JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Article

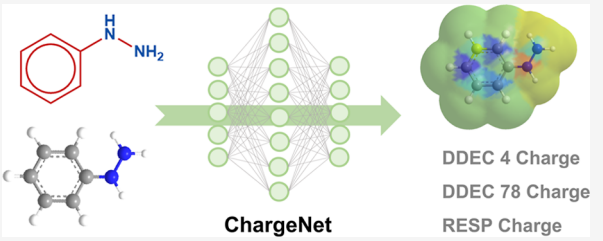
ChargeNet: E(3) Equivariant Graph Attention Network for Atomic Charge Prediction

Qiaolin Gou,‖ Qun Su,‖ Jike Wang,‖ Hui Zhang, Huiyong Sun, Xujun Zhang, Linlong Jiang, Meijing Fang, Yu Kang, Huanxiang Liu,* Tingjun Hou,* and Chang-Yu Hsieh*

Cite This: J. Chem. Inf. Model. 2025, 65, 10364−10374 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Atomic charge is a fundamental quantum chemical property essential for advancing drug design and discovery. Although quantum mechanics (QM) methods offer the highest level of accuracy, their computational demands scale quadratically with the number of atoms, limiting their practicality for large-scale applications. In light of this, empirical and semiempirical methods have been introduced to improve computational efficiency, albeit often at the expense of accuracy. The advent of artificial intelligence has witnessed a growing application of machine learning (ML) techniques to accelerate atomic charge predictions. However, existing ML models often suffer from low accuracy and limited generalization capabilities. To address these challenges, we introduce an advanced equivariant graph attention neural network specifically engineered to model long-range atomic electrostatic interactions with high precision. This model introduces a sophisticated global graph attention mechanism, enabling it to capture charge contributions across multiple scales. By utilizing a combination of structural symmetry-preserving transformations and multiscale attention, our approach not only preserves the inherent symmetries of molecular structures but also substantially improves the model's accuracy, generalization, and robustness in complex scenarios. Our empirical analyses demonstrate that, compared to leading baseline models, the proposed model improves charge prediction accuracy by over 40% on average across various charge-calculation schemes. Remarkably, the model achieves superior performance on the external RESP (restrained electrostatic potential) test data sets, with a 54.6% improvement over the baseline. Additionally, we evaluated our charge model under the setting of virtual screening, where it outperforms both the OPLS3 charges and baseline deep learning models across all evaluation metrics, highlighting its extensive potential for scientific discovery.



INTRODUCTION

Atomic charge is a fundamental and essential property of molecules, describing electrostatic potential and intermolecular electrostatic interactions. For both drug and material molecules, the distribution of charges directly influences key physicochemical properties such as chemical reactivity, electronic transport properties, and band structure. In computational chemistry, atomic charges are widely employed in molecular dynamics simulations, molecular docking, and HOMO−LUMO gap calculations,[1−4] to name a few prominent applications. Accurately calculating atomic charges typically requires high-precision quantum mechanics (QM) methods, such as restrained electrostatic potential (RESP) fitting.[5] However, the computational complexity of these methods often scales exponentially with the number of atoms,[6] making them impractical for many large-scale systems of real-world interests. To improve the computational efficiency, some semiempirical methods have been developed. For example, AM1-BCC charges[7] are obtained using the Austin Model 1 (AM1) method to calculate the molecular orbitals and molecular charge distribution, followed by corrections using empirical parameters in bond charge correction (BCC).[8,9] However, its accuracy heavily relies on correction parameters, limiting its applicability to complex or uncommon molecular structures. Similarly, the OPLS3 charge calculation adopts a force-field-based model,[10,11] achieving moderate improvements in computational efficiency. Clearly, while these methods provide speed advantages over QM methods, they still partially rely on QM calculations and remain computationally intensive. On the other hand, non-QM methods, such as Gasteiger charges,[12] provide only qualitative estimates with limited accuracy. Balancing computational efficiency and accuracy remains a perennial challenge, significantly hindering the practicality of atomic charge calculations.

The recent trend in machine learning (ML) applications for molecular property prediction has driven the development of innovative data-driven approaches,[13−15] presenting promising

Received: March 19, 2025
Revised: June 8, 2025
Accepted: June 9, 2025
Published: June 18, 2025

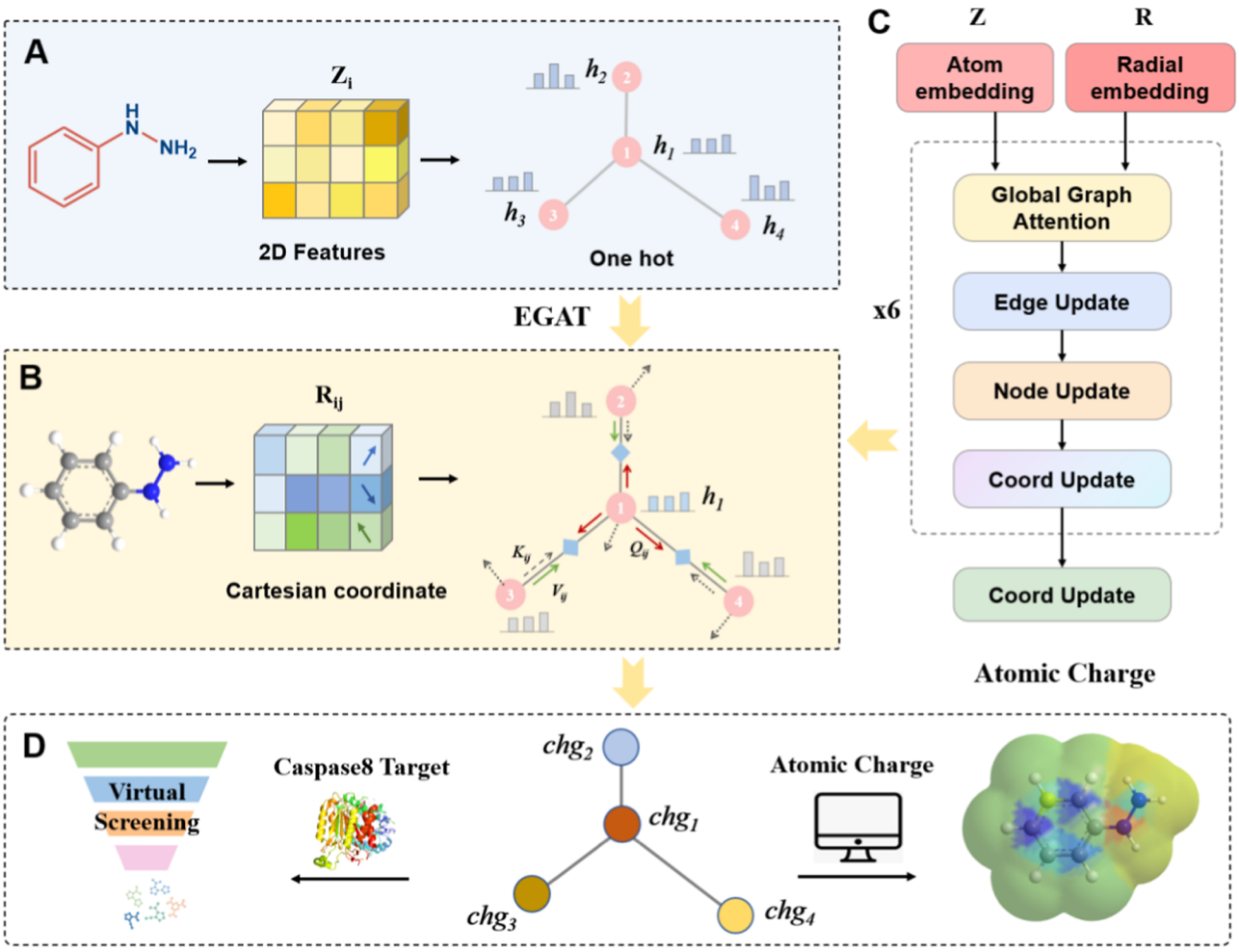


Figure 1. Overview of the ChargeNet framework. (A) 2D atomic descriptors and (B) 3D geometric features are extracted and embedded as node (Z) and radial (R) features, respectively. (C) These features are processed through an Equivariant Graph Attention Network (EGAT), which performs stacked edge, node, and coordinate updates. A fully connected molecular graph enables the model to capture both bonded and nonbonded interactions. (D) The predicted atomic charges serve as input for downstream tasks such as virtual screening.

