cite This: *J. Phys. Chem. C* 2025, 129, 14346–14355



Read Online

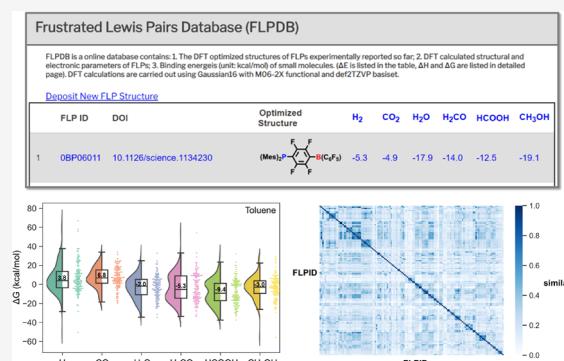
ACCESS |

Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Frustrated Lewis Pairs (FLPs) are a burgeoning field in chemistry, with novel FLPs and surprising new reactions continually being discovered. However, experimentally examining each FLP's activity toward all possible small molecules (SMs) is impractical due to time constraints and potential dangers from toxic, flammable, or explosive SMs. Here, we developed the first open-access Frustrated Lewis pairs Database (FLPDB), with DFT-optimized atomic structures and computed binding free energies and electronic properties data for each FLP toward SMs, including H₂, CO₂, H₂O, H₂CO, HCOOH, and CH₃OH, via high-throughput computational screening. Machine learning was employed to predict the small molecule binding free energies and identify the most important feature governing the binding free energies. Further integrating molecular fingerprints, the H₂ binding free energies can be predicted with a single feature—hydride affinity—for FLPs with molecular similarity scores ≥ 0.55 . The FLPDB and the identified structure–activity relationships in this work will enable countless applications for novel FLP design, new reaction discovery, or toxic gas sensing and detection.



INTRODUCTION

Frustrated Lewis Pairs (FLPs) are combinations of bulky Lewis acids (LAs) and bases (LBs) that are sterically hindered from forming classical Lewis acid–base adducts (Figure 1, top center). This leaves the unquenched LA and LB sites available to accept and donate electrons, enabling unique routes for the activation and conversion of small molecules (SMs),^{1–4} which are crucial for a sustainable and clean energy future.^{5–9} In 2006, Stephan et al. reported an FLP that can reversibly activate H₂ across sterically encumbered Lewis acidic boron and Lewis basic phosphorus sites.¹⁰ For the first time, this FLP shifted the paradigm by demonstrating that H₂ activation, previously considered the realm of transition metals for the past 100 years, can be achieved by nonmetal systems. To date, numerous FLPs have been reported to activate SMs (e.g., H₂, CO, CO₂, N₂O, NO, SO₂, H₂O, HCOOH, olefins, acetylenes, and carbonyl compounds),^{11–15} and serve as active metal-free catalysts for small molecule (SM) conversion under mild conditions, including hydrogenation, reduction of CO₂, hydrosilylation, transformations of alkynes to organic derivatives, C–H bond borylation, and polymerization.^{16–22}

FLPs are a burgeoning field in chemistry; novel FLPs and surprising new reactions catalyzed by FLPs are increasingly being discovered. However, experimentally examining the activity of each FLP for all possible small molecules is impractical due to the time, expense, and potential danger from

toxic, flammable, or explosive SMs, such as acetylene, diborane, phosphine, and hydrogen cyanide. Therefore, experimental data on the activity of FLPs are often limited. Unlike experiments, density functional theory (DFT) enables the direct investigation of reaction mechanisms and the identification of activity descriptors at the atomistic scale. In addition, catalyst design via the combination of emerging data science, cheminformatics, machine learning, and the DFT database is ever increasing.^{23–33}

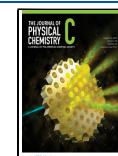
Although density functional theory has been widely applied to study the mechanisms of reactions catalyzed by FLPs,^{20,34–40} research on the exploration of the structure–activity relationships is limited and has been conducted for only a small number of FLPs.^{41–45} The lack of computational data and inconsistent theoretical methods (e.g., different density functionals, basis sets, or convergence criteria^{43,45–49}) make great challenges to systematically analyzing and rationalizing the structure–activity relationships (SARs). Therefore,

Received: May 6, 2025

Revised: June 24, 2025

Accepted: July 17, 2025

Published: July 31, 2025



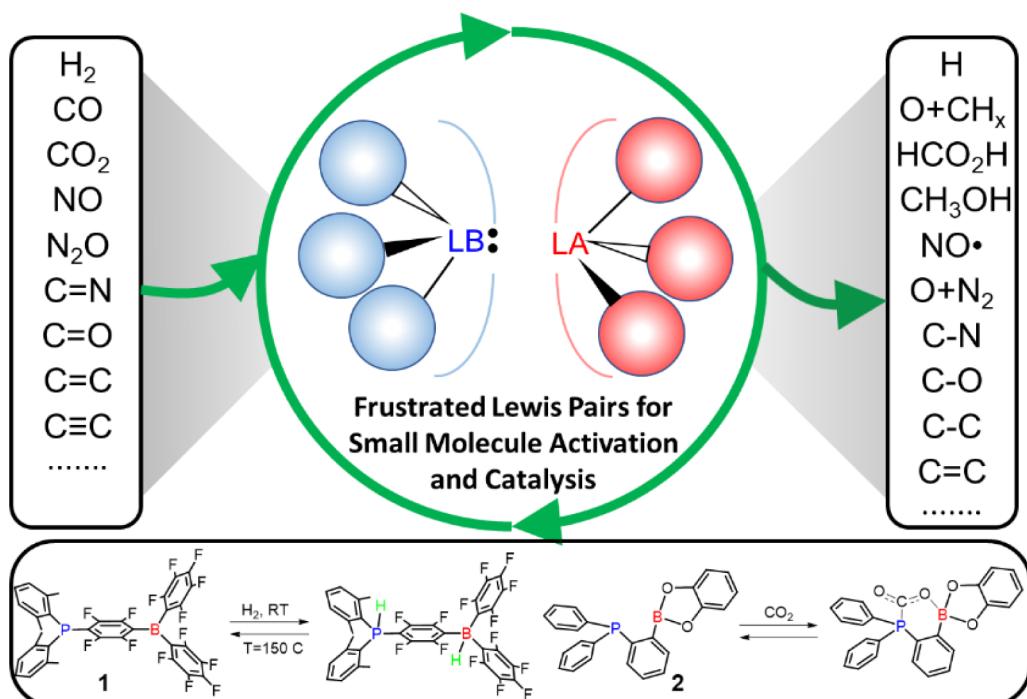


Figure 1. Frustrated Lewis pairs chemistry (top) and the activation of H_2 and CO_2 over FLPs (bottom).

developing a DFT database with structures and properties of all experimentally reported FLPs, along with their binding and activation energy data for various SMs, will provide a comprehensive data set to correlate properties with structures, understand FLP chemistry, and enable FLP design and property forecasting through data science and machine learning.

In this work, we report the first open-access “Frustrated Lewis Pairs Database” (FLPDB), compiling 146 intramolecular FLPs reported experimentally from 2006 to 2023, where LA and LB are covalently connected within a single molecule. We optimized geometries, computed structural and electronic properties, and SM binding energy data using DFT. This first version of FLPDB includes three categories of SMs: (1) H_2 , the clean energy carrier in the foreseen sustainable hydrogen economy, and the primary hydrogen source for many practical hydrogenation reactions in which H_2 activation is the primary step. (2) CO_2 , the most abundant (80%) greenhouse gas in the atmosphere, causing global warming and severe climate change. Studies on the binding and activation of CO_2 at FLPs can accelerate the discovery of promising materials for carbon capture and conversion. (3) H_2O , H_2CO , $HCOOH$, and CH_3OH , the potential products from CO_2 hydrogenation which compete with H_2 and CO_2 activation and influence selectivity and stability.

