Synergistic Application of Molecular Docking and Machine Learning for Improved Protein-Ligand Binding Pose Prediction

He Yang*
DP Technology
yanghe@dp.tech

Hongrui Lin*
DP Technology
linhongrui@dp.tech

Yannan Yuan
DP Technology
yuanyn@dp.tech

Yaqi Li
DP Technology
liyq@dp.tech

Rongfeng Zou
DP Technology
zourongfeng@dp.tech

Gengmo Zhou
Renmin University of China
DP Technology
zgm2015@ruc.edu.cn

Linfeng Zhang†
DP Technology
zhanglf@dp.tech

Hang Zheng†
DP Technology
zhengh@dp.tech

September 9, 2023

Abstract

Accurate prediction of protein-ligand complex structures is a crucial step in structure-based drug design. Traditional molecular docking methods exhibit limitations in terms of accuracy and sampling space, while relying on machine-learning approaches may lead to invalid conformations. In this study, we propose a novel strategy that combines molecular docking and machine learning methods. Firstly, the protein-ligand binding poses are predicted using the Uni-Mol Docking machine learning approach. Subsequently, position-restricted docking(PR Docking) on predicted binding poses is performed using Uni-Dock, generating physically constrained and valid binding poses. Finally, the binding poses are re-scored and ranked using machine learning scoring functions. This strategy harnesses the predictive power of machine learning and the physical constraints advantage of molecular docking. Evaluation experiments on multiple datasets demonstrate that, compared to using molecular docking or machine learning methods alone, our proposed strategy can significantly improve the success rate and accuracy of protein-ligand complex structure predictions. This strategy is avaliable at https://github.com/dptech-corp/Uni-Dock.

Keywords: Binding Pose, Molecular Docking, Machine Learning

1 Introduction

Protein-ligand complex structure prediction is one of the essential aspects of drug design. Accurately predicting protein-ligand complex structure can provide a basis for structure-based drug design, thereby facilitating the design and selection of potential drug molecules. Furthermore, reasonable complex structures can help medicinal chemists understand the binding mechanism of small molecules with target proteins, laying the foundation for structure-activity relationship analysis and rational drug design [1, 2]. Consequently, developing accurate protein-ligand complex structure prediction methods is of great importance for structure-based drug design.

Existing molecular docking schemes, such as AutoDock Vina [3], Uni-Dock [4], and LeDock [5], typically rely on conformational sampling algorithms and empirical scoring functions to search for protein and ligand binding poses and predict ligand conformations at the target protein binding site

*Equal contribution.
†Corresponding author

based on factors such as ligand internal energy and protein-ligand interaction energy [6]. However, these methods struggle to accurately describe various interaction forms between proteins and ligands, mainly due to simplified scoring functions for ensuring computational speed. Moreover, the search complexity of conformational sampling algorithms limits their coverage of chemical space [7]. These factors result in limited capabilities of traditional molecular docking software in protein-ligand complex structure prediction.

In recent years, machine learning-based scoring functions, such as GNINA [8] and RTMScore [9], have gained wide attention. These methods establish more refined and accurate scoring functions by learning from protein-ligand complex structures and affinity data. Studies have shown that compared to traditional empirical scoring functions, machine learning models can improve docking prediction success rates [10]. However, machine learning model inference speeds are relatively slow and are generally only used for re-scoring and ranking several protein-ligand binding poses obtained through molecular docking to select the optimal structure [11]. If the preceding molecular docking step fails to sample conformations close to the crystal structure, machine learning scoring functions may become ineffective.

On the other hand, some deep learning models, such as DeepDock [12] and Uni-Mol Docking [13], attempt to directly predict protein-ligand complex structures end-to-end without explicit conformational search, achieving promising performance. These methods can avoid the limited sampling space size of traditional molecular docking algorithms and exhibit higher prediction success rates. However, the lack of physical constraints on chemical structures in deep learning models may result in predicted conformations that do not adhere to basic physical laws, such as invalid bond lengths and angles, and protein collisions [14].

To fully exploit the advantages of traditional molecular docking methods and machine learning approaches while avoiding their respective shortcomings, we propose a novel strategy that combines molecular docking and machine learning approaches for more accurate and valid protein-ligand complex structure prediction. Firstly, we use Uni-Mol Docking to predict the binding poses of protein-ligand complexes. Subsequently, based on the predicted binding poses, we perform PR Docking using the Uni-Dock molecular docking software, generating a series of physically constrained docked binding poses. Finally, we employ scoring functions such as GNINA, RTMScore, and Vinardo to re-score and rank the docked binding poses, yielding the optimal protein-ligand complex conformation. This approach combines traditional molecular docking methods with machine learning techniques, providing an efficient means for structure-based drug design.

