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

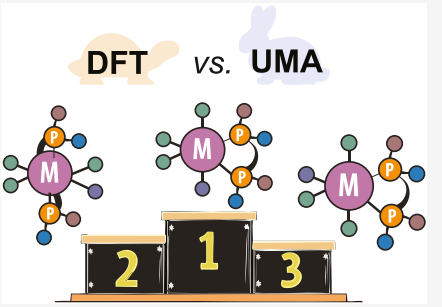
pubs.acs.org/JPCA
Article

Performance of Meta's Universal Model for Atoms across the Conformational and Configurational Space of Diverse Transition-Metal Catalysts

Adarsh V. Kalikadien and Evgeny A. Pidko*

Cite This: J. Phys. Chem. A 2026, 130, 1897−1904
Read Online

ABSTRACT: Machine Learning Interatomic Potentials (MLIPs) promise to transform computational catalysis by delivering near-density functional theory (DFT) accuracy at a fraction of the computational cost. Here, we evaluate the Universal Machine Learning Potential for Atoms (UMA) on two data sets of transition-metal complexes. UMA enables high-throughput evaluations in seconds per structure on consumer-grade GPUs. Analysis of per-ligand Spearman rank correlations ($\rho > 0.6$, $p < 0.05$) reveals variability in ranking reliability that is not captured by aggregate metrics such as R^2 or RMSE. However, these inaccuracies are shown to mainly occur in the near-DFT accuracy regime where these complexes are practically indistinguishable. For square-planar Ni complexes, reliable rankings are obtained for 84% of ligands in rigid Ni−Cl_2 complexes and drop to 53% for flexible asymmetric coordination environments, particularly only when conformers differ by <2 kJ/mol. Data set 2 shows a similar trend, with 61% and 44% reliability for Ru(II) and Mn(I) complexes, respectively, and, as expected, challenges for fluxional systems with small (<5 kJ/mol) relative energy gaps. These findings highlight the promise of MLIPs for both rigid, well-defined systems and highly flexible or fluxional catalysts, while underscoring the need to combine the speed of ML with validation and domain expertise to ensure robust and meaningful chemical insights.



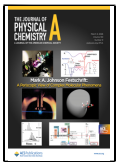
1. INTRODUCTION

Density functional theory (DFT) is indispensable for modeling catalytic reactions.[1−3] In homogeneous catalysis, energetic barriers and molecular properties are typically derived from optimized geometries and their electronic structure.[1,4] Although the molecular structures of these catalysts are relatively well-defined, their accurate calculations using even most modern quantum chemical methods come at a considerable computational cost.[5] In fact, exhaustive computational sampling of the conformational and configurational space of a metal−ligand complex with DFT calculations can incur costs similar to the operating costs of wet-lab experimentation.

To illustrate this, consider a simple thought experiment. In high-performance computing (HPC), Standard Billing Units (SBUs) are used to quantify computational usage and cost, defined as the product of the number of CPU cores and the number of wall-time hours. In our experience, a typical TM complex requires approximately 4 h of computation using 32 CPU cores to optimize structure and carry out frequency analysis with hybrid DFT functional and double-ζ quality basis set. Given an optimistic estimate cost of €0.01 per SBU in The Netherlands, optimizing 10,000 such complexes would amount to roughly €12,800. In high-throughput screening efforts, such scales are common.[6,7] It is, therefore, essential to develop reductionist approaches that maintain chemical accuracy while significantly lowering computational demand.[6,8]

Machine Learning Interatomic Potentials (MLIPs) offer a promising solution, enabling the approximation of DFT-level energies within seconds.[9−12] In the early seminal papers, MLIPs were limited to generating interatomic potentials for highly specific systems.[12,13] Unfortunately, to this day, a central challenge in developing MLIPs still lies in achieving sufficient generalization across the diverse domains and tasks for which DFT is employed.[14−16] Recently, Meta released large, chemically diverse data sets designed to support general-purpose models.[17,18] Alongside these data sets, a family of Universal Models for Atoms (UMA) was presented.[18] These general-purpose models have demonstrated competitive or superior performance in terms of accuracy, inference speed, and memory efficiency when benchmarked against specialized models across a wide range of molecules.[18] An exciting feature of UMA is that automated workflows were used to include an

Received: October 14, 2025
Revised: January 6, 2026
Accepted: January 7, 2026
Published: February 18, 2026



