THE JOURNAL OF PHYSICAL CHEMISTRY LETTERS
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JPCL Letter

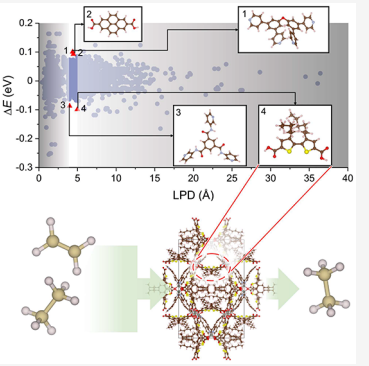
Data-Driven Discovery of Gas-Selective Organic Linkers in Metal−Organic Frameworks for the Separation of Ethylene and Ethane

Mingzheng Zhang,[▽] Qiming Xie,[▽] Zhuozheng Wang, Wentao Zhang, Yawen Bo, Zhiying Zhang, Hao Li, Yi Luo, Qihan Gong,* Shunning Li,* and Feng Pan*

Cite This: *J. Phys. Chem. Lett.* 2024, 15, 4815−4822 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

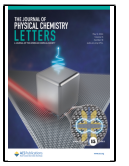
ABSTRACT: Metal−organic frameworks (MOFs) are potential candidates for gas-selective adsorbents for the separation of an ethylene/ethane mixture. To accelerate material discovery, high-throughput computational screening is a viable solution. However, classical force fields, which were widely employed in recent studies of MOF adsorbents, have been criticized for their failure to cover complicated interactions such as those involving π electrons. Herein, we demonstrate that machine learning force fields (MLFFs) trained on quantum-chemical reference data can overcome this difficulty. We have constructed a MLFF to accurately predict the adsorption energies of ethylene and ethane on the organic linkers of MOFs and discovered that the π electrons from both the ethylene molecule and the aromatic rings in the linkers could substantially influence the selectivity for gas adsorption. Four kinds of MOF linkers are identified as having promise for the separation of ethylene and ethane, and our results could also offer a new perspective on the design of MOF building blocks for diverse applications.



Ethylene (C_2H_4) is the most important olefin raw material for petrochemical production.[1] Current industrial technologies for producing C_2H_4 are based on either steam cracking of naphtha or thermal decomposition of ethane (C_2H_6), which necessitates a purification process to remove C_2H_6. Because of the similar physical properties of both hydrocarbon molecules, such purification typically involves a costly cryogenic distillation at >23 bar and below −25 °C.[2] In the past decade, metal−organic frameworks (MOFs) with their large surface areas, customizable pore geometry, and versatile functional components have opened a new avenue in the field of separation of C_2H_4 and C_2H_6.[3−11] Through a careful combination of metal-ion nodes and organic ligands, MOFs could serve as solid adsorbents to capture one of the gases while releasing the other under ambient conditions, thus rendering the separation process much less energy-intensive than the conventional distillation method.[12] However, most of the MOFs suitable for gas separation were historically discovered via a laborious trial-and-error process and could be considered as serendipitous results. This has led researchers to seek efficient *in silico* approaches to guide the design of MOFs, which is generating more and more research interest given the recent advancements in high-throughput computational techniques.[13,14]

Pioneering studies have employed grand canonical Monte Carlo (GCMC) simulations with standard force fields to predict the adsorption isotherms of gas molecules in MOFs.[15−17] While these predictions were demonstrated to be in good agreement with the experimental findings, especially for gases such as CH_4, H_2, and CO_2, whether the force fields can correctly describe the interaction of π electrons from aromatic rings in MOFs and from the C═C bond in the C_2H_4 molecule is still debatable.[18,19] Hence, quantum-chemical calculations are necessary to ensure accuracy, but their substantial computational cost has prevented us from performing a high-throughput screening over all existing MOFs (there are ∼100 000 MOF materials in the Cambridge Structural Database).[20,21] In this context, the recently developed data-driven machine learning (ML) models have captivated researchers due to the promise of their high prediction accuracy and low level of computational requirement.[22−28] Some successes have been achieved for the application of ML force fields (MLFFs) in different ranges of materials, covering organic and inorganic categories.[29−35] MLFFs trained on reference data from quantum-chemical calculations could potentially allow for the better description of intra- and intermolecular interactions than the classical force fields and may therefore overcome the deficiency of previous GCMC studies in predicting the guest−host interactions involving π electrons.[36]

Received: March 22, 2024
Revised: April 19, 2024
Accepted: April 23, 2024
Published: April 26, 2024



https://doi.org/10.1021/acs.jpclett.4c00860
J. Phys. Chem. Lett. 2024, 15, 4815−4822

Moreover, most of the efforts in computational screening of MOFs for gas separation are not directly aimed at the optimization of individual building blocks in MOFs.[37] In previous studies, the entire framework structures of MOFs were adopted for simulation, which leads to a great difficulty in deconvoluting the respective roles of inorganic nodes, organic linkers, and framework topologies (Figure 1). So far, there is

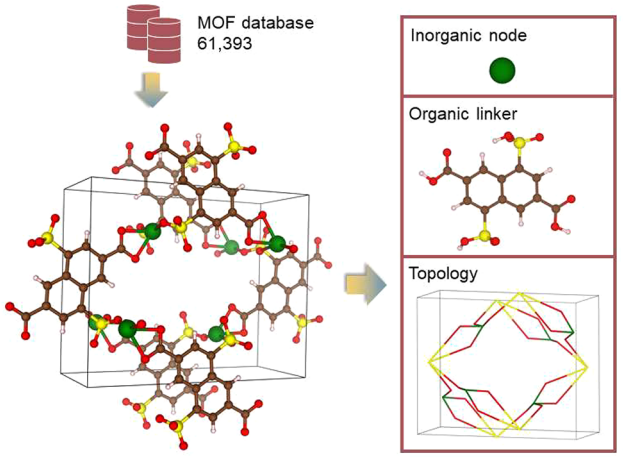


Figure 1. Extraction of building blocks (inorganic nodes and organic linkers) and topology of the MOF structures.

