Acta Pharmaceutica Sinica B 2025;15(11):5758−5771

Chinese Pharmaceutical Association
Institute of Materia Medica, Chinese Academy of Medical Sciences

Acta Pharmaceutica Sinica B

www.elsevier.com/locate/apsb
www.sciencedirect.com

TOOLS

CarsiDock-Cov: A deep learning-guided approach for automated covalent docking and screening

Chao Shen[a,b,c], Hongyan Du[b,d], Xujun Zhang[b,d], Shukai Gu[b], Heng Cai[e], Yu Kang[b,d], Peichen Pan[b,d], Qingwei Zhao[a,c], Tingjun Hou[a,b,d,*]

[a]*Department of Clinical Pharmacy, the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China*
[b]*College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China*
[c]*Zhejiang Provincial Key Laboratory for Drug Evaluation and Clinical Research, Hangzhou 310003, China*
[d]*Zhejiang Provincial Key Laboratory for Intelligent Drug Discovery and Development, Jinhua 321016, China*
[e]*Hangzhou Carbonsilicon AI Technology Co., Ltd., Hangzhou 310018, China*

Received 7 January 2025; received in revised form 24 April 2025; accepted 6 May 2025

KEY WORDS

Covalent docking;
Deep learning;
Covalent binders;
Binding pose generation;
Virtual screening;
Covalent drug design and discovery;
Covalent bond constraints;
Non-covalent scoring

Abstract The interest in covalent drugs has resurged in recent decades, spurring the development of numerous specialized computational docking tools to facilitate covalent ligand design and screening. Herein, we present CarsiDock-Cov, a new paradigm distinguishing itself as the first deep learning (DL)-guided approach for covalent docking. CarsiDock-Cov retains the core components of its non-covalent predecessor, leveraging a DL model pretrained on millions of docking complexes to predict protein−ligand distance matrices, along with a dedicated-designed geometric optimization procedure to convert these distances into refined binding poses. Additionally, it incorporates several key enhancements specifically tailored to optimize the protocol for covalent docking applications. Our approach has been extensively validated on multiple public datasets regarding the docking and screening of covalent ligands, and the results indicate that our approach not only achieves comparably improved applicability compared to its non-covalent predecessor, but also exhibits competitive performance against various state-of-the-art covalent docking tools. Collectively, our approach represents a significant advance in covalent docking methodology, offering an automated and efficient solution that shows considerable promise for accelerating covalent drug discovery and design.

*Corresponding author.
E-mail address: tingjunhou@zju.edu.cn (Tingjun Hou).

Peer review under the responsibility of Chinese Pharmaceutical Association and Institute of Materia Medica, Chinese Academy of Medical Sciences.

https://doi.org/10.1016/j.apsb.2025.07.043

1. Introduction

The last three decades have witnessed a resurgence of interest in developing covalent inhibitors, which are deliberately avoided before owing to their potential off-target reactivity and safety concerns[1-3]. Unlike conventional ligands solely relying on non-covalent interactions to bind to the desired protein target, a typical covalent inhibitor not only retains the non-bonded interactions to selectively recognize the binding region, but also contains a reactive functional group termed "warhead", capable of the covalent attachment with the nucleophilic residue (mostly cysteine or serine) within the pocket. The formation of covalent bonds with those rare and non-conserved residues significantly improves the binding potency and selectivity, and even remarkably facilitates the ligand discovery for those "undruggable" targets that usually own a shallow binding site[4-6]. Until 2019, a total of 117 (7%) covalent drugs have been approved by FDA, highlighting its critical role in the market[7].

Elucidating protein−ligand interactions has long been a grand challenge in the context of structure-based drug design (SBDD), and due to high cost and technical difficulties of experimental approaches, computational solutions like molecular docking have emerged as an effective alternative[8]. Typically relying on a search engine to provide various binding poses and a scoring function to quantify their binding strengths toward the target, molecular docking has made a tremendous contribution to early drug discovery, especially for the stages including hit identification and hit/lead optimization[9,10]. When it comes to covalent inhibitors, conventional docking algorithms usually fail to simulate the covalent binding between the ligand and protein, thus leading to a variety of docking variants customized for covalent ligands, *e.g.*, CovDock[11] built on Glide[12,13], Dockcovalent[14] and HcovDock[15] derived from DOCK[16,17], and CovalentDock[18] modified from AutoDock[19]. Additionally, some programs have also developed their specialized modules for covalent docking, represented by AutoDock[20], GOLD[21], MOE[22], FITTED[23,24] and ICM-Pro[25]. However, while some of these tools have succeeded in the recognition of new hits/leads in real-world applications[26-30], a retrospective benchmark study conducted by Scarpino et al.[31] has pointed out that most of them only exhibit moderate docking performance (re-docking success rates of 40%−60%), suggesting their vast room for improvement. Of course, recent efforts have also been made to further improve the docking accuracy by using quantum mechanical (QM) approaches to take the binding contributions of covalent attachment into account (*e.g.*, Cov_Dox[32]), or extensive sampling both the non-covalent and covalent binding poses relying on physics-based force field (*e.g.*, Attracting Cavities (AC)[33]), but these approaches are extremely time-consuming, which means that they may be more suitable for elaborate lead optimization while hardly applicable to high-throughput virtual screening (VS).

The emergence of machine learning (ML) and artificial intelligence (AI) in the past decade has dramatically invigorated the docking realm. Notable progress includes the development of a series of ML-based scoring functions for binding strength estimation[34-38], and the recent surge of deep learning (DL)-powered docking approaches that is capable of predicting protein−ligand binding poses with an end-to-end manner[39-43]. While a few doubts have been raised toward the applicability of most of current AI tools[44-47], they surely represent a promising strategy for further enhancement of docking and screening performance. But to the best of our knowledge, few of these DL models have been specifically adapted for covalent binders. Even if some approaches have the capability to occasionally obtain the correct locations and orientations for covalent binders just regarding them as non-covalent ones, they cannot automatically recognize the structural transformation of the ligand before and after the reaction, suggesting their inapplicability to accurately predict the true covalent binding complexes. The official release of AlphaFold3 very recently has further evoked a tremendous splash[48]. It is claimed that this diffusion-based model that is derived from high-profile protein structure prediction tool AlphaFold2[49] has the capability to predict the joint complex structures for almost all biomolecules, including proteins, nucleic acids, small molecules, ions and even covalent modifications. Moreover, it has exhibited substantially improved accuracy over almost all state-of-the-art docking tools for ligand binding pose prediction, which has somewhat overturned the small-molecule docking field. Nevertheless, it requires a quite large amount of prior knowledge to artificially define the reaction atoms for different ligands, making it inapplicable to the automated docking and screening of large-scale covalent compound library.

In this study, we intend to explore the application potential of our recently-released CarsiDock[50], a DL-guided non-covalent docking tool that comprises of a DL model to forecast the protein−ligand atomic distance matrices and a geometric optimization stage to reshape the distances into a credible binding pose, for covalent docking. Based on the elaborate design of CarsiDock, we further propose a specialized variant named CarsiDock-Cov, which is able to conduct automated covalent docking by integrating multiple specific functions, such as warhead detection, conformer generation, covalent attachment, binding pose prediction and binding strength estimation. Compared with CarsiDock and a variety of covalent docking predecessors, the primary methodological contributions of CarsiDock-Cov include: (1) creatively leveraging the DL models trained for non-covalent complexes to handle covalent binders, thus ensuring model generalizability; (2) incorporating covalent bond formation directly into pose reconstruction stage to enforce geometric constraints; and (3) integrating the core DL model with multiple auxiliary modules into a unified pipeline for streamlined automation. These customized settings enable our approach to achieve excellent performance on several well-established datasets regarding covalent docking and screening tasks, showing great promise for guiding the design and discovery of novel covalent binders.