2 Methods

The proposed workflow that combines machine learning methods and molecular docking is shown in figure 1, consisting of three steps:

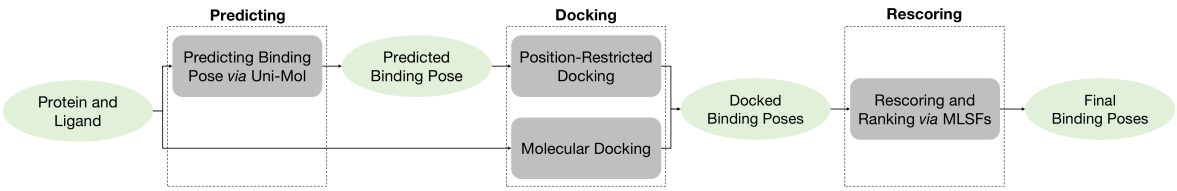

Figure 1: Workflow that combines machine learning methods and molecular docking.

Predicting Utilize the Uni-Mol Docking to predict protein-ligand binding poses. Uni-Mol is a universal 3D molecular representation learning framework based on SE(3)-equivariant transformers [13]. We use the Uni-Mol Docking module and over 70,000 protein-ligand complexes from the MOAD dataset as training data, obtaining a prediction model for protein-ligand binding poses. During prediction, we extract amino acid residues within a 10Å range of the ligand as the binding pocket and re-generate the ligand's 3D conformation using RDKit, which is then input into the trained Uni-Mol Docking model along with the binding pocket to obtain the "predicted binding poses".

Docking Perform PR Docking using Uni-Dock based on the "predicted binding poses". Uni-Dock is a high-performance GPU-accelerated molecular docking software [4] that can search a larger conformational space. We use the coordinates of each heavy atom in "predicted binding poses" as position-restricted offsets for PR Docking (specific details are in the appendix A.2), guiding Uni-Dock to focus on the binding pose region for conformation sampling, and generating "docked binding poses" that comply with the position constraints.

Rescoring Apply a machine learning scoring function to re-score the "docked binding poses". We use various traditional and machine learning scoring functions, including GNINA [8] and RTMScore [9], to score the "docked binding poses". The conformations are then ranked, and the highest-ranked binding pose is selected as the "final binding pose".

This strategy takes full advantage of the machine learning conformation prediction capabilities and the physical constraints of traditional molecular docking, avoiding their respective limitations, and is expected to effectively improve the success rate and accuracy of protein-ligand complex structure prediction.

This strategy is avaliable at https://github.com/dptech-corp/Uni-Dock.

3 Datasets

To evaluate the performance of the proposed method, we used several commonly used protein-ligand complex datasets, including Astex Diverse set [15], CASF-2016 [16], PoseBusters [14], and PDBbind Refined set [17]. Due to the significant differences in docking sampling space brought by varying numbers of rotatable bonds in ligands, we classified the test sets based on the number of rotatable bonds in ligands into different difficulty levels: ligands with 0-5 rotatable bonds were classified as "easy", 6-12 as "medium", and ligands with more than 12 rotatable bonds as "difficult".

We performed the following preparation steps for the proteins and ligands in the datasets. After obtaining the protein structures from the RCSB database [18] based on the PDB code, we retained the crystal waters and cofactors that affect the binding mode and completed missing protein side chains and lost hydrogen atoms. For ligands, we searched the RCSB database for the isomer SMILES corresponding to the PDB code and determined the correct protonation state according to the receptor pocket environment. Then, we generated 3D conformations for each ligand. After excluding systems with failed preparation and those with large natural products or polypeptide ligands, 70 systems from the Astex Diverse set, 258 systems from CASF-2016, 386 systems from PoseBusters, and 4831 systems from the PDBbind Refined Set were used as test sets.

The key statistical information is summarized in Table 1. These datasets broadly represent protein-small molecule systems of varying difficulty levels and complexities.

Table 1: Datasets uesd as test sets.