no consensus about what kind of metal-ion coordination configuration would be favorable for selective separation of a C_2H_4/C_2H_6 mixture, as well as what kind of functional group on the organic linker and what kind of pore topology are optimal. Suitable isolation of these components is a prerequisite for the rational design of MOFs, and we here chose to isolate and focus on the organic linkers, for which the molecular interactions with C_2H_4 and C_2H_6 can be effectively reproduced by MLFFs in the polarizable atom interaction neural network (PaiNN) architecture.[38,39]

In this work, we have extracted 15 361 organic linkers from a total of 61 393 MOF structures stored in the Cambridge Structural Database.[20] Quantum-chemical calculations were conducted on a small subset of the organic linkers, from which the trajectories during structural optimization of the adsorption of C_2H_4 and C_2H_6 on the linkers were obtained and used as the training data set. The gas adsorption properties calculated from the obtained MLFF have demonstrated promising accuracy with respect to the quantum-chemical results. From the computational screening of the linkers, we have discovered that C_2H_4 is preferentially captured by the polar functional groups (e.g., with amine or carboxyl moieties) due to the charge transfer of the π electrons on the C═C bond, while C_2H_6 tends to be attracted by aromatic rings that exhibit relatively strong dispersion interaction with the H atoms of the gas molecule. This work constitutes the first theoretical study that unravels the individual role of organic linkers in facilitating the separation of C_2H_4 and C_2H_6 and will pave the way for future development of MOF adsorbents in the petrochemical industry.

The data set of this work is based on the Cambridge Structural Database, and by discarding error entries as well as disordered MOF materials, we retain a total of 61 393 unique structures. For each structure, neighboring atoms were searched using the covalent radii for all of the main group elements, from which we could easily identify the organic building blocks. To achieve charge balance, a suitable number of hydrogen atoms were added to the functional groups that serve as the coordinating ligands to the inorganic nodes of the MOF. A single C_2H_4 or C_2H_6 molecule was placed near this organic linker to evaluate the adsorption properties. To identify a better initial configuration for structural optimization, we have relied on the van der Waals (VDW) potential using a Xe atom as the probe (Figure 2a), which can roughly estimate the weak intermolecular interactions between the linker and the gas.[40,41] Three minimum-energy configurations were selected, with a criterion that the corresponding locations of the probe are separated by ≥4 Å between the different configurations. We also avoided the cases in which the probe is too close to the coordinating ligands of the linker (the threshold is 4 Å, as well). The center of the gas molecule was then situated at the site of the probe, and we rotated the molecule in different directions, from which the orientation that corresponds to the minimum VDW potential energy was selected.

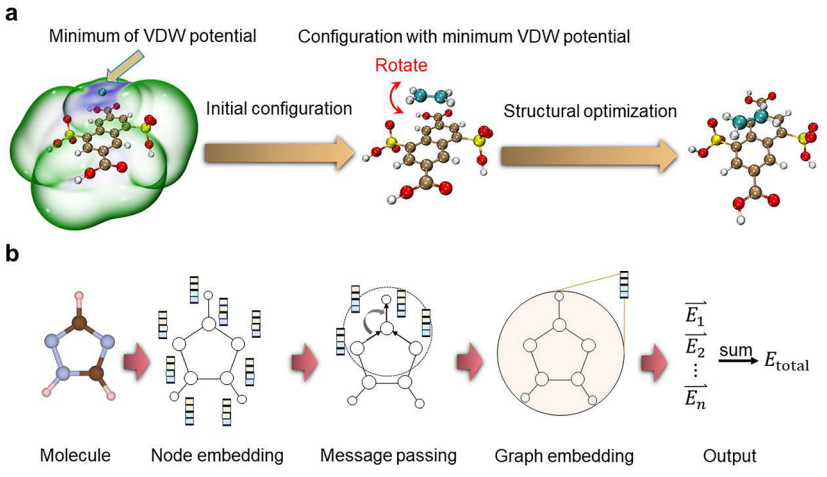


Figure 2. Construction of the initial configuration for simulation and scheme of feature extraction for MLFF. (a) Identification of minimum-energy configurations according to the VDW potential. These configurations are used as the initial setting in the quantum-chemical calculations. (b) PaiNN architecture for the construction of a MLFF. This model predicts molecular properties through equivariant message passing.

https://doi.org/10.1021/acs.jpclett.4c00860
J. Phys. Chem. Lett. 2024, 15, 4815−4822