2. Materials and methods

2.1. Overview of CarsiDock-Cov

Fig. 1 depicts the overall workflow of CarsiDock-Cov, which can be summarized into four key steps, including protein preparation,

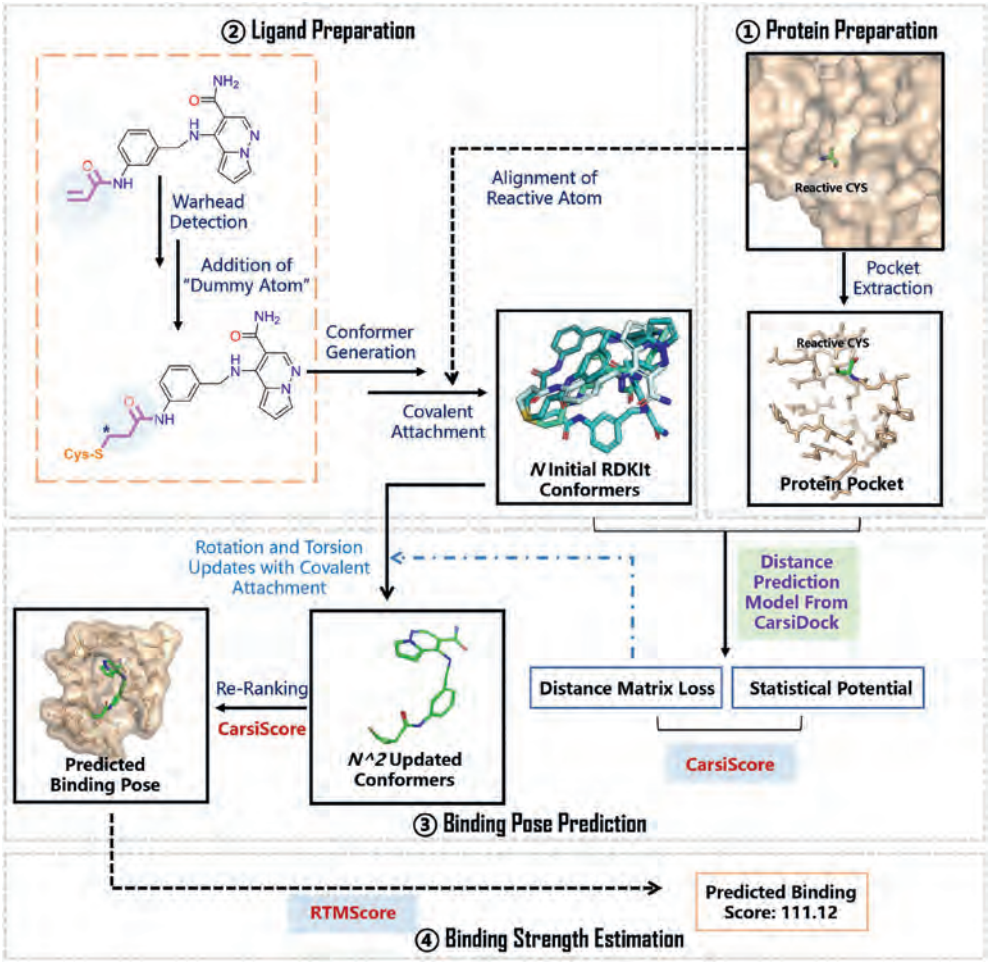

Figure 1 Overview of CarsiDock-Cov. The whole workflow includes four primary steps, *i.e.*, protein preparation, ligand preparation, binding pose prediction, and binding strength estimation. The protein and ligand should first be preprocessed to satisfy the requirements for covalent docking, and next the prepared ligand and protein are fed into a DL model inherited from CarsiDock, to learn the protein-ligand distance matrices. Then a covalent bond-constrained geometric optimization procedure is conducted to obtain multiple potential binding poses, followed by rescoring through CarsiScore and RTMScore, sequentially to obtain the final binding pose and binding strength.

ligand preparation, binding pose prediction, and binding strength estimation. CarsiDock-Cov inherits the core components of CarsiDock[50], but makes a few changes to facilitate the applicability of the whole protocol for covalent docking. Specifically, for a given ligand−protein pair, similar to most mainstream covalent docking programs, the desired nucleophilic residue in protein and the reaction type that may undergo between the ligand and the desired residue should be appointed beforehand, and then the protein and ligand are accordingly prepared to satisfy the requirements for covalent docking. In protein preparation step, the pocket defined as the residues within 6 Å around the reference ligand is extracted according to the strategy adopted in CarsiDock, along with the recording of the coordinates for the reactive atom in nucleophilic residue. As for the processing of ligand, a SMARTS string-guided substructure searching procedure is first conducted to detect potential warheads, followed by the attachment of a "dummy atom" in reactive site, whose atomic properties are set as the same as those of the reactive atom in nucleophilic residue. This procedure mimics the reaction between the ligand and the nucleophilic residue, whose transformation rule (Fig. 2) has been predefined. Multiple diverse initial conformers are then yielded for each processed ligand, with the position of "dummy atom" exactly aligned with that for reactive atom in protein, thus forcing the formation of the covalent bond between the ligand and protein.

Given that a typical covalent binder usually relies on non-bonded interactions to recognize the binding pose, and considering that retraining the model on limited covalent binding complexes may introduce biases, here we directly fetch the DL model employed in CarsiDock for the prediction of distance matrices and the parameters that determine the mixture density models. For protein pocket, all atoms are involved for the calculation of atomic and pair representations, whereas for ligand, the "dummy atom" that has the same coordinates as the reactive atom in protein is specifically excluded. Next, with the guidance of a distance matrix-derived loss function, a covalent bond-constrained geometric optimization procedure is carried out by tuning the rotation and torsion angles of the initial ligand conformers, whose objective is to force the updated distance loss to approach the one predicted by the DL model. Of note, each individual ligand conformer can lead to a specific set of distance matrices and mixture density model, and here we follow the pairwise

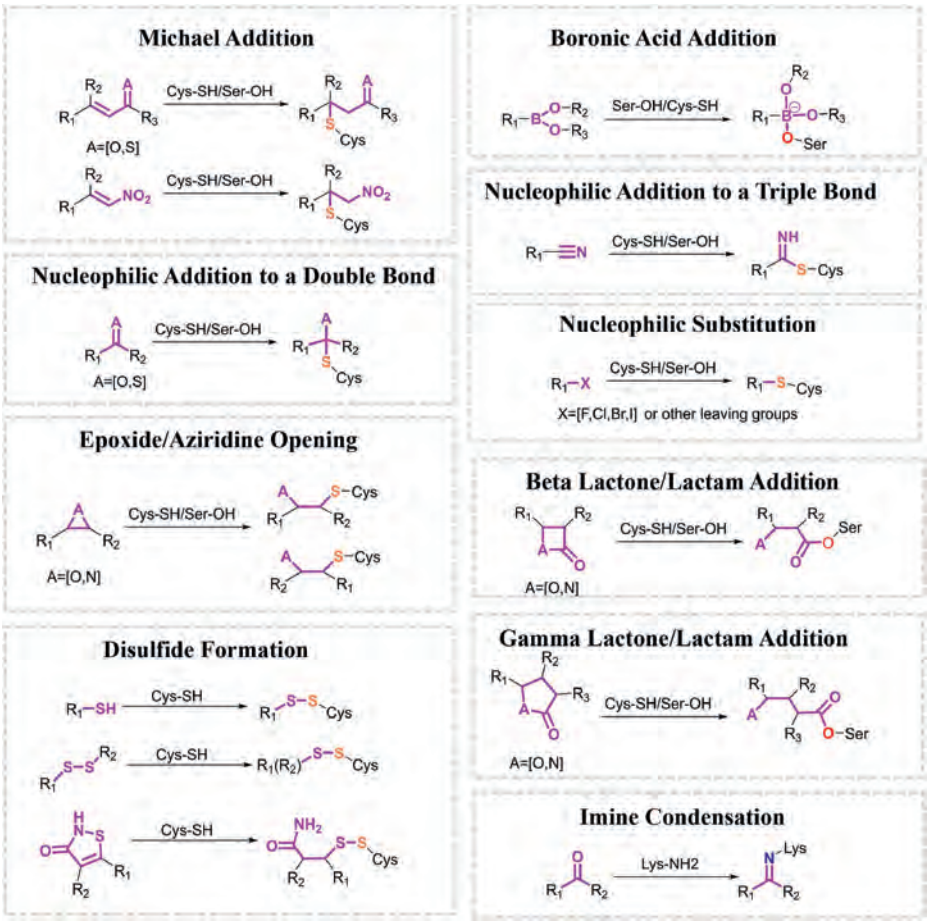

Figure 2 The routine reaction mechanisms of common reaction types supported in CarsiDock-Cov.

combination strategy proposed in CarsiDock, thus leading to *N2* binding poses with *N* initial conformers. The produced binding poses are then re-ranked by built-in CarsiScore (weighed sum of distance loss and statistical potential), with the final binding strength estimated by RTMScore[51].