Dataset	Total Num.	Num. of Easy Cases	Num. of Medium Cases	Num. of Hard Case
CASF-2016	231	127	92	12
PoseBusters	386	158	197	37
Astex Diverse	70	41	29	0
PDBbind Refined Set	4831	1904	2103	824

4 Results and Discussion

4.1 Predicting protein-ligand complex binding poses by Uni-Mol Docking

We first evaluated the performance of Uni-Mol Docking in predicting protein-ligand binding poses on four datasets (success was defined as the root-mean-square deviation (RMSD) between the predicted pose and the crystal pose below a certain threshold). As shown in Figure 2, the structures with RMSD

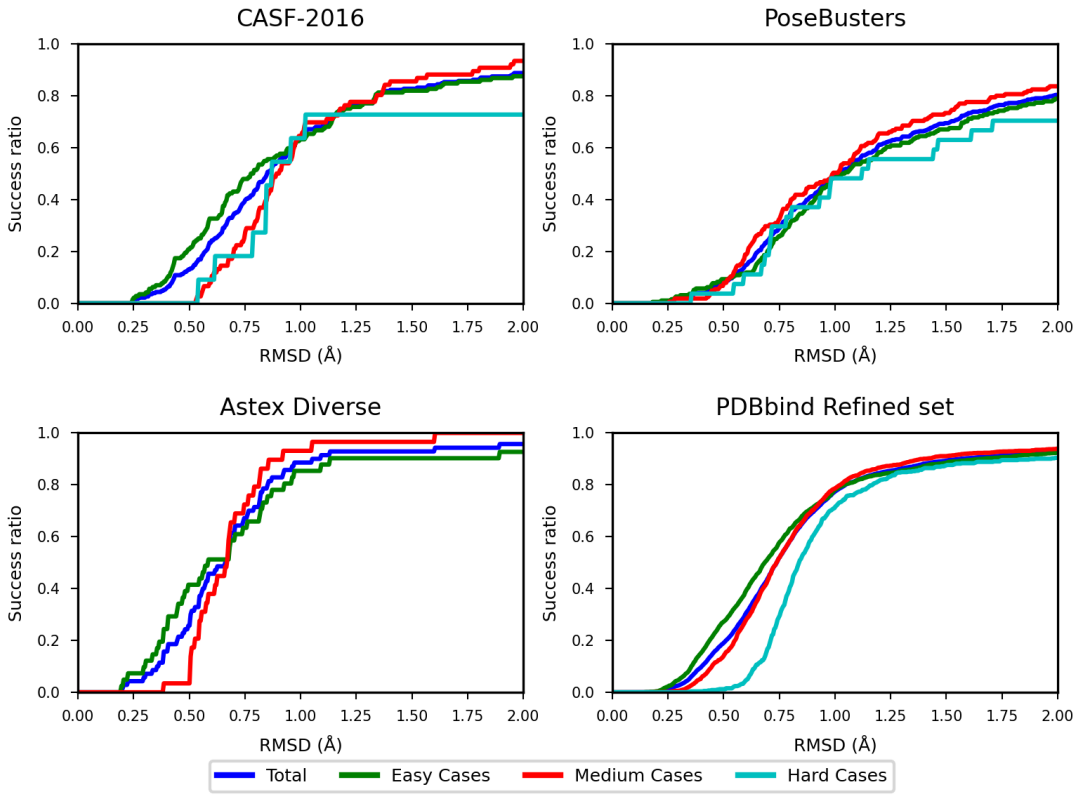

Figure 2: Uni-Mol Docking performance on test sets.

less than 2Å in the Uni-Mol Docking prediction results accounted for over 80% for all the datasets. This demonstrates that the Uni-Mol Docking has a strong ability to predict binding poses. It is particularly noteworthy that the samples in the PoseBusters dataset did not appear in the training data of Uni-Mol Docking (as described at Appendix A.1), yet its prediction results were similar to those of other datasets, indicating that the model has a certain generalization ability. We also observed that as the number of rotatable bonds in the ligand increased, the prediction difficulty increased, and the success rate of Uni-Mol Docking decreased.

Interestingly, we found that Uni-Mol Docking had a very low success rate in predicting detailed structures, specifically in the regions with smaller RMSD values. Therefore, we selected a few representative systems for demonstration in Figure 3. From the overlay of the predicted binding poses and crystal structures, we can see that Uni-Mol Docking can accurately predict the overall trend of the molecules. However, in the prediction of symmetric structures, such as phenyl rings and isopropyl groups, Uni-Mol Docking exhibited non-physical bond lengths and angles.

Therefore, we subsequently employed Uni-Dock, a physics-based molecular docking method, to refine and optimize the predicted binding poses obtained from Uni-Mol Docking.

4.2 Getting Binding Poses by PR Docking

We utilized the predicted binding poses generated by Uni-Mol Docking as a basis and transformed them into position-restricted bias potentials (Figure 4) during the docking process. Uni-Dock was then employed for PR Docking to generate more reasonable binding poses. During docking processing, when the atoms of the ligand molecule enter the range of the bias potential, the binding pose score receives a reward. Consequently, Uni-Dock makes the final docked binding pose more inclined towards the parts with bias potential, as shown in Figure 4. Since Uni-Dock explicitly avoids physical clash, such as ligand-protein proximity, and generates conformations based on rotatable bonds, this workflow can effectively leverage the binding structure prediction ability of Uni-Mol Docking while ensuring the physical reliability of binding poses.