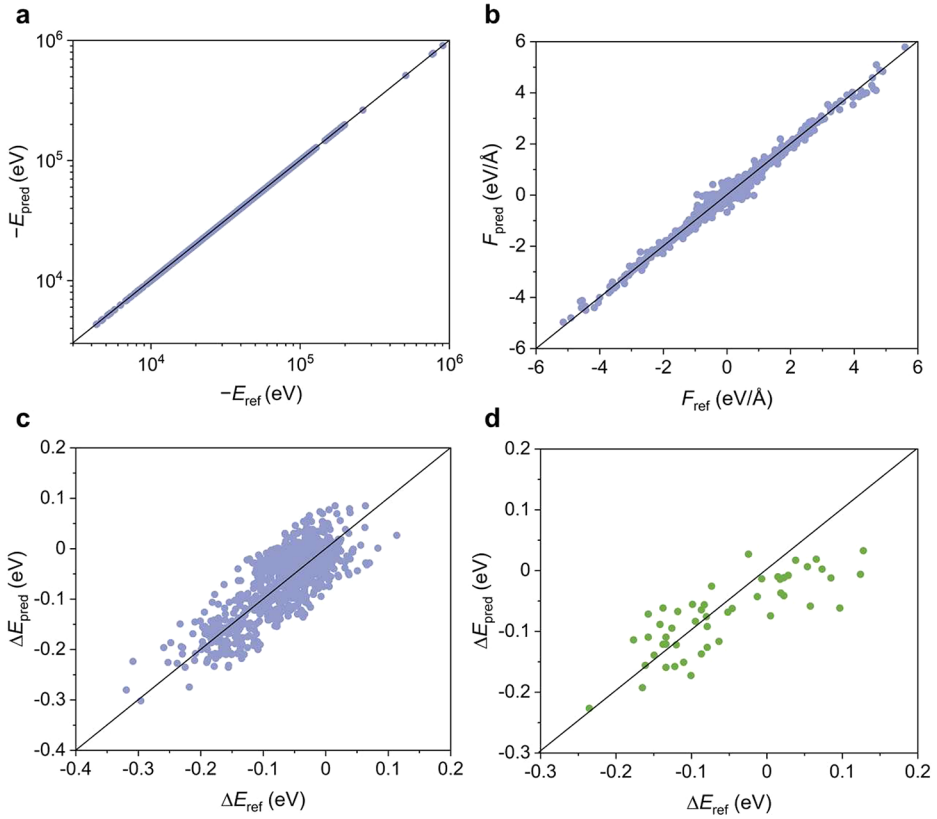


Figure 3. Performance of the MLFF in PaiNN architecture. (a) Energy of the adsorption configurations in the training set. (b) Force on atoms in the MOF structures and the gas molecules in the training set. Differences in the adsorption energies between C_2H_4 and C_2H_6 in the (c) training set and (d) test set. The quantum-chemical calculation results are used as the reference data.

The final configurations were employed for quantum-chemical calculations using Gaussian 16 at the B3LYP/6-31G* level of theory with dispersion correction.[42−46] The trajectories during structural optimization of the adsorption of C_2H_4 and C_2H_6 were then used as the training data set for our MLFF, which can help achieve a balance between sample size and feature space.[47] Figure 2b illustrates the PaiNN architecture, in which a molecule is represented by a graph and the elemental information on atoms is stored in its nodes. The information about bond lengths and angles will be extracted upon the repeated exchange of messages followed by updates of node features. The obtained MLFF was finally employed to calculate the adsorption energies of C_2H_4 and C_2H_6, with their energy difference taken as the metric for selective separation of the gas mixture.

Figure 3a shows a good linear relationship between the energies predicted by the MLFF and the reference data from Gaussian in the training set. The R^2 is 0.99, and the error in energy is 5 meV/atom. The forces are also predicted accurately, as shown in Figure 3b, with an R^2 of 0.97 and an error of 0.2 meV/Å. The mean absolute errors of energy and forces over all 62 epochs are provided in Figure S1. The performance in predicting the difference in adsorption energies between C_2H_4 and C_2H_6 (Figure 3c) is noticeably worse than directly predicting the energy of the adsorption configuration because the energy difference involves the subtraction of energy values, which will significantly amplify the relative error. This is also observed on the test set, as displayed in Figure 3d. The calculated prediction error implies that the MLFF in this work is still unable to reach the accuracy of quantum-chemical calculations for intermolecular interactions, which is probably due to the intrinsically short cutoff radius in the model for searching the interacting atoms.[33,48−50] Nevertheless, the MLFF could still serve as a promising screening tool for the identification of organic linkers that exhibit superior ability for gas-selective adsorption. This enables a proper evaluation of the influence of specific functional groups on the separation of the C_2H_4/C_2H_6 mixture.

On the basis of the MLFF-predicted adsorption energies of C_2H_4 and C_2H_6, a gradient boosting regression model[S1] combined with RDKit descriptors was constructed (Figure S2), followed by SHapley Additive exPlanations (SHAP) to unravel the potential role of functional groups for the separation of C_2H_4 and C_2H_6.[S2,S3] A total of 80 RDKit features from fragment descriptors and Lipinski parameters for machine learning have been taken into consideration, and the most important ones are displayed in Figure 4 for the prediction of adsorption energies of C_2H_4 and C_2H_6, respectively. The color of the data points in the right panels indicates the adsorption energy (blue for more negative values, corresponding to stronger adsorption), which shows some signs of scaling correlation with the SHAP value for all of the displayed features. The normalized mean value of SHAP shown in the left panels reflects the contribution of the corresponding feature to the output of the adsorption energy. Notably, the numbers of aromatic rings (n_{AR}), NH/OH species (n_{NOH}), and sulfonamide groups (n_S) are the top three decisive factors for both C_2H_4 and C_2H_6. The contribution of aromatic rings is potentially related to the interaction between their π electrons and the hydrogen atoms in the gas molecule. To

https://doi.org/10.1021/acs.jpclett.4c00860
J. Phys. Chem. Lett. 2024, 15, 4815−4822