2.2. *Reaction types supported by CarsiDock-Cov*

A total of 10 common reaction types are currently supported in CarsiDock-Cov, including Michael addition, nucleophilic addition to a double bond, nucleophilic addition to a triple bond, nucleophilic substitution, boronic acid addition, epoxide/aziridine opening, imine condensation, beta lactone/lactam addition, gamma lactone/lactam addition, and disulfide formation. The basic rules to convert the pre-reaction ligands into their post-reaction counterparts are summarized in Fig. 2. However, it should be noted that the products may be various for a specific pair of reagents under different protein environments, and besides the covalent attachment with the nucleophilic residue, some ligands may undergo extra reactions such as hydrolysis. The current version of CarsiDock-Cov cannot automatically detect these special reaction mechanisms, and only the most frequently-occurred one is adopted as the standard transformation rule for each reaction type. However, users can implement custom transformation rules through RDKit toolkit when necessary, enabling post-reaction structures to be directly used as input.

2.3. *Framework of distance prediction model*

CarsiDock-Cov directly utilizes the pretrained DL model embedded in CarsiDock for the prediction of protein−ligand distance matrices, which could be considered as the core step for maintaining its excellent docking accuracy. Specifically, the input ligand conformer and protein pocket are individually characterized as the form of atomic type token and intramolecular atomic distance matrix, and then fed into a Transformer-based model, which consists of two independent embedding blocks to obtain the initial atom and pair representations, two independent encoder blocks to update the representations within the ligand or protein, an interactive encoder block to further update the representation between the ligand and protein interactively, a distance prediction block to obtain both the final protein−ligand and intra-ligand atomic distance matrices, as well as an auxiliary mixture density network (MDN) block to learn the parameters that determine the mixture density models. The algorithmic details of the model architecture could refer to the original paper of CarsiDock[50].

2.4. *Covalent bond-constrained geometry optimization*

Unlike CarsiDock that allows the ligand to freely move within the pocket by tuning the parameters regarding the translations, rotations and torsions of the ligand, CarsiDock-Cov constrains the

freedoms by forcing the "dummy atom" of the ligand exactly at the position of the reactive atom in nucleophilic residue. Centered on the "dummy atom", each ligand conformer can be characterized as a vector with a length of $3 + k$, where the first three represent the Euler angles while the latter k denote the dihedral angles of the k rotatable bonds in a ligand. Then the optimization is carried out by minimizing the SmoothL1 loss between the updated distance matrices and the corresponding one output from the model for both the protein−ligand and intra-ligand atomic pairs. The optimization is executed relying on the LBFGS[52] algorithm, and it proceeds unless the loss is no longer improved in 5 successive epochs. The coordinates with the optimal loss can then be obtained by updating the rotation and torsion parameters of the specific initial conformer, thus resulting in the produced docking pose. Initialized with N ligand conformers, the optimization procedure can yield $N2$ intermediate binding poses by taking a pairwise combination of N initial conformers and N resulting distance matrices. These poses are next re-ranked by CarsiScore, a built-in scoring function in CarsiDock that can be described as the weighted sum of above loss and the statistical potential resulting from the mixture density model with the weights of 5 and -1. The final binding strength is given by RTMScore, a residue-atom distance likelihood potential that has proved its effectiveness in multiple public benchmarks[50,51,53]. Of note, neither CasiScore nor RTMScore specifically takes the contributions of covalent attachment into account, thus struggling to estimate the absolute binding energy between the ligand and protein. Nevertheless, the use of non-covalent scoring functions remains methodologically justified for covalent docking, especially for their simple uses for pose ranking or high-throughput VS. This strategy finds support in established covalent docking tools like CovDock[11] and DOCKTITE[22], whose effectiveness has been well-documented through numerous retrospective and prospective validation studies.

2.5. Details of test sets

To better compare our methodology with other covalent docking algorithms, here we employed two well-established datasets as the core docking sets, *i.e.*, the CS405 set curated by Wei et al.[32] and the CSKDE304 set proposed by Goullieux et al.[33], which consisted of 405 and 304 diverse covalent binding complexes, respectively. The former could be considered as one of the largest datasets yet for benchmarking the docking performance of covalent docking approaches, whereas the latter had undergone a more rigorous filtering and checking procedure, thus ensuring the overall higher structural quality. Of note, some complexes had been treated as non-covalent ones and included in PDBbind-v2020 general set[54], which had been widely used for the training of AI-powered docking and scoring approaches including CarsiDock and RTMScore. To evaluate our approach more objectively, we manually eliminated the duplicated PDB entries, thus leaving 207 and 166 complexes for two sets, respectively. In addition, we also retrieved a target-specific dataset named SARS-MP-76 from the study conducted by Goullieux et al.[33] This set comprised of a total of 76 crystal complexes for SARS-Cov-2 main protease with diverse ligands and reaction mechanisms, and more importantly, none of them had been collected in PDBbind-v2020, thereby serving as an extra resource to investigate the generalization of our protocol in more complex cross-docking scenarios. The detailed information of above three sets such as PDB entries, reaction types and reactive residues could be found in Supporting Information

As for screening power, due to the lack of gold benchmarks for covalent binders, here we just involved a small VS dataset originally proposed for the validation of CovDock[11]. This dataset included 3 kinase targets and 10 PDB entries, sharing10 known inhibitors and 100 decoys. The active compounds were all acrylamide Michael acceptors verified as covalent binders to the kinase domains of all three targets, and the decoys also contained the acrylamide group, thus making the dataset more challenging. Among 100 decoys, 44 were selected from the Schrödinger internal database according to the fingerprint Tanimoto similarity, while the remaining mimicked the property space by integrating 8 common molecular descriptors. The molecular structures were extracted from the literature by using the Structure Extraction module in DrugFlow[55].

2.6. Structure preparation

Since most covalent docking datasets only provided PDB entries, we directly retrieved the corresponding complex structures from the Protein Data Bank (PDB)[56] database. The crystal structures were carefully examined through the Maestro interface, and then pretreated using the Protein Preparation Wizard[57] module in Schrödinger 2021, including assigning bond orders, adding hydrogens, filling in missing side chains and loops using Prime, removing unwanted chains and waters, optimizing hydrogen bond networks, and minimizing the whole structure until the RMSD of heavy atoms averaged at 0.30 Å using the OPLS-2005[58] force field. The proteins and co-crystallized ligands were protonated using the embedded PropKa[59] and Epik[60] utilities, respectively with the pH value set to 7.0. Of note, the prepared crystal ligands were in their post-reaction forms, so we further obtained the SMILES (Simplified Molecular Input Line Entry System) strings of the original ligands by querying the CovBinderInPDB[61] and CovalentInDB[62,63] databases, thus serving as the starting point for CarsiDock-Cov.

2.7. Parameter settings for CarsiDock and CarsiDock-Cov

For both CarsiDock and CarsiDock-Cov, 5 diverse initial conformers were produced for each ligand, which meant that a total of 25 intermediate binding poses would be generated. For CarsiDock, to better calculate the docking metrics, the crystalized ligands with post-reaction forms (whose coordinates would be re-initialized using the built-in conformer generation module) were directly fetched as inputs when conducting the docking experiments. As for CarsiDock-Cov, we fed the pre-reaction forms into our protocol for most ligands, while for the ones that violated the common reaction mechanisms or whose warheads failed to be automatically recognized, we manually skipped the assumed reaction procedures and directly fed the post-reaction forms as inputs. However, some molecules that succeeded in docking might also yield multiple initial inputs due to the ambiguous reaction products (*e.g.*, some might have multiple potential warheads), and for such a special scenario, each of them was fed into the docking protocol and only the pose with the best score estimated by CarsiScore was considered as the final output.