https://doi.org/10.1021/acs.jpca.5c07061
J. Phys. Chem. A 2026, 130, 1897−1904

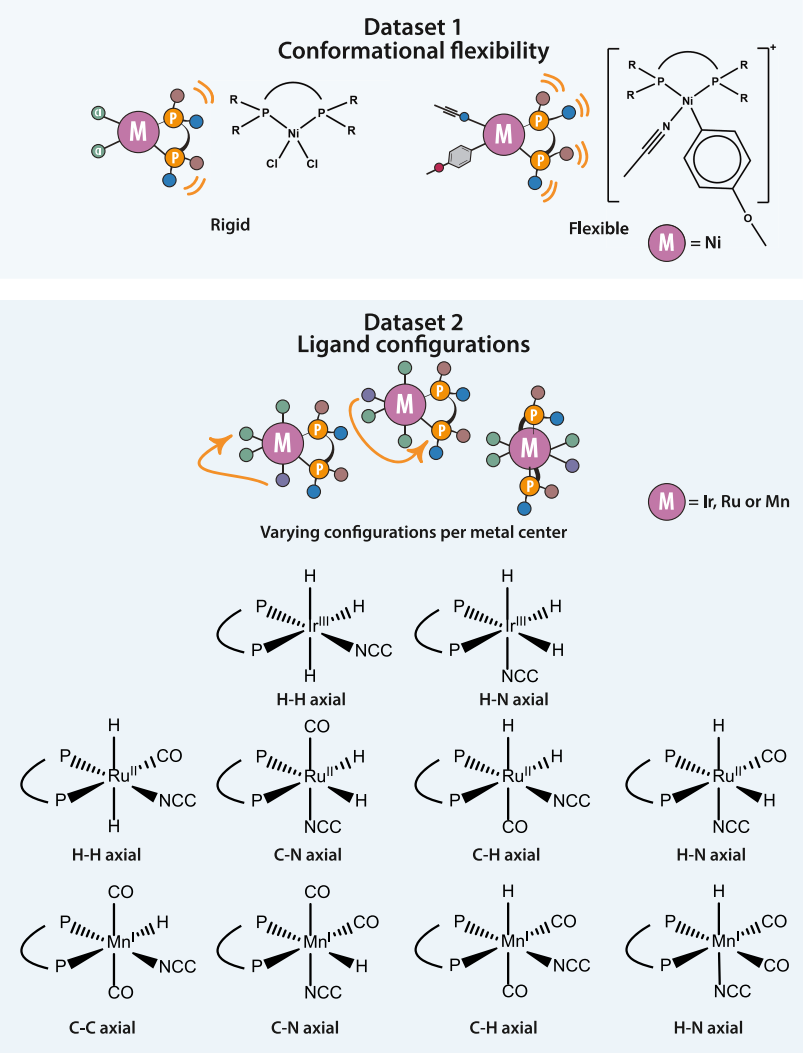


Figure 1. Overview of the two data sets used in this study. Top: Data set 1 contains conformers of square-planar Ni-based rigid and flexible model structures relevant to nitrile arylation, each with one of 25 bisphosphine ligands. Bottom: Data set 2 includes octahedral Ir, Ru, and Mn complexes relevant to hydrogenation chemistry, each with one of 88 bisphosphine ligands. Details on Data set 2 are available in our previous work.[22] Figure adapted from ref 22 by A.V Kalikadien, N.J. van der Lem, C. Valsecchi, L. Lefort and E.A. Pidko. Available under a CC-BY 3.0 license. Copyright RSC.

extremely large number of TM complexes in the Open Molecules 2025 (OMOL25) data set, allowing for extensive sampling of conformer space and configurational flexibility across different DFT data sets and tasks.[18] However, the ranking of conformers has only been evaluated for a subset of the data set (GEOM) aimed at drug-like molecules.[18,19]

An ideal long-term goal would be to employ MLIPs directly in geometry optimization, where both energies and forces are evaluated by the MLIP, bypassing the need for DFT calculations. However, while active research is being devoted to this area, the methodology is not yet stable enough for routine application to complex systems.[20,21] In the present work, therefore, we restrict our focus to assessing how well UMA can reproduce DFT-calculated energies on DFT-optimized geometries for catalytically relevant organometallic complexes with a particular focus on the correct description of conformational and configurational ensembles.

Specifically, we investigate the ability of UMA to rank the relative stability of TM complexes with varying conformational and configurational flexibility. Given the high inference speed of UMA, we assessed whether the predicted relative stabilities of different configurations of diverse transition-metal complexes are accurate enough for practical use cases. To address this question, we evaluate the performance of the smallest UMA model using our previously published data sets containing DFT-optimized geometries and energies for a broad array of TM complexes with varying bisphosphine ligands.

2. COMPUTATIONAL METHODS

In this study, the latest version of the small UMA model (UMA-s-1.1) pretrained on the OMol25 data set with 150 M total parameters was used. This model was chosen because we were mainly interested in a single-point energy calculation of structures with less than 1k atoms. Single-point energy calculations were performed via the Atomic Simulation Environment (ASE) in Python on a Dell X-ray photoelectron spectroscopy (XPS) 15 9520 laptop with a 12th generation Intel Core i5 processor, 32GB of RAM, and a Nvidia RTX 3050 GPU.

All generated DFT data sets focus on transition-metal based catalysts with bisphosphine ligands (Figure 1). Data set 1 consists of

https://doi.org/10.1021/acs.jpca.5c07061
J. Phys. Chem. A 2026, 130, 1897−1904

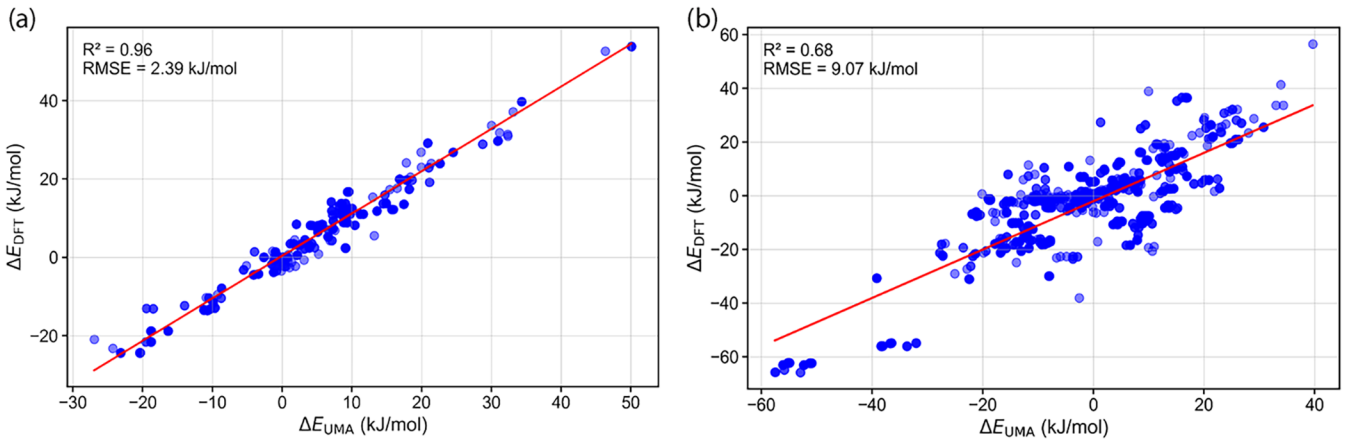


Figure 2. Correlation between DFT and UMA relative single-point energies for Data set 1 conformers: (a) rigid and (b) flexible model structures.

