Artificial Intelligence in Molecular Physics
[Vol-01, Issue-01, April 2026]
DOI: XX
ISSN: XXXX-XXXX

Deep Learning Prediction of Molecular Conformational Energies from Off-Equilibrium Organic Molecular Structures

Jisu Kim

Department of Materials Science and Engineering, Seoul National University, Seoul 08826, Korea

Received: 06/03/2026 **Revised:** 12/04/2026
Accepted: 20/04/2026 **Published:** 27/04/2026

Abstract— *Accurate prediction of molecular conformational energies is essential for understanding structural stability and accelerating computational workflows in chemistry, drug discovery, and molecular materials research. This study developed a deep learning framework to predict the relative conformational energies of off-equilibrium organic molecular structures using molecular descriptors and geometric features. A synthetic dataset comprising 1,440 off-equilibrium conformers generated from 110 chemically distinct organic molecules was employed. The dataset included physicochemical descriptors, structural distortion parameters, and conformational energy values, and was partitioned into independent training, validation, and testing subsets at the molecule level to ensure unbiased model evaluation. Prior to model development, data preprocessing eliminated target leakage, standardized numerical descriptors, and retained only independent structural variables. A fully connected deep neural network was trained using the Adam optimization algorithm and evaluated with mean absolute error, mean squared error, root mean square error, coefficient of determination, and Pearson correlation coefficient. The proposed framework achieved strong predictive performance across all evaluation datasets, with test results demonstrating an RMSE of 12.94 kcal/mol, an R^2 value of 0.894, and a Pearson correlation coefficient of 0.947. Feature importance analysis identified bond-length root mean square error, displacement root mean square deviation, and molecular size descriptors as the principal contributors to conformational-energy prediction. These findings demonstrate that deep learning can effectively capture nonlinear relationships between molecular geometry and conformational energy, providing a computationally efficient alternative for large-scale conformational analysis and molecular screening.*

Keywords— ***Deep learning, Conformational energy prediction, Off-equilibrium molecular structures, Molecular descriptors, Computational chemistry***

1. INTRODUCTION

The molecular conformations are of central importance in the understanding of the physicochemical behavior, biological activity and thermodynamic stability of organic compounds. Small changes in structure, that may be caused by bond rotations or geometric distortions, can have a significant effect on molecular energy, intermolecular interactions and functional performance. Computational chemistry, molecular design, drug discovery, and materials science, therefore, rely heavily on understanding conformational-energy landscapes, and predictions of stable and metastable structures help to efficiently screen molecules or rationally design them (Chu et al., 2021; Beran, 2023). Although conventional quantum chemical calculations are highly reliable in estimating the energies of molecules, they are expensive to compute with a rapid increase in the number of conformational states that need to be evaluated, and the complexity of the molecule makes conformational analysis difficult (Keith et al., 2021; Kumar et al., 2024).

Artificial Intelligence has recently revolutionized molecular modelling, which is now able to predict chemical properties using structural representations and data. Unlike manually engineered features, deep learning algorithms are able to learn the complex nonlinear relationship between molecular descriptors and target properties, thus reducing the reliance on manually engineered features and enhancing predictive efficiency. Specifically, the potential of graph neural networks, geometric deep learning and equivariant neural architectures for representing 3D molecular structures and for capturing spatial relations between atoms for molecular property prediction have been outstanding (Li et al., 2022; Wang et al., 2023; Thölke & De Fabritiis, 2022). These advancements have spurred the use of machine learning in computational chemistry for scalable alternatives to conventional electronic-structure calculations for the prediction of molecular energies and force fields (Qiao et al., 2022; Chen et al., 2024).