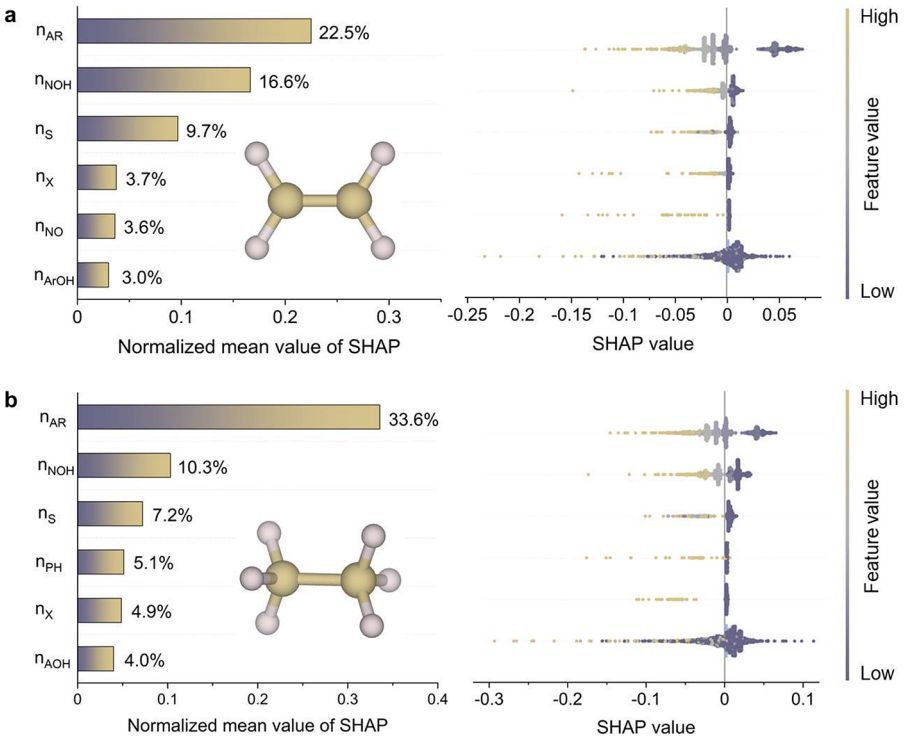


Figure 4. Correlation analysis for gas adsorption energy via SHAP values. Adsorption energies of (a) C_2H_4 and (b) C_2H_6 interpreted by the gradient boosting regression model with SHAP values. RDKit features are sorted in descending order according to global parameter importance (left panels), and the beeswarm plots are displayed for each gas molecule (right panels). The horizontal axis indicates the influence on the model output. n_{AR}, n_{NOH}, n_S, n_X, n_{NO}, n_{ArOH}, n_{PH}, and n_{AOH} correspond to the numbers of aromatic rings, NH/OH species, sulfonamide groups, halogen atoms, hydroxylamine groups, aromatic hydroxyl groups, *para* hydroxylation sites, and aliphatic hydroxyl groups, respectively.

justify this, we have added another feature, the shortest distance from the H atom in the gas molecule to the center of the aromatic ring of the ligand, which is designated as d_{AR}. For C_2H_6, the normalized mean SHAP value for d_{AR} is 21.1%, considerably higher than the second most important feature, n_{NOH} (10.3%), which demonstrates that C_2H_6 adsorption can greatly benefit from the π electrons of aromatic rings in the organic linker. On the contrary, for C_2H_4, the contribution of n_{AR} is much less dominating, and other features generally involve polar functional groups that point toward the C═C bond of C_2H_4. We therefore assume that the π electrons of the C═C bond in C_2H_4 could play a role and have added the following feature for evaluation, the distance from the middle point of the C═C bond in C_2H_4 to the nearest H atom on the functional group in the organic linker, designated as d_H. For C_2H_4, the normalized mean SHAP value for d_H reaches 41.0%, indicating the indispensable role of the π electrons of the C═C bond in analyzing C_2H_4 adsorption. This interaction could be ascribed to a slight electron transfer from C_2H_4 to the ligand, as will be discussed below.

According to the calculated difference in adsorption energy between C_2H_4 and C_2H_6, along with the pore size of the corresponding MOF, we can screen for potential MOF adsorbents that are promising in the separation of C_2H_4 and C_2H_6. We note that the kinetic diameters of C_2H_4 and C_2H_6 are 4.163 and 4.443 Å, respectively.[54] A limiting pore diameter (LPD) closer to both values would likely offer a more pronounced molecular sieving effect. As shown in Figure 5a, we have pinpointed four candidate MOFs with the largest difference in adsorption energy and with a proper LPD in the unshaded area. The structures of their corresponding linkers are provided in the insets, and the framework structures[55−58] are displayed in Figure 5b. We believe that these materials are worth the experimental effort to explore in future studies. Moreover, because the contribution of inorganic nodes has not been evaluated in this work, it is possible that some MOFs could achieve high selectivity from open metal sites while their organic linkers show inferior ability for gas-selective adsorption. In this work, we do not intend to screen for the optimal MOF adsorbent with all factors treated simultaneously but instead aim to identify what kind of organic linker could facilitate the separation of C_2H_4 and C_2H_6, directly from a pool of 15 361 species found as the building blocks of existing MOFs. This could offer a strategy for the rational design of the functional components of MOF adsorbents without the interference of other factors.

To further understand the effect of π electrons from a quantum-chemical perspective, we have selected six linkers from the training data set, all of which exhibit relatively large adsorption energy differences between C_2H_4 and C_2H_6. Two of them are displayed in Figure 5c, while the others are shown in Figures S3 and S4. The independent gradient model based on Hirshfeld partition of molecular density (IGMH) analysis is conducted to visualize the regions of intermolecular interactions.[59−61] C_2H_4 preferentially adsorbs on linker 1, while the interaction between C_2H_6 and linker 1 is relatively weak. In the configuration of C_2H_4 on linker 1, the energy decomposition suggests a ≲38.3% contribution from the electrostatic energy and a 17.0% contribution from the orbital energy (Table S1). We note that the sum of both values surpasses the contribution from the dispersion correction energy (44.8%), in agreement with our assumption that the

https://doi.org/10.1021/acs.jpclett.4c00860
J. Phys. Chem. Lett. 2024, 15, 4815−4822