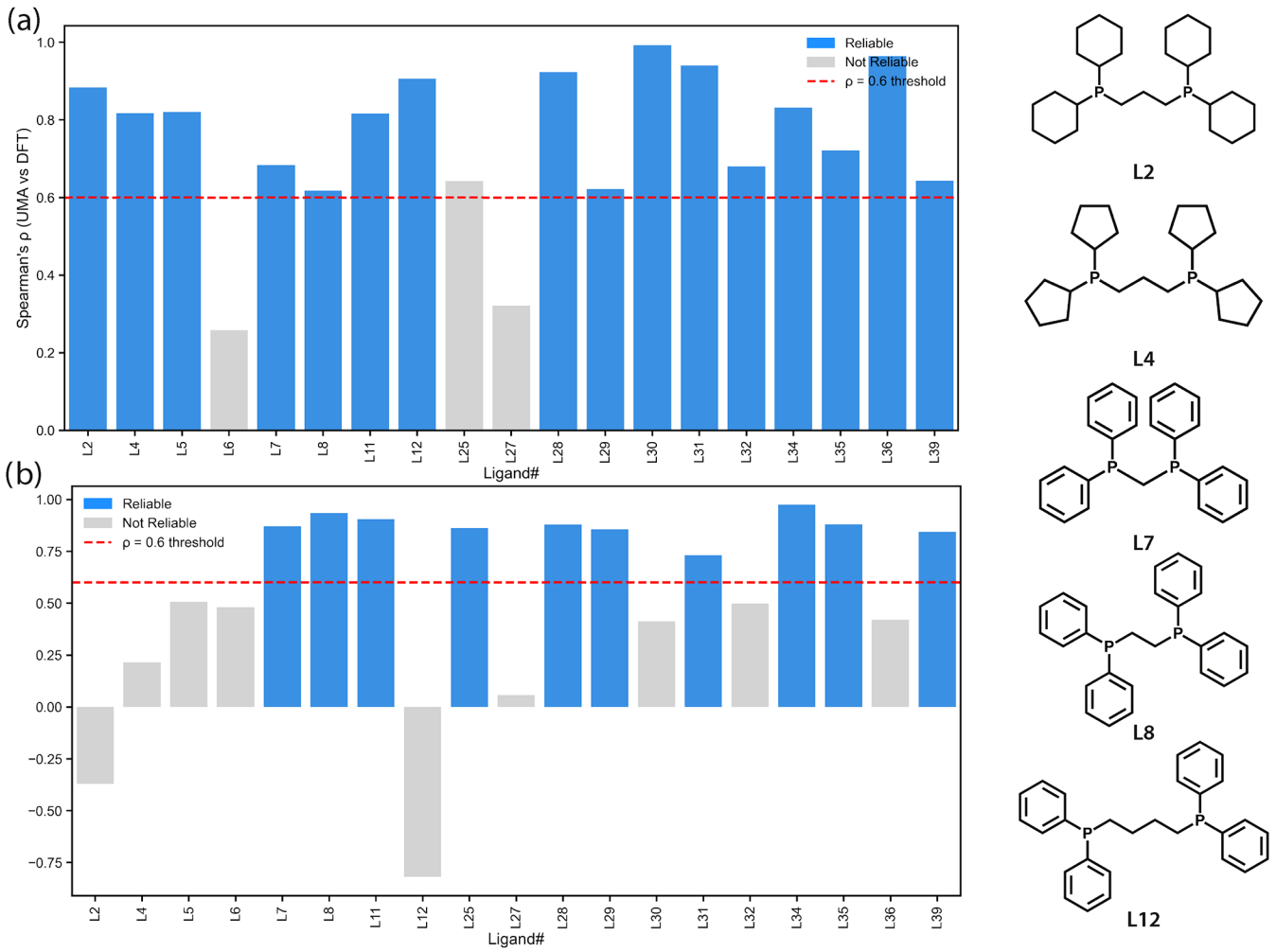


Figure 3. Spearman rank correlation (ρ) between UMA and DFT conformer energy rankings for Data set 1: (a) rigid and (b) flexible model structures. Bars are colored by reliability based on $\rho > 0.6$ (blue = reliable, $p < 0.05$ and $\rho > 0.6$; gray = not reliable). The red dashed line indicates the reliability threshold. The right-hand side shows representative bisphosphine ligands L2 (dCyhpp), L4 (Dcyppp), L7 (dppm), L8 (dppe), and L12 (dppb), illustrating differences in phosphorus substituents (phenyl, cyclohexyl, or pentyl) and linker length between donor atoms, which influence conformational flexibility and energy ranking performance.

conformers for square-planar Ni-based catalyst structures with 25 varying bisphosphine ligands. Two surrogate "model" structures were used to represent the catalyst, each exhibiting different degrees of conformational flexibility: a rigid [ligand]-Ni(II)-Cl_2 complex (top left), in which Cl_2 creates a symmetric coordination environment, and a more flexible [ligand]-Ni(II)-(CH_3CN)(−pOMe(C_6H_4)) complex (top right), where asymmetry is introduced by the coordination environment which increases conformational freedom.[2,3] These complexes are viewed as representative models of the precatalysts and relevant intermediate in the Ni-catalyzed arylation of nitriles.