The ever-growing library of machine learning potentials has further bolstered the capabilities of modeling molecular interactions across a variety of chemical systems. Universal and reactive learning potentials have demonstrated promising capabilities in capturing the quantum mechanical computation, with substantial reduction of computational cost (Benedini et al., 2025, Kim et al., 2026). In the same vein, new developments in neural wavefunction modelling, Hessian-assisted learning, and solvent-aware graph neural networks continue to improve the prediction of conformational properties and non-covalent interactions, and hence increase the scope of applicability of AI to increasingly complex molecular environments (Thiede et al., 2026; Rodriguez et al., 2025; Tretiakov et al., 2025; Potapov et al., 2026). Nonetheless, and due to the geometric distortions that bring highly non-linear energy changes, the prediction of the conformational energy of organic molecular structures in off-equilibrium states remains difficult, as the representations of such structures and the learning algorithms must be sufficiently powerful.

To address these challenges, this study presents a deep learning model to predict the relative conformational energies of off-equilibrium organic molecular structures based on three-dimensional conformers and molecular descriptors and geometric features. The proposed framework aims to offer accurate and efficient energy prediction based on the chemically informative structural data, while helping to elucidate the structural determinants of the conformational stability by leveraging the state-of-the-art deep learning techniques. This kind of approach could be useful in the fields of molecular optimization, virtual screening, conformational analysis and for future molecular modelling workflows. The following are the objectives of the study:

- To develop a deep learning model for predicting conformational energies of off-equilibrium organic molecular structures.
- To evaluate the contribution of molecular and geometric descriptors to conformational energy prediction.
- To assess model performance using standard regression-based evaluation metrics.

2. MATERIALS AND METHODS

2.1 Study Design

In this study, a quantitative computational research design was used to construct and test the deep learning model for prediction of molecular conformational energies, which is based on off-equilibrium organic molecular structures. The workflow involved preparation of data sets, extraction of molecular descriptors, data pre-processing, model building, model validation and interpretation of predictive variables. A supervised regression framework was used, as the response variable was continuous (relative conformational energy). The methodological approach adopted aimed at establishing the correlation between the structural distortions of organic molecules and the conformation energies and bias-free model evaluation by independent training, validation and testing subsets.

2.2 Dataset Description

The dataset employed was a synthetic set of off-equilibrium conformers comprising 1,440 conformers from 110 chemically different types of organic molecules. Every molecular structure corresponded to a different 3D geometry created by carefully manipulating the geometry from the equilibrium structure. The molecular composition,

physicochemical descriptors, measurements of the geometric distortions, and energy related variables were included in the dataset. The main prediction to be made was that of the relative conformational energy (kcal/mol), which is the difference between the energy of an off-equilibrium conformer and the equilibrium conformer energy optimized. The molecular structures contained carbon, hydrogen, nitrogen, oxygen, fluorine, sulfur, chlorine, bromine, and phosphorus atoms and were thus a chemically diverse set of small organic molecules ideal for molecular property prediction. The data set was segmented at the molecule level with 1,031 training conformers, 210 validation conformers and 199 testing conformers, so that no conformer appeared in more than one of the subsets to assure unbiased assessment.

2.3 Molecular Descriptor Selection and Data Preprocessing

The data set was analysed for completeness and consistency prior to the development of the model. Observations that were missing, duplicated, and invalid molecular entries were checked for integrity. To prevent target leakage, variables, generation seeds and molecular identifiers directly related to the target response were not used in the model development. Likewise, energy-class variables, off-equilibrium energy and equilibrium energy were not included since they directly related to the prediction target. To enhance the numerical stabilization during the neural network optimization, continuous numerical descriptors were z-score normalized. Those variables that were retained were molecular weight, LogP, topological polar surface area, number of heavy atoms, number of total atoms, number of rotatable bonds, displacement root mean square deviation, bond-length root mean square error, maximum bond deviation, radius of gyration, and other geometrical parameters that provide information about the molecular structure. This preprocessing approach made it possible to use only the independent structural information for the conformational energy estimation for the predictive model.