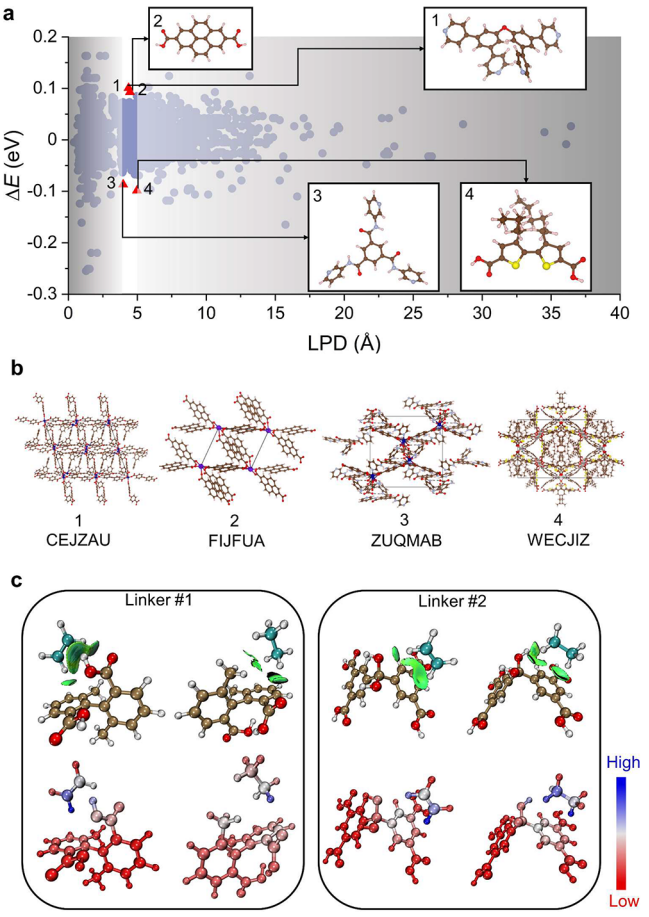


Figure 5. Computational screening of MOF materials for the separation of C_2H_4 and C_2H_6. (a) MLFF-predicted difference in adsorption energy (ΔE) between C_2H_4 and C_2H_6 vs the limiting pore diameter (LPD) of the MOF structure. The insets show the corresponding organic linkers. (b) Most promising MOF adsorbents identified from the screening of both ΔE and LPD. (c) Configurations of the adsorption of C_2H_4 and C_2H_6 on two representative organic linkers. The green isosurfaces in the top panels indicate the regions of intermolecular interactions based on IGMH analysis. The color of atoms in the bottom panels indicates the contribution of individual atoms to the adsorption of gas molecules.

electron transfer from C_2H_4 to the ligand is nonnegligible. The corresponding regions of interactions are localized at the side of the C_2H_4 molecule near the C═C bond, corresponding to the π electron region of C_2H_4. When C_2H_6 is adsorbed on linker 1, the dispersion interaction will contribute to 71.0% of the adsorption energy (Table S2), demonstrating that the electron transfer is quite inferior in C_2H_6 adsorption as compared to that in C_2H_4 adsorption. In comparison with linker 1, the gas selectivity is reversed for linker 2; that is, C_2H_6 adsorption is more energetically favorable than C_2H_4 adsorption. We find that in the adsorption configurations of both molecules on linker 2, the dispersion correction energy always contributes ~64%. On the basis of the correlation analysis using SHAP values presented above, this interaction is mainly ascribed to the aromatic rings in the linker, which can attract the H atoms in the gas molecule. This is consistent with the IGMH analysis in Figure 5c showing localized interaction regions between the aromatic ring of linker 2 and the H atoms in the molecules.

In conclusion, we have used a MLFF trained on quantum-chemical calculation results to investigate the molecular interactions of MOF linkers with C_2H_4 and C_2H_6. The overall good agreement with the reference data suggests the capability of MLFF in high-throughput screening for MOF adsorbents with the aim of separating the C_2H_4/C_2H_6 mixture. Our results indicate that C_2H_4 is preferentially captured by the polar functional groups due to charge transfer of the π electrons, while C_2H_6 could instead be captured by aromatic rings in the organic linker. The insights in this work will not only facilitate the development of MLFFs for describing intermolecular interactions but also may stimulate new avenues for future research of MOF adsorbents in the field of petrochemical production.

■ ASSOCIATED CONTENT

Data Availability Statement

The trained MLFF model used to replicate all of the results is available at https://github.com/PKUsam2023/MOFlinker_MLFF.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpclett.4c00860.

Computational methods, mean absolute errors of energy and forces, representative organic linkers, and decomposition of the interaction energy (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Qihan Gong − *Fundamental Science & Advanced Technology Lab, PetroChina Petrochemical Research Institute, China National Petroleum Corporation, Beijing 102200, China*; Email: gongqihan@petrochina.com.cn

Shunning Li − *School of Advanced Materials, Shenzhen Graduate School, Peking University, Shenzhen 518055, China*; orcid.org/0000-0002-5381-6025; Email: lisn@pku.edu.cn

Feng Pan − *School of Advanced Materials, Shenzhen Graduate School, Peking University, Shenzhen 518055, China*; orcid.org/0000-0002-8216-1339; Email: panfeng@pkusz.edu.cn

Authors

Mingzheng Zhang − *School of Advanced Materials, Shenzhen Graduate School, Peking University, Shenzhen 518055, China*

Qiming Xie − *School of Advanced Materials, Shenzhen Graduate School, Peking University, Shenzhen 518055, China*

Zhuozheng Wang − *Fundamental Science & Advanced Technology Lab, PetroChina Petrochemical Research Institute, China National Petroleum Corporation, Beijing 102200, China*

Wentao Zhang − *School of Advanced Materials, Shenzhen Graduate School, Peking University, Shenzhen 518055, China*

Yawen Bo − *Fundamental Science & Advanced Technology Lab, PetroChina Petrochemical Research Institute, China National Petroleum Corporation, Beijing 102200, China*

Zhiying Zhang − *School of Advanced Materials, Shenzhen Graduate School, Peking University, Shenzhen 518055, China*

https://doi.org/10.1021/acs.jpclett.4c00860
J. Phys. Chem. Lett. 2024, 15, 4815−4822

Hao Li − *Fundamental Science & Advanced Technology Lab, PetroChina Petrochemical Research Institute, China National Petroleum Corporation, Beijing 102200, China*

Yi Luo − *Fundamental Science & Advanced Technology Lab, PetroChina Petrochemical Research Institute, China National Petroleum Corporation, Beijing 102200, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpclett.4c00860

Author Contributions

[▽]M.Z. and Q.X. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge financial support from CNPC Science and Technology Management Department Funding (2022DJ6113), the Basic and Applied Basic Research Foundation of Guangdong Province (2023A1515011391), the Soft Science Research Project of Guangdong Province (2017B030301013), the National Natural Science Foundation of China (22109003), and the Major Science and Technology Infrastructure Project of Material Genome Big-science Facilities Platform supported by Municipal Development and Reform Commission of Shenzhen.

■ REFERENCES

(1) Spallina, V.; Velarde, I. C.; Jimenez, J. A. M.; Godini, H. R.; Gallucci, F.; Van Sint Annaland, M. Techno-Economic Assessment of Different Routes for Olefins Production through the Oxidative Coupling of Methane (Ocm): Advances in Benchmark Technologies. *Energ. Convers. Manage.* **2017**, *154*, 244−261.