https://doi.org/10.1021/acs.jpca.5c07061
J. Phys. Chem. A 2026, 130, 1897−1904

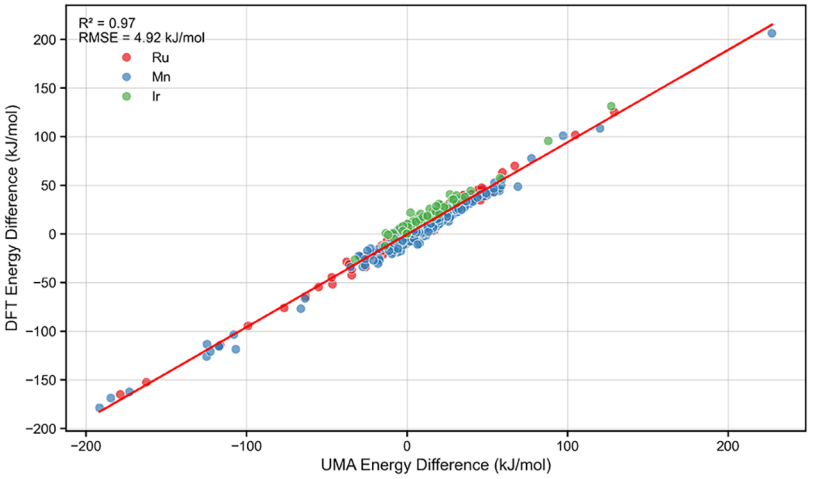


Figure 4. Correlation between UMA and DFT relative energies for Data Set 2 (configurations). Points are colored by metal center: Ir (green), Ru (red), and Mn (blue).

Data set 2 contains various octahedral complexes with Ir, Ru, or Mn metal centers in combination with 88 bisphosphine ligands relevant for homogeneous hydrogenation catalysis presented in our previous research.[22] The studied ligand configurations (bottom) for the Ir, Ru, and Mn complexes are named according to the donor atoms of ligands in the axial position.[22] For both data sets, the atomic structures of the coordination complexes including all possible stereoisomers were generated with our MACE Python package,[24] followed by exhaustive conformer search. For Data set 1, this conformer search was performed with CREST at the GFN2-xTB level of theory,[25−27] after which all resulting structures were subjected to density functional theory calculations. For Data set 2, conformers were generated using RDkit[28] and the lowest-energy conformer per ligand was selected for further analysis. Density functional theory calculations in gas phase were then used to optimize all of the resulting structures at the PBE0-D3(BJ)/def2-SVP level of theory.[29−32] Normal mode analysis was carried out to confirm that the optimized geometries correspond to local minima on the potential energy surface. For structures with imaginary frequencies, the PyQRC python package was used to remove these imaginary frequencies and restart geometry optimizations.[33,34]

Our case studies focus on accurately ranking the relative stabilities of either conformers (Data set 1) or ligand configurations (Data set 2). This ranking is performed by calculating the energy of the i-th conformer or configuration relative to a reference structure for the same ligand: $\Delta E_{\mathrm{DFT/UMA}} = E_i - E_{\mathrm{ref}}$. All relative energies are evaluated on DFT-optimized geometries to ensure consistency in structural input. To assess how well UMA reproduces DFT-based relative stabilities, conventional statistical metrics are employed, in particular, the Pearson correlation coefficient (R^2) and the root-mean-square error (RMSE). The strength of the linear correlation between DFT and UMA energies was quantified using Pearson's correlation coefficient (r), and its square (R^2) is reported as a measure of how well the data follow a linear trend. with values closer to 1 indicating better agreement. The RMSE provides a direct measure of the average deviation between the UMA and DFT energies, expressed in kJ/mol.

3. RESULTS AND DISCUSSION

Data set 1 consists of 23 ligands with 2100 conformers of the rigid dichloride model structure and 21 ligands with 3505 conformers of the flexible model structure. For comparative analysis, only ligands for which both the rigid and flexible model structures had fully converged conformer geometries were considered. This filtering resulted in 19 ligands, comprising 746 conformers for the rigid model structure and 1260 conformers for the flexible model structure. A comparison of relative energies computed by DFT and UMA on these structures (Figure 2) reveals excellent agreement, particularly for the rigid model's conformers (R^2 = 0.96, RMSE = 2.4 kJ/mol), and reasonable correlation for the flexible model structures (R^2 = 0.68, RMSE = 9.1 kJ/mol). As expected, these results indicate that UMA has difficulties capturing energetic trends when the flexibility of complexes increase. While R^2 and RMSE capture the overall quality of the linear correlation and average energy deviation, they do not directly assess whether the relative ranking of conformers per ligand is preserved, which is an essential criterion for a reliable conformational analysis. Therefore, to more directly evaluate the ability of UMA to rank conformers correctly, we also computed Spearman's rank-order correlation coefficient (ρ), a nonparametric metric that measures the monotonic relationship between two variables. In contrast to parametric measures such as Pearson's correlation coefficient, which assumes linearity and normally distributed variables, Spearman's rank-order correlation coefficient (ρ) operates solely on the ranked values. It, therefore, assesses whether two variables exhibit a monotonic relationship, regardless of the shape of their distributions. To ensure statistical robustness, we considered only ligands containing at least four conformer structures and required that the correlation be statistically significant (p-value <0.05). In addition, we adopted a threshold of $\rho > 0.6$ to define a "trusted" correlation, reflecting a moderate to strong monotonic agreement between UMA and DFT rankings. This cutoff is chosen to distinguish meaningful performance from weak or inconsistent ordering, as lower ρ values may indicate substantial deviations in the predicted energy ranking. By combining statistical significance with a minimum ranking strength criterion, this approach highlights cases where UMA is reliably predictive of the DFT-level ranking of relative stabilities.