solutions for efficiently and accurately estimating atomic charges. The first generation of charge prediction models predominantly relied on traditional ML algorithms. In 2013, Rai et al.[16] developed a partial charge distribution model based on symmetry functions using a random forest (RF) algorithm, achieving a mean unsigned error (MUE) of 0.03 e on a test set of 5000 molecules. Similarly, in 2018, Bleiziffer et al.[17] collected a data set of 130,000 drug-like molecules and calculated DDEC charges to train an RF model, obtaining R^2 values of 0.983 and 0.997 on two external test sets. Since then, interests in atomic charge prediction has steadily increased, with numerous studies exploring similar methodologies.[18−20] In 2020, Wang et al.[18] proposed a novel molecular descriptor, Atom-Path-Descriptor (APD), which characterizes the chemical environment around atoms through atomic pair paths. Based on APD, they trained the RF and XGBoost models for atomic charge prediction. In 2021, Kancharlapalli et al.[19] selected a subset from the computation-ready experimental metal−organic framework (CoREMOF-2019) database and employed density functional theory (DFT) to construct a data set of partial atomic charges for MOFs, subsequently training an RF model. While these works have provided new avenues for atomic charge estimation, significant room for improvement remains. A major limitation of traditional ML algorithms lies in their reliance on predefined molecular descriptors or fingerprints as input features. These descriptors are often manually crafted and may fail to fully capture the intricate structural and electronic properties of molecules. This constraint prevents traditional models from effectively representing both fine-grained and global graph-level molecular information, which can limit their predictive accuracy for complex tasks.[21]

This limitation has motivated the development of the second generation of atomic charge prediction models based on deep learning, leveraging the inherent advantages of graph neural networks (GNNs) in processing molecular graphs and effectively capturing molecular structures.[22−24] GNNs are particularly well-suited to address the limitations of traditional methods as they can model complex relationships and topologies within molecular graphs. In these models, nodes represent atoms, while edges denote the chemical bonds. Each central atom aggregates information from its neighboring atoms, enabling automatic learning of its chemical environment.[25,26] To advance atomic charge prediction, Wang et al. (2021) developed DeepAtomicCharge,[27] a GNN-based model designed to predict atomic charges of small molecules under varying dielectric constants. This model aimed to offer a fast and accurate approach for determining partial charges,

https://doi.org/10.1021/acs.jcim.5c00602
J. Chem. Inf. Model. 2025, 65, 10364−10374

facilitating molecular simulation and reactivity analysis. Building on this, Jiang et al. (2022) introduced SuperAtomicCharge,[28] an enhanced model leveraging transfer learning to better capture molecular structure and improve accuracy across diverse molecules. After training on molecules with varying dielectric constants, SuperAtomicCharge demonstrated high predictive accuracy. However, atomic charge prediction shares characteristics with other tasks like point cloud modeling, 3D molecular structure prediction, and N-body simulations. All these tasks require careful handling of 3D translational and rotational symmetries.[29,30] Classical GNN models inherently preserve translational invariance but struggle to cover the full set of symmetries, particularly under 3D rotations and reflections, limiting their ability to distinguish isomorphic from nonisomorphic structures and surpass the 1-WL test.[31,32] This limits their ability to fully capture the chemical environment of atomic charges, highlighting the need for further advancements to enhance model generalization and accuracy in predicting atomic charge distributions across diverse molecular structures.[33,34]

To address the aforementioned challenges, we propose ChargeNet, an equivariant graph atomic charge prediction framework that deeply integrates the principles of equivariance. ChargeNet offers several key advantages. (1) Symmetry preservation and robustness. ChargeNet ensures that the symmetry and transformation properties of the input data are preserved. When handling transformations such as rotation, translation, and scaling, the model automatically adjusts while retaining the corresponding transformations in its output. This property enhances the robustness and generalization capabilities of the framework. (2) Efficient global information capture. Classical GNNs are limited in their receptive fields, often requiring multiple layers to capture global information, leading to information loss and increased computational cost. To address this, ChargeNet equips a GNN with the attention mechanism,[35] enabling the model to adaptively learn relationships and connection weights between nodes, effectively combining local and global information. Furthermore, using multiple attention heads to process input data in parallel enables the model to learn different relationships and feature representations, enhancing its ability to perceive information at different scales and abstraction levels. This adaptive attention mechanism allows ChargeNet to accurately focus on and utilize useful information, extracting key features and information, thus improving its expressiveness and generalization. (3) Modeling long-range interactions. Representing covalent bonds as edges may not be optimal for atomic charge prediction, as electronic interactions are long-range interactions involving all pairs of atoms, not just those connected by covalent bonds.[36,37] In our approach, we employ a fully connected adjacency matrix to represent atomic interactions, enabling the model to directly perceive global information. We evaluated ChargeNet on three types of atomic charges, including DDEC4 (ε = 4), DDEC78 (ε = 78), and RESP charges, and demonstrated its superiority over the state-of-the-art GNN-based predictors and machine learning-based predictors, achieving improvements exceeding 40% on average across all charge types. Furthermore, the model demonstrates outstanding performance on external RESP test data sets, achieving a 54.6% improvement over the baseline models. Additionally, ChargeNet excelled in structure-based virtual screening applications, outperforming commonly used atomic charge methods. This highlights its exceptional scoring and screening capabilities, making it a powerful tool for computational chemistry and drug discovery.

RESULTS AND DISCUSSION

The overall framework of ChargeNet is illustrated in Figure 1. The model begins by embedding atomic features[38] (e.g., nuclear charge, formal charge, and the number of radical electrons) and geometric information (e.g., interatomic distances) into learnable node (Z) and radial (R) representations. These embeddings are then processed by an Equivariant Graph Attention Network (EGAT), which iteratively updates node and edge features through attention-based message passing. Each update block employs directional attention via query-key-value mechanisms to capture spatial anisotropy. While the graph edges are typically defined based on a cutoff radius in interatomic distances, a fully connected adjacency matrix can also be used. This design ensures that both covalent and noncovalent interactions are captured. Given that partial atomic charges are sensitive to long-range, nonbonded interactions, a fully connected graph is critical for accurate modeling. Furthermore, the use of multihead attention facilitates the extraction of hierarchical and multiscale features, enhancing the overall predictive accuracy. The final atomic charges predicted by ChargeNet can be directly applied to downstream tasks such as virtual screening.[39]

A core innovation of ChargeNet lies in its careful handling of long-range electronic effects and geometric symmetries. Long-range electronic interactions are inherently nonlocal and cannot be captured effectively by models limited to local neighborhoods. By leveraging a fully connected graph structure and attention-based message passing, ChargeNet explicitly models these interactions across molecular space. Simultaneously, the model is designed to be $E(3)$-equivariant, meaning that its predictions respond predictably and consistently to global transformations such as rotations and translations of the input geometry. This is achieved by allowing both node and edge features to be dynamically updated during message passing in a symmetry-aware manner. As a result, ChargeNet can effectively distinguish conformers with different 3D geometries while preserving the transformation consistency between inputs and outputs. A formal proof of this equivariance property is provided in the Methods section.

Impact of Graph Connectivity on Atomic Charge Prediction Tasks. In the SuperAtomicCharge model, the significance of 3D information for graph representation and properties has been well demonstrated. The spatial positions and relative distances between atoms effectively convey the importance of atomic interactions. Within graph neural networks, information propagation primarily relies on edges, so edge features generally need to be fully considered, such as torsion angles. However, in atomic charge prediction, these additional edge-specific features are generally not essential. The task's intrinsic simplicity allows models to capture sufficient structural features using only the most fundamental inputs: atomic coordinates and node features. Since atomic charges represent a relatively straightforward atomic property, the model can effectively learn the necessary interactions and relationships without relying on complex edge-based features.

Nevertheless, it is essential to account for the fact that electrostatic interactions are a type of long-range force with non-negligible influence over considerable distances. The message-passing process in graph-based models heavily relies on the edge connection information we provide. Should we

https://doi.org/10.1021/acs.jcim.5c00602
J. Chem. Inf. Model. 2025, 65, 10364−10374

directly follow the common practice of using the topological graph identified by RDKit, which generally connects atoms based on covalent bonds? To investigate this, we conducted experiments on our proposed EGAT (equivariant graph attention transformer) module using two distinct graph representations: a fully connected molecular graph and a molecular graph based on covalent bond adjacency matrices. These experiments were designed to evaluate how the choice of edge connectivity impacts the model's ability to capture electrostatic interactions and predict atomic charges effectively.