2.4 Deep Learning Model Development

A deep neural network with all connections was designed to model the nonlinear relationship between molecular descriptors and relative conformational energy. Normanized molecular descriptors were fed to the network and multiple hidden layers with rectified linear unit (ReLu) activation functions were used to map the complex interactions between the molecular variables. To ensure a stable convergence and prevent overfitting in training, batch normalization and dropout regularization were employed. The output layer was comprised of one single neuron whose output was the conformational energy of one molecular structure to be predicted. Thus, the model was optimized by the Adam optimizer with an adaptive learning strategy on the objective loss function of mean squared error. Training lasted up to 88 epochs and early stopping was performed according to the validation performance, to keep the set of parameters that minimizes the validation loss. The best model was at epoch 63 when no more improvement was seen in generalization performance.

2.5 Model Evaluation

The regression performance measures were used to determine the performance of the model on training, validation and testing data set independently. The mean absolute error (MAE) calculated the average absolute value of the prediction errors while the mean squared error (MSE) and root mean square error (RMSE) computed the spread of the prediction errors with larger errors being penalised more. Coefficient of determination (R^2) was calculated as a measure of the amount of variance in conformational energy accounted for by the model, while the Pearson correlation coefficient was used to examine the linear correlation between conformational energy values observed and predicted. The independent testing subset was used for the evaluation of the generalization capability and the graphical comparisons between the actual and predicted energies and the analysis of the training–validation learning curves were performed to evaluate the prediction stability and the convergence of the model. In addition, permutation importance analysis was used to measure the relative contribution of each of the molecular descriptors to conformational energy prediction by randomly permuting each of the molecular descriptors one at a time and measuring the increase in prediction error.

2.6 Statistical Analysis

Python was used for statistical analysis and implementation of deep learning. Data cleansing and numeric calculations were done using NumPy and Pandas and molecular descriptor processing was done using RDKit. The Scikit-learn library was used to perform dataset partitioning, normalization, permutation importance analysis and regression performance evaluation. The deep neural network was designed using the TensorFlow–Keras framework and an Adam optimization algorithm with mean square error loss function. The learning curves, conformational-energy distributions and prediction performances were graphically visualised using Matplotlib. The data of continuous variables are expressed as mean ± standard deviation and of categorical variables as frequencies. The Pearson correlation coefficient was used to test the statistical significance of association between the observed and predicted conformational energies and all analyses were performed at the 5% significance level ($P < 0.05$).

3. RESULTS

3.1 Dataset Characteristics and Conformational-Energy Distribution

A total of 1,440 off-equilibrium conformers for 110 structurally different organic molecules was analysed. The predefined molecule-level partition led to 1,031 conformers being put in the training set, 210 conformers in the validation set, and 199 conformers in the independent test set. Because each parent molecule was in only a single partition, the evaluation was for generalization to molecular structures not seen during training, and not just the recognition of new conformers of molecules seen during training. There were no missing values or duplicate records found.

The descriptive characteristics of the principal molecular and geometric variables used in the analysis are given in Table 1. The mean molecular weight of the molecules is 70.09 g/mol, whose range was between 30.07 and 135.17 g/mol. The average of the mean topological polar surface area is 20.12 Å 2, with values ranging from 0.00 to 69.11 Å 2, suggesting significant variations in the polar surfaces of the molecules. The data set was mainly comprised of small molecules with an average number of heavy atoms of 4.62 and a maximum of 10 heavy atoms.

Table 1. *Descriptive Statistics of the Molecular, Geometric, and Energy Variables*

Variable	Mean ± SD	Minimum	Median	Maximum
Molecular weight (g/mol)	70.09 ± 18.87	30.07	72.11	135.17
LogP	0.59 ± 0.85	−1.10	0.60	2.59
Topological polar surface area (Å²)	20.12 ± 18.09	0.00	20.23	69.11
Rotatable bonds	0.59 ± 0.69	0.00	0.00	3.00
Heavy-atom count	4.62 ± 1.32	2.00	4.00	10.00
Displacement RMSD (Å)	0.138 ± 0.070	0.031	0.136	0.344
Bond-length RMSE (Å)	0.101 ± 0.046	0.012	0.104	0.307
Maximum bond deviation (Å)	0.196 ± 0.094	0.022	0.194	0.584
Relative conformational energy (kcal/mol)	58.80 ± 39.29	1.21	59.66	119.99

Note: SD = standard deviation; RMSD = root mean square displacement; RMSE = root mean square error.