Our results show that for the rigid model structures, UMA reliably predicts the ranking for 84% of the ligands (Figure 3). However, this performance declines when applied to the flexible model structures, where 53% of the ligands exhibit a reliable ranking correlation. Notably, in two cases, a negative Spearman ρ was observed, indicating that UMA predicts the reverse energy ranking compared to DFT.

To illustrate these trends, the right-hand side of Figure 3 presents five representative bisphosphine ligands (L2, L4, L7,

https://doi.org/10.1021/acs.jpca.5c07061
J. Phys. Chem. A 2026, 130, 1897−1904

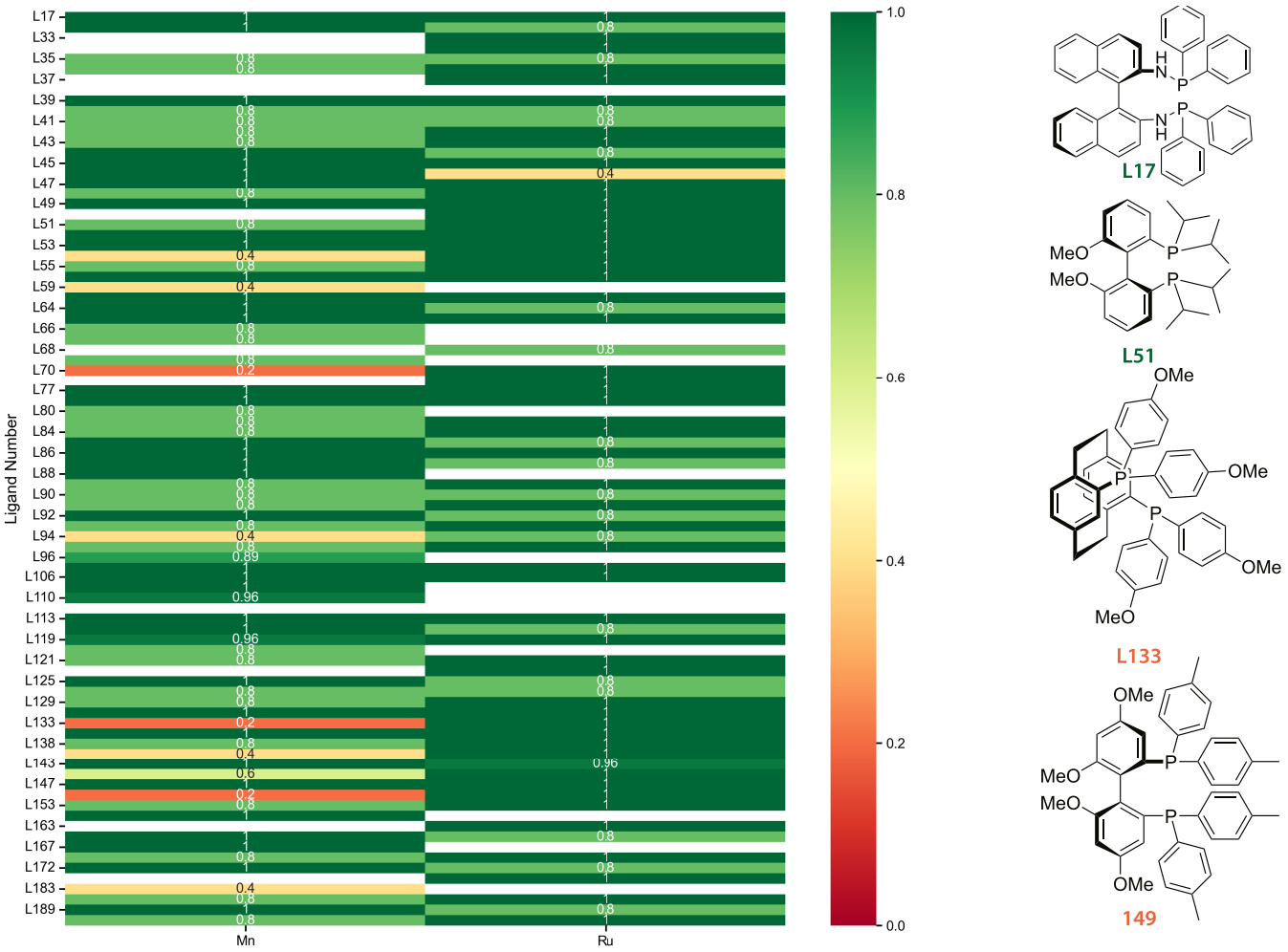


Figure 5. Spearman rank correlation (ρ) between UMA and DFT energy rankings for Data set 2: Mn(I)- and Ru(II)-based complexes. Cells are colored from red (low ρ) to green (high ρ). Ir(III) complexes are excluded due to insufficient configurations for ranking. The right panels show representative chiral bisphosphine ligands: L17 ((R)-BINAM-P) and L51 ((S)-iPr-BIPHEP) (green) achieve perfect agreement for both metals, while L133 ((R)-An-PhanePhos) and L149 ((R)-Tol-GarPhos) (orange) are reliable for Ru(II) but not for Mn(I).

L8, and L12), which differ primarily in the nature of the substituents on the phosphorus donors (phenyl, cyclohexyl, or pentyl) and in the length of the carbon linker between them. L2 (dCyhpp) and L4 (Dcyppp) are reliably predicted for the rigid model structure but not for the flexible model structure, with L2 even exhibiting a negative ρ. However, only eight conformers were identified for L2 in the flexible model structure with relative DFT energies differing by less than 2 kJ/mol. As expected, these are conditions under which UMA struggles to reproduce the DFT-level precision. L7 (dppm) and L8 (dppe), which have short linkers of one and two carbons, respectively, perform well for both rigid and flexible models, owing to their restricted conformational space. L12 (dppb) performs well in the rigid model structure but yields a negative ρ for the flexible model structure. In this case, only six conformers were found, again with relative DFT energies within 2 kJ/mol.