METHODS

Data Collection. We collected literature published since 2006 containing the phrase “Frustrated Lewis Pairs” in the title or/and abstract. We identified 647 research articles from the Web of Science between 2006 and 2023. Excluding papers that focus on solid materials, compounds involving the FLP concept, and FLPs with metals, 146 intramolecular metal-free FLPs were identified. We reviewed the literature and collected the following data from each paper: (1) the title, DOI, and year of the paper that reported an FLP for the first time; (2) the

Table 1. Features Used for Machine Learning Model Construction

feature	category	definition
1	d_{A-B}	Distance of LA and LB sites
2	$\%V_{bur}$	Percent buried volume
3	ΔG_{H^-}	Hydride affinity, $\Delta G_{H^-} = G_{FLPH^-} - G_{FLP} - G_{H^-}$
4	ΔG_{H^+}	Proton affinity, $\Delta G_{H^+} = G_{FLPH^+} - G_{FLP} - G_{H^+}$
5	χ	Electronegativity, $\chi \approx -1/2(E_{LUMO} + E_{HOMO})$
6	η	Global Chemical hardness, $\eta \approx 1/2(E_{LUMO} - E_{HOMO})$
7	S	Chemical softness, $S \approx 1/(E_{LUMO} - E_{HOMO})$
8	ω	Electrophilicity, $\omega \approx \chi^2/2\eta$
9	E_{HOMO}	HOMO energy
10	E_{LUMO}	LUMO energy
11	E_{gap}	HOMO–LUMO gap, $E_{gap} = (E_{LUMO} - E_{HOMO})$
12	ΔE_{prep}	Preparation energy, $\Delta E_{prep} = E_{FLP(FLP,H2)} - E_{FLP}$
13	q_A	CMS charge of LA site
14	q_B	CMS charge of LB site
15	f_A^+	LA site Fukui function, $f_A^+ = q_A(N+1) - q_A(N)$
16	f_B^-	Local LB site Fukui function, $f_B^- = q_B(N) - q_B(N-1)$
17	Δf	Difference between f_A^+ and f_B^- , $\Delta f_k = f_k^+ - f_k^-$ ($k = A$ or B)
18	ω_A^+	Electrophilicity of LA site, $\omega_A^+ = \omega f_k^+$
19	ω_B^-	Electrophilicity of LB site, $\omega_k^- = \omega f_k^-$
20	LA/LB	Lewis acid atom (B/P, B/N, Si/P, Si/N, Al/P, B/S, P/N)
21	LA – LB	“open”: (no dative bond) or “close”: (dative bond)

structures of FLPs submitted as cif or xyz files, either from the Cambridge Crystallographic Data Centre or in the Supporting

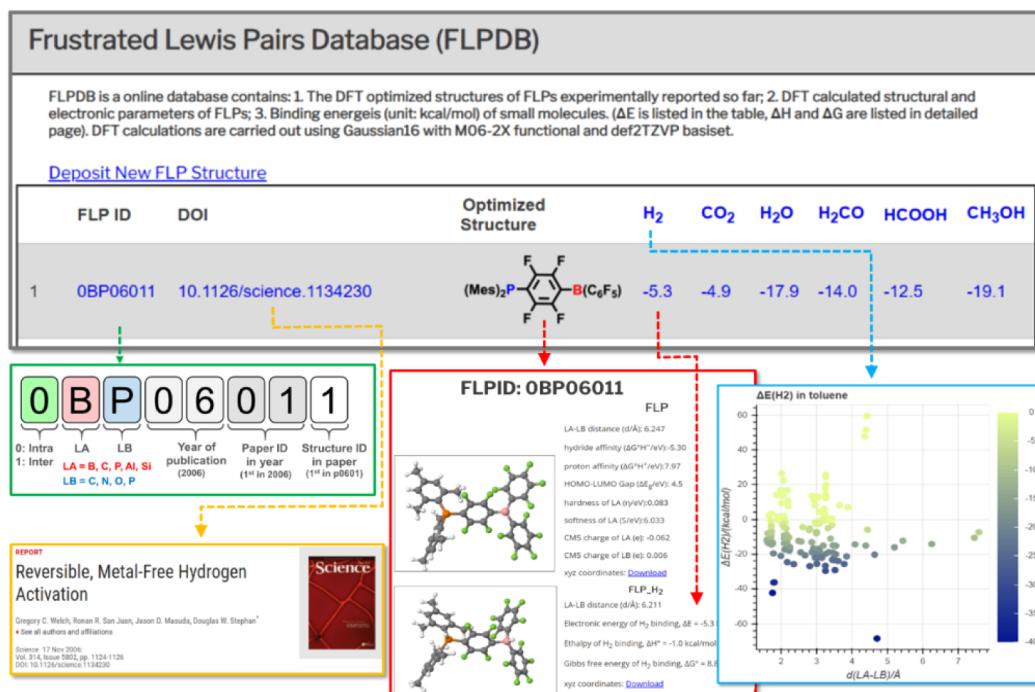


Figure 2. Frustrated Lewis pairs database. FLP ID: FLPs are labeled using 8-character strings that encode information about their structure (intramolecular or intermolecular FLPs, LA and LB center atom types) as well as their provenance (publication year, article ID, and structure ID). For example, **OBP06011** represents an intramolecular (0) FLP with an LA center of B and an LB center of P. The first four digits “0601” following the letters are the reference index; “06” represents the publication year (2006), “01” indicates it is the first publication in 2006 with B/P as LA/LB centers, and the final digit “1” signifies that it is the first FLP structure reported in paper 0601.

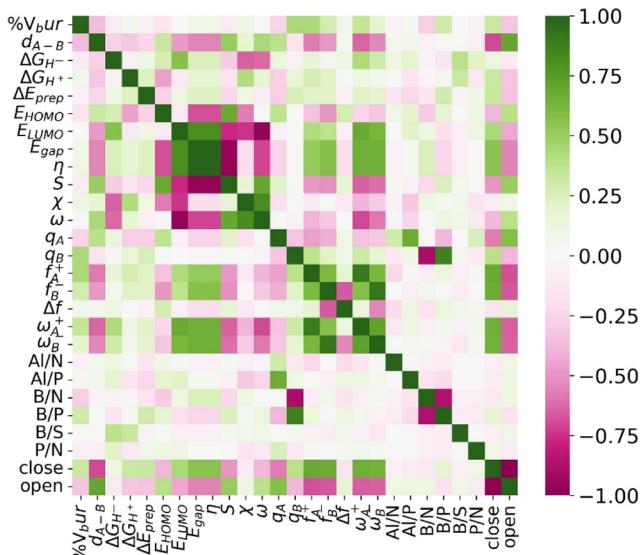


Figure 3. Pearson correlations of features.

Information: (3) the geometries and electronic properties of FLPs optimized and calculated using DFT with Gaussian 16; the optimized geometries of FLPs with H₂, CO₂, H₂O, H₂CO, HCOOH, and CH₃OH binding at LA and LB sites, and their calculated binding energies in the gas phase and toluene.