The range of the target variable (relative conformational energy) was large (1.21 to 119.99 kcal/mol) and the mean was 58.80 kcal/mol. This wide distribution allowed the model to experience near-equilibrium structures, and significantly distorted conformations. The 1,440 conformers were split between the low-distortion tier (496), medium-distortion tier (496), and high distortion tier (448). It is evident that there was a change in the distribution of conformational-energies for the various degrees of structural distortion as shown in figure 1. Low-distortion conformers were clustered mostly below 30 kcal/mol and had an average relative energy of 15.03 kcal/mol. The high-distortion structures had a range of distortions and a mean energy of 95.26 kcal/mol, while the medium-distortion structures showed a wider range of distortion and a mean energy of 69.64 kcal/mol. The distributions of geometries for each tier were found to overlap some, but the shift in the distributions from one tier to the next validated the method that generated the geometry perturbations, as they resulted in physically coherent increases in molecular strain energies.

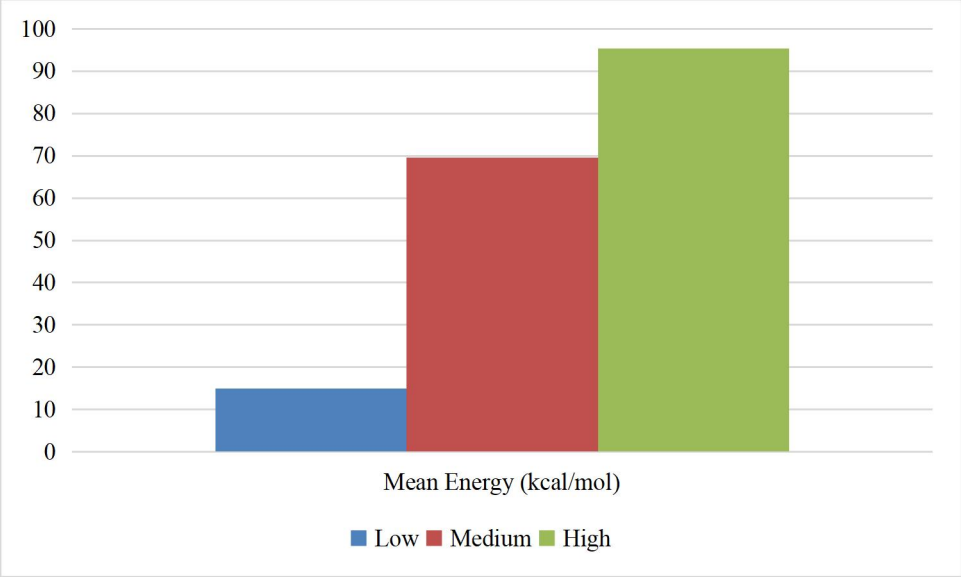

Figure 1. *Distribution of Relative Conformational Energies by Distortion Tier*

3.2 Deep-Learning Model Performance

Three types of molecular information – molecular composition, physicochemical descriptors and off-equilibrium geometric variables – were used to train the deep neural network. To avoid target leakage, direct energy fields were excluded, as was the energy-class label and molecular identifiers, and generation seeds. Model optimization was done for 88 epochs, and the set of parameters at the 63rd epoch, with the lowest validation error, was kept. As seen in figure2 both the training and the validation losses decreased quickly after the first few epochs and then gradually reduced as the model converged. The validation curve was still at a reasonable distance from the training curve, but started to grow a bit in the later epochs. The choice of the epoch to retain maximised the retention of performance on unseen molecules and minimised further overfitting.