The tight energy ranges between conformers within Data set 1 provided a relatively strict test of UMA's ranking capability, as energy differences between conformers are low and often approach chemical accuracy of the DFT method. To examine UMA's performance under conditions where structural and energetic variations are greater, we next turned to Data set 2, which contains ligand configurations that span a much broader range of relative energies.

Data set 2 features 88 chiral bisphosphine ligands coordinated to various transition-metal complexes. After sampling stereoisomerism, 909 geometries were obtained. The metal centers, Ir(III), Ru(II), and Mn(I), are stabilized with different auxiliary ligands, resulting in three distinct classes of complexes: [ligand]-$IrH_3(CH_3CN)$, [ligand]-$RuH_2(CO)(CH_3CN)$, and [ligand]-$MnH(CO)_2(CH_3CN)$. A comparison of relative energies computed by DFT and UMA across this data set (Figure 4) again reveals excellent agreement (R^2 = 0.97, RMSE = 4.9 kJ/mol), highlighting UMA's general applicability across multiple transition metals and ligand environments.

To more closely examine the reliability of energy rankings for individual ligands, we excluded the Ir(III)-based complexes due to the presence of only two configurations per ligand, which is insufficient for rank correlation analysis. For the Mn(I)- and Ru(II)-based complexes, we again applied the significance and strength thresholds (p-value <0.05, ρ > 0.6) to define a "trusted" ranking. Based on this analysis, UMA is found to be reliably predictive for 53% of the ligand−metal combinations across the two metal centers (Figure 5). A closer examination per metal center reveals that 61% of the ligands

https://doi.org/10.1021/acs.jpca.5c07061
J. Phys. Chem. A 2026, 130, 1897−1904

are reliably predicted for Ru(II)-based complexes, whereas this drops to 44% for Mn(I)-based complexes. This difference is rooted in fundamental organometallic differences between the metal centers, where Mn(I) complexes are known to exhibit more fluxional behavior. This can lead to a more complex and shallow potential energy surface, making the energy landscape harder to learn and predict reliably, especially with a general-purpose machine-learned potential, such as UMA. In contrast, Ru(II) complexes tend to be more rigid and structurally well-defined under the same ligand field, contributing to the improved energy ranking accuracy observed for this class.

The right-hand side of Figure 5 highlights four representative bisphosphine ligands that exemplify these trends. L17 ((R)-BINAM-P) and L51 ((S)-iPr-BIPHEP) are colored green because they achieve a near-perfect ρ of 1.0 for both Mn(I) and Ru(II) complexes. In both cases, the relative energies between configurations are substantial, with L17 showing differences of at least 6 kJ/mol (up to 43 kJ/mol for the H−H configuration and 34 kJ/mol for the C−N configuration in Ru complexes) and L51 showing differences of at least 4 kJ/mol (C−N configuration in Mn) and typically exceeding 15 kJ/mol for both metals. These large energy differences appear to be easily predicted by UMA.

In contrast, L133 ((R)-An-PhanePhos) and L149 ((R)-Tol-GarPhos) illustrate cases in which UMA performs well for Ru(II) complexes but fails for Mn(I). For L133, Ru complexes display relative energy differences of at least 6 kJ/mol, whereas Mn complexes that feature C−C and C−H configurations differ by less than 2 kJ/mol from the reference. This aligns with our earlier observation that these Mn complexes can exhibit substantial structural isomerism, presenting multiple ligand configurations within 10 kJ/mol.[22] For L149, the Ru C−N configuration is separated from the reference by only 2.3 kJ/mol, yet is still ranked correctly, while the Mn C−C configuration differs by just 0.2 kJ/mol and the remaining Mn configurations lie within 5 kJ/mol, leading to inaccurate rankings. However, these examples reinforce the observation that only small energy differences that are near the DFT accuracy pose a challenge for UMA.

4. CONCLUSIONS

Our results demonstrate that UMA represents a significant step forward in the development of MLIPs. With near-DFT accuracy in energy prediction across a wide range of ligand conformers and configurations, UMA enables a rapid and scalable analysis in computational chemistry. For example, single-point energy evaluations using UMA can be performed in seconds on consumer-grade GPUs, in stark contrast with the CPU time required for equivalent DFT calculations. This speedup offers clear advantages in high-throughput screening and early stage catalyst design workflows. However, the current generation of MLIPs, including UMA, is not without limitations. While they are highly effective at predicting relative energies and forces, they do not yet provide access to electronic structure information such as the electron density, which remains essential for understanding charge distribution, reactivity, and spectroscopic properties. In addition, UMA is trained exclusively on gas-phase data and does not currently incorporate solvation effects. Because the solvent can substantially influence conformational energetics and catalytic behavior, extending UMA with implicit- or data-driven solvent models represents an important direction for future development. Such advances would broaden the applicability of MLIPs to solution-phase catalysis and other realistic chemical environments.

In Data set 1, reliable rankings are achieved for 84% of ligands in the rigid [ligand]-Ni(II)-Cl_2 model structures, but this drops to only 53% for the more flexible [ligand]-Ni(II)-(CH_3CN)(−pOMe(C_6H_4)) model structures, where asymmetric coordination and increased conformational freedom pose additional challenges. Ligands such as L7 (dppm) and L8 (dppe) are well-predicted in both rigid and flexible cases due to their short linkers and restricted conformational space, whereas L2 (dCyhpp) and L12 (dppb) are ranked less accurately for the flexible model structures, often when relative DFT energies between conformers differ by less than 2 kJ/mol. This is a regime where errors in ranking relative stabilities may have limited impact on equilibrium ground-state populations, and similar differences in ranking could appear from any DFT methodology by changing basis sets or functionals within the same rung on Jacob's ladder.