(2) Sholl, D. S.; Lively, R. P. Seven Chemical Separations to Change the World. *Nature* **2016**, *532* (7600), 435−437.

(3) Ye, Y.; Xie, Y.; Shi, Y.; Gong, L.; Phipps, J.; Al-Enizi, A. M.; Nafady, A.; Chen, B.; Ma, S. A Microporous Metal-Organic Framework with Unique Aromatic Pore Surfaces for High Performance C2h6/C2h4 Separation. *Angew. Chem., Int. Ed.* **2023**, *62* (21), No. e202302564.

(4) Yang, H.; Wang, Y.; Krishna, R.; Jia, X.; Wang, Y.; Hong, A. N.; Dang, C.; Castillo, H. E.; Bu, X.; Feng, P. Pore-Space-Partition-Enabled Exceptional Ethane Uptake and Ethane-Selective Ethane−Ethylene Separation. *J. Am. Chem. Soc.* **2020**, *142* (5), 2222−2227.

(5) Xu, Z.; Xiong, X.; Xiong, J.; Krishna, R.; Li, L.; Fan, Y.; Luo, F.; Chen, B. A Robust Th-Azole Framework for Highly Efficient Purification of C2h4 from a C2h4/C2h2/C2h6Mixture. *Nat. Commun.* **2020**, *11* (1), 3163.

(6) Lin, R.-B.; Wu, H.; Li, L.; Tang, X.-L.; Li, Z.; Gao, J.; Cui, H.; Zhou, W.; Chen, B. Boosting Ethane/Ethylene Separation within Isoreticular Ultramicroporous Metal−Organic Frameworks. *J. Am. Chem. Soc.* **2018**, *140* (40), 12940−12946.

(7) Wang, Y.; Hao, C.; Fan, W.; Fu, M.; Wang, X.; Wang, Z.; Zhu, L.; Li, Y.; Lu, X.; Dai, F.; Kang, Z.; Wang, R.; Guo, W.; Hu, S.; Sun, D. One-Step Ethylene Purification from an Acetylene/Ethylene/Ethane Ternary Mixture by Cyclopentadiene Cobalt-Functionalized Metal−Organic Frameworks. *Angew. Chem., Int. Ed.* **2021**, *60* (20), 11350−11358.

(8) Bereciartua, P. J.; Cantín, Á.; Corma, A.; Jordá, J.; Palomino, M.; Rey, F.; Valencia, S.; Corcoran, E. W.; Kortunov, P.; Ravikovitch, P. I. J. S.; et al. Control of Zeolite Framework Flexibility and Pore Topology for Separation of Ethane and Ethylene. *Science* **2017**, *358* (6366), 1068.

(9) Li, L.; Lin, R.-B.; Krishna, R.; Li, H.; Xiang, S.; Wu, H.; Li, J.; Zhou, W.; Chen, B. Ethane/Ethylene Separation in a Metal-Organic Framework with Iron-Peroxo Sites. *Science* **2018**, *362* (6413), 443−446.

(10) Yang, Y.; Li, L.; Lin, R.-B.; Ye, Y.; Yao, Z.; Yang, L.; Xiang, F.; Chen, S.; Zhang, Z.; Xiang, S.; Chen, B. Ethylene/Ethane Separation in a Stable Hydrogen-Bonded Organic Framework through a Gating Mechanism. *Nat. Chem.* **2021**, *13*, 933.

(11) Ma, X.; Kumar, P.; Mittal, N.; Khlyustova, A.; Daoutidis, P.; Mkhoyan, K. A.; Tsapatsis, M. Zeolitic Imidazolate Framework Membranes Made by Ligand-Induced Permselectivation. *Science* **2018**, *361*, 1008.

(12) Wang, Y.; Peh, S. B.; Zhao, D. Alternatives to Cryogenic Distillation: Advanced Porous Materials in Adsorptive Light Olefin/Paraffin Separations. *Small* **2019**, *15* (25), No. 1900058.

(13) Daglar, H.; Gulbalkan, H. C.; Avci, G.; Aksu, G. O.; Altundal, O. F.; Altintas, C.; Erucar, I.; Keskin, S. Effect of Metal−Organic Framework (Mof) Database Selection on the Assessment of Gas Storage and Separation Potentials of Mofs. *Angew. Chem., Int. Ed.* **2021**, *60* (14), 7828−7837.

(14) Colón, Y. J.; Snurr, R. Q. High-Throughput Computational Screening of Metal−Organic Frameworks. *Chem. Soc. Rev.* **2014**, *43* (16), 5735−5749.

(15) Bae, Y.-S.; Mulfort, K. L.; Frost, H.; Ryan, P.; Punnathanam, S.; Broadbelt, L. J.; Hupp, J. T.; Snurr, R. Q. Separation of Co2 from Ch4 Using Mixed-Ligand Metal−Organic Frameworks. *Langmuir* **2008**, *24* (16), 8592−8598.

(16) Al-Janabi, N.; Fan, X.; Siperstein, F. R. Assessment of Mof's Quality: Quantifying Defect Content in Crystalline Porous Materials. *J. Phys. Chem. Lett.* **2016**, *7* (8), 1490−1494.

(17) Keskin, S.; Liu, J.; Rankin, R. B.; Johnson, J. K.; Sholl, D. S. Progress, Opportunities, and Challenges for Applying Atomically Detailed Modeling to Molecular Adsorption and Transport in Metal−Organic Framework Materials. *Ind. Eng. Chem. Res.* **2009**, *48* (5), 2355−2371.

(18) Bobbitt, N. S.; Shi, K.; Bucior, B. J.; Chen, H.; Tracy-Amoroso, N.; Li, Z.; Sun, Y.; Merlin, J. H.; Siepmann, J. I.; Siderius, D. W.; Snurr, R. Q. Mofx-Db: An Online Database of Computational Adsorption Data for Nanoporous Materials. *Journal of Chemical & Engineering Data* **2023**, *68* (2), 483−498.