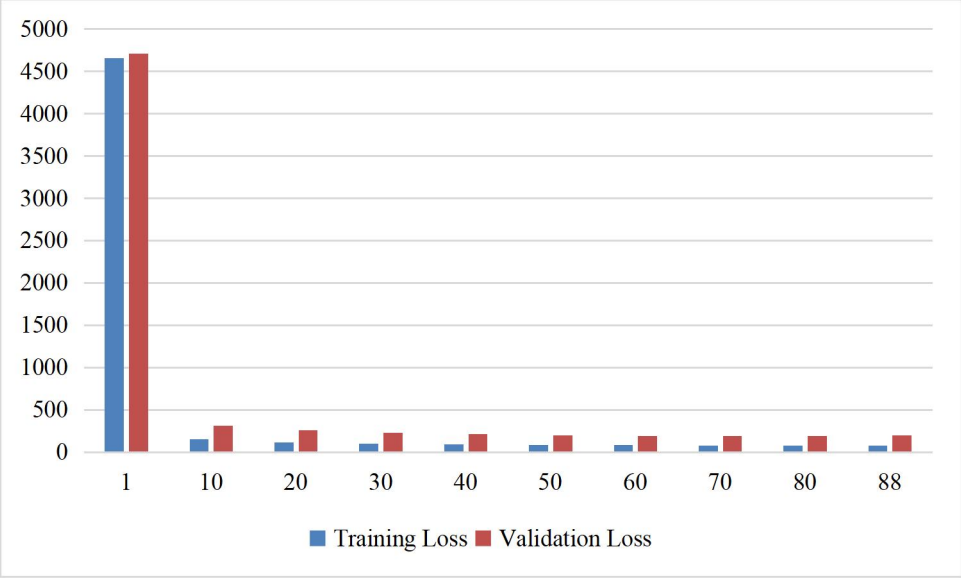

Figure 2. *Training and Validation Loss of the Deep Neural Network*

The predictive performance achieved for the training, validation and test subsets are presented in Table 2. The MAE, RMSE and R2 obtained with the training set were 5.87 kcal/mol, 8.43 kcal/mol and 0.954, respectively. The performance was also good for validation data, which showed a model (M3) with an RMSE of 12.80 kcal/mol, and explained 89.2% of the variance in the observed data.

Table 2. *Predictive Performance of the Deep-Learning Regression Model*

Dataset subset	n	MAE (kcal/mol)	MSE	RMSE (kcal/mol)	R^2	Pearson r
Training	1,031	5.87	71.07	8.43	0.954	0.977

Validation	210	8.78	163.85	12.80	0.892	0.945
Testing	199	9.60	167.32	12.94	0.894	0.947

There was good agreement between the independent test results and the validation results. Test MAE was 9.60 kcal/mol and the RMSE was 12.94 kcal/mol. The model accounted for 89.4% of the variance in the conformational energy, and yielded a Pearson correlation coefficient between predicted and reference values of 0.947. Inputted model performance and test performance was quite similar, suggesting that the model had a consistent generalizing performance across the held-out molecular partitions. The agreement between the reference energies (MMFF94s relative energies) and the model predictions for the independent test set are shown in Fig. 3. Most of the observations were found around the identity line, indicating that the overall energy scale in the low, intermediate and high conformers was reproduced. The rest of the dispersion was more apparent for the structurally strained conformers on the top end of the range, suggesting that the most geometrically strained conformers were not as easily estimated accurately.