A similar pattern is observed in Data set 2, where UMA reliably ranks configurations for 61% of Ru(II)-based and 44% of Mn(I)-based complexes. Ligands with large relative configuration energy gaps, such as L17 ((R)-BINAM-P) and L51 ((S)-iPr-BIPHEP), achieve near-perfect correlations for both metals, while ligands with small energy separations within the error range of the DFT perform worse. These findings underscore that UMA as a general MLIP can be a powerful tool to achieve near-DFT accuracy for both rigid and structurally well-defined systems as well as fluxional and highly flexible complexes.

Finally, it is important to acknowledge the paradigm shift that MLIPs introduce. Although DFT itself is an approximation, the behavior and limitations of exchange-correlation functionals, basis sets, and dispersion corrections have been extensively studied and understood over decades. These methods are rooted in physics. In contrast, general-purpose ML models such as UMA are trained as black boxes on vast data sets, and their performance is not easily interpretable. This abstraction risks concealing underlying failures if used uncritically. As the field transitions into a post-UMA era, it will be crucial to combine the speed of ML with validation and domain expertise to ensure robust and meaningful chemical insights. Importantly, the instances where UMA fails to reproduce DFT rankings occur predominantly in near-degenerate energy regimes, where DFT itself cannot provide a uniquely reliable ranking. This exemplifies that expert judgment remains essential when interpreting such cases, even as MLIPs like UMA enable rapid and accurate exploration of broader regions of chemical space.

■ ASSOCIATED CONTENT

Data Availability Statement

The data sets, an overview of ligands and the code for reproducing the analysis presented in this study are available with an extensive readme via 4TU.ResearchData at 10.4121/6b178daf-e1c0-4c99-840f-06d382f37945.

■ AUTHOR INFORMATION

Corresponding Author

Evgeny A. Pidko − *Inorganic Systems Engineering, Department of Chemical Engineering, Faculty of Applied Sciences, Delft University of Technology, 2629 HZ Delft, The*

https://doi.org/10.1021/acs.jpca.5c07061
J. Phys. Chem. A 2026, 130, 1897−1904

Netherlands; orcid.org/0000-0001-9242-9901; Email: e.a.pidko@tudelft.nl

Author

Adarsh V. Kalikadien − *Inorganic Systems Engineering, Department of Chemical Engineering, Faculty of Applied Sciences, Delft University of Technology, 2629 HZ Delft, The Netherlands*; orcid.org/0000-0002-5414-3424

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.5c07061

Author Contributions

A.V.K.: Investigation, methodology, conceptualization, software, validation, data curation, formal analysis, visualization, writing—original draft, writing—review & editing, project administration E.A.P.: Supervision, conceptualization, resources, funding acquisition, writing—review & editing, project administration

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by Janssen Pharmaceutica NV, a Johnson & Johnson company. The authors thank the NWO Domein Exacte en Natuurwetenschappen for the use of the national supercomputer, Snellius.

■ REFERENCES

(1) Nandy, A.; Duan, C.; Taylor, M. G.; Liu, F.; Steeves, A. H.; Kulik, H. J. Computational Discovery of Transition-metal Complexes: From High-throughput Screening to Machine Learning. *Chem. Rev.* **2021**, *121*, 9927−10000.

(2) Kalikadien, A. V.; Mirza, A.; Hossaini, A. N.; Sreenithya, A.; Pidko, E. A. Paving the road towards automated homogeneous catalyst design. *ChemPlusChem* **2024**, *89*, No. e202300702.

(3) Foscato, M.; Jensen, V. R. Automated in Silico Design of Homogeneous Catalysts. *ACS Catal.* **2020**, *10*, 2354−2377.

(4) Butera, V. Density functional theory methods applied to homogeneous and heterogeneous catalysis: a short review and a practical user guide. *Phys. Chem. Chem. Phys.* **2024**, *26*, 7950−7970.

(5) Bellonzi, N.; Kunitsa, A.; Cantin, J. T.; Campos-Gonzalez-Angulo, J. A.; Radin, M. D.; Zhou, Y.; Johnson, P. D.; Martínez-Martínez, L. A.; Jangrouei, M. R.; Brahmachari, A. S.; Wang, L.; Patel, S.; Kodrycka, M.; Loaiza, I.; Lang, R. A.; Aspuru-Guzik, A.; Izmaylov, A. F.; Fontalvo, J. R.; Cao, Y. Feasibility of accelerating homogeneous catalyst discovery with fault-tolerant quantum computers. *Mar. Technol. Soc. J.* **2024**, *59*, 62−65.

(6) Matsuoka, W.; Harabuchi, Y.; Maeda, S. Virtual Ligand-Assisted Screening Strategy to Discover Enabling Ligands for Transition Metal Catalysis. *ACS Catal.* **2022**, *12*, 3752−3766.

(7) Gensch, T.; Gomes, G. D. P.; Friederich, P.; Peters, E.; Gaudin, T.; Pollice, R.; Jorner, K.; Nigam, A.; Lindner-D'Addario, M.; Sigman, M. S.; Aspuru-Guzik, A. A Comprehensive Discovery Platform for Organophosphorus Ligands for Catalysis. *J. Am. Chem. Soc.* **2022**, *144*, 1205−1217.

(8) Pidko, E. A. Toward the Balance between the Reductionist and Systems Approaches in Computational Catalysis: Model versus Method Accuracy for the Description of Catalytic Systems. *ACS Catal.* **2017**, *7*, 4230−4234.

(9) Behler, J. Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17930−17955.

(10) Behler, J. Constructing high-dimensional neural network potentials: A tutorial review. *Int. J. Quantum Chem.* **2015**, *115*, 1032−1050.