(19) Kanal, I. Y.; Keith, J. A.; Hutchison, G. R. A Sobering Assessment of Small-Molecule Force Field Methods for Low Energy Conformer Predictions. *Int. J. Quantum Chem.* **2018**, *118* (5), No. e25512.

(20) Moghadam, P. Z.; Li, A.; Wiggin, S. B.; Tao, A.; Maloney, A. G. P.; Wood, P. A.; Ward, S. C.; Fairen-Jimenez, D. Development of a Cambridge Structural Database Subset: A Collection of Metal−Organic Frameworks for Past, Present, and Future. *Chem. Mater.* **2017**, *29* (7), 2618−2625.

(21) Chung, Y. G.; Haldoupis, E.; Bucior, B. J.; Haranczyk, M.; Lee, S.; Zhang, H.; Vogiatzis, K. D.; Milisavljevic, M.; Ling, S.; Camp, J. S.; Slater, B.; Siepmann, J. I.; Sholl, D. S.; Snurr, R. Q. Advances, Updates, and Analytics for the Computation-Ready, Experimental Metal−Organic Framework Database: Core Mof 2019. *Journal of Chemical & Engineering Data* **2019**, *64* (12), 5985−5998.

(22) Keith, J. A.; Vassilev-Galindo, V.; Cheng, B.; Chmiela, S.; Gastegger, M.; Müller, K.-R.; Tkatchenko, A. Combining Machine Learning and Computational Chemistry for Predictive Insights into Chemical Systems. *Chem. Rev.* **2021**, *121* (16), 9816−9872.

(23) Sanchez-Lengeling, B.; Aspuru-Guzik, A. Inverse Molecular Design Using Machine Learning: Generative Models for Matter Engineering. *Science* **2018**, *361* (6400), 360−365.

(24) Boyd, P. G.; Chidambaram, A.; García-Díez, E.; Ireland, C. P.; Daff, T. D.; Bounds, R.; Gładysiak, A.; Schouwink, P.; Moosavi, S. M.; Maroto-Valer, M. M.; Reimer, J. A.; Navarro, J. A. R.; Woo, T. K.; Garcia, S.; Stylianou, K. C.; Smit, B. Data-Driven Design of Metal−Organic Frameworks for Wet Flue Gas Co2 Capture. *Nature* **2019**, *576* (7786), 253−256.

(25) Rosen, A. S.; Iyer, S. M.; Ray, D.; Yao, Z.; Aspuru-Guzik, A.; Gagliardi, L.; Notestein, J. M.; Snurr, R. Q. Machine Learning the Quantum-Chemical Properties of Metal−Organic Frameworks for Accelerated Materials Discovery. *Matter* **2021**, *4* (5), 1578−1597.