As shown in Table 1, employing a fully connected adjacency matrix for the atomic charge prediction task significantly outperforms the covalent bond-based adjacency matrix across all three data sets, with improvements ranging from 20% to 30%. These results highlight that fully connected approach provides the model with better representation capabilities. Therefore, it also suggests that using covalent bonds as the adjacency matrix for message passing results in some degree of information loss. In reality, the bond length of covalent bonds typically ranges from 1 to 3 Å, but the range of electrostatic interactions far exceeds this distance, which undoubtedly causes some discrepancies. To address this, an alternative approach involves setting a distance threshold, where pairs of atoms within this threshold but not covalently bonded are treated as noncovalent connections. By using one-hot encoding to distinguish between covalent and noncovalent bonds, the model can more effectively differentiate between covalent interactions and noncovalent interactions. This approach enhances the interpretability of the model and provides a more accurate representation of molecular interactions.

Table 1. Impact of Different Versions of Graph Connectivity on Prediction Results across the RESP, DDEC4, and DDEC78 Datasets (Randomly Split 80:10:10 into Training, Validation, and Test Sets)

type of charge	feature	RMSE/e		
		train	valid	test
DDEC4	full connection	0.00476	0.00433	0.00457
	bond edge	0.00678	0.00712	0.00694
DDEC78	full connection	0.00584	0.00585	0.00597
	bond edge	0.00661	0.00673	0.00634
RESP	full connection	0.03273	0.03556	0.03574
	bond edge	0.03741	0.03912	0.03897

Impact of Equivariant Graph Neural Networks on Atomic Charge Representation. In atomic charge prediction tasks, our analysis of the data set indicates that over 90% of the molecules contain ring structures, with a notable presence of macrocycles and polycyclic systems. This inherent structural complexity often leads to high molecular symmetry, posing a challenge for effective feature extraction. The design of equivariant graph neural networks (GNNs) allows them to be sensitive to these symmetry transformations, which helps in more effectively capturing the characteristics of molecular structures. By learning the intricate relationships among nodes, these networks can adapt to the complex interactions between atoms. In this regard, ChargeNet demonstrates a significant advantage over SuperAtomicCharge. The introduction of multihead attention mechanism equips equivariant GNNs with a broader global perspective, which is particularly crucial for predicting atomic charges. Since atomic charges are strongly influenced by surrounding atoms (both near and far), so the network's ability to comprehensively capture both local and global interactions between atoms leads to more accurate charge predictions. As shown in Figure S1, we visualized both global and local attention weights, which emphasize key short-range interactions while maintaining sensitivity to long-range dependencies. Equivariant GNNs are naturally suited for handling graph structures with geometric information. In atomic charge prediction, 3D coordinates provide spatial relationships between atoms, which are essential for understanding atomic interactions and charge distributions. Equivariant GNNs leverage this geometric information effectively, making them highly advantageous in this domain.

As illustrated in Table 2, the prediction performance on DDEC4 and DDEC78 charges are quite similar, whereas the performance on RESP charges is noticeably lower. This discrepancy arises because RESP charges rely on fitting to the molecule's electrostatic potential (ESP) and introduce constraints to balance atomic charges, thus making RESP well-suited for capturing subtle variations in molecular charge distribution and reliable in describing molecular electrostatic interactions. In contrast, the DDEC methods emphasize molecular polarization effects and electronic density distributions, which may sacrifice some ESP fitting accuracy. In this sense, RESP charges are more precise and require finer-grained information, making them more challenging to fit accurately. Furthermore, we applied a post hoc correction to the predicted atomic charges to ensure total charge conservation. As shown in Figure S2, the R^2 values before and after correction both remain at 0.9996 when compared with the reference charges, indicating that this correction preserves the model's high predictive accuracy while enforcing chemically consistent charge distributions.

Table 2. Impact of Different Models on Prediction Results Across the RESP, DDEC4, and DDEC78 Datasets

type of charge	model	RMSE/e		
		train	valid	test
DDEC4	AP-RF	0.01433	0.01671	0.01659
	ADP-RF	0.00688	0.01038	0.01043
	DeepAtomicCharge[27]	0.01127	0.01142	0.01148
	SuperAtomicCharge[28]	0.00912	0.00935	0.00942
	ChargeNet	**0.00476**	**0.00433**	**0.00457**
DDEC78	AP-RF	0.01640	0.01914	0.01886
	ADP-RF	0.00819	0.01223	0.01208
	DeepAtomicCharge[27]	0.01382	0.01413	0.01404
	SuperAtomicCharge[28]	0.01134	0.01197	0.01197
	ChargeNet	**0.00584**	**0.00585**	**0.00597**
RESP	AP-RF	0.05541	0.06349	0.06353
	ADP-RF	0.03500	0.05509	0.05472
	DeepAtomicCharge[27]	0.05041	0.05481	0.05469
	SuperAtomicCharge[28]	0.04286	0.04829	0.04805
	ChargeNet	**0.03271**	**0.03551**	**0.03570**

Furthermore, as shown in Table 2, the ChargeNet model demonstrates a significant advantage in charge prediction across different types. In the prediction of DDEC4 charges, ChargeNet outperforms the second-best model by 51.4%; for DDEC78 charges, the improvement is 50.1%; and for RESP charges, it is 25.7%. These results indicate that ChargeNet

https://doi.org/10.1021/acs.jcim.5c00602
J. Chem. Inf. Model. 2025, 65, 10364−10374

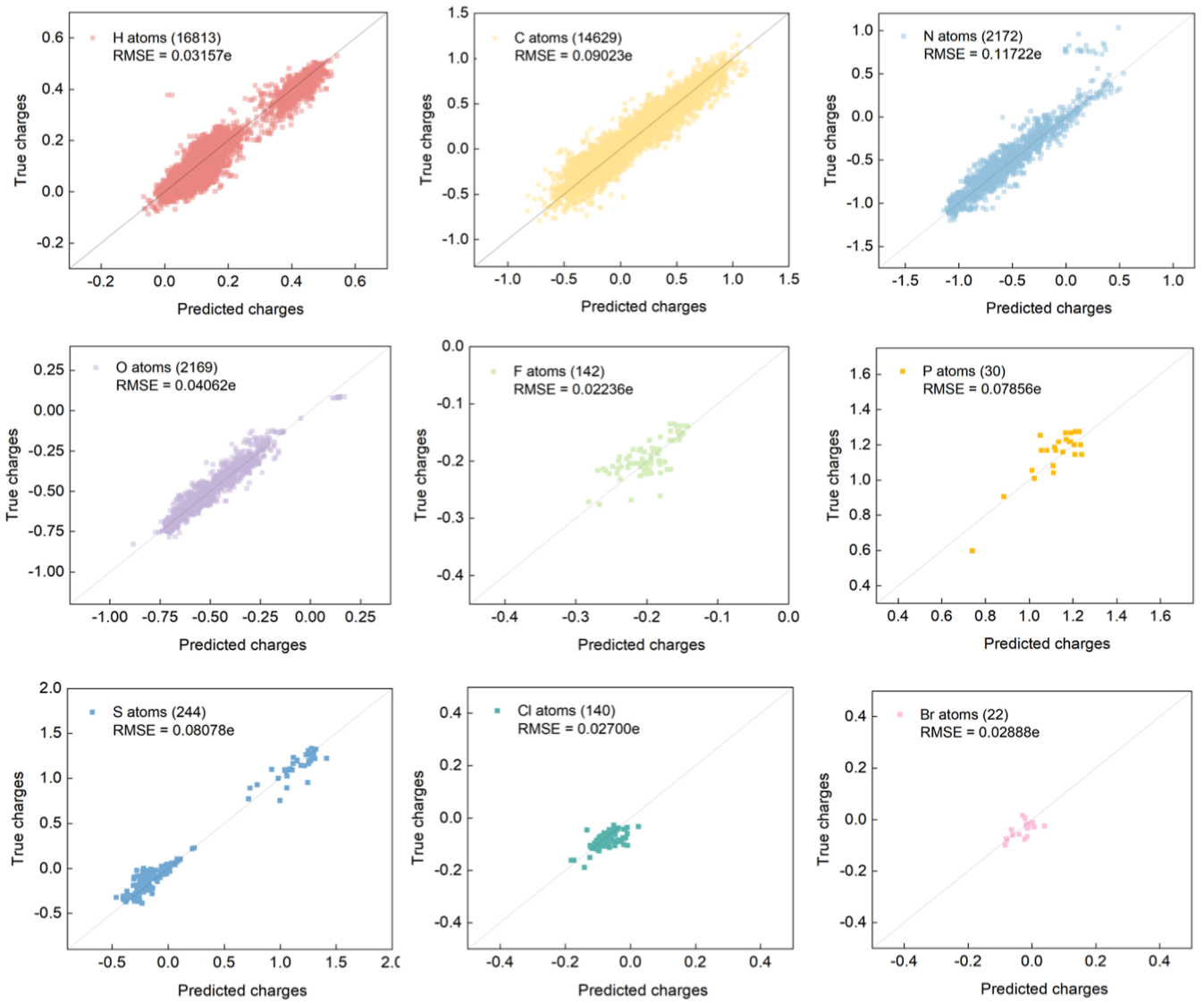


Figure 2. Distribution of predicted and calculated RESP charges for different elements in the external test set of 1000 molecules.

exhibits high accuracy and strong generalization capability in charge prediction tasks.