Density Functional Theory Calculations. Gaussian 16⁵⁰ calculations were performed using the hybrid meta exchange-correlation functional M06-2X⁵¹ with a def2-TZVP basis set for atoms.^{52,53} M06-2X shows accurate results, particularly for main-group chemistry, thermochemistry, kinetics, and non-covalent interactions.⁵¹ The structures of all species were

optimized in the gas phase, and harmonic vibrational frequencies were computed to confirm the nature of all intermediates (no imaginary frequencies). The gas-phase Gibbs free energies, G , were calculated at $T = 298.15$ K and 1 atm pressure using the harmonic approximation for the optimized structures. The solvation effect of toluene, the solvent most commonly used in FLP experiments,⁵⁴ was included by performing single-point energy calculations at the gas-phase geometries using the SMD solvation model.⁵⁵ The relative solution-phase Gibbs free energies were calculated by adding solvation energies to the gas-phase Gibbs free energies. The Cartesian coordinates of all structures and their associated electronic energies, enthalpies, and Gibbs free energies in both the gas phase and in solution are given in the **Supporting Information**. The energy values reported in the main text are Gibbs free energies (298.15 K, standard state of 1 atm for gases, and 1 M for solutes), including the solvent effect of toluene. Single-point DFT calculations were performed for the optimized FLPs to calculate the structure and electronic features listed in **Table 1**. Partial atomic charges were calculated for the gas-phase molecules using the CMS charge model developed by Truhlar and coworkers.⁵⁶ The buried volumes were calculated using the SEQ CROW^{57,58} toolset in ChimeraX.⁵⁹ The occupied volume of a 3.5 Å radius sphere was calculated by defining the center of the sphere as the center of the LA/LB site of the FLPs or the center of the carbon–carbon triple bond in the alkynes.

FLP Database. The first version of FLPDB is available at <https://jingyun-ye.github.io/FLPDB/>. **Figure 2** is a screenshot of the database website with detailed descriptions. Each entry contains the identification of the FLP (FLP ID), the digital object identifier (DOI), the molecular graph, and the binding energies of the SMs.

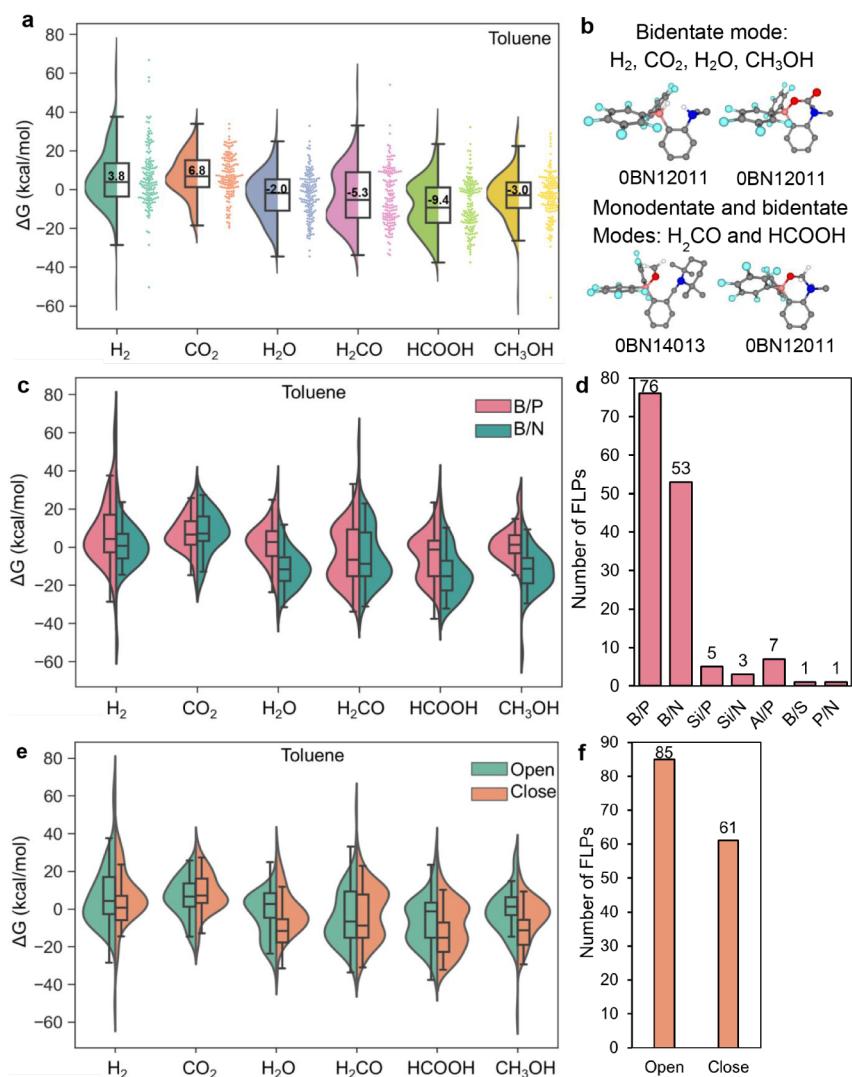


Figure 4. (a) Violin, box, and swarm plots of the DFT-computed binding free energies (ΔG) of H_2 , CO_2 , H_2O , H_2CO , $HCOOH$, and CH_3OH at FLPs in toluene; (b) the structures of monodentate and bidentate binding modes for SMs; (c) and (e) the numbers of FLPs with respect to LA/LB pairs and LA-LB linkage; (d) and (f) violin and box plots of ΔG with respect to B/P and B/N FLPs and the linkage of LA-LB. The box plots show the extrema (whisker tails), interquartile range (box boundaries), and median (horizontal line with the value labeled).

Each DOI links to the original literature (orange box). The FLP ID and molecular graph link to the FLP page (red box), which includes the detailed structure and property data of each FLP. As depicted in the red box in Figure 2, the 3D structures of the optimized FLP (**06BP06011**), without and with H_2 , CO_2 , H_2O , H_2CO , $HCOOH$, or CH_3OH , can be visualized via the embedded Jmol program. Their xyz coordinates can be downloaded via the “Download” link.

The LA-LB distance, hydride affinity, proton affinity, HOMO–LUMO gap, hardness of LA, softness of LA, CMS charge of LA and LB sites, and the electronic energies, enthalpies, and Gibbs free energies of SM binding at FLP are included on the FLP page as well. The column name of each SM (e.g., “ H_2 ”, “ CO_2 ”, “ H_2O ”, “ H_2CO ”, “ $HCOOH$ ”, and “ CH_3OH ”) links to its respective SM page (blue box), containing interactive plots of SM binding energies as a function of structural and DFT-calculated properties. The dotted plot in the blue box of Figure 2 shows electronic binding energies of H_2 in toluene as a function of LA-LB distance ($d_{A-B}/\text{\AA}$).