2.8. Evaluation metrics

The quality of a docking pose was primarily measured by the routine root-mean-square deviation (RMSD) between the predicted and the native binding poses, which were calculated based

on the symmetry-corrected algorithms in spyrmsd[64] utility. If the RMSD value was within a predefined threshold (usually 2.0 Å), the sample could be marked as a successful prediction; then the docking accuracy could be quantified through the success rate (SR) by calculating the percentage of the successful cases among all the samples. Given that DL-guided docking approaches like CarsiDock could hardly produce sufficiently diverse binding poses as traditional methods, here we computed the metrics only based on the top-ranked pose predicted by each program, which shall gain the most attention in practice. As for VS, considering that the dataset involved here were small in size, we followed Zhu et al.[11] to adopt the area under the receiver operating characteristic curve (AUROC) as the major indicator, and further presented the ROC curve to intuitively demonstrate the enrichment performance, which could somewhat reflect the screening power.

3. Results and discussion

3.1. General assessment of docking accuracy

The applicability of CarsiDock to covalent binders is first estimated by assessing its capability to reproduce the near-native binding poses on two public docking datasets composed of diverse covalent binding complexes, *i.e.*, CS405 and CSKDE304 sets curated by Wei et al.[32] and Goullieux et al.[33], respectively, with the results for baseline tools directly retrieved from the corresponding literature reports. It should be noted that some PDB entries have been handled as non-covalent binding structures and collected in PDBbind-v2020 general set[54], which serves as a crucial resource for the training of CarsiDock and RTMScore. To assess our approaches more fairly, we have manually eliminated those duplicated entries, thus leaving 207 and 166 samples for two datasets, respectively. As shown in Table 1 and Fig. 3, while not specifically designed for covalent ligands, CarsiDock still demonstrates impressive docking performance in terms of both the top1 success rates and overall RMSD values. On CS405 set (Table 1 and Fig. 3A and B), CarsiDock could reproduce 30.42% and 76.81% of the top-ranked poses using an RMSD cutoff of 1.0 and 2.0 Å, respectively with the mean/median RMSD value across all the samples as low as 1.91/1.36 Å, which is dramatically superior to several specialized covalent docking programs such as CovDock[11], GOLD[21] and ICM-Pro[25]. Even in comparison to Cov_DOX[32] that relies on multiple time-consuming stages of QM calculations to extensively sample the binding poses, CarsiDock performs only a little poorer in terms of the top1 success rate with a 2.0 Å RMSD cutoff, but remains competitive regarding the other indicators. As for CarsiDock-Cov, while we can still observe its modest improvements over CarsiDock in RMSD indicators (1.75/1.28 Å *vs.* 1.91/1.36 Å), its overall superiority on this dataset is less pronounced. This limited performance gap likely reflects the inherent characteristics of the benchmark set composition. Notably, since CarsiDock already demonstrates exceptional performance on this dataset, achieving further significant enhancement with CarsiDock-Cov proves particularly challenging. These findings motivated our evaluation using additional datasets to more comprehensively assess the method's capabilities.

Table 1 Docking accuracy in terms of the top 1 success rates within different RMSD thresholds and mean/median RMSD values on the CS405 and CSKDE304 sets. The results for the approaches except for CarsiDock and CarsiDock-Cov on two datasets are retrieved from the study conducted by Wei et al.[32] and Goullieux et al.[33], respectively.

Method	Success rate (%)		RMSD (Å)	
	<1.0 Å	<2.0 Å	Mean	Median
CS405 (207)				
CovDock	23.67	54.11	2.96	1.76
GOLD	16.91	46.38	3.30	2.26
ICM-Pro	21.26	44.93	2.94	2.12
MOE	11.11	47.83	2.82	2.03
Cov_DOX	21.26	80.68	1.94	1.55
CarsiDock	30.43	76.81	1.91	1.36
CarsiDock-Cov	30.43	78.74	1.75	1.28
CSKDE304 (166)				
GOLD (PLP/R)	22.89	59.64	2.46	1.62
AutoDock (R)	9.04	30.72	4.42	3.57
AC (S/R)	48.80	70.48	1.71	1.02
CarsiDock	26.51	72.89	1.91	1.40
CarsiDock-Cov	32.53	77.11	1.70	1.26

The results on CSKDE304 set demonstrates that the involvement of covalent attachment in CarsiDock-Cov indeed improves the docking accuracy for CarsiDock (Table 1 and Fig. 3C and D). Using a RMSD threshold of 1.0 and 2.0 Å, the top1 success rates increase from 26.51% and 72.89% to 32.53% and 77.11%, respectively, and the corresponding mean/median RMSD value decreases from 1.91/1.40 Å to 1.70/1.26 Å. The superior performance could be further proved through the cumulative distribution plot (Fig. 3C), where CarsiDock-Cov consistently runs ahead of CarsiDock under almost all RMSD thresholds. In comparison to the results reported by Goullieux et al.[33], CarsiDock-Cov performs remarkably superior to the covalent docking variants implemented in commonly-used GOLD and AutoDock[20], but no better than their dedicatedly-designed AC, a protocol that relies on physics-based force field to extensively sample both the non-covalent and covalent binding poses. Further analysis in terms of the cumulative distribution plot (Fig. 3C) indicates that AC tends to predict the binding poses more precisely (better performance under low RMSD cutoff), whereas our CarsiDock-Cov has the capability to provide acceptable poses (*e.g.*, RMSD<2.0 Å) for more diverse cases. Given the significantly higher computational costs needed for AC, CarsiDock-Cov may be a promising complement, serving as a more general protocol for covalent drug design and discovery.

3.2. Analysis of distance between the two reactive atoms

While the RMSD indicator could well reflect the overall docking quality for a binding pose, it is not comprehensive enough for covalent binders since it could hardly recognize the status for the reactive atoms, which seem subtle but are rather crucial for covalent binding. A pivotal premise for the formation of a covalent bond is that two reactive atoms shall be close enough, so we specifically analyze the distance between the two reactive atoms for the poses produced by CarsiDock and CarsiDock-Cov. Here we merge the samples in two datasets together and collect a total of 248 complexes with non-repetitive PDB entries. As depicted in Fig. 4A, thanks to the inheriting of bond length from RDKit-initialized ligand conformers, almost all the poses predicted by CarsiDock-Cov could produce an appropriate distance between the two reactive atoms, which is well consistent with that observed