Validation Testing on External Test Set. For models predicting basic properties, generalization ability is a crucial metric. The data we collect for training represents only a tiny fraction of the vast molecular space. Therefore, it is desirable for the model to maintain predictive accuracy while exhibiting strong generalization ability for property prediction on unseen data. This is a key advantage of many traditional tools, which, despite not being highly accurate, maintain a stable performance across different parts of the chemical space. To evaluate this, we randomly selected 1000 molecules with 40−60 atoms from the ChEMBL database[40] and compared their chemical space distribution with that of the training set, as shown in Figure S3. We then calculated their RESP charges using the Gaussian program to construct an external test set. We compared our ChargeNet model with both the SuperAtomicCharge and DeepAtomicCharge models on the external data set. ChargeNet achieved an RMSE of 0.0608 e and an R^2 of 0.947, significantly outperforming the SuperAtomicCharge model (RMSE: 0.1339 e, R^2: 0.800) and the DeepAtomicCharge model (RMSE: 0.1691 e, R^2: 0.782). This represents a 54.6% improvement in accuracy over SuperAtomicCharge and a 64.0% improvement over DeepAtomicCharge, highlighting the superior performance of ChargeNet. In addition to its superior accuracy, ChargeNet demonstrated strong generalization ability, as evidenced by the consistent performance observed when transitioning from the RESP training and testing sets to this external data set, with no significant drop in predictive accuracy. Moreover, compared to RESP—which relies on computationally intensive quantum mechanical (QM) calculations—ChargeNet offers orders-of-magnitude faster predictions. Together, these results highlight ChargeNet as a robust and efficient model for charge prediction across diverse data sets.

As shown in Figure 2, we evaluated the charge prediction accuracy of the model for each type of element within the molecules, focusing on elements commonly involved in drug design, such as hydrogen (H), carbon (C), nitrogen (N), oxygen (O), fluorine (F), and phosphorus (P). The model demonstrates high accuracy across these elements, with particularly strong performance for hydrogen atoms, which is critical for accurately capturing hydrogen bonding interactions—key determinants of molecular stability and binding

https://doi.org/10.1021/acs.jcim.5c00602
J. Chem. Inf. Model. 2025, 65, 10364−10374

specificity. For electronegative heteroatoms like O and N, the model achieves precise charge predictions essential for modeling electrostatic interactions and polar binding sites, which are directly linked to solubility, reactivity, and binding affinity in protein−ligand complexes. Furthermore, the model performs well on biologically relevant but less common atoms, such as P and S, which are crucial for understanding interactions in enzyme activity, phosphorylation, and disulfide bonding.[41] This robust prediction capability ensures accurate calculations of hydrogen bonds, van der Waals forces, hydrophobic interactions, and electrostatic forces, highlighting the model's potential as a reliable tool for guiding molecular design and optimizing interactions critical to drug discovery.

ChargeNet Applied in Protein−Ligand Virtual Screening. Atomic charges are critical in drug design as they determine key intermolecular interactions such as electrostatic interaction and hydrogen bonding. Accurate atomic charges prediction models can significantly influence the enrichment efficiency of active compounds in virtual screening, ultimately shaping the efficiency of drug discovery. To comprehensively validate the performance of ChargeNet, we designed a scenario to evaluate the model's screening power in virtual screening. Caspase-8, a cysteine protease involved in the death receptor-mediated apoptosis pathway, is strongly linked to the progression of various inflammatory conditions, such as immune-related disorders, neurodegenerative diseases, and cancer, positioning it as a highly promising therapeutic target.[42,43] Furthermore, the active site of Caspase-8 contains a wealth of hydrogen bond donor and acceptor residues. In the crystal structures of known inhibitor complexes, a complex hydrogen bond network and significant ionic interactions have been observed (Figure 3). This

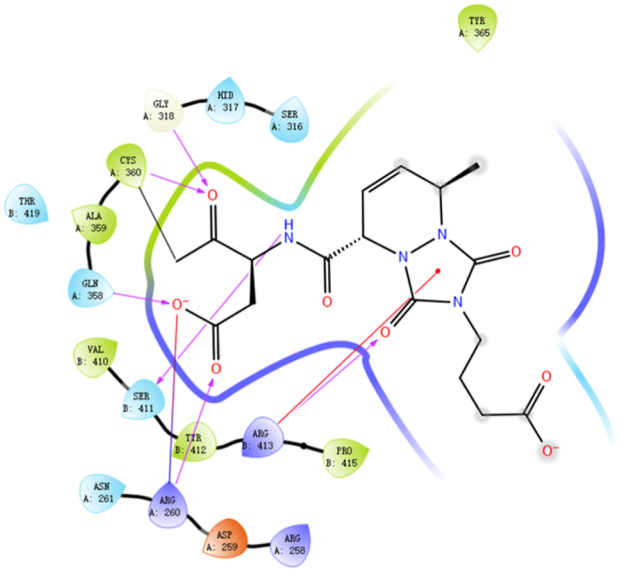


Figure 3. Protein−Ligand interactions diagram for Caspase-8 with inhibitor (PDB code: 3KJQ).

structural feature suggests that Caspase-8 exhibits a high sensitivity to the charge distribution of its ligands. In light of the considerations mentioned above, we selected Caspase 8 (PDB code: 3H11)[44] as the drug target, with the data set sourced from Jiang et al.[28] This data set comprises 217 actives with IC_{50} < 10 μM, retrieved from BindingDB,[45] and 7000 decoys randomly selected from the ChemBridge database, achieving an active-to-inactive ratio of approximately 1:30. Protein and ligand preprocessing was performed using Schrödinger's Protein Preparation Wizard and LigPrep module with default parameters. The receptor grid was centered on the centroid of the cocrystallized ligand, with dimensions of 20 × 20 × 20 Å. All compounds were docked using Glide[46] with standard precision (SP) scoring mode, and the effects of three types of atomic charges on virtual screening performance were evaluated: the default OPLS3e charges, along with the RESP and DDEC4 charges predicted by ChargeNet, with normalization applied to maintain charge conservation. The results were evaluated using three metrics: the p-value of the Mann−Whitney U test, the enrichment factor, and the area under the curve (AUC) value of the receiver operating characteristic (ROC) curve.

As shown in Figure 4, based on the three metrics (P-value, AUC, and enrichment factor), the predicted DDEC4 charges exhibited the best virtual screening performance. For caspase-8, the p-value for DDEC4 charges (Figure 4C) was 2.50 × 10^{-80}, and the AUC (Figure 4D) was 0.878, significantly outperforming the results obtained with the original OPLS3e charges (Figure 4A,D, p-value = 2.22 × 10^{-50}, AUC = 0.798). Compared to random selection, the probability of identifying actives was nearly 16 times higher with DDEC4 charges, while with OPLS3e charges, this increase was only 4 times. Furthermore, although RESP charges (Figure 4B) slightly underperformed OPLS3e charges in their ability to distinguish active from inactive peaks (with a p-value of 5.84 × 10^{-45} and an AUC of 0.780), the early portion of the ROC curve was higher than that of OPLS3e, indicating superior early enrichment power. The enrichment factor shown in Figure 4E further confirmed that the early enrichment rate for RESP charges was nearly twice that of OPLS3e charges. QM methods for RESP charges calculation typically require a day on a single-core CPU, while ChargeNet accelerates this by several orders of magnitude. These results highlight the significant potential and value of ChargeNet in large-scale virtual screening, particularly in enhancing early enrichment efficiency and improving screening accuracy.