Primary Features. DFT enables the quantification of the structural and electronic properties of both known and hypothetical systems, including those that are experimentally inaccessible. The features, including the distance between LA and LB sites, buried volume,^{60,61} hydride and proton affinity,^{46,62} chemical hardness and softness,^{63–66} electronegativity and electrophilicity,^{64–66} HOMO–LUMO gap energy, preparation energy,⁴⁶ LA and LB combination (LA/LB), the dative bond type of LA-LB, atomic charge of LA and LB sites,^{67,68} and the Fukui function,^{66,69–75} were selected based on previous studies^{13,23,27–34} for the machine learning model construction. Our previous work revealed that catalytic activity is governed not only by electronic effects but also by steric effects.⁷⁶ The buried volumes of the local LA-LB site of FLPs were calculated for a 3.5 Å radius sphere centered at the LA/LB site of the FLPs. The features are classified into two categories: global features and local features (see Table 1), enabling evaluation of FLPs as a whole and the local properties of LA and LB centers. The local features also included three categorical variables. Further details regarding the features are given in Supporting Information. To evaluate the correlations

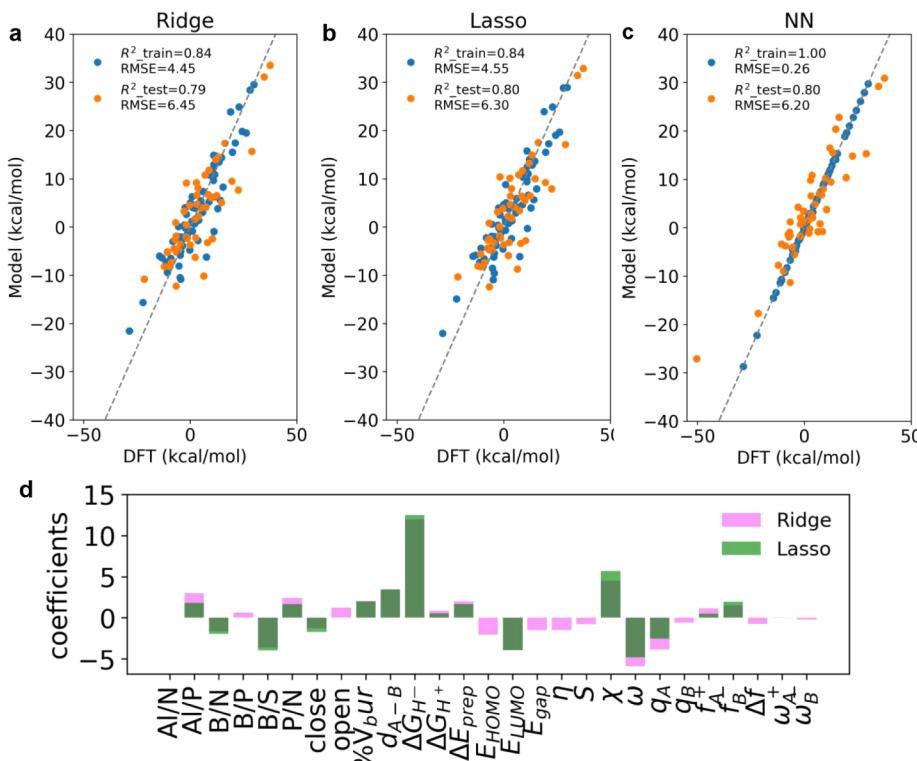


Figure 5. Comparison of the model performance using (a) RidgeCV, (b) LassoCV, (c) multilayer perceptron neural network, and (d) feature importance analysis.

between features, Pearson correlations were analyzed, as shown in Figure 3. This measures the strength of the linear relationship between two variables, with values ranging from -1 to 1 . A value of -1 indicates a total negative linear correlation, 0 indicates no correlation, and $+1$ indicates a total positive correlation. We observed strong linear correlations between certain features, such as electrophilicity (ω) and LUMO orbital energy (E_{LUMO}), HOMO–LUMO gap (E_{gap}) and chemical hardness (η), and chemical softness (S), η , and S .

RESULTS AND DISCUSSION

Binding of Small Molecules. Prior to highlighting how these data can be used in practice, we first investigated several properties of the FLPs database. We began by comparing the DFT-calculated binding free energies of SMs for FLPs in FLPDB. As shown in Figure 4a, the violin and swarm plots display the probability density of the binding free energies for each FLP. We observed quite different binding free energy distribution for each SM. From the box plot, we found that FLPs generally bind H_2O , H_2CO , HCOOH , and CH_3OH more strongly than H_2 and CO_2 . This is because H_2O (1.857 D), H_2CO (2.332 D), HCOOH (1.410 D), and CH_3OH (1.672 D) are polar molecules where the positively charged atom (H or C) attracts the LA site (the electron acceptor), while the negatively charged atom (O) attracts the LB site (the electron donor). In contrast, H_2 and CO_2 are nonpolar molecules with a net dipole moment of 0 D. Therefore, the binding of H_2O , H_2CO , HCOOH , and CH_3OH at FLPs is generally stronger than that of H_2 and CO_2 .

We observed a monomodal binding energy distribution for H_2 , CO_2 , H_2O , and CH_3OH , while H_2CO and HCOOH displayed a bimodal distribution. The monomodal energy distribution for H_2 , CO_2 , H_2O and CH_3OH results from a

single binding mode. As shown in Figure 4b, H_2 binding leads to heterolytic splitting, with one hydride binding at LA and one proton binding at LB. CO_2 binds to FLPs in a bidentate mode, with O binding at LA and C binding at LB. The strong attractions between polar molecules and LA/LB sites cause O–H bond cleavage, particularly for H_2O and CH_3OH , resulting in heterolytic splitting with a proton binding to the LB and OH/ CH_3O binding to the LA. The bimodal energy distribution for H_2CO and HCOOH occurs because they bind to FLPs in either a bidentate mode or a monodentate mode.

Among 146 intramolecular FLPs, 88% are either B/P FLPs (76) or B/N FLPs (53), and 12% comprise other LA/LB pairs (5 Si/P, 3 Si/N, 7 Al/P, 1B/S, and 1P/N), as shown in Figure 4d. We found that B/N FLPs generally bind SMs more strongly than B/P FLPs, except for CO_2 (Figure 4c), which can be attributed primarily to the higher electronegativity of nitrogen (3.0) compared to phosphorus (2.1), leading to a stronger attraction toward polar molecules. This principle holds true for H_2 , as its binding involves heterolytic dissociation with the generation of polarized products—proton, and hydride. However, it does not apply to CO_2 , as CO_2 binds in a bidentate mode at FLPs molecularly.

The intramolecular FLPs in FLPDB can be categorized into two types: open and closed. “Open” FLPs have LA and LB sites separated, while “closed” FLPs have LA and LB sites datively bound to each other. **OBP06011** ($\text{Mes}_2\text{P}-\text{C}_6\text{F}_4-\text{B}(\text{C}_6\text{F}_5)_2$) is an “open” FLP where B and P are separated by a phenylene group (see Figure 2). **OBP07011** ($\text{Mes}_2\text{P}-\text{C}_2\text{H}_2-\text{B}(\text{C}_6\text{F}_5)_2$) is a “closed” FLP where B and P are linked by the $-\text{C}_2\text{H}_2-$ group and form a dative bond, leading to a four-membered ring structure. As shown in Figure 4f, there are 85 “open” and 61 “closed” FLPs in the FLPDB, which account for 58% and 42% of the total number of FLPs, respectively. We