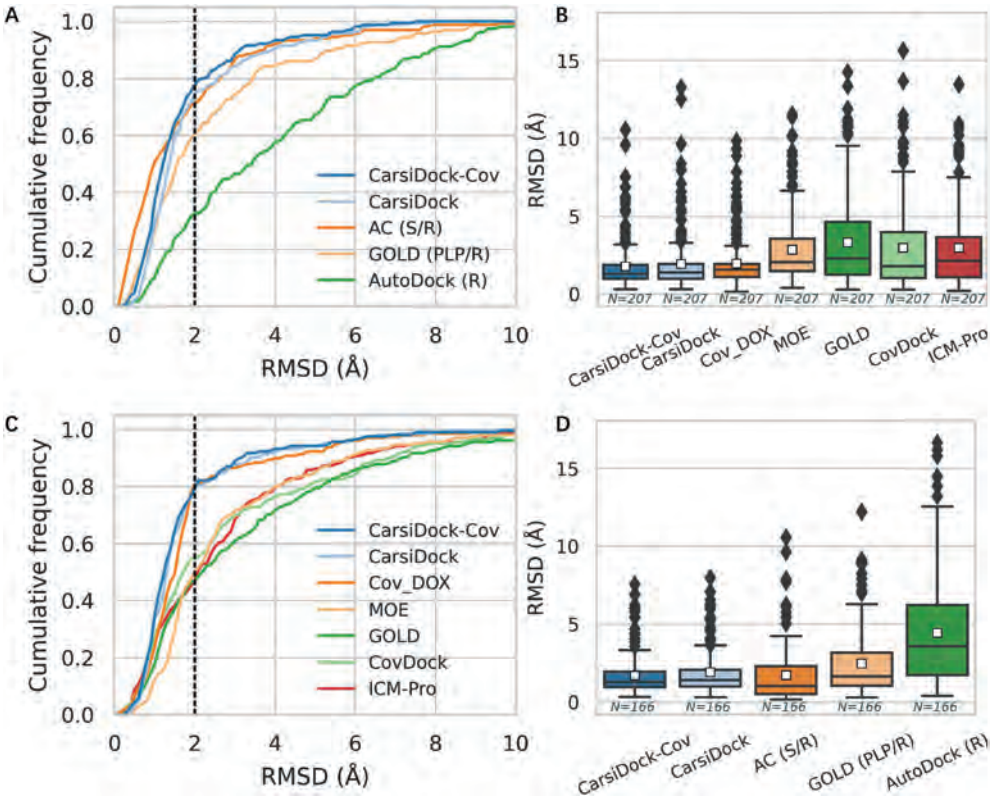

Figure 3 Prediction accuracy of several covalent docking programs based on (A, B) CS405 and (C, D) CSKDE304 sets in terms of (A, C) cumulative distributions of the RMSD values and (B, D) overall RMSD distributions. The dotted lines in the cumulative distribution plots denote a RMSD cutoff of 2.0 Å, whereas the white square in the box plot represents the average of the statistic. The data except for CarsiDock-Cov and CarsiDock on these two datasets are collected from Refs. 32 and 33, respectively.

in crystalized binding poses. But for the poses generated by CarsiDock, the corresponding distances are varying, and the average distances for several atomic pairs (*e.g.*, S−C, S−S and N−C) are remarkably higher than those in native binding poses. These findings may somewhat account for the slightly worse docking accuracy of CarsiDock, and on the other hand, also demonstrate the importance of the covalent bond-constrained geometric optimization procedure involved in CarsiDock-Cov.

Our inference could be further exemplified through three typical cases, as shown in Fig. 4B−G. For 4m1y (Fig. 4C), while the benzothiadiazole moiety of the ligand could be successfully reproduced by CarsiDock, the orientation of the alkylsulfonyl tail is completely opposite; but for CarsiDock-Cov, the formation of covalent bond could force the tail to approach the reactive Cys12, thus enhancing the alignment between the predicted and crystalized binding poses (Fig. 4B). A similar phenomenon could also be observed for 3pdf, where the benzene ring of the ligand could be well predicted by CarsiDock whereas the other more flexible region is poorly aligned (Fig. 4E). Similarly, covalent attachment could fix this flexible region, facilitating the superior matching of the two poses (Fig. 4D). The pose produced by CarsiDock for 4hax is more extreme, in which the whole structure is in an opposite orientation due to the near-symmetric overall structure of the ligand (Fig. 4G). The constraint induced by covalent bond drives the pose to turn around, contributing to its more reasonable orientation and location.

3.3. Impact of multiple factors on docking accuracy

We further investigate the impact of three potentially important factors, *i.e.*, reaction type, type of reactive residue, and number of rotatable bonds in ligand on docking accuracy based on the aforementioned merged dataset. Here we only focus on the differences between CarsiDock-Cov and CarsiDock, while other approaches such as Cov_DOX and AC are excluded for comparison due to their missing data for part of the entries.

In principle, CarsiDock-Cov and CarsiDock shall be insusceptible to reaction type since reaction has occurred before the formal docking procedure, but we can still observe the varying performance among different reaction types (Fig. 5A and B). For the types that contain more than 10 samples, the cases that undergo a reaction like addition to carbonyl (double bond), addition to nitrile (triple bond) and beta lactam addition seem much easier to be predicted, whereas the types like disulfide formation and nucleophilic substitution are rather harder. The fluctuation in performance may be majorly attributed to the limited samples and various reaction environments. But fortunately, the top1 success rates for all the cases are still above 50%, with the corresponding median RMSD values all below 2.0 Å, which could somewhat indicate the robustness of our approaches. As for the relative performance between the two methods, except for the types including nucleophilic substitution, disulfide formation and imine condensation, CarsiDock-Cov consistently

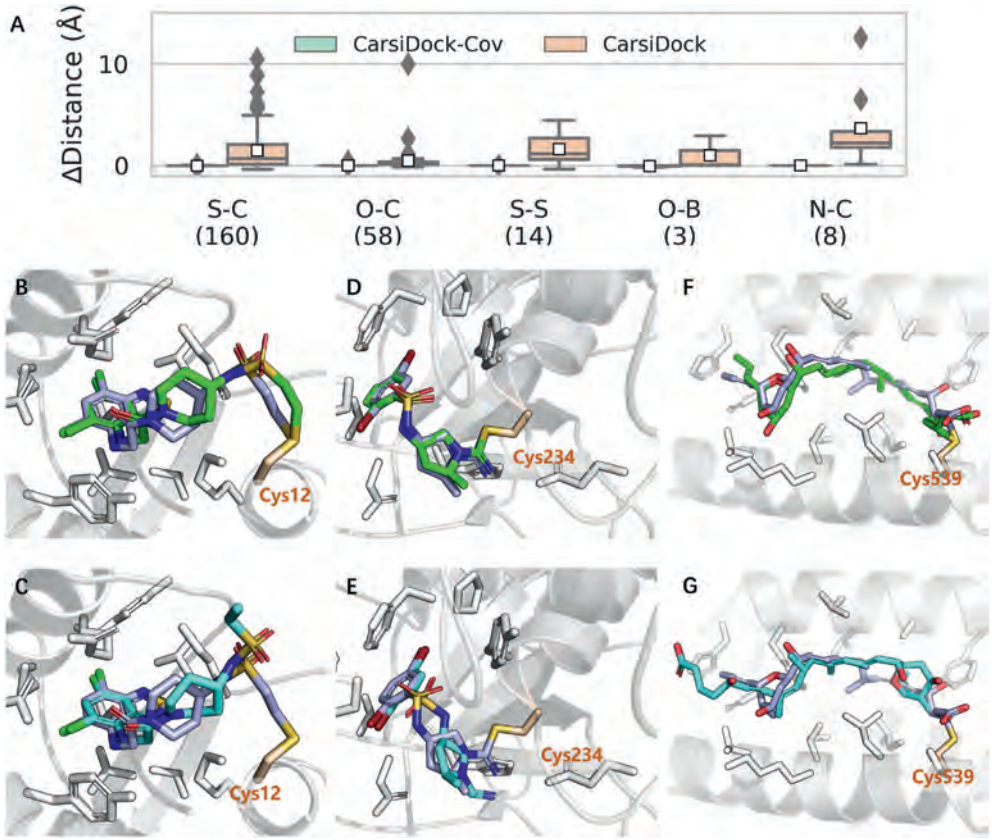

Figure 4 Analysis of the distance between the two reactive atoms that are ready to form a covalent bond. (A) The distance difference between the crystal binding poses and the poses predicted by different docking programs. (B) Binding pose for 4mly predicted by CarsiDock-Cov. (C) Binding pose for 4m1y predicted by CarsiDock. (D) Binding pose for 3pdf predicted by CarsiDock-Cov. (E) Binding pose for 3pdf predicted by CarsiDock. (F) Binding pose for 4hax predicted by CarsiDock-Cov. (G) Binding pose for 4hax predicted by CarsiDock. The white square in the box plot represents the average of the statistic, and the crystalized binding poses, the poses predicted by CarsiDock, and the poses generated by CarsiDock-Cov are colored in light blue, cyan and green, respectively.

outperforms CarsiDock, highlighting the critical role of covalent constraints in docking covalent binders.

The type of reactive residues also exerts a large influence on docking performance, illustrated as Fig. 5C and D. Overall, the poses reacting with Ser and Cys have larger chance to be successfully reproduced while the result for rare reactive residues such as Lys is relatively unsatisfactory. Additionally, here the superiority of CarsiDock-Cov over CarsiDock is majorly embodied in Cys, while for Ser and Lys, the improvement is not so distinct. As for these, we should recognize that Ser is dominated by beta lactam addition and Lys only undergoes imine condensation, so the impact of residue types is actually consistent with that for reaction types.