CONCLUSION

In this study, we propose a novel framework, ChargeNet, designed to predict atomic charges in drug molecules—a critical parameter involved in computational and analytical in drug discovery. Our model is meticulously crafted to incorporate long-range electrostatic interactions and molecular geometric symmetry, leveraging a graph attention mechanism to capture interaction information effectively. When evaluated on benchmark data sets, ChargeNet demonstrated over 40% on average improvement in charge prediction accuracy across various charge types compared to state-of-the-art baselines. Moreover, on the RESP external test set, the model surpassed the top-performing existing model by 54.6%. This marks a significant breakthrough in atomic charge prediction. To further assess its applicability in real-world scenarios, we conducted a virtual screening study targeting Caspase 8, where the atomic charge predictions from ChargeNet substantially enhanced screening performance, indicating its potential for large-scale batch processing. ChargeNet strikes an optimal balance between prediction accuracy and computational efficiency. For broader applications, such as in biomaterials and synthetic chemistry, we plan to expand the data set and refine the model to address domain-specific requirements.

https://doi.org/10.1021/acs.jcim.5c00602
J. Chem. Inf. Model. 2025, 65, 10364−10374

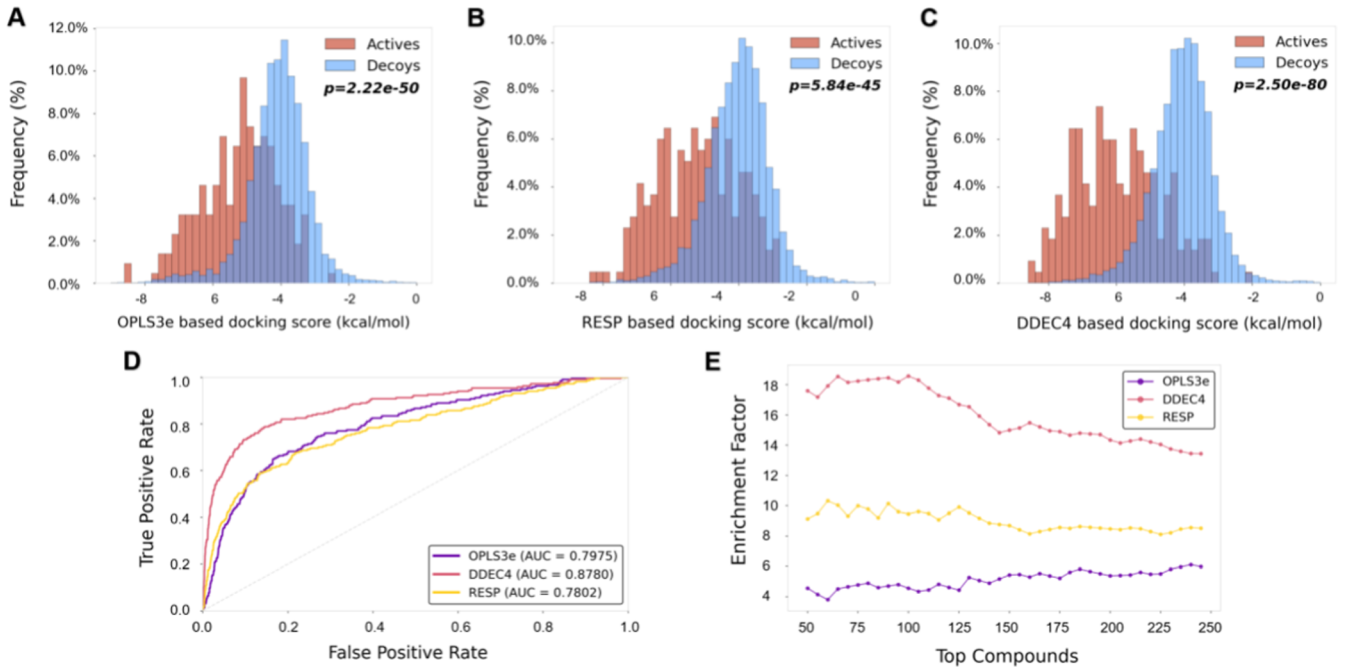


Figure 4. Performance of OPLS3e, RESP, and DDEC4 charges on virtual screening. The screening power (A−C), AUC value under ROC curve (D), and the enrichment factor (E) are used as the metrics for comparison.

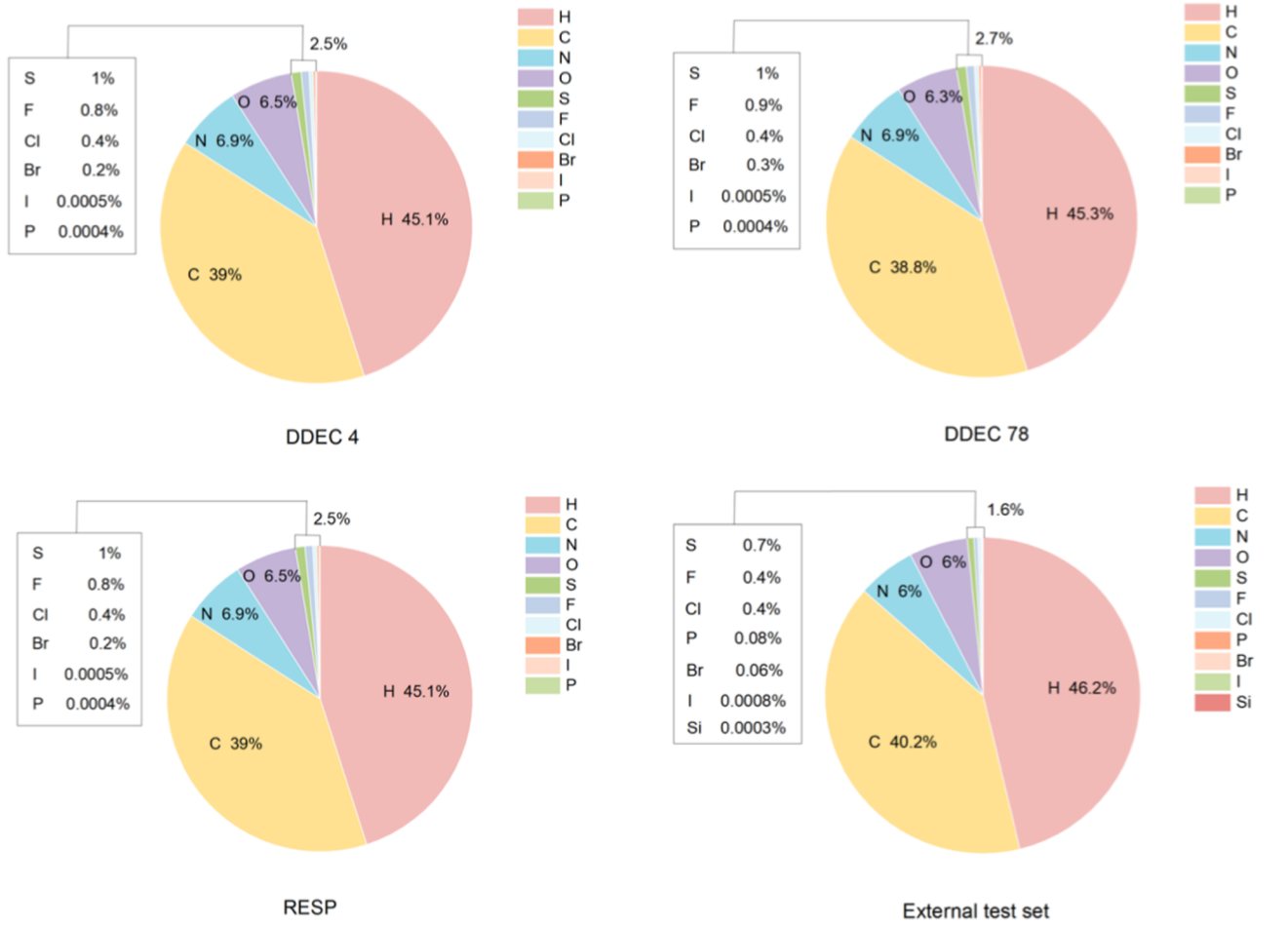


Figure 5. Elemental composition distributions of the molecules in the DDEC4, DDEC78, RESP data sets, and external test set.