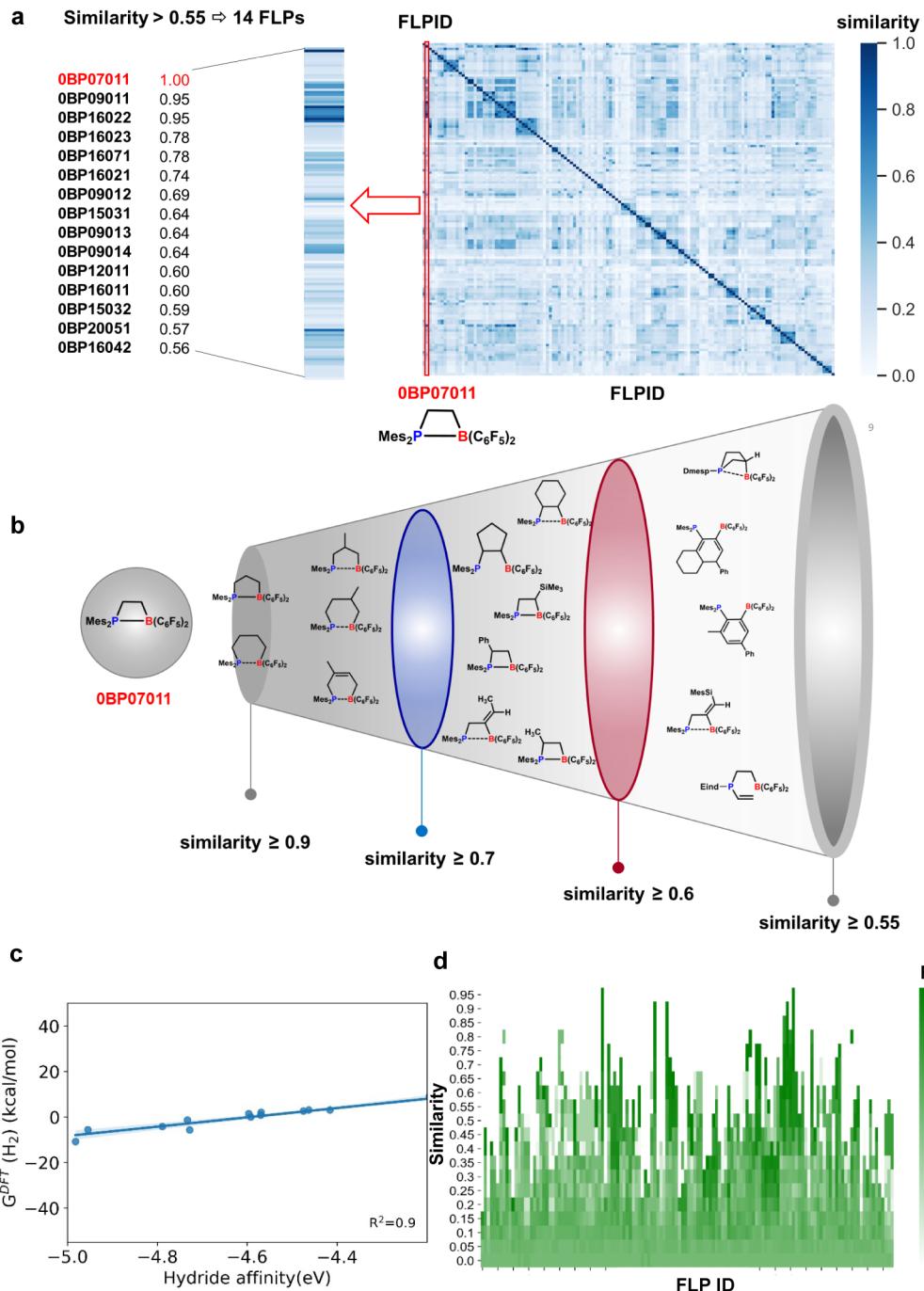


Figure 6. (a) Molecular similarity heatmap; (b) FLP structure derivation as molecular similarity decreases, (c) H_2 binding free energies as a function of hydride affinity for FLPs with a similarity score ≥ 0.55 compared to OBP07011; and (d) R^2 heatmap.

found that “closed” FLPs generally bind H_2O , $HCOOH$, and CH_3OH more strongly than “open” FLPs. This is due to the chemical dissociation (O–H bond cleavage) of the SMs at “closed” FLPs, which leads to stronger binding compared to the molecular binding at “open” FLPs. The binding of H_2 , CO_2 , and H_2CO at “closed” FLPs and “open” FLPs is competitive (Figure 4e). Figure S1 shows a pair plot of binding free energies for different small molecules at FLPs, aiding in the identification of candidates for selective small molecule capture.

Machine Learning. To minimize the number of DFT calculations required for high-throughput screening, we

evaluated seven machine learning (ML) models—Ridge Regression, Lasso Regression, Kernel Ridge, Decision Tree, Random Forest, Gradient Boosting, and multilayer perceptron neural network (5 layers, with each layer containing 100, 80, 50, 30, and 20 nodes, “relu” activation function, “lbgfs” solver, and 10000 iterations)—for predicting small molecule binding energies on FLPs. These predictions were based on FLP features using the Scikit-learn library. Categorical features, including the LA/LB combination type and the LA-LB dative bond type, were converted into numerical data via one-hot encoding. Hyperparameters were optimized using GridSearchCV method with 5-fold cross-validation. The best

hyperparameters, root mean squared error (RMSE), and R^2 values for the training and test data sets for H_2 binding free energies are summarized in Table S1. Ridge and Lasso regression models exhibit similar performance, with train R^2 of 0.84 and a test R^2 around 0.80 (Figure 5a,b). The multilayer perceptron neural network (MLP) shows overfitting, with a train R^2 of 1.00 and a test R^2 of 0.80. We focus on H_2 binding free energies in this work, and the model performance for other small molecules is summarized in Table S2–S6.

Feature importance analysis from Ridge and Lasso regression models indicates that hydride affinity (ΔG_{H^-}) significantly influences predictions of H_2 binding free energies, scoring much higher than other features. However, the linear correlation between H_2 binding free energies and hydride affinity is poor, with an R^2 of 0.5 (see Figure S2). This is expected given the significant variability in FLP structures within FLPDB, including differences in topologies, LA and LB ligands, and linkers. This observation suggests the need to explore correlations among FLPs with similar structures.

Molecular Fingerprints. Molecular fingerprint (FP) is an essential cheminformatics tool for various computational tasks, including similarity searching, clustering, and machine learning. It typically encodes molecule's chemical structure into a list of unique patterns or vectors that capture key features of the molecule.^{77–81} Morgan FPs, also known as extended connectivity fingerprints (ECFPs), are a class of topological fingerprints and are among the most commonly used representations.⁸² We used Morgan FPs implemented in RDKit⁸³ to scan FLPDB and identify the molecular similarities between FLPs. The molecular similarity heatmap is shown in Figure 6a. In this heatmap, the similarity data are organized into a grid, with each cell representing a specific similarity data point. The position of each cell corresponds to its respective FLPs in the data set. Dark blue indicates high similarity, while light blue represents low similarity. For example, the similarity data for all the FLPs compared to FLP **0BP07011** is highlighted in the red box. We identified 14 FLPs with a similarity score of ≥ 0.55 to **0BP07011**. As illustrated in Figure 6b, two FLPs are identified with a similarity score of ≥ 0.9 , and their structures are very similar to **0BP07011**, differing only by the organic linker chain length increase from C2 to C3 and C4. As similarity decreases, FLP structural diversity increases. Three FLPs have similarity scores between 0.7 and 0.9, where the organic linkers start to have branches. With further decreases in similarity, FLPs exhibit more complex branches on the organic linker, completely different linkers, or different Lewis base sites. We plotted the H_2 binding free energies for these 15 FLPs (including **0BP07011**) as a function of hydride affinity, resulting in an R^2 value of 0.9. The H_2 binding free energies scale linearly with the hydride affinity, but only for FLPs with similar chemical structures. Therefore, defining a reasonable molecular similarity threshold to classify similar FLPs is crucial for identifying structure–property relationships and guiding further FLP design. The R^2 heatmap is shown in Figure 6d. In this heatmap, the R^2 data are organized into a grid, with each cell representing a specific FLP with a specific molecular similarity threshold. Dark green indicates a greater R^2 value, while light green represents a smaller R^2 value. On average, a similarity score of 0.55 works well for structure–property relationship identification. For CO_2 , H_2O , HCOOH , CH_3OH , FLPs with similar structures exhibit better prediction performance using only one feature—hydride affinity—than

machine learning models using all the features (see Figure S1–S8).