Ligand flexibility, which is explicitly reflected through the number of rotatable bonds or torsion angles, directly determines the searching freedoms for those traditional search-based docking approaches, and their docking accuracy decreases remarkably when the ligands contain excessive rotatable bonds (*e.g.*, >15)[32,50]. Nevertheless, while CarsiDock and CarsiDock-Cov also need to tune the torsion angles at the stage of binding pose reconstruction, their performance on covalent binders seems not so sensitive to ligand flexibility (Fig. 5E and F), which may be attributed to the fact that they have learned sufficient binding information from large-scale non-covalent binding training data. As for the comparison of CarsiDock-Cov and CarsiDock, except for the extremely flexible cases whose rotatable bonds are more than 15 (just one successful case less), CarsiDock-Cov always takes the lead, which further demonstrates its wide applicability.

3.4. Docking performance toward SARS-CoV-2 main protease

The performance of our approaches is also estimated on an external dataset named SARS-MP-76[33], which consists of a total of 76 covalent-binding crystal complexes for SARS-Cov-2 main protease with diverse ligands and reaction mechanisms. Since the protein pocket is incredibly shallow and solvent exposed (Fig. 6A), a variety of ligands have been designed to covalently bind to the reactive Cys145. In addition, the protease comprises of four sub-pockets for different ligands to occupy, thus making the docking extremely challenging.

We first conduct a routine redocking experiment by redocking each covalent ligand into its corresponding pocket with the same PDB entry. As summarized in Table 2 and Fig. 6B, while substantial decrease is observed for both CarsiDock and CarsiDock-Cov in comparison to the results on CS405 and CSKDE304 sets (Table 1 and Fig. 3A−D), CarsiDock-Cov still consolidates its advantages over CarsiDock, whether in terms of top1 success rates (34.21/64.47% *vs* 22.37/56.58%) or overall RMSD values (1.91/1.26 Å *vs* 2.41/1.82 Å). Moreover, CarsiDock-Cov even

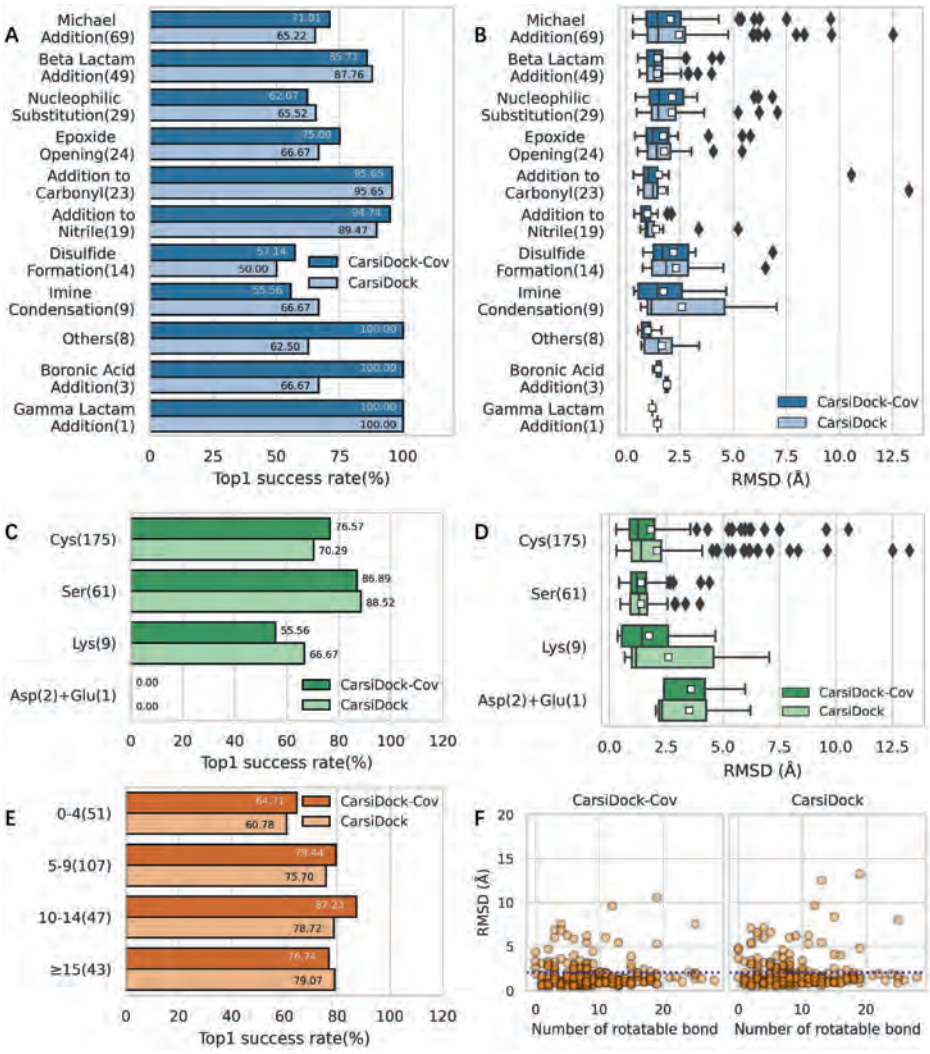

Figure 5 The impact of (A, B) reaction type, (C, D) type of reactive residue, and (E, F) number of rotatable bonds in ligand on (A, C, E) the top1 success rates with a RMSD cutoff of 2.0 Å and (B, D, F) overall RMSD distributions based on a dataset involving the samples from both the CS405 and CSKDE304 sets. The white square in the box plot represents the average of the statistics, and the dotted lines in the scatter plots denote a cutoff of 2.0 Å.

outperforms AC regarding all the four indicators, suggesting its excellent generation on this specific target.

We then design a more challenging scenario by docking the specific ligand into the pockets of the other 75 PDB entries. Goullieux et al.[33] have also conducted such a cross-docking experiment for GOLD and AC, but only used 7c7p as docking template. This test may be biased since different approaches may obtain varying results based on different PDB entries. As outlined in Table 2[33] both CarsiDock and CarsiDock-Cov cannot obtain satisfactory result based on this specific PDB entry (*i.e.*, 7c7p), but remain competitive if a more suitable PDB entry (*e.g.*, 5rfw) is adopted. To make a fairer comparison between CarsiDock-Cov and CarsiDock, we depict the results for all the 76 groups in Fig. 6C−E. Specifically, CarsiDock-Cov outperforms CarsiDock for 53 of 76 cases in terms of the top1 success rate (RMSD <2.0 Å), and the lead is even extended to 67 and 68 cases, respectively when the mean and median RMSD values are employed as the final indicators. Even if using the average result of all the 76 groups to indicate the performance (Table 2), CarsiDock-Cov is still far ahead of CarsiDock (1.81/14.05% *vs* 1.23/11.04% in terms of top1 success rates and 4.81/4.59 Å *vs* 5.37/5.18 Å regarding the mean/median RMSD values).