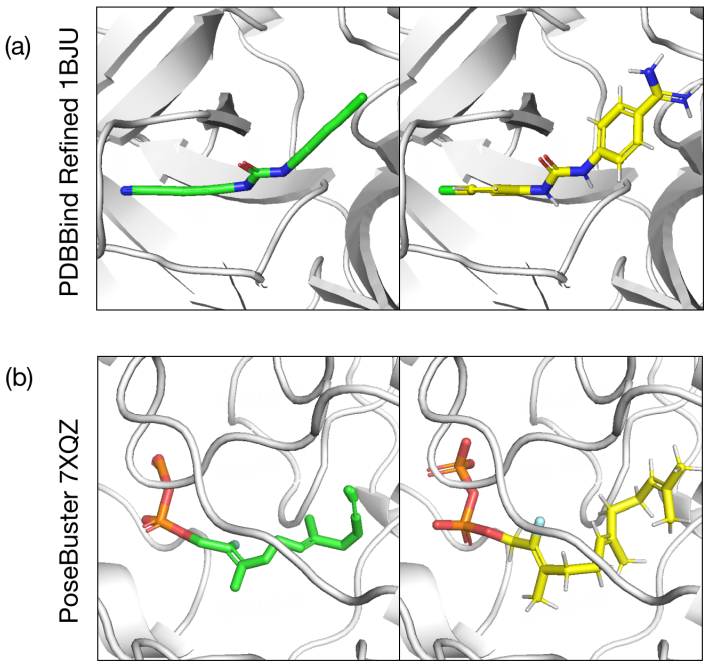

Figure 3: The binding conformation predicted by Uni-Mol Docking(left) and the crystal conformation(right).

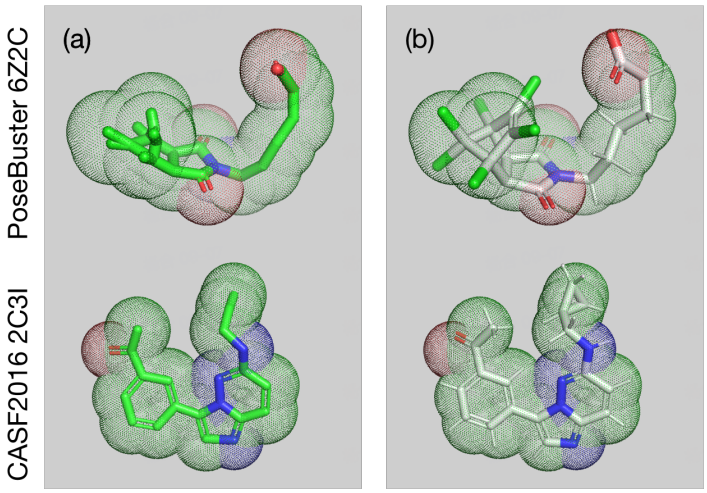

Figure 4: (a) Conversion of Uni-Mol Docking predicted conformations to bias potentials. (b) Results obtained by PR Docking with Uni-Mol Docking prediction

We applied this workflow to four test datasets. In addition, we conducted PR Docking using crystal structures, which can be considered as an upper bound for this workflow. On the other hand, we conducted unbiased molecular docking as a lower bound. The results are shown in Figure 5.

We observed that the PR Docking results based on Uni-Mol Docking prediction consistently improved the success rate of binding conformation prediction compared to Uni-Dock molecular docking. In particular, for systems with a higher number of rotatable bonds in the ligand, this combined method had a more significant improvement in prediction accuracy, indicating that PR Docking effectively reduced the complexity of searching in chemical space, helping the molecular docking method to converge rapidly around the true structure position. Compared to Uni-Mol Docking's results, this combined method significantly increased the success rate for RMSD less than 1Å, proving that Uni-Dock can effectively correct structures that do not conform to physical constraints and improve the local structure prediction accuracy.