CONCLUSION

In summary, we developed the first open-access Frustrated Lewis Pairs Database with DFT-optimized atomic structures, computed binding free energies for SMs, including H_2 , CO_2 , H_2O , H_2CO , HCOOH , and CH_3OH , and electronic and structural features for each FLP. We found that employing molecular fingerprints to identify FLPs with similar structures and then correlating the binding free energies with a single feature, outperformed machine learning models fed with all features. The FLPDB provides highly accurate computational data for the development of density functionals, basis sets, and new ML algorithms. The computed descriptors and discovered structure–property relationships will open up novel strategies for the design of different catalysts and reactions. We are upgrading FLPDB by including intermolecular FLPs and expanding to cover toxic, flammable, and explosive SMs.

ASSOCIATED CONTENT

Data Availability Statement

The xyz coordinates, electronic energies, enthalpies, and Gibbs free energies of all the structures in both the gas phase and benzene are downloaded from FLPDB at <https://jingyun-ye.github.io/FLPDB/>.

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcc.5c02882>.

The binding free energies and best hyper-parameters via GridSearchCV with five-fold (PDF)

AUTHOR INFORMATION

Corresponding Author

Jingyun Ye – Department of Chemistry and Biochemistry, Duquesne University, Pittsburgh, Pennsylvania 15282, United States; Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, New York 13699, United States;  orcid.org/0000-0003-4373-9625; Email: yej1@duq.edu, jye@clarkson.edu

Authors

Chathura Wijethunga – Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, New York 13699, United States

Megan McEwen – Department of Chemistry and Biomolecular Science, Clarkson University, Potsdam, New York 13699, United States

Complete contact information is available at: <https://pubs.acs.org/10.1021/acs.jpcc.5c02882>

Author Contributions

J.Y. collected the FLP structures, performed the solvent calculations, feature calculations, data analysis, fingerprint analysis, and machine learning; wrote the manuscript; and created and managed the database. C.W. and M.M. optimized the geometries of FLPs and small molecule binding at FLPs.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by the U.S. National Science Foundation under award NSF 2247481. The DFT computations were performed at the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) under project TG-CHE200106. Megan McEwen is supported in part by the Walsh Undergraduate Research Fellowship of Clarkson University.