3.5. *Assessment of screening power*

One of the most practical application of docking programs in real-world drug design campaigns is to identify potential hits/leads from chemical library, so whether the approach can distinguish active compounds from those inactive ones also warrants attention. Hence, we further test CarsiDock-Cov and CarsiDock on a miniature VS dataset originally employed for the validation of CovDock[11], and summarize the results in terms of AUROC in Table 3[11,15]. Additionally, the results for HCovDock[15] and CovDock are also retrieved from the original papers for comparison. Interestingly, even if directly feeding the pre-reaction forms of the ligands as inputs (*i.e.*, completely relying on non-covalent version of CarsiDock for VS), CarsiDock can still obtain an average AUROC of 0.895, which is significantly superior to HCovDock (AUROC = 0.798) and CovDock (AUROC = 0.801). Converting the ligands into their post-reaction forms further improves the

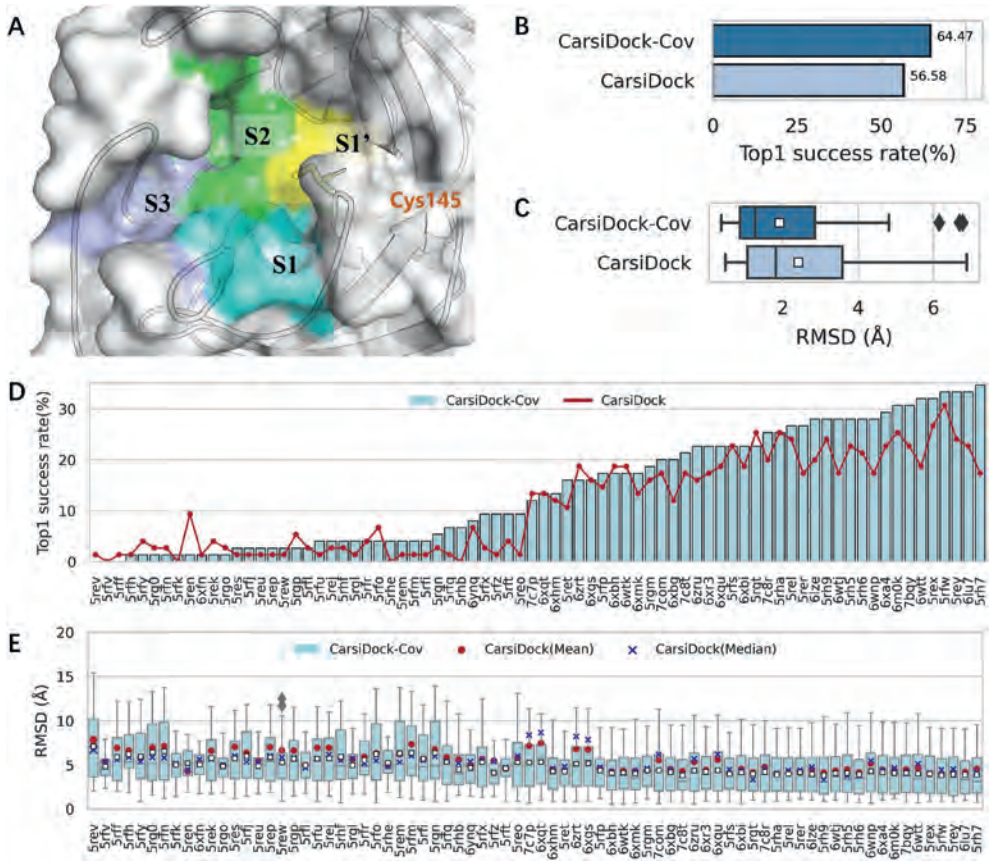

Figure 6 Docking performance toward SARS-CoV-2 Main Protease. (A) The binding site of SARS-CoV-2 main protease (PDB entry: 7c7p) with its 4 subpockets highlighted. (B) Re-docking results on SARS-MP-76 set in terms of the top1 success rates with a RMSD cutoff of 2.0 Å. (C) Re-docking results on SARS-MP-76 set in terms of the overall RMSD distributions. (D) Cross-docking results on SARS-MP-76 set in terms of the top1 success rates with a RMSD cutoff of 2.0 Å. (E) Cross-docking results on SARS-MP-76 set in terms of the overall RMSD distributions. The white square in the box plot represents the average of the statistic, and red circles and blue " × " denote the mean and median statistics of CarsiDock, respectively.

Table 2 Docking accuracy in terms of the top 1 success rates within different RMSD thresholds and mean/median RMSD values on the SARS-MP-76 set regarding both the re-docking and cross-docking scenarios. The results for GOLD and AC are retrieved from the study conducted by Goullieux et al.[33]

Method	PDB templates	Success rate (%)		RMSD (Å)	
		<1.0 Å	<2.0 Å	Mean	Median
Re-Docking					
GOLD	/	19.74	44.74	2.63	2.58
AC	/	30.26	57.89	2.25	1.62
CarsiDock	/	22.37	56.58	2.41	1.82
CarsiDock-Cov	/	34.21	64.47	1.91	1.26
Cross-Docking					
GOLD	7c7p	1.33	16.00	5.28	6.28
AC	7c7p	2.67	26.67	4.45	4.21
CarsiDock	7c7p	0.00	13.33	7.17	8.38
CarsiDock-Cov	7c7p	5.33	12.00	5.27	5.33
CarsiDock	5rfw	5.33	30.67	4.00	4.42
CarsiDock-Cov	5rfw	9.33	33.33	3.81	3.69
CarsiDock	Average	1.23	11.04	5.37	5.18
CarsiDock-Cov	Average	1.81	14.05	4.87	4.59

Table 3 Screening performance of multiple approaches on a dataset consisting of 3 kinase targets and 10 PDB entries.

PDB	Target	CarsiDock-Cov	CarsiDock-post[a]	CarsiDock[a]	HCovDock[b]	CovDock-pre[c]	CovDock-post[c]	CovDock-average[c]
2hwo	cSrc	0.988	0.937	0.975	0.736	0.677	0.626	0.780
2hwp	cSrc	0.979	0.957	0.961	0.756	0.804	0.873	0.858
2qlq	cSrc	0.993	0.948	0.884	0.878	0.802	0.835	0.825
2qq7	cSrc	0.981	0.963	0.958	0.753	0.732	0.824	0.836
3lok	cSrc	0.980	0.942	0.923	0.785	0.956	0.816	0.864
2j5e	EGFR	0.996	0.902	0.869	0.888	0.862	0.955	0.950
2j5f	EGFR	0.993	0.939	0.943	0.845	0.764	0.842	0.829
2jiv	EGFR	0.982	0.768	0.819	0.701	0.942	0.906	0.941
3ika	EGFR	0.985	0.789	0.740	0.875	0.614	0.735	0.701
3v6s	JNK	0.892	0.937	0.880	0.758	0.448	0.395	0.428
Average		0.977	0.908	0.895	0.798	0.760	0.781	0.801

[a]CarsiDock-post means that the warheads for ligands have been correspondingly prepared according to the reaction mechanism, while CarsiDock employs the original ligands for docking.
[b]Results for HCovDock are retrieved from Ref. 15.
[c]Results for CovDock are retrieved from Ref. 11.

indicator to 0.908, and the involvement of covalent bond-constrained geometric optimization (*i.e.*, CarsiDock-Cov) even enhances the VS power to a new level, achieving the optimal performance in 9 of 10 PDB entries with the average AUROC high up to 0.977. Further analysis on the ROC curves (Fig. 7) indicates that the enhancement of performance not only comes from the better distinguishing of negative samples but also from the enrichment of active compounds. These findings suggest that our approaches are indeed applicable to docking-based VS, showing great promise for the screening of potential covalent binders.