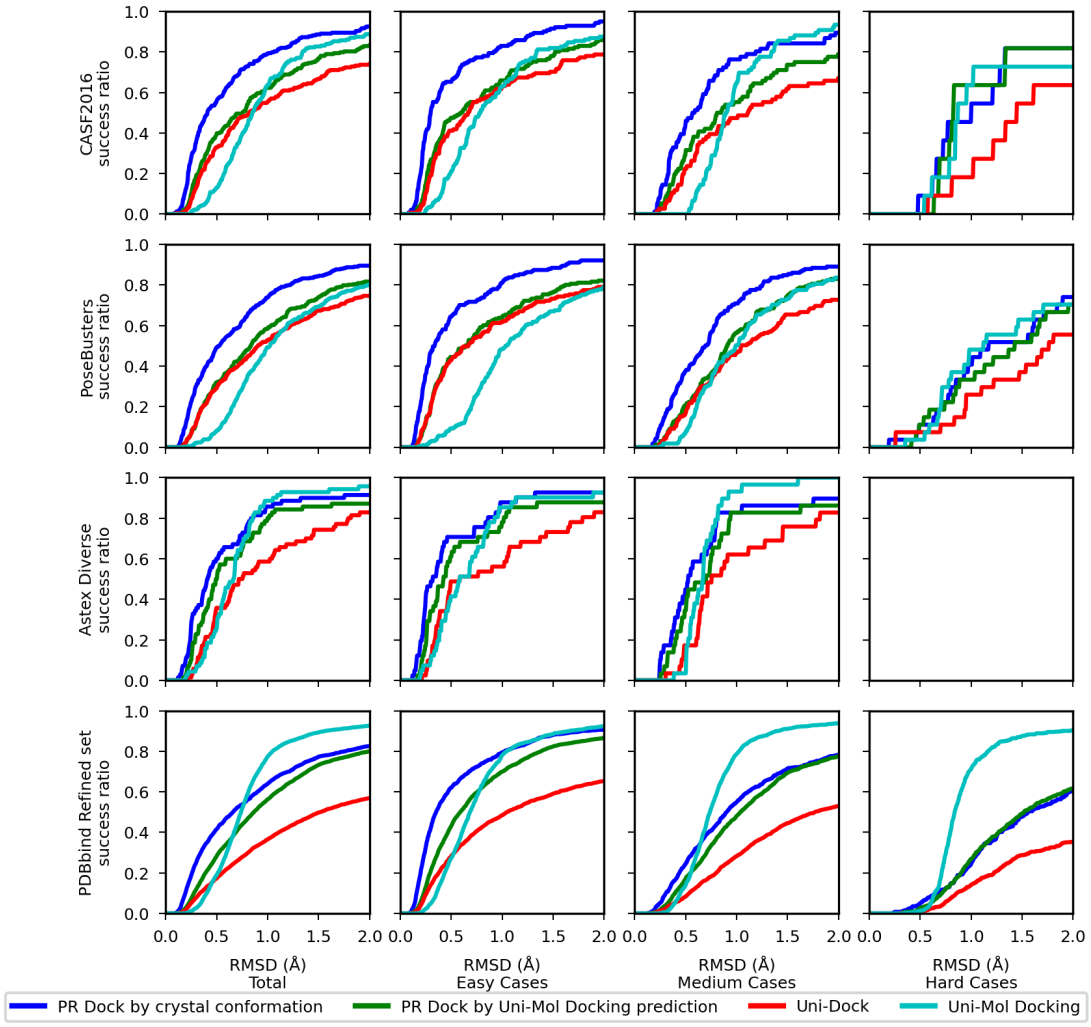

Figure 5: The Results of PR Docking with Uni-Mol Docking prediction

Although the protein-ligand complex prediction success rate of Uni-Mol Docking in some systems with RMSD less than 2Å is even higher than that of PR Docking which uses crystal structures, we found that the accuracy of PR Docking with Uni-Mol docking predicted structures did not exceed that of PR Docking using crystal structures, and there was even a significant gap. This indicates that the predicted structures of Uni-Mol Docking cannot yet serve as a perfect solution to guide molecular docking in conformation search.

4.3 Machine learning scoring functions re-rank docking poses

We further investigated the binding poses obtained by PR Docking and conventional molecular docking to assess their actual sampling capabilities. The ability of these methods to reproduce crystal structures when retaining a certain number of docking poses is shown in Table 2.

We observed that as the number of considered conformations increases, the probability of finding a conformation with an RMSD less than 2.0 Å from the ligand's crystal conformation also increases for both PR Docking and conventional molecular docking. Additionally, when considering all possible docking poses, the success rates of PR Docking and conventional molecular docking are comparable. We also calculated the results considering both PR Docking and conventional molecular docking together. We found that considering both methods simultaneously leads to higher success rates than either method alone, implying good complementarity between Uni-Mol Docking and Uni-Dock in cases where Uni-Mol Docking incorrectly predicts the conformation.

Table 2: Using Uni-Dock and PR Docking with Uni-Mol Docking prediction to perform molecular docking, and comparing the RMSD success ratio of top 1, top 3, top 5, top 10 and all ligand poses with crystal conformations.