(26) Boyd, P. G.; Lee, Y.; Smit, B. Computational Development of the Nanoporous Materials Genome. *Nat. Rev. Mater.* **2017**, *2* (8), No. 17037.
(27) Jablonka, K. M.; Ongari, D.; Moosavi, S. M.; Smit, B. Big-Data Science in Porous Materials: Materials Genomics and Machine Learning. *Chem. Rev.* **2020**, *120* (16), 8066−8129.
(28) Yao, Z.; Sánchez-Lengeling, B.; Bobbitt, N. S.; Bucior, B. J.; Kumar, S. G. H.; Collins, S. P.; Burns, T.; Woo, T. K.; Farha, O. K.; Snurr, R. Q.; Aspuru-Guzik, A. Inverse Design of Nanoporous Crystalline Reticular Materials with Deep Generative Models. *Nature Machine Intelligence* **2021**, *3* (1), 76−86.
(29) Hu, Q.; Weng, M.; Chen, X.; Li, S.; Pan, F.; Wang, L.-W. Neural Network Force Fields for Metal Growth Based on Energy Decompositions. *J. Phys. Chem. Lett.* **2020**, *11* (4), 1364−1369.
(30) Friederich, P.; Häse, F.; Proppe, J.; Aspuru-Guzik, A. Machine-Learned Potentials for Next-Generation Matter Simulations. *Nat. Mater.* **2021**, *20* (6), 750−761.
(31) Galib, M.; Limmer, D. T. Reactive Uptake of N2O5 by Atmospheric Aerosol Is Dominated by Interfacial Processes. *Science* **2021**, *371* (6532), 921−925.
(32) Ackland, G. J.; Dunuwille, M.; Martinez-Canales, M.; Loa, I.; Zhang, R.; Sinogeikin, S.; Cai, W.; Deemyad, S. Quantum and Isotope Effects in Lithium Metal. *Science* **2017**, *356* (6344), 1254−1259.
(33) Ko, T. W.; Finkler, J. A.; Goedecker, S.; Behler, J. A Fourth-Generation High-Dimensional Neural Network Potential with Accurate Electrostatics Including Non-Local Charge Transfer. *Nat. Commun.* **2021**, *12* (1), 398.
(34) Smith, J. S.; Nebgen, B. T.; Zubatyuk, R.; Lubbers, N.; Devereux, C.; Barros, K.; Tretiak, S.; Isayev, O.; Roitberg, A. E. Approaching Coupled Cluster Accuracy with a General-Purpose Neural Network Potential through Transfer Learning. *Nat. Commun.* **2019**, *10* (1), 2903.
(35) Takamoto, S.; Shinagawa, C.; Motoki, D.; Nakago, K.; Li, W.; Kurata, I.; Watanabe, T.; Yayama, Y.; Iriguchi, H.; Asano, Y.; Onodera, T.; Ishii, T.; Kudo, T.; Ono, H.; Sawada, R.; Ishitani, R.; Ong, M.; Yamaguchi, T.; Kataoka, T.; Hayashi, A.; Charoenphakdee, N.; Ibuka, T. Towards Universal Neural Network Potential for Material Discovery Applicable to Arbitrary Combination of 45 Elements. *Nat. Commun.* **2022**, *13* (1), 2991.
(36) Unke, O. T.; Chmiela, S.; Sauceda, H. E.; Gastegger, M.; Poltavsky, I.; Schütt, K. T.; Tkatchenko, A.; Müller, K.-R. Machine Learning Force Fields. *Chem. Rev.* **2021**, *121* (16), 10142−10186.
(37) Demir, H.; Daglar, H.; Gulbalkan, H. C.; Aksu, G. O.; Keskin, S. Recent Advances in Computational Modeling of Mofs: From Molecular Simulations to Machine Learning. *Coord. Chem. Rev.* **2023**, *484*, No. 215112.
(38) Zhang, W.; Weng, M.; Zhang, M.; Ye, Y.; Chen, Z.; Li, S.; Li, S.; Pan, F.; Wang, L.-w. Revealing Morphology Evolution of Lithium Dendrites by Large-Scale Simulation Based on Machine Learning Force Field. *Adv. Energy Mater.* **2023**, *13* (4), No. 2202892.
(39) Schütt, K.; Unke, O.; Gastegger, M. Equivariant Message Passing for the Prediction of Tensorial Properties and Molecular Spectra. *arXiv* **2021**, DOI: 10.48550/arXiv.2102.03150.
(40) Lu, T.; Chen, Q. Van Der Waals Potential: An Important Complement to Molecular Electrostatic Potential in Studying Intermolecular Interactions. *J. Mol. Model.* **2020**, *26* (11), 315.
(41) Lu, T.; Chen, F. Multiwfn: A Multifunctional Wavefunction Analyzer. *J. Comput. Chem.* **2012**, *33* (5), 580−592.
(42) Becke, A. D. Density-Functional Thermochemistry. Iii. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648−5652.
(43) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98* (45), 11623−11627.
(44) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self—Consistent Molecular Orbital Methods. Xii. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56* (5), 2257−2261.
(45) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (Dft-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), No. 154104.
(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; 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, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16*, rev. B.01; Gaussian, Inc.: Wallingford, CT, 2016.
(47) Liu, Y.; Yang, Z.; Zou, X.; Ma, S.; Liu, D.; Avdeev, M.; Shi, S. Data Quantity Governance for Machine Learning in Materials Science. *Natl. Sci. Rev.* **2023**, *10* (7), No. nwad125.
(48) Chmiela, S.; Vassilev-Galindo, V.; Unke, O. T.; Kabylda, A.; Sauceda, H. E.; Tkatchenko, A.; Müller, K.-R. Accurate Global Machine Learning Force Fields for Molecules with Hundreds of Atoms. *Sci. Adv.* **2023**, *9* (2), No. eadf0873.
(49) Unke, O. T.; Chmiela, S.; Gastegger, M.; Schütt, K. T.; Sauceda, H. E.; Müller, K.-R. Spookynet: Learning Force Fields with Electronic Degrees of Freedom and Nonlocal Effects. *Nat. Commun.* **2021**, *12* (1), 7273.
(50) Kabylda, A.; Vassilev-Galindo, V.; Chmiela, S.; Poltavsky, I.; Tkatchenko, A. Efficient Interatomic Descriptors for Accurate Machine Learning Force Fields of Extended Molecules. *Nat. Commun.* **2023**, *14* (1), 3562.
(51) 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, É. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(52) Tosco, P.; Stiefl, N.; Landrum, G. Bringing the MMFF Force Field to the RDKit: Implementation and Validation. *J. Cheminf.* **2014**, *6*, 37.
(53) Lundberg, S. M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. *arXiv* **2017**, DOI: 10.48550/arXiv.1705.07874.
(54) Li, J.-R.; Kuppler, R. J.; Zhou, H.-C. Selective Gas Adsorption and Separation in Metal−Organic Frameworks. *Chem. Soc. Rev.* **2009**, *38* (5), 1477−1504.
(55) Earl, L. D.; Patrick, B. O.; Wolf, M. O. Synthesis, Structure, and Luminescent Properties of Oligothiophene-Containing Metal−Organic Frameworks. *CrystEngComm* **2012**, *14* (18), 5801−5808.
(56) Luo, F.; Meng, P. P.; Feng, X. F.; Dang, L. L.; Zhang, X. J. A Complex Self-Catenated Coordination Framework with a Rare (3,12)-Connected Underlying Net Showing Selective Adsorption of CO2. *Eur. J. Inorg. Chem.* **2015**, *2015* (28), 4633−4637.
(57) Qin, L.; Hu, J.-S.; Zhang, M.-D.; Guo, Z.-J.; Zheng, H.-G. Structure−Property Relationship of Homochiral and Achiral Supramolecular Isomers Obtained by One-Pot Synthesis. *Chem. Commun.* **2012**, *48* (87), 10757−10759.
(58) Rosi, N. L.; Kim, J.; Eddaoudi, M.; Chen, B.; O'Keeffe, M.; Yaghi, O. M. Rod Packings and Metal−Organic Frameworks Constructed from Rod-Shaped Secondary Building Units. *J. Am. Chem. Soc.* **2005**, *127* (5), 1504−1518.
(59) Lu, T.; Chen, Q. Independent Gradient Model Based on Hirshfeld Partition: A New Method for Visual Study of Interactions in Chemical Systems. *J. Comput. Chem.* **2022**, *43* (8), 539−555.
(60) Lefebvre, C.; Rubez, G.; Khartabil, H.; Boisson, J.-C.; Contreras-García, J.; Hénon, E. Accurately Extracting the Signature of Intermolecular Interactions Present in the Nci Plot of the Reduced

Density Gradient Versus Electron Density. *Phys. Chem. Chem. Phys.* **2017**, *19* (27), 17928−17936.
(61) Lu, T.; Chen, Q. Simple, Efficient, and Universal Energy Decomposition Analysis Method Based on Dispersion-Corrected Density Functional Theory. *J. Phys. Chem. A* **2023**, *127* (33), 7023−7035.

https://doi.org/10.1021/acs.jpclett.4c00860
J. Phys. Chem. Lett. 2024, 15, 4815−4822