(11) Kocer, E.; Ko, T. W.; Behler, J. Neural Network Potentials: A Concise Overview of Methods. *Annu. Rev. Phys. Chem.* **2022**, *73*, 163−186.

(12) Eyert, V.; Wormald, J.; Curtin, W. A.; Wimmer, E. Machine-learned interatomic potentials: Recent developments and prospective applications. *J. Mater. Res.* **2023**, *38*, 5079−5094.

(13) Handley, C. M.; Popelier, P. L. A. Potential Energy Surfaces Fitted by Artificial Neural Networks. *J. Phys. Chem. A* **2010**, *114*, 3371−3383.

(14) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chem. Rev.* **2021**, *121*, 10073−10141.

(15) Li, Y.; Zhang, X.; Liu, M.; Shen, L. A critical review of machine learning interatomic potentials and Hamiltonian. *J. Mater. Inf.* **2025**, *5*, No. 43.

(16) Kulichenko, M.; Nebgen, B.; Lubbers, N.; Smith, J. S.; Barros, K.; Allen, A. E. A.; Habib, A.; Shinkle, E.; Fedik, N.; Li, Y. W.; Messerly, R. A.; Tretiak, S. Data Generation for Machine Learning Interatomic Potentials and Beyond. *Chem. Rev.* **2024**, *124*, 13681−13714.

(17) Levine, D. S.; Shuaibi, M.; Spotte-Smith, E. W. C.; Taylor, M. G.; Hasyim, M. R.; Michel, K.; Batatia, I.; Csányi, G.; Dzamba, M.; Eastman, P.; Frey, N. C.; Fu, X.; Gharakhanyan, V.; Krishnapriyan, A. S.; Rackers, J. A.; Raja, S.; Rizvi, A.; Rosen, A. S.; Ulissi, Z.; Vargas, S.; Zitnick, C. L.; Blau, S. M.; Wood, B. M. The Open Molecules 2025 (OMol25) Dataset, Evaluations, and Models, 2025 arXiv preprint. arXiv.org e-Printarchive. arXiv:2505.08762 https://doi.org/10.48550/arXiv.2505.08762.

(18) Wood, B. M.; Dzamba, M.; Fu, X.; Gao, M.; Shuaibi, M.; Barroso-Luque, L.; Abdelmaqsoud, K.; Gharakhanyan, V.; Kitchin, J. R.; Levine, D. S.; Michel, K.; Sriram, A.; Cohen, T.; Das, A.; Rizvi, A.; Sahoo, S. J.; Ulissi, Z. W.; Zitnick, C. L. UMA: A Family of Universal Models for Atoms, 2025 arXiv preprint. arXiv.org e-Printarchive. arXiv:2506.23971 https://doi.org/10.48550/arXiv.2506.23971.

(19) Axelrod, S.; Gómez-Bombarelli, R. GEOM, energy-annotated molecular conformations for property prediction and molecular generation. *Sci. Data* **2022**, *9*, No. 185.

(20) Wu, Z.; Zhou, L.; Hou, P.; Liu, Y.; Guo, T.; Liu, J.-C. Catalytic Large Atomic Model (CLAM): A Machine-Learning-Based Interatomic Potential Universal Model, ChemRxiv preprint, 2023.

(21) Goldsmith, B. R.; Esterhuizen, J.; Liu, J.-X.; Bartel, C. J.; Sutton, C. Machine learning for heterogeneous catalyst design and discovery. *AIChE J.* **2018**, *64*, 2311−2323.

(22) Kalikadien, A. V.; van der Lem, N. J.; Valsecchi, C.; Lefort, L.; Pidko, E. A. Unveiling the impact of ligand configurations and structural fluxionality on virtual screening of transition-metal complexes. *Digital Discovery* **2025**, *4*, 2033−2044.

(23) Baidun, M. S.; Kalikadien, A. V.; Lefort, L.; Pidko, E. A. Impact of Model Selection and Conformational Effects on the Descriptors for In Silico Screening Campaigns: A Case Study of Rh-Catalyzed Acrylate Hydrogenation. *J. Phys. Chem. C* **2024**, *128*, 7987−7998.

(24) Chernyshov, I. Y.; Pidko, E. A. MACE: Automated Assessment of Stereochemistry of Transition Metal Complexes and Its Applications in Computational Catalysis. *J. Chem. Theory Comput.* **2024**, *20*, 2313−2320.

(25) 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.

(26) Bannwarth, C.; Caldeweyher, E.; Ehlert, S.; Hansen, A.; Pracht, P.; Seibert, J.; Spicher, S.; Grimme, S. Extended tight-binding quantum chemistry methods. *WIREs Comput. Mol. Sci.* **2021**, *11*, No. e1493.

(27) Grimme, S. Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. *J. Chem. Theory Comput.* **2019**, *15*, 2847−2862.

(28) Landrum, G. RDKit: Open-source cheminformatics 2020; http://www.rdkit.org/ (access: August 08, 2025).
(29) 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-Young, D.; 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 Inc 16 Revision C.01*; Gaussian Inc.: Wallingford CT, 2016.
(30) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158−6170.
(31) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A generally applicable atomic-charge dependent London dispersion correction. *Chem. Phys.* **2019**, *150*, No. 154122.
(32) 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*, 3297−3305.
(33) Silva, M. A.; Goodman, J. M. Aziridinium ring opening: a simple ionic reaction pathway with sequential transition states. *Tetrahedron Lett.* **2005**, *46*, 2067−2069.
(34) Goodman, J. M.; Silva, M. A. QRC: a rapid method for connecting transition structures to reactants in the computational analysis of organic reactivity. *Tetrahedron Lett.* **2003**, *44*, 8233−8236.

https://doi.org/10.1021/acs.jpca.5c07061
J. Phys. Chem. A 2026, 130, 1897−1904