https://doi.org/10.1021/acs.jcim.5c00602
J. Chem. Inf. Model. 2025, 65, 10364−10374

METHODS

Data Collection. This study utilized approximately 130,000 organic small molecules to train ChargeNet models, encompassing three types of atomic charge: DDEC4 (ε = 4), DDEC78 (ε = 78), and RESP. These molecules were sourced from the data set curated by Bleiziffer et al.[17] This work directly utilized the high-level quantum mechanical calculations conducted by Bleiziffer et al. for DDEC4 (ε = 4) and DDEC78 (ε = 78) atomic charges. Specifically, ε = 4 was employed to simulate atomic charges in protein environments, while ε = 78 was used to represent those in solvent environments. Since RESP atomic charges could not be directly obtained for the organic small molecules, we first optimized all molecules in the data set using the semiempirical AM1 method.[7] Subsequently, the RESP atomic charges were calculated by deriving each molecule's electrostatic potential using the Hartree−Fock self-consistent field (SCF) method at the 6−31G* level of theory[47] implemented in the Gaussian16[48] program. Additionally, using the same computational setup, we computed RESP charges for an external test set comprising 1000 randomly selected molecules extracted from the ChEMBL database.[40] As shown in Figure 5, the elemental compositions across the DDEC4, DDEC78, RESP, and external test set are largely consistent, with hydrogen (H), carbon (C), nitrogen (N), and oxygen (O) being the predominant elements, while other elements such as sulfur (S), fluorine (F), chlorine (Cl), bromine (Br), iodine(I), and phosphorus (P) appear in much smaller proportions. Notably, silicon (Si) is present only in the external test set. These calculations ensure consistency in charge derivation and provide a robust data set for model evaluation.

Data Preprocessing. First, we calculated the features of each atom and bond using RDKit. Considering atoms as nodes, to maintain the same length of molecules (containing the same number of atoms), pseudoatoms with zero features were added to each molecule. Each atom (including pseudoatoms) was represented by a 35-bit vector. Therefore, each molecule could ultimately be represented by an $n \times 35$ matrix (where n is the number of atoms in the molecule). The value of n depends on the maximum number of atoms per batch in the data set. Meanwhile, atomic coordinates were input separately into the model as atomic 3D information for coordinate updates, maintaining network equivariance.

Graph Representation and Feature Extraction. Molecular descriptors play a crucial role in predicting molecular properties. By selecting appropriate descriptors, considering diversity and complementarity, performing dimensionality reduction, data normalization, feature engineering, and data preprocessing, the expressiveness and predictive performance of the descriptors can be significantly enhanced, thereby achieving more accurate molecular property predictions. The design of this module was based on the studies by Jiang et al.[28] and Wang et al.[27] In this section, each molecule was abstracted as an undirected graph $G = (V, E)$, where V represents atoms and E represents edges, with $(i, j) \in E$ indicating the existence of a chemical bond between atoms i and j. Since the electrostatic force is a long-range interaction, each atom pair was assigned an edge, forming a fully connected graph. The fully connected graph allows each atom to interact with all other atoms, thereby enhancing the model's ability to capture long-range interactions. Many recent physics-based neural networks also avoid relying solely on covalent bonds to construct molecular graphs. Instead, they commonly apply distance-based cutoffs, typically in the range of 4−6 Å, to incorporate long-range interactions. Our fully connected graph can be viewed as a generalized form of this strategy. For large molecular systems, where computational cost becomes a concern, we plan to integrate distance-based cutoffs in future work to balance physical expressiveness and efficiency. In this work, only node feature descriptors were utilized (as shown in Table 3). These features include common atomic information,

Table 3. Node Feature Descriptors Used in this Study

type	attributes name	descriptions
2D	atom type	numeric encoding of atom types
	num atom	number of atoms
	IsAromatic	whether the molecule is aromatic [0, 1]
	hybridization	hybridization mode
	NumHs	number of hydrogen atoms
	FormalCharge	FormalCharge
	ExplicitValence	ExplicitValence
	ImplicitValence	ImplicitValence
	NumExplicitHs	NumExplicitHs
	NumRadicalElectrons	NumRadicalElectrons
3D	Coordination	Coordination

with the atomic coordinates being the sole 3D-specific feature. Spatial distances are essential for describing atomic geometries, emphasizing the importance of including coordinate information in molecular representations.

Equivariant Graph Neural Networks. Equivariance refers to the property where the output of a mapping transformation changes predictably with the input transformation. If the mapping transformation Ψ satisfies the following equivariance condition for all u belonging to the transformation group G

$$\forall\, u \in G,\ \Psi[T_{u[f]}] = T_o^u[\Psi[f]] \tag{1}$$

For the transformation group G, input feature map f, and group actions T_u and T_o^u. Here, T_u represents the action of transformation u on the input features, while T_o^u represents the action of transformation u on the output features. Additionally, since it is desired that the operations of consecutive transformations u and v on the feature map are equivalent to directly transforming the feature map by the combined transformation uv, it is required that $T_u T_v = T_{uv}$, where uv is the group product of transformations u and v. The same requirement applies to T_o^u. Specifically, in equivariant graph neural networks, their information propagation and update processes are as follows.

$$m_{ij} = \varphi_e(h_i^l, h_j^l, \|x_i^l - x_j^l\|^2, a_{ij}) \tag{2}$$

$$x_i^{l+1} = x_i^l + C\sum_{j\neq i}(x_i^l - x_j^l)\varphi_x(m_{ij}) \tag{3}$$

$$m_i = \sum_{i\in N(i)} m_{ij} \tag{4}$$

$$h_i^{l+1} = \varphi_h(h_i^l, m_i) \tag{5}$$

For a graph $G = (V, E)$ with node v_{ij} and edge e_{ij}, where each node feature is encoded as h_i and the n-dimensional coordinates associated with each graph node are denoted as

https://doi.org/10.1021/acs.jcim.5c00602
J. Chem. Inf. Model. 2025, 65, 10364−10374

x_i as input. It differs from the original graph neural networks (GNNs) mainly in eqs 2 and 3. In eq 2, the squared relative distance between two coordinates ($\|x_i^l - x_j^l\|^2$) is passed as input to the edge information e_{ij}. The node features h_i^l,h_j^l, and edge attribute a_{ij} are also provided as input to edge information as in the case of GNNs.

In eq 3, the position of each particle xi is updated by the weighted sum of all relative differences ($x_i^l - x_j^l$). The weight of this sum is provided by the output of a function, which accepts the edge embedding mij from the previous edge operation as input and outputs a scalar value. This is also the reason why it can maintain equivariance.

Attention Mechanism. The standard attention mechanism can be considered as consisting of three components: a set of query vectors $Q_i \in R^r$, where $i = 1\cdots m$, a set of key vectors $K_j \in R^p$, where $j = 1\cdots n$, and a set of value vectors $V_j \in R^p$, where $j = 1\cdots n$, where r and p are the dimensions of low-dimensional embeddings. Keys K_j and values V_j are typically interpreted as being associated with the same point j. For a given query Q_i, the attention mechanism can be written as

$$\mathrm{Attn}(Q_i, K_j, V_j) = \sum_{j=1}^{n} \mathrm{Score}_{ij} V_j \tag{6}$$

$$\mathrm{Score}_{ij} = \frac{\exp(Q_i^T K_j)}{\sum_{j'=1}^{n} \exp(Q_i^T K_{j'})} \tag{7}$$

where we use softmax as the nonlinearity applied to the weights. Generally, the number of query vectors does not have to be equal to the number of input points. In the case of self-attention, the query, key, and value vectors are embeddings of input features.

$$Q = h_Q(f) \tag{8}$$

$$K = h_K(f) \tag{9}$$

$$V = h_V(f) \tag{10}$$

where Q, K, V are the output values of each node, and Score_{ij} reflects the similarity between the query Q and the key K. For clarity, we only represent the calculation of single-head attention and omit the bias terms.

ChargeNet. The ChargeNet model takes graph node features and 3D coordinates as input to construct a fully connected molecular graph, integrating equivariant graph neural networks and graph attention mechanisms, fully utilizing global information and symmetry to effectively aggregate features of molecules for predicting atomic charges.