	CASF-2016			PoseBusters		
	Uni-Dock	PR Docking	Uni-Dock & PR Docking*	Uni-Dock	PR Docking	Uni-Dock & PR Docking
top 1	0.741	0.833	0.877	0.747	0.817	0.877
top 3	0.798	0.886	0.908	0.809	0.851	0.893
top 5	0.811	0.895	0.917	0.843	0.872	0.909
top 10	0.846	0.921	0.925	0.862	0.885	0.914
all poses	0.917	0.934	**0.939**	0.924	0.914	**0.940**

	Astex Diverse			PDBbind Refined set		
	Uni-Dock	PR Docking	Uni-Dock & PR Docking	Uni-Dock	PR Docking	Uni-Dock & PR Docking
top 1	0.823	0.871	0.886	0.567	0.800	0.901
top 3	0.843	0.900	0.914	0.661	0.837	0.934
top 5	0.886	0.900	0.914	0.705	0.852	0.943
top 10	0.886	0.926	0.928	0.759	0.864	0.954
all poses	0.957	0.957	**0.971**	0.865	0.882	**0.971**

[1] The Uni-Dock & PR Docking method considers the conformations of both the Uni-Dock method and the PR Docking method. For example, the top10 conformations of the Uni-Dock & PR Docking method are composed of the Uni-Dock method's top10 conformations and the PR Docking method's top10 conformations, with a total of 20 conformations.

The information above indicates that molecular docking methods can effectively collect conformations close to the crystal structure of the ligand. The challenge lies in selecting excellent conformations and placing them in the forefront. In particular, when the top-ranked complex conformation structures given by Uni-Dock and PR Docking with Uni-Mol Docking prediction are inconsistent, it is challenging to determine which structure is better. Therefore, we subsequently tested the rescoring and re-ranking performance of machine learning scoring functions GNINA, RTMScore, and physics-based scoring func-

tion Vinardo, by assessing whether rescoring the top 1, top 10, and all docking conformations could improve the prediction success rate of binding conformations. The results are shown in Table 3.

We can see that, overall, the machine learning scoring functions GNINA and RTMScore exhibit better rescoring performance than the physics-based scoring function Vinardo. Specifically, after rescoring the Top1 structure, the success rates based on GNINA and RTMScore are better than those obtained by using Uni-Dock's Top1 structure alone or by using PR Docking with Uni-Mol Docking prediction. This confirms that machine learning scoring functions help select superior binding conformations. Furthermore, when rescoring the top 10 docking poses obtained from both molecular docking and PR Docking, machine learning scoring functions can significantly improve the overall success rate of protein-ligand binding mode prediction, especially GNINA, which demonstrates robust improvement capabilities on larger datasets (PoseBusters and PDBbind Refined Set).

However, it is worth noting that when rescoring all docking poses, the performance of machine learning scoring functions GNINA and RTMScore declines significantly, while the performance decline of the physics-based scoring function Vinardo is relatively smaller. This may suggest that the training data for machine learning scoring functions such as GNINA and RTMScore may be biased and might not adequately cover the entire conformational space, while physics-based scoring functions, due to the existence of physical constraints, perform more robustly in evaluating uncommon structures.

In summary, by using machine learning scoring functions for rescoring, we can further make reasonable selections and obtain better docking structures based on the results of molecular docking and PR Docking with Uni-Mol Docking prediction.

Table 3: Using vinardo, GNINA, RTMScore to re-score top 1, top 10 or all conformations obtained by Uni-Dock and PR Docking with Uni-Mol Docking prediction.

	CASF-2016			PoseBusters		
	Uni-Dock	PR Docking	Uni-Dock & PR Docking*	Uni-Dock	PR Docking	Uni-Dock & PR Docking
vinardo rescoring on top 1 poses	\	\	0.768	\	\	0.799
vinardo rescoring on top 10 poses	0.711	0.803	0.811	0.744	0.796	0.838
vinardo rescoring on all poses	0.715	0.737	0.776	0.742	0.749	0.802
GNINA rescoring on top 1 poses	\	\	0.838	\	\	0.843
GNINA rescoring on top 10 poses	0.750	0.820	0.846	0.773	0.809	**0.864**
GNINA rescoring on all poses	0.732	0.776	0.789	0.723	0.752	0.796
RTMScore rescoring on top 1 poses	\	\	0.846	\	\	0.812
RTMScore rescoring on top 10 poses	0.750	0.816	**0.852**	0.710	0.762	0.820
RTMScore rescoring on all poses	0.737	0.794	0.825	0.646	0.666	0.720