REFERENCES

- (1) Stephan, D. W. The Broadening Reach of Frustrated Lewis Pair Chemistry. *Science* **2016**, *354*, aaf7229.
- (2) Stephan, D. W. Frustrated Lewis Pairs. *J. Am. Chem. Soc.* **2015**, *137*, 10018–10032.
- (3) Stephan, D. W. Frustrated Lewis Pairs: From Concept to Catalysis. *Acc. Chem. Res.* **2015**, *48* (2), 306–316.
- (4) Stephan, D. W. *Frustrated Lewis Pair Catalysis: an Introduction*; Springer: Cham, 2021; pp. 1–28.
- (5) Taylor, L. J.; Kays, D. L. Low-Coordinate First-Row Transition Metal Complexes in Catalysis and Small Molecule Activation. In *Dalton Transactions*; Royal Society of Chemistry, 2019; pp. 12365–12381.
- (6) Gardner, B. M.; Liddle, S. T. Small-Molecule Activation at Uranium(III). *Eur. J. Inorg. Chem.* **2013**, *2013*, 3753–3770.
- (7) Liu, X.; He, L.; Liu, Y. M.; Cao, Y. Supported Gold Catalysis: From Small Molecule Activation to Green Chemical Synthesis. *Acc. Chem. Res.* **2014**, *47* (3), 793–804.
- (8) Su, Y.; Kinjo, R. Small Molecule Activation by Boron-Containing Heterocycles. *Chem. Soc. Rev.* **2019**, *48*, 3613–3659.
- (9) Wang, Q.; Brooks, S. H.; Liu, T.; Tomson, N. C. Tuning Metal-Metal Interactions for Cooperative Small Molecule Activation. *Chem. Commun.* **2021**, *57* (23), 2839–2853.
- (10) Welch, G. C.; Juan, R. R. S.; Masuda, J. D.; Stephan, D. W. Reversible, Metal-Free Hydrogen Activation. *Science* **2006**, *314* (5802), 1124 LP–1126.
- (11) Stephan, D. W. Discovery of Frustrated Lewis Pairs: Intermolecular FLPs for Activation of Small Molecules. *Frustrated Lewis Pairs I* 2012, Springer, Berlin, Heidelberg, pp. 1–44.
- (12) Stephan, D. W.; Erker, G. Frustrated Lewis Pair Chemistry of Carbon, Nitrogen and Sulfur Oxides. *Chem. Sci.* **2014**, *5*, 2625–2641.
- (13) Stephan, D. W. Frustrated Lewis Pairs: A New Strategy to Small Molecule Activation and Hydrogenation Catalysis. *J. Chem. Soc. Dalt. Trans.* **2009**, No. 17, 3129–3136.
- (14) Chen, D.; Klankermayer, J. Frustrated Lewis Pairs: From Dihydrogen Activation to Asymmetric Catalysis. *Top Curr. Chem.* **2013**, *334*, 1–26.
- (15) Stephan, D. W. Frustrated Lewis Pair” Hydrogenations Organic and Biomolecular Chemistry. Royal Society of Chemistry, 2012, pp. 5740–5746
- (16) Sumerin, V.; Chernichenko, K.; Schulz, F.; Leskelä, M.; Rieger, B.; Repo, T. *Amine-Borane Mediated Metal-Free Hydrogen Activation and Catalytic Hydrogenation*; Springer: Berlin, Heidelberg, 2012; pp. 111–155.
- (17) Feng, X.; Meng, W.; Du, H. *Frustrated Lewis Pair Catalyzed Asymmetric Reactions*; Springer: Cham, 2021; pp. 29–86.
- (18) Wang, T.; Daniliuc, C. G.; Kehr, G.; Erker, G. *FLP Reduction of Carbon Monoxide and Related Reactions*; Springer: Cham, 2021; pp. 87–112.
- (19) Ashley, A. E.; O’Hare, D. FLP-Mediated Activations and Reductions of CO₂ and CO. *Top Curr. Chem.* **2012**, *334*, 191–217.
- (20) Courtemanche, M. A.; Légaré, M. A.; Maron, L.; Fontaine, F. G. Reducing CO₂ to Methanol Using Frustrated Lewis Pairs: On the Mechanism of Phosphine-Borane-Mediated Hydroboration of CO₂. *J. Am. Chem. Soc.* **2014**, *136* (30), 10708–10717.
- (21) Soltani, Y.; Fontaine, F.-G. *FLP-Mediated C–H-Activation*; Springer: Cham, 2021; pp. 113–166.
- (22) Li, N.; Zhang, W. X. Frustrated Lewis Pairs: Discovery and Overviews in Catalysis. *Chin. J. Chem.* **2020**, *38*, 1360–1370.
- (23) Gu, G. H.; Choi, C.; Lee, Y.; Situmorang, A. B.; Noh, J.; Kim, Y. H.; Jung, Y. Progress in Computational and Machine-Learning Methods for Heterogeneous Small-Molecule Activation. *Adv. Mater.* **2020**, *32*, 1907865.
- (24) Yang, W.; Fidelis, T. T.; Sun, W. H. Machine Learning in Catalysis, from Proposal to Practicing. *ACS Omega* **2020**, *5*, 83–88.
- (25) Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Cent. Sci.* **2018**, *4* (2), 268–276.
- (26) Andersen, M.; Levchenko, S. V.; Scheffler, M.; Reuter, K. Beyond Scaling Relations for the Description of Catalytic Materials. *ACS Catal.* **2019**, *9* (4), 2752–2759.
- (27) Xu, J.; Cao, X. M.; Hu, P. Improved Prediction for the Methane Activation Mechanism on Rutile Metal Oxides by a Machine Learning Model with Geometrical Descriptors. *J. Phys. Chem. C* **2019**, *123* (47), 28802–28810.
- (28) Tran, K.; Ulissi, Z. W. Active Learning across Intermetallics to Guide Discovery of Electrocatalysts for CO₂ Reduction and H₂ Evolution. *Nat. Catal.* **2018**, *1* (9), 696–703.
- (29) Banerjee, S.; Sreenithya, A.; Sunoj, R. B. Machine Learning for Predicting Product Distributions in Catalytic Regioselective Reactions. *Phys. Chem. Chem. Phys.* **2018**, *20* (27), 18311–18318.
- (30) Li, Z.; Ma, X.; Xin, H. Feature Engineering of Machine-Learning Chemisorption Models for Catalyst Design. *Catal. Today* **2017**, *280*, 232–238.
- (31) Krenn, M.; Häse, F.; Nigam, A. K.; Friederich, P.; Aspuru-Guzik, A. Self-Referencing Embedded Strings (SELFIES): A 100% Robust Molecular String Representation. *Mach. Learn.: sci. Technol.* **2020**, *1* (4), 045024.
- (32) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chem. Mater.* **2019**, *31* (9), 3564–3572.
- (33) Guimaraes, G. L.; Sanchez-Lengeling, B.; Outeiral, C.; Farias, P. L. C.; Aspuru-Guzik, A. Objective-Reinforced Generative Adversarial Networks (ORGAN) for Sequence Generation Models. *arXiv* **2018**.
- (34) Holschumacher, D.; Bannenberg, T.; Hrib, C. G.; Jones, P. G.; Tamm, M. Heterolytic Dihydrogen Activation by a Frustrated Carbene-Borane Lewis Pair. *Angew. Chem., Int. Ed.* **2008**, *47* (39), 7428–7432.
- (35) Hamza, A.; Stirling, A.; Rokob, T. A.; Pápai, I. Mechanism of Hydrogen Activation by Frustrated Lewis Pairs: A Molecular Orbital Approach. *Int. J. Quantum Chem.* **2009**, *109* (11), 2416–2425.
- (36) Camaioni, D. M.; Ginoska-Pangovska, B.; Schenter, G. K.; Kathmann, S. M.; Autrey, T. Analysis of the Activation and Heterolytic Dissociation of H₂ by Frustrated Lewis Pairs: NH₃/BX₃ (X = H, F, and Cl). *J. Phys. Chem. A* **2012**, *116* (26), 7228–7237.
- (37) Jiang, B.; Zhang, Q.; Dang, L. Theoretical Studies on Bridged Frustrated Lewis Pair (FLP) Mediated H₂ Activation and CO₂ Hydrogenation. *Org. Chem. Front.* **2018**, *5* (12), 1905–1915.
- (38) Daru, J.; Bakó, I.; Stirling, A.; Pápai, I. Mechanism of Heterolytic Hydrogen Splitting by Frustrated Lewis Pairs: Comparison of Static and Dynamic Models. *ACS Catal.* **2019**, *9* (7), 6049–6057.
- (39) Grimme, S.; Kruse, H.; Goerigk, L.; Erker, G. The Mechanism of Dihydrogen Activation by Frustrated Lewis Pairs Revisited. *Angew. Chem., Int. Ed.* **2010**, *49* (8), 1402–1405.
- (40) Das Neves Gomes, C.; Blondiaux, E.; Thuéry, P.; Cantat, T. Metal-Free Reduction of CO₂ with Hydroboranes: Two Efficient Pathways at Play for the Reduction of CO₂ to Methanol. *Chem. - A Eur. J.* **2014**, *20* (23), 7098–7106.
- (41) Ananthraj, S. Ru-Tweaking of Non-Precious Materials: The Tale of a Strategy That Ensures Both Cost and Energy Efficiency in Electrocatalytic Water Splitting. *J. Mater. Chem. A* **2021**, *9* (11), 6710–6731.
- (42) Li, H.; Zhao, L.; Lu, G.; Mo, Y.; Wang, Z. X. Insight into the Relative Reactivity of “Frustrated Lewis Pairs” and Stable Carbenes in