In this model, based on the framework of equivariant graph neural networks (EGNN), a graph attention mechanism is used as the message passing strategy for information aggregation. It mainly consists of the following modules: Embedding layer: Node and edge features are mapped through fully connected layers for representation learning and transformation into continuous hidden layer space. Equivariant graph attention module: The features h_i, where $i = 1\cdots m$, entering the hidden space are transformed to obtain a set of query vectors (query), key vectors (key), and value vectors (value). The attention weights obtained from the calculation of Q and K are multiplied with edge information to obtain feature information for each edge.

$$\mathrm{Score}_{ij}^l = \mathrm{Softmax}\left(\left(\frac{q_i^l \cdot k_j^l}{\sqrt{d_k}}\right) \cdot e_{ij}^l\right) \tag{11}$$

Node feature update module: node value vectors are concatenated with edge information Score_{ij}^l weighted by attention weights to obtain fused features. Linear operations are performed on the fused features through fully connected layers (linear) for weight adjustment and dimension transformation, resulting in updated node features

$$h_i^{l+1} = h_i^l + w_{ij}(\mathrm{Score}_{ij}^l \oplus v_j^l) \tag{12}$$

Thus, node updates are achieved by integrating value vectors and edge information weighted by attention weights, preserving node information while considering associated information with other nodes, resulting in more representative updated node features.

Edge feature update module: Input information h_i^l is processed through an edge operation function to generate edge information Score_{ij}^l, which is fused with current-level information to obtain the representation of the next layer.

$$e_{ij}^{l+1} = e_{ij}^l + w_{ij}(\mathrm{Score}_{ij}^l) \tag{13}$$

This involves capturing relationships between nodes, connectivity, and other edge-related information. The purpose of this stage is to provide important information about edges for subsequent node feature updates, ensuring that the network better understands local relationships and connectivity in the graph structure.

Coordinate update module: The coordinate update module is responsible for ensuring the equivariance of the neural network. In the process of coordinate updating, considering that a_{ij} is an element in the score matrix Score_{ij}^l, to maintain equivariance, we only need to replace a_{ij} in eq 3 with a new term representing edge information. This is a crucial part for maintaining coordinate equivariance. Through this replacement, the equivariance of the network is guaranteed, and the specific coordinate update process can be referred to from eqs 2 and 3. This ensures that the changes in coordinates are consistent with the changes in edge features, thereby maintaining the sensitivity of the network to input data equivariance.

Model Training. We built the model based on the PyTorch[49] framework. The DDEC4 ($\varepsilon = 4$), DDEC78 ($\varepsilon = 78$), and RESP charge data sets are split into a 80:10:10 ratio for training, validation, and validation sets, respectively.

Here, we optimized using gradient descent in the Adam optimizer, with Mean Squared Error (MSE) Loss used as the loss function for the charge prediction task. During training, early stopping,[50] dropout,[51] and regularization were employed to prevent overfitting and reduce training time. The best validation performance was generally achieved around 50 epochs. In addition to a systematic hyperparameter search, the random search was employed to explore optimal combinations due to the large number of hyperparameters. A 10-fold cross-validation strategy was utilized for model construction, leading to the identification of three key hyperparameters, including the optimal number of layers.

https://doi.org/10.1021/acs.jcim.5c00602
J. Chem. Inf. Model. 2025, 65, 10364−10374

■ ASSOCIATED CONTENT

Data Availability Statement

The code used in the study is publicly available from the GitHub repository:https://github.com/cholin01/ChargeNet.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.5c00602.

Overview of the ChargeNet framework. Distribution of predicted and calculated RESP charges for different elements in the external test set of 1000 molecules (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Huanxiang Liu − *Faculty of Applied Science, Macao Polytechnic University, Macao 999078, China;* orcid.org/0000-0002-9284-3667; Email: kimhsieh@zju.edu.cn

Tingjun Hou − *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China;* orcid.org/0000-0001-7227-2580; Email: tingjunhou@zju.edu.cn

Chang-Yu Hsieh − *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China;* orcid.org/0000-0002-6242-4218; Email: hxliu@mpu.edu.mo

Authors

Qiaolin Gou − *Faculty of Applied Science, Macao Polytechnic University, Macao 999078, China*

Qun Su − *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China*

Jike Wang − *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China;* orcid.org/0000-0002-8118-8572

Hui Zhang − *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China*

Huiyong Sun − *Department of Medicinal Chemistry, China Pharmaceutical University, Nanjing, Jiangsu 210009, China;* orcid.org/0000-0002-7107-7481

Xujun Zhang − *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China*

Linlong Jiang − *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China*

Meijing Fang − *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China*

Yu Kang − *College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China;* orcid.org/0000-0002-0999-8802

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jcim.5c00602

Author Contributions

‖Equivalent authors.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work has been supported in part by the National Key R&D Program of China (2024YFA1307501), and the National Natural Science Foundation of China (22303081).

■ REFERENCES

(1) Chen, F.; Ma, T.; Zhang, T.; Zhang, Y.; Huang, H. Atomic-Level Charge Separation Strategies in Semiconductor-Based Photocatalysts. *Adv. Mater.* **2021**, *33* (10), 2005256.

(2) He, X.; Man, V. H.; Yang, W.; Lee, T.-S.; Wang, J. A fast and high-quality charge model for the next generation general AMBER force field. *J. Chem. Phys.* **2020**, *153* (11), 114502.

(3) Zhao, J.; Zhu, Z.-W.; Zhao, D.-X.; Yang, Z.-Z. Atomic charges in molecules defined by molecular real space partition into atomic subspaces. *Phys. Chem. Chem. Phys.* **2023**, *25* (13), 9020−9030.

(4) Alibakhshi, A.; Schäfer, L. V. Theoretical Evaluation of Radii of Atoms in Molecules and their Dependence on Atomic Partial Charge, **2024**; chemrxiv-2024-5qz9b. .

(5) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97* (40), 10269−10280.

(6) Huggins, D. J. Comparing the Performance of Different AMBER Protein Forcefields, Partial Charge Assignments, and Water Models for Absolute Binding Free Energy Calculations. *J. Chem. Theory Comput.* **2022**, *18* (4), 2616−2630.

(7) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23* (16), 1623−1641.

(8) Petris, P. C.; Becherer, P.; Fraaije, J. G. E. M. Alkane/Water Partition Coefficient Calculation Based on the Modified AM1Method and Internal Hydrogen Bonding Sampling Using COSMO-RS. *J. Chem. Inf. Model.* **2021**, *61* (7), 3453−3462.

(9) Orr, A. A.; Sharif, S.; Wang, J.; MacKerell, A. D., Jr. Preserving the Integrity of Empirical Force Fields. *J. Chem. Inf. Model.* **2022**, *62* (16), 3825−3831.

(10) Kurki, M.; Poso, A.; Bartos, P.; Miettinen, M. S. Structure of POPC Lipid Bilayers in OPLS3e Force Field. *J. Chem. Inf. Model.* **2022**, *62* (24), 6462−6474.

(11) Jorgensen, W. L.; Ghahremanpour, M. M.; Saar, A.; Tirado-Rives, J. OPLS/2020 Force Field for Unsaturated Hydrocarbons, Alcohols, and Ethers. *J. Phys. Chem. B* **2024**, *128* (1), 250−262.

(12) Gasteiger, J.; Marsili, M. Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges. *Tetrahedron* **1980**, *36* (22), 3219−3228.

(13) Walters, W. P.; Barzilay, R. Applications of Deep Learning in Molecule Generation and Molecular Property Prediction. *Acc. Chem. Res.* **2021**, *54* (2), 263−270.

(14) Deng, J.; Yang, Z.; Wang, H.; Ojima, I.; Samaras, D.; Wang, F. A systematic study of key elements underlying molecular property prediction. *Nat. Commun.* **2023**, *14* (1), 6395.

(15) Gou, Q.; Liu, J.; Su, H.; Guo, Y.; Chen, J.; Zhao, X.; Pu, X. Exploring an accurate machine learning model to quickly estimate stability of diverse energetic materials. *iScience* **2024**, *27* (4), 109452.

(16) Rai, B. K.; Bakken, G. A. Fast and accurate generation of ab initio quality atomic charges using nonparametric statistical regression. *J. Comput. Chem.* **2013**, *34* (19), 1661−1671.

(17) Bleiziffer, P.; Schaller, K.; Riniker, S. Machine Learning of Partial Charges Derived from High-Quality Quantum-Mechanical Calculations. *J. Chem. Inf. Model.* **2018**, *58* (3), 579−590.