	Astex Diverse			PDBbind Refine Set		
	Uni-Dock	PR Docking	Uni-Dock & PR Docking	Uni-Dock	PR Docking	Uni-Dock & PR Docking
vinardo rescoring on top 1 poses	\	\	0.829	\	\	0.707
vinardo rescoring on top 10 poses	0.743	0.814	0.814	0.599	0.698	0.853
vinardo rescoring on all poses	0.700	0.771	0.771	0.596	0.611	0.814
GNINA rescoring on top 1 poses	\	\	0.871	\	\	0.755
GNINA rescoring on top 10 poses	0.814	0.886	0.886	0.631	0.732	**0.881**
GNINA rescoring on all poses	0.786	0.843	0.871	0.634	0.660	0.847
RTMScore rescoring on top 1 poses	\	\	0.870	\	\	0.765
RTMScore rescoring on top 10 poses	0.696	0.851	**0.910**	0.458	0.681	0.824
RTMScore rescoring on all poses	0.638	0.742	0.848	0.357	0.633	0.754

[1] The Uni-Dock & PR Docking method considers the conformations of both the Uni-Dock method and the PR Docking method. For example, the top10 conformations of the Uni-Dock & PR Docking method are composed of the Uni-Dock method's top10 conformations and the PR Docking method's top10 conformations, with a total of 20 conformations.

5 Conclusion

In this paper, we propose a novel method that combines molecular docking and machine learning to enhance the accuracy of protein-ligand complex structure prediction. First, we employ the machine learning model Uni-Mol Docking to predict protein-ligand complex conformations; next, we use PR Docking with Uni-Mol Docking prediction to perform molecular docking, generating physically constrained docking conformations. Finally, we re-score multiple conformations generated by molecular docking using a machine learning scoring function to identify the best-scoring conformation as the final predicted protein-ligand complex structure.

Evaluation experiments on multiple benchmark datasets demonstrate that, compared to using traditional docking or machine learning methods alone, this combined strategy significantly improves the success rate and accuracy of binding conformation prediction, particularly for systems with high ligand flexibility. This shows that machine learning-predicted conformations can effectively guide molecular docking searches, while the physical constraints provided by molecular docking prevent the generation of non-physically plausible conformations.

However, our work also reveals some limitations of the current methods: 1) the prediction accuracy of the machine learning model Uni-Mol in terms of structural plausibility still needs improvement, especially for symmetric structures; 2) the re-scoring by machine learning scoring functions did not bring significant improvement, suggesting potential issues in the training process or evaluation methods of the current scoring functions. Based on these findings, we will attempt to incorporate more physical constraints into the Uni-Mol Docking process and test various combinations of machine learning scoring functions and workflows to further enhance the prediction ability of protein-ligand complex structures.

References

[1] Rob LM Van Montfort and Paul Workman. Structure-based drug design: aiming for a perfect fit. *Essays in biochemistry*, 61(5):431–437, 2017.

[2] Xin Wang, Ke Song, Li Li, and Lijiang Chen. Structure-based drug design strategies and challenges. *Current Topics in Medicinal Chemistry*, 18(12):998–1006, 2018.

[3] Jerome Eberhardt, Diogo Santos-Martins, Andreas F Tillack, and Stefano Forli. Autodock vina 1.2.0: New docking methods, expanded force field, and python bindings. *Journal of chemical information and modeling*, 61(8):3891–3898, 2021.

[4] Yuejiang Yu, Chun Cai, Jiayue Wang, Zonghua Bo, Zhengdan Zhu, and Hang Zheng. Uni-dock: Gpu-accelerated docking enables ultralarge virtual screening. *Journal of Chemical Theory and Computation*, 19(11):3336–3345, 2023.

[5] Hongtao Zhao. Ledock. http://www.lephar.com/software.htm.

[6] Nataraj S Pagadala, Khajamohiddin Syed, and Jack Tuszynski. Software for molecular docking: a review. *Biophysical reviews*, 9:91–102, 2017.

[7] Zhe Wang, Huiyong Sun, Xiaojun Yao, Dan Li, Lei Xu, Youyong Li, Sheng Tian, and Tingjun Hou. Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. *Physical Chemistry Chemical Physics*, 18(18):12964–12975, 2016.

[8] Andrew T McNutt, Paul Francoeur, Rishal Aggarwal, Tomohide Masuda, Rocco Meli, Matthew Ragoza, Jocelyn Sunseri, and David Ryan Koes. Gnina 1.0: molecular docking with deep learning. *Journal of cheminformatics*, 13(1):1–20, 2021.