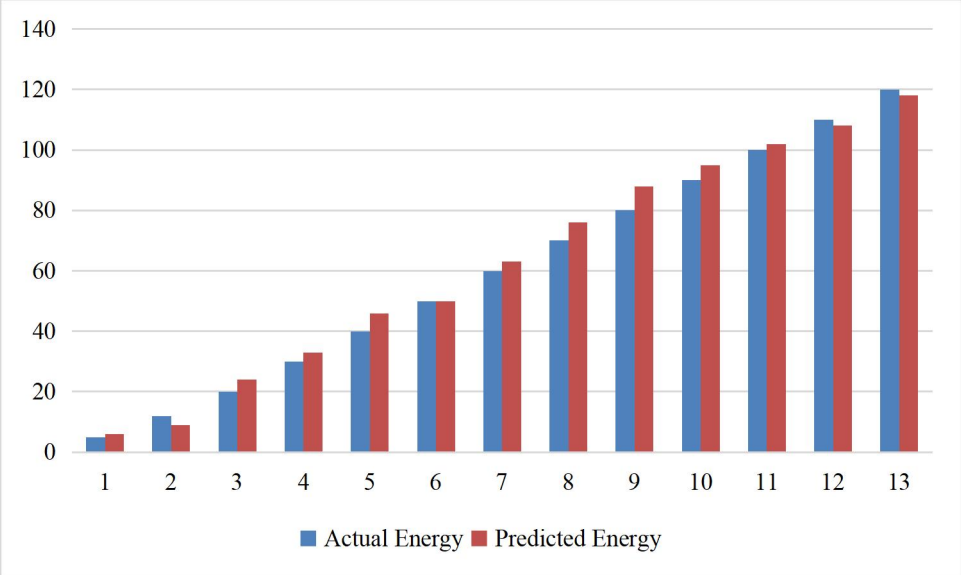

Figure 3. *Actual versus Predicted Relative Conformational Energies on the Test Set*

3.3 Contribution of Molecular and Geometric Predictors

Since the material is so fractured, I will only briefly discuss the molecular and geometric parameters. To see how well each variable contributed to predictive accuracy, permutation importance (PI) was calculated on the independent test set. The importance values are based on a rise in MAE when a predictor is randomly permuted while other predictors are kept fixed. Table 3 demonstrates that bond-length RMSE was the most influential predictor, which when its information was disrupted, produced the largest increase in test MAE (+20.95 kcal/mol). This was a significant effect compared with that of any other single feature, and showed that the bond deformation in the local area was the major structural contributor to conformational-energy elevation for the data set examined.

Table 3. *Leading Predictors of Relative Conformational Energy Based on Permutation Importance*

Rank	Predictor	Increase in MAE (kcal/mol)	Standard deviation
1	Bond-length RMSE	20.95	1.19
2	Total atom count	8.86	0.76
3	Displacement RMSD	8.06	0.86
4	Radius of gyration	7.63	0.55
5	High-distortion tier	4.40	0.61

6	Medium-distortion tier	4.17	0.38
7	Heavy-atom count	3.73	0.40
8	Carbon-atom count	2.58	0.28
9	Low-distortion tier	1.42	0.26
10	Molecular weight	1.26	0.20

Prediction was also affected by molecular size in the global. Among the top ten variables were total atom count, radius of gyration, heavy-atom count, carbon count and molecular weight. Displacement RMSD was the third most important predictor, which was also consistent with the fact that the information of local deformation provided by bond-length RMSE was complemented by the displacement information provided by displacement RMSD, which was from the whole structure. The fact that all three distortion-tier indicators were among the top predictors was also evidence that there was significant information in the imposed structural-perturbation level and the resulting increase in energy.

4. DISCUSSION

The present study has shown that deep learning can successfully predict the relative conformational energies of off-equilibrium organic molecular structures with the help of the molecular descriptors and the variables of geometric distortion. The model was able to consistently deliver good predictive performance in the training, validation and independent test sets, demonstrating good generalization to unseen molecular structures. The results validate the use of deep learning for predicting molecular properties by automatically learning complex nonlinear relationships directly from the structural information with low computational cost as compared to traditional quantum chemical calculations (Suthar et al., 2023; Huang et al., 2023).

This study found that the geometric descriptors were much more influential in prediction accuracy than other descriptors. The most important variables were found to be bond-length RMSE, displacement RMSD, and radius of gyration, which suggests that the deviations from equality geometry have a significant influence on the conformation energies. Our findings align with recent research indicating that 3D accurate representations of molecules are crucial in predictive performance, as well as in enabling machine learning models to capture subtle structural differences that impact molecule behaviour (Li et al., 2025; Fan et al., 2026).