(18) Wang, J.; Cao, D.; Tang, C.; Chen, X.; Sun, H.; Hou, T. Fast and accurate prediction of partial charges using Atom-Path-Descriptor-based machine learning. *Bioinformatics* **2020**, *36* (18), 4721−4728.

(19) Kancharlapalli, S.; Gopalan, A.; Haranczyk, M.; Snurr, R. Q. Fast and Accurate Machine Learning Strategy for Calculating Partial Atomic Charges in Metal−Organic Frameworks. *J. Chem. Theory Comput.* **2021**, *17* (5), 3052−3064.

(20) Marchenko, E. I.; Fateev, S. A.; Petrov, A. A.; Korolev, V. V.; Mitrofanov, A.; Petrov, A. V.; Goodilin, E. A.; Tarasov, A. B. Database of Two-Dimensional Hybrid Perovskite Materials: Open-Access Collection of Crystal Structures, Band Gaps, and Atomic Partial Charges Predicted by Machine Learning. *Chem. Mater.* **2020**, *32* (17), 7383−7388.

(21) Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. *AI Open* **2020**, *1*, 57−81.
(22) Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Yu, P. S. A Comprehensive Survey on Graph Neural Networks. *IEEE Transact. Neural Networks Learn. Syst.* **2021**, *32* (1), 4−24.
(23) Wu, L.; Cui, P.; Pei, J.; Zhao, L.; Guo, X. Graph Neural Networks: Foundation, Frontiers and Applications. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*; Association for Computing Machinery: Washington DC, USA, 2022; pp 4840−4841.
(24) Su, Q.; Wang, J.; Gou, Q.; Hu, R.; Zhang, H.; Jiang, L.; Wang, T.; Liu, Y.; Shen, C.; Kang, Y.; et al. Robust Protein-Ligand Interaction Modeling Integrating Physical Laws and Geometric Knowledge for Absolute Binding Free Energy Calculation. *arXiv Preprint* **2024**, chemrxiv-2024-9k6z7.
(25) Dwivedi, V. P.; Joshi, C. K.; Luu, A. T.; Laurent, T.; Bengio, Y.; Bresson, X. Benchmarking graph neural networks. *J. Mach. Learn. Res.* **2023**, *24* (43), 1−48.
(26) You, J.; Gomes-Selman, J. M.; Ying, R.; Leskovec, J. Identity-aware Graph Neural Networks. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35* (12), 10737−10745.
(27) Wang, J.; Cao, D.; Tang, C.; Xu, L.; He, Q.; Yang, B.; Chen, X.; Sun, H.; Hou, T. DeepAtomicCharge: a new graph convolutional network-based architecture for accurate prediction of atomic charges. *Briefings Bioinf.* **2020**, *22* (3), bbaa183.
(28) Jiang, D.; Sun, H.; Wang, J.; Hsieh, C.-Y.; Li, Y.; Wu, Z.; Cao, D.; Wu, J.; Hou, T. Out-of-the-box deep learning prediction of quantum-mechanical partial charges by graph representation and transfer learning. *Briefings Bioinf.* **2022**, *23* (2), bbab597.
(29) Townshend, R. J.; Vögele, M.; Suriana, P.; Derry, A.; Powers, A.; Laloudakis, Y.; Balachandar, S.; Jing, B.; Anderson, B.; Eismann, S. Atom3d: Tasks on molecules in three dimensions. *arXiv Preprint* **2020**, arXiv:2012.04035.
(30) Zhang, X.; Wang, L.; Helwig, J.; Luo, Y.; Fu, C.; Xie, Y.; Liu, M.; Lin, Y.; Xu, Z.; Yan, K. Artificial intelligence for science in quantum, atomistic, and continuum systems. *arXiv Preprint* **2023**, arXiv:2307.08423.
(31) Xu, K.; Hu, W.; Leskovec, J.; Jegelka, S. How powerful are graph neural networks? *arXiv Preprint* **2018**, arXiv:1810.00826.
(32) Chen, Z.; Chen, L.; Villar, S.; Bruna, J. Can graph neural networks count substructures? *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 10383−10395.
(33) Zopf, M. In *1-WL Expressiveness Is (Almost) All You Need. 2022 International Joint Conference on Neural Networks (IJCNN), 18−23 July 2022*, 2022; pp 1−8.
(34) Wang, Y.; Zhang, M. *An Empirical Study of Realized GNN Expressiveness, Forty-first International Conference on Machine Learning*, 2024.
(35) Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48−62.
(36) Li, W. H.; Yang, J.; Wang, D. Long-range interactions in diatomic catalysts boosting electrocatalysis. *Angew. Chem.* **2022**, *134* (52), No. e202213318.
(37) Li, M.; Zhirihin, D.; Gorlach, M.; Ni, X.; Filonov, D.; Slobozhanyuk, A.; Alù, A.; Khanikaev, A. B. Higher-order topological states in photonic kagome crystals with long-range interactions. *Nat. Photonics* **2020**, *14* (2), 89−94.
(38) Mejia-Rodriguez, D.; Aprà, E.; Autschbach, J.; Bauman, N. P.; Bylaska, E. J.; Govind, N.; Hammond, J. R.; Kowalski, K.; Kunitsa, A.; Panyala, A.; Peng, B.; Rehr, J. J.; Song, H.; Tretiak, S.; Valiev, M.; Vila, F. D. NWChem: Recent and Ongoing Developments. *J. Chem. Theory Comput.* **2023**, *19* (20), 7077−7096.
(39) Vaswani, A. Attention is All you Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.
(40) Zdrazil, B.; Felix, E.; Hunter, F.; Manners, E. J.; Blackshaw, J.; Corbett, S.; de Veij, M.; Ioannidis, H.; Lopez, D. M.; Mosquera, J.; Magarinos, M.; Bosc, N.; Arcila, R.; Kizilören, T.; Gaulton, A.; Bento, A. P.; Adasme, M.; Monecke, P.; Landrum, G.; Leach, A. The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res.* **2023**, *52* (D1), D1180−D1192.
(41) Bilbrough, T.; Piemontese, E.; Seitz, O. Dissecting the role of protein phosphorylation: a chemical biology toolbox. *Chem. Soc. Rev.* **2022**, *51* (13), 5691−5730.
(42) Orning, P.; Lien, E. Multiple roles of caspase-8 in cell death, inflammation, and innate immunity. *J. Leukocyte Biol.* **2021**, *109* (1), 121−141.
(43) Zhang, W.; Zhu, C.; Liao, Y.; Zhou, M.; Xu, W.; Zou, Z. Caspase-8 in inflammatory diseases: a potential therapeutic target. *Cell. Mol. Biol. Lett.* **2024**, *29* (1), 130.
(44) Soisson, S. M.; Parthasarathy, G.; Adams, A. D.; Sahoo, S.; Sitlani, A.; Sparrow, C.; Cui, J.; Becker, J. W. Identification of a potent synthetic FXR agonist with an unexpected mode of binding and activation. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (14), 5337−5342.
(45) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: a web-accessible database of experimentally determined protein−ligand binding affinities. *Nucleic Acids Res.* **2007**, *35* (Database), D198−D201.
(46) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47* (7), 1739−1749.
(47) Kwiatkowski, J. S.; Leszczyński, J. Ab initio Hartree—Fock and post-Hartree—Fock studies on molecular structures and vibrational spectra of formamide and formamidic acid. *J. Mol. Struct.* **1992**, *270*, 67−86.
(48) 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, Jr. J. A.; 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. C.01: Wallingford, CT, 2016.
(49) Imambi, S.; Prakash, K. B.; Kanagachidambaresan, G. R.. In *PyTorch. In Programming with TensorFlow: Solution for Edge Computing Applications*; Prakash, K. B., Kanagachidambaresan, G. R., Eds.; Springer International Publishing: Cham, 2021; pp 87−104.
(50) Li, M.; Soltanolkotabi, M.; Oymak, S. Gradient Descent with Early Stopping is Provably Robust to Label Noise for Over-parameterized Neural Networks. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*; Silvia, C., Roberto, C., Eds.; PMLR: Proceedings of Machine Learning Research, 2020; Vol. *108*, pp 4313−4324.
(51) Hinton, G. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv Preprint* **2012**, arXiv:1207.0580.

https://doi.org/10.1021/acs.jcim.5c00602
J. Chem. Inf. Model. 2025, 65, 10364−10374