[9] Chao Shen, Xujun Zhang, Yafeng Deng, Junbo Gao, Dong Wang, Lei Xu, Peichen Pan, Tingjun Hou, and Yu Kang. Boosting protein–ligand binding pose prediction and virtual screening based on residue–atom distance likelihood potential and graph transformer. *Journal of Medicinal Chemistry*, 65(15):10691–10706, 2022.

[10] Chao Shen, Junjie Ding, Zhe Wang, Dongsheng Cao, Xiaoqin Ding, and Tingjun Hou. From machine learning to deep learning: Advances in scoring functions for protein–ligand docking. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 10(1):e1429, 2020.

[11] Qifeng Bai, Shuo Liu, Yanan Tian, Tingyang Xu, Antonio Jesús Banegas-Luna, Horacio Pérez-Sánchez, Junzhou Huang, Huanxiang Liu, and Xiaojun Yao. Application advances of deep learning methods for de novo drug design and molecular dynamics simulation. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 12(3):e1581, 2022.

[12] Francesco Gentile, Vibudh Agrawal, Michael Hsing, Anh-Tien Ton, Fuqiang Ban, Ulf Norinder, Martin E Gleave, and Artem Cherkasov. Deep docking: a deep learning platform for augmentation of structure based drug discovery. *ACS central science*, 6(6):939–949, 2020.

[13] Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng Zhang, and Guolin Ke. Uni-mol: A universal 3d molecular representation learning framework. In *The Eleventh International Conference on Learning Representations*, 2023.

[14] Martin Buttenschoen, Garrett M Morris, and Charlotte M Deane. Posebusters: Ai-based docking methods fail to generate physically valid poses or generalise to novel sequences. *arXiv preprint arXiv:2308.05777*, 2023.

[15] Michael J Hartshorn, Marcel L Verdonk, Gianni Chessari, Suzanne C Brewerton, Wijnand TM Mooij, Paul N Mortenson, and Christopher W Murray. Diverse, high-quality test set for the validation of protein- ligand docking performance. *Journal of medicinal chemistry*, 50(4):726–741, 2007.

[16] Minyi Su, Qifan Yang, Yu Du, Guoqin Feng, Zhihai Liu, Yan Li, and Renxiao Wang. Comparative assessment of scoring functions: the casf-2016 update. *Journal of chemical information and modeling*, 59(2):895–913, 2018.

[17] Zhihai Liu, Minyi Su, Li Han, Jie Liu, Qifan Yang, Yan Li, and Renxiao Wang. Forging the basis for developing protein–ligand interaction scoring functions. *Accounts of chemical research*, 50(2):302–309, 2017.

[18] Stephen K Burley, Helen M Berman, Gerard J Kleywegt, John L Markley, Haruki Nakamura, and Sameer Velankar. Protein data bank (pdb): the single global macromolecular structure archive. *Protein crystallography: methods and protocols*, pages 627–641, 2017.

A Appendix

A.1 Overlap of test data sets and Uni-Mol Docking training data

Table 4: Overlap of test data sets and Uni-Mol Docking training data

Dataset	Total Num.	Num. of Complexes in Uni-Mol Training Set	Coverage (%)
Astex Diverse	70	30	43
CASF-2016	231	189	82
PoseBusters	386	0	0
PDBbind Refined Set	4831	3816	79

These datasets are available at:
CASF-2016: https://bohrium-api.dp.tech/ds-dl/casf2016-fyfy-v1
PoseBustsers: https://bohrium-api.dp.tech/ds-dl/posebuster-5f7t-v1
Astext Diverse: https://bohrium-api.dp.tech/ds-dl/astex-i1rs-v1
PDBbind Refined Set: https://bohrium-api.dp.tech/ds-dl/pdbbind-refined-set-7db0-v1

A.2 Parameter settings for PR Docking

In conventional molecular docking, we set the center of mass of the ligand in the protein-ligand crystal conformation as the center coordinate of the docking pocket, and use the space within 10 angstroms from the ligand serves as the docking box size. the *energy_range* is set to 9 kal/mol, and the *search_mode* method is detail.

In PR Docking with Uni-Mol Docking prediction, we use the predicted conformation of the Uni-Mol Docking as the reference conformation, and use its heavy atoms' coordinates and atom type information to add bias potential according to equation (1). Then, the PR Docking is carried out to obtain the binding poses.

$$\Delta E \ = \ Vset * \ e^{-r^2/r0} \tag{1}$$

In equation (1), Vset represents the energy value which is added at the bias center, r0 represents the radius distribution size of the bias, r represents the distance from the point to be modified to the bias center position, Vset is equal to -0.8 kcal/mol, and r0 is equal to 1.2 Å.