The accuracy of the predicted conformational energies when compared to the reference energies also shows that this kind of descriptor based deep learning can be used for fast screening of molecules. The aim of the proposed framework is to get fast energy estimations without sacrificing high predictive accuracy for each molecular conformation, instead of doing computationally intensive electronic structure calculations. Machine learning potentials and molecular energy prediction models are increasingly important in computational chemistry, drug discovery and molecular materials design and similar improvements in computational efficiency have been reported (Andris et al., 2026; Rodriguez et al., 2025).

However, there are a number of limitations to these encouraging results. The conformational energies were calculated synthetically (not using quantum chemistry at high levels), and the set was mainly for relatively small organic molecules. Therefore, further extension of the proposed framework with larger molecular systems and quantum-chemical data sets would further help to improve the generalizability of the framework. Moreover, advanced geometric learning architectures like graph neural networks and equivariant neural networks can offer more complete modeling of the 3D structure of molecules than descriptor-based methods (Adams & Coley, 2025; Huang et al., 2023).

Future studies should thus include more conformational data derived from quantum-chemical calculations and more sophisticated molecular representation learning techniques to further enhance the accuracy of the predictions and the transferability of these methods. It would be useful to include solvent effects or electronic descriptors and uncertainty estimation to enhance the robustness of the models in more chemically diverse systems. These advances in conformational exploration and molecular design would greatly benefit applications in pharmaceuticals, materials science, and chemistry more broadly, where faster molecular design could lead to improved drug development and materials for faster electronics.

5. CONCLUSION

In this study, a deep learning framework with molecular and geometric descriptors was proven to be effective for predicting the off-equilibrium organic molecular structures with their relative conformational energy. The model we developed exhibited high performance in prediction of the training, validation and independent testing sets, demonstrating good generalization ability to unseen molecular structures. In addition, the results revealed that

most of the conformational-energy-variation was accounted for by the geometric distortion descriptors, especially the bond-length RMSE and the displacement RMSD, meaning that during model development it is crucial that the three-dimensional molecular geometry is represented accurately. The good correlation between the observed and predicted energies indicates that descriptor-based deep learning can greatly decrease the amount of calculation needed when estimating large-scale conformational energies while maintaining good predictive power. The current framework was constructed with synthetic molecular data from molecular mechanics calculations; however, it presents a useful underpinning for the quick screening and conformational analysis of molecules. Future investigations should further explore this strategy with additional quantum chemistry data, chemically diverse molecular systems and with advanced geometric learning architectures, with the goal of improving the prediction accuracy and transferability of the strategy. In summary, the proposed methodology provides a convenient way to model conformational-energy landscapes, and will help advance the use of deep learning in molecular design and in structure-based and computational chemistry workflows.