- Activating H₂ and CH₄: A Comparative Computational Study. *Phys. Chem. Chem. Phys.* **2010**, *12* (20), 5268–5275.
- (43) Mane, M. V.; Vanka, K. Less Frustration, More Activity—Theoretical Insights into Frustrated Lewis Pairs for Hydrogenation Catalysis. *ChemCatChem* **2017**, *9* (15), 3013–3022.
- (44) Bistoni, G.; Auer, A. A.; Neese, F. Understanding the Role of Dispersion in Frustrated Lewis Pairs and Classical Lewis Adducts: A Domain-Based Local Pair Natural Orbital Coupled Cluster Study. *Chem. - A Eur. J.* **2017**, *23* (4), 865–873.
- (45) Zhu, J. Rational Design of a Carbon-Boron Frustrated Lewis Pair for Metal-Free Dinitrogen Activation. *Chem. - An Asian J.* **2019**, *14* (9), 1413–1417.
- (46) Rokob, T. A.; Hamza, A.; Pápai, I. Rationalizing the Reactivity of Frustrated Lewis Pairs: Thermodynamics of H₂ Activation and the Role of Acid-Base Properties. *J. Am. Chem. Soc.* **2009**, *131* (30), 10701–10710.
- (47) Cabrera-Trujillo, J. J.; Fernández, I. Aromaticity Can Enhance the Reactivity of P-Donor/Borole Frustrated Lewis Pairs. *Chem. Commun.* **2019**, *55* (5), 675–678.
- (48) Zhuang, D.; Li, Y.; Zhu, J. Antiaromaticity-Promoted Activation of Dihydrogen with Borole Fused Cyclooctatetraene Frustrated Lewis Pairs: A Density Functional Theory Study. *Organometallics* **2020**, *39* (14), 2636–2641.
- (49) Sharma, G.; Newman, P. D.; Melen, R.; Platts, J. A. Computational Design of an Intramolecular Frustrated Lewis Pair Catalyst for Enantioselective Hydrogenation. *J. Theor. Comput. Chem.* **2020**, *19* (2), 2050009.
- (50) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A., et al.. *Gaussian 09 Revision B.01*, 2009.
- (51) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120* (1–3), 215–241.
- (52) Weigend, F. Accurate Coulomb-Fitting Basis Sets for H to Rn. *Phys. Chem. Chem. Phys.* **2006**, *8* (9), 1057–1065.
- (53) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297–3305.
- (54) Chernichenko, K.; Madarász, A.; Pápai, I.; Nieger, M.; Leskelä, M.; Repo, T. A Frustrated-Lewis-Pair Approach to Catalytic Reduction of Alkynes to Cis-Alkenes. *Nat. Chem.* **2013**, *5* (8), 718–723.
- (55) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. 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* (18), 6378–6396.
- (56) Marenich, A. V.; Jerome, S. V.; Cramer, C. J.; Truhlar, D. G. Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8* (2), 527–541.
- (57) Ingman, V. M.; Schaefer, A. J.; Andreola, L. R.; Wheeler, S. E. QChASM: Quantum Chemistry Automation and Structure Manipulation. *Wiley Interdiscip. Rev.: comput. Mol. Sci.* **2021**, *11* (4), No. e1510.
- (58) Schaefer, A. J.; Ingman, V. M.; Wheeler, S. E. SEQCROW: A ChimeraX Bundle to Facilitate Quantum Chemical Applications to Complex Molecular Systems. *J. Comput. Chem.* **2021**, *42* (24), 1750–1754.
- (59) Meng, E. C.; Goddard, T. D.; Pettersen, E. F.; Couch, G. S.; Pearson, Z. J.; Morris, J. H.; Ferrin, T. E. UCSF ChimeraX: Tools for Structure Building and Analysis. *Protein Sci.* **2023**, *32* (11), No. e4792.
- (60) Shao, Y.; Zhang, J.; Li, Y.; Liu, Y.; Ke, Z. Frustrated Lewis Pair Catalyzed C-H Activation of Heteroarenes: A Stepwise Carbene Mechanism Due to Distance Effect. *Org. Lett.* **2018**, *20* (4), 1102–1105.
- (61) Heshmat, M.; Ensing, B. Optimizing the Energetics of FLP-Type H₂Activation by Modulating the Electronic and Structural Properties of the Lewis Acids: A DFT Study. *J. Phys. Chem. A* **2020**, *124* (32), 6399–6410.
- (62) Mao, H.; Chen, Z.; Cheng, L.; Wang, K. The Activity and Reversibility of Intramolecular B/N Frustrated Lewis Pairs in the Hydrogenation. *Int. J. Hydrogen Energy* **2021**, *46* (1), 110–118.
- (63) Ayers, P. W.; Parr, R. G.; Pearson, R. G. Elucidating the Hard/Soft Acid/Base Principle: A Perspective Based on Half-Reactions. *J. Chem. Phys.* **2006**, *124* (19), 194107.
- (64) Koopmans, T. Ordering of Wave Functions and Eigenenergies to the Individual Electrons of an Atom. *Physica* **1934**, *1* (1–6), 104–113.
- (65) Pearson, R. G. Hard and Soft Acids and Bases—the Evolution of a Chemical Concept. *Coord. Chem. Rev.* **1990**, *100* (C), 403–425.
- (66) Parr, R. G.; Yang, W. Density Functional Approach to the Frontier-Electron Theory of Chemical Reactivity. *J. Am. Chem. Soc.* **1984**, *106* (14), 4049–4050.
- (67) Skara, G.; De Vleeschouwer, F.; Geerlings, P.; De Proft, F.; Pinter, B. Heterolytic Splitting of Molecular Hydrogen by Frustrated and Classical Lewis Pairs: A Unified Reactivity Concept. *Sci. Rep.* **2017**, *7* (1), 1–15.
- (68) Zhuang, D.; Rouf, A. M.; Li, Y.; Dai, C.; Zhu, J. Aromaticity-Promoted CO₂ Capture by P/N-Based Frustrated Lewis Pairs: A Theoretical Study. *Chem. - An Asian J.* **2020**, *15* (2), 266–272.
- (69) Yang, W.; Mortier, W. J. The Use of Global and Local Molecular Parameters for the Analysis of the Gas-Phase Basicity of Amines. *J. Am. Chem. Soc.* **1986**, *108* (19), 5708–5711.
- (70) Wu, D.; Jia, D.; Liu, A.; Liu, L.; Guo, J. Theoretical Study on the Reactivity of Lewis Pairs PR 3/B(C 6F 5) 3 (R = Me, Ph, TBu, C 6F 5). *Chem. Phys. Lett.* **2012**, *541*, 1–6.
- (71) Christopher Jeyakumar, T.; Deepa, M.; Baskaran, S.; Sivasankar, C. Construction of Frustrated Lewis Pair from Nitride and Phosphine for the Activation and Cleavage of Molecular Hydrogen. *Appl. Organom. Chemis* **2020**, *34* (10), No. e5811.
- (72) Méndez, M.; Cedillo, A. Chemical Reactivity of the Frustrated Lewis Pairs in Borophosphines: A Theoretical Analysis of Their Lewis Acidity, Lewis Basicity and Fukui Function. *J. Mol. Model.* **2018**, *24* (9), 238.
- (73) Wu, D.; Liu, A.; Jia, D. Density Functional Reactivity Theory Characterizing the Reactivity of Frustrated Lewis Pairs. *Comput. Theor. Chem.* **2018**, *1131*, 33–39.
- (74) Morell, C.; Grand, A.; Toro-Labbé, A. Theoretical Support for Using the Δf(r) Descriptor. *Chem. Phys. Lett.* **2006**, *425* (4–6), 342–346.
- (75) Morell, C.; Grand, A.; Toro-Labbé, A. New Dual Descriptor for Chemical Reactivity. *J. Phys. Chem. A* **2005**, *109* (1), 205–212.
- (76) Zeiss, A.; Hartman, J.; Kelemen, S.; Ye, J. Theoretical Investigation of the Steric Effects on Alkyne Semihydrogenation Catalyzed by Frustrated Lewis Pairs. *J. Phys. Chem. C* **2024**, *128*, 128.
- (77) Shilpa, S.; Kashyap, G.; Sunoj, R. B. Recent Applications of Machine Learning in Molecular Property and Chemical Reaction Outcome Predictions. *J. Phys. Chem. A* **2023**, *127* (40), 8253–8271.
- (78) Yang, J.; Cai, Y.; Zhao, K.; Xie, H.; Chen, X. Concepts and Applications of Chemical Fingerprint for Hit and Lead Screening. *Drug Discov. Today* **2022**, *27* (11), 103356.
- (79) Raghunathan, S.; Priyakumar, U. D. Molecular Representations for Machine Learning Applications in Chemistry. *Int. J. Quantum Chem.* **2022**, *122* (7), No. e26870.
- (80) Gallegos, L. C.; Luchini, G.; St John, P. C.; Kim, S.; Paton, R. S. Importance of Engineered and Learned Molecular Representations in Predicting Organic Reactivity, Selectivity, and Chemical Properties. *Acc. Chem. Res.* **2021**, *54* (4), 827–836.
- (81) Willett, P. Similarity Searching Using 2D Structural Fingerprints. *Methods Mol. Biol.* **2010**, *672*, 133–158.
- (82) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50* (5), 742–754.

(83) Landrum, G. RDKit: Open-source cheminformatics. 2016, RDKit Project. <https://www.rdkit.org>.



CAS BIOFINDER DISCOVERY PLATFORM™

BRIDGE BIOLOGY AND CHEMISTRY FOR FASTER ANSWERS

Analyze target relationships,
compound effects, and disease
pathways

Explore the platform



A division of the
American Chemical Society