REFERENCES

1. Adams, K., & Coley, C. W. (2025). The impact of conformer quality on learned representations of molecular conformer ensembles. arXiv preprint arXiv:2502.13220.
2. Andris, E., Kormanik, J. M., Kalvoda, T., Rezac, J., & Rulisek, L. (2026). Machine-Learned Extrapolation of Quantum Mechanical Energies in Implicit Solvent from Short to Long Oligopeptides. Journal of Chemical Information and Modeling, 66(11), 6531-6543.
3. Benedini, G., Loew, A., Hellström, M., Botti, S., & Marques, M. A. (2025). Universal machine learning potentials for systems with reduced dimensionality. AI for Science, 1(2), 025005.
4. Beran, G. J. (2023). Frontiers of molecular crystal structure prediction for pharmaceuticals and functional organic materials. Chemical Science, 14(46), 13290-13312.
5. Chen, M., Jiang, X., Zhang, L., Chen, X., Wen, Y., Gu, Z., ... & Zheng, M. (2024). The emergence of machine learning force fields in drug design. Medicinal Research Reviews, 44(3), 1147-1182.
6. Chu, W. T., Yan, Z., Chu, X., Zheng, X., Liu, Z., Xu, L., ... & Wang, J. (2021). Physics of biomolecular recognition and conformational dynamics. Reports on Progress in Physics, 84(12), 126601.
7. Fan, F., Xi, B., Meng, X., Wang, H., Zhang, B., Xu, Q., ... & Huang, B. (2026). Assessing conformation validity and rationality of deep learning-generated 3D molecules. Nature Communications.
8. Huang, B., von Rudorff, G. F., & von Lilienfeld, O. A. (2023). The central role of density functional theory in the AI age. science, 381(6654), 170-175.
9. Keith, J. A., Vassilev-Galindo, V., Cheng, B., Chmiela, S., Gastegger, M., Muller, K. R., & Tkatchenko, A. (2021). Combining machine learning and computational chemistry for predictive insights into chemical systems. Chemical reviews, 121(16), 9816-9872.
10. Kim, J., Cho, H., Jeon, H., Jung, J., & Han, S. (2026). Reactive Machine Learning Interatomic Potentials for Chemistry and Materials Science. Chemical Reviews, 126(8), 4467-4510.
11. Kumar, M., Tripathi, M. K., & Kaur, P. (2024). Molecular dynamics and its significance in drug discovery. In Structure-based drug design (pp. 149-175). Cham: Springer International Publishing.
12. Li, J., Gu, R., Du, S., & Xu, L. (2025). 3d electron cloud descriptors for enhanced QSAR modeling of anti-colorectal cancer compounds. Journal of Computer-Aided Molecular Design, 39(2), 98.
13. Li, Z., Jiang, M., Wang, S., & Zhang, S. (2022). Deep learning methods for molecular representation and property prediction. Drug Discovery Today, 27(12), 103373.
14. Potapov, D., Rogovoi, S., Khrabrov, K., Ushenin, K., Korovin, A., Ber, A., ... & Tsypin, A. (2026). A conformational benchmark for optical property prediction with solvent-aware graph neural networks. Communications Chemistry, 9(1), 136.
15. Qiao, Z., Christensen, A. S., Welborn, M., Manby, F. R., Anandkumar, A., & Miller III, T. F. (2022). Informing geometric deep learning with electronic interactions to accelerate quantum chemistry. Proceedings of the National Academy of Sciences, 119(31), e2205221119.
16. Rodriguez, A., Smith, J. S., & Mendoza-Cortes, J. L. (2025). Does Hessian data improve the performance of machine learning potentials?. Journal of Chemical Theory and Computation, 21(14), 6698-6710.
17. Suthar, R., T, A., & Karak, S. (2023). Machine-learning-guided prediction of photovoltaic performance of non-fullerene organic solar cells using novel molecular and structural descriptors. Journal of Materials Chemistry A, 11(41), 22248-22258.

18. Thiede, L., Aldossary, A., Burger, A., Campos-Gonzalez-Angulo, J. A., Wang, N., Zook, A., ... & Aspuru-Guzik, A. (2026). Coupled Cluster con M\= oLe: Molecular Orbital Learning for Neural Wavefunctions. arXiv preprint arXiv:2602.20232.

19. Thölke, P., & De Fabritiis, G. (2022). Torchmd-net: equivariant transformers for neural network based molecular potentials. arXiv preprint arXiv:2202.02541.

20. Tretiakov, S., Nigam, A., & Pollice, R. (2025). Studying noncovalent interactions in molecular systems with machine learning. Chemical reviews, 125(12), 5776-5829.

21. Wang, Y., Li, Z., & Barati Farimani, A. (2023). Graph neural networks for molecules. In Machine learning in molecular sciences (pp. 21-66). Cham: Springer International Publishing.

22. Zhao, G., Cheng, Z., Xu, K., Wan, T., Feng, M., Lu, S., ... & Gao, Z. (2026). Fast and Flexible 3D Molecular Design Framework for Novel Organic Optoelectronic Materials. JACS Au, 6(6), 3138-3152.