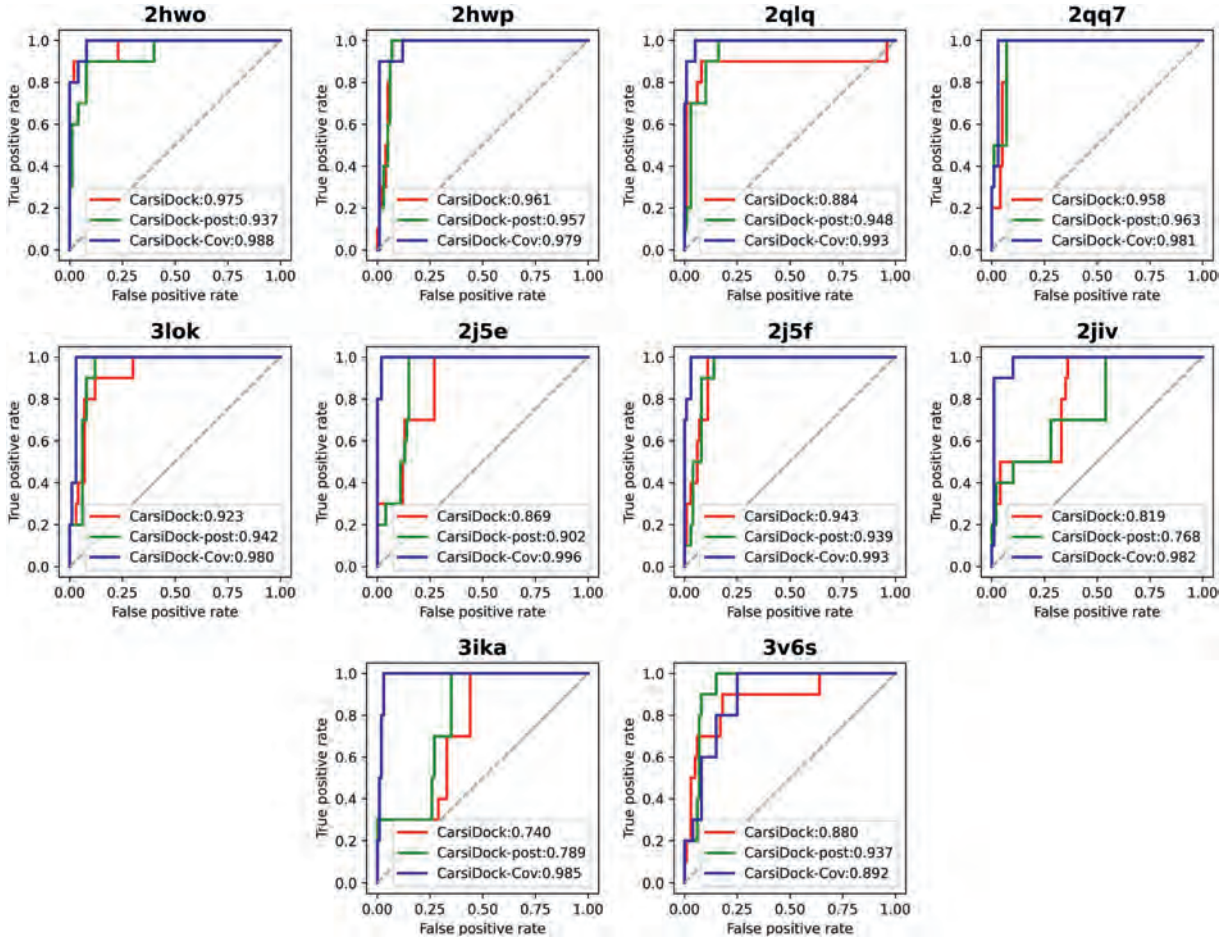

Figure 7 ROC curves of different VS strategies on 10 datasets based on different PDB templates.

4. Conclusions

In this work, we present a DL-guided covalent docking approach adapted from our recently-developed CarsiDock, where a covalent bond-constrained geometric optimization procedure is proposed to restrict the search freedoms when reconstructing the distance matrices output from the DL model into the final binding poses. By integrating non-covalent scoring with covalent constraints, our method demonstrates significantly enhanced applicability for covalent binders, particularly in more challenging cross-docking and VS scenarios. Additionally, the involvement of multiple customized functions such as warhead detection, protein processing and covalent attachment streamlines the whole workflow, making it feasible for automated covalent docking. However, we need to clarify that the current framework has limitations in accurately predicting absolute binding affinities, primarily due to incomplete treatment of covalent bond contributions in scoring. A potential solution could involve combining our pose predictions with quantum mechanics/molecular mechanics (QM/MM)-based binding affinity calculations[65-67], which could be a direction for future research.

Acknowledgments

This study was supported by National Natural Science Foundation of China (22220102001, 22303081), China Postdoctoral Science Foundation (2024M762886) and Zhejiang Provincial Natural Science Foundation of China (LMS25B030003). Additionally, the authors would like to thank Prof. Lei Xu at the Jiangsu University of Technology for preparing the structures employed in this study based on the Schrödinger software, which has made a significant contribution to our research.

Author contributions

Tingjun Hou and Chao Shen designed the research study. Chao Shen developed the method and wrote the code. Chao Shen, Hongyan Du, Xujun Zhang, Sukai Gu, Heng Cai, Yu Kang, Peichen Pan and Qingwei Zhao performed the analysis. Chao Shen and Tingjun Hou wrote the paper. All authors read and approved the manuscript.

Data and code availability

The prepared datasets employed for the validation of CarsiDock-Cov are available at https://zenodo.org/records/14064834, and the codes of CarsiDock-Cov are available at https://github.com/sc8668/CarsiDock-Cov.

Conflicts of interest

There are no conflicts to declare.

References

1. Singh J, Petter RC, Baillie TA, Whitty A. The resurgence of covalent drugs. *Nat Rev Drug Discov* 2011;**10**:307−17.
2. De Cesco S, Kurian J, Dufresne C, Mittermaier AK, Moitessier N. Covalent inhibitors design and discovery. *Eur J Med Chem* 2017;**138**:96−114.
3. Boike L, Henning NJ, Nomura DK. Advances in covalent drug discovery. *Nat Rev Drug Discov* 2022;**21**:881−98.
4. Xie X, Yu T, Li X, Zhang N, Foster LJ, Peng C, et al. Recent advances in targeting the "undruggable" proteins: from drug discovery to clinical trials. *Signal Transduct Tar* 2023;**8**:335.
5. Moore AR, Rosenberg SC, McCormick F, Malek S. RAS-targeted therapies: is the undruggable drugged?. *Nat Rev Drug Discov* 2020;**19**:533−52.
6. Ostrem JM, Peters U, Sos ML, Wells JA, Shokat KM. K-Ras (G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* 2013;**503**:548−51.
7. Singh J. The ascension of targeted covalent inhibitors. *J Med Chem* 2022;**65**:5886−901.
8. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov* 2004;**3**:935−49.
9. Ballante F, Kooistra AJ, Kampen S, de Graaf C, Carlsson J. Structure-based virtual screening for ligands of G protein-coupled receptors: what can molecular docking do for you?. *Pharmacol Rev* 2021;**73**:1698−736.
10. Sabe VT, Ntombela T, Jhamba LA, Maguire GE, Govender T, Naicker T, et al. Current trends in computer aided drug design and a highlight of drugs discovered *via* computational techniques: a review. *Eur J Med Chem* 2021;**224**:113705.
11. Zhu K, Borrelli KW, Greenwood JR, Day T, Abel R, Farid RS, et al. Docking covalent inhibitors: a parameter free approach to pose prediction and scoring. *J Chem Inf Model* 2014;**54**:1932−40.
12. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 2004;**47**:1739−49.
13. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, et al. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein−ligand complexes. *J Med Chem* 2006;**49**:6177−96.
14. London N, Miller RM, Krishnan S, Uchida K, Irwin JJ, Eidam O, et al. Covalent docking of large libraries for the discovery of chemical probes. *Nat Chem Biol* 2014;**10**:1066−72.
15. Wu Q, Huang S-Y. HCovDock: an efficient docking method for modeling covalent protein−ligand interactions. *Briefings Bioinf* 2023;**24**:bbac559.
16. Coleman RG, Carchia M, Sterling T, Irwin JJ, Shoichet BK. Ligand pose and orientational sampling in molecular docking. *PLoS One* 2013;**8**:e75992.
17. Ewing TJ, Makino S, Skillman AG, Kuntz ID. D 4.0: search strategies for automated molecular docking of flexible molecule databases. *J Comput Aided Mol Des* 2001;**15**:411−28.
18. Ouyang X, Zhou S, Su CTT, Ge Z, Li R, Kwoh CK. CovalentDock: automated covalent docking with parameterized covalent linkage energy estimation and molecular geometry constraints. *J Comput Chem* 2013;**34**:326−36.
19. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem* 2009;**30**:2785−91.
20. Bianco G, Forli S, Goodsell DS, Olson AJ. Covalent docking using autodock: two-point attractor and flexible side chain methods. *Protein Sci* 2016;**25**:295−301.
21. Jones G, Willett P, Glen RC, Leach AR, Taylor R. Development and validation of a genetic algorithm for flexible docking. *J Mol Biol* 1997;**267**:727−48.
22. Scholz C, Knorr S, Hamacher K, Schmidt B. DOCKTITE—a highly versatile step-by-step workflow for covalent docking and virtual screening in the molecular operating environment. *J Chem Inf Model* 2015;**55**:398−406.
23. Labarre A, Stille JK, Patrascu MB, Martins A, Pottel J, Moitessier N. Docking ligands into flexible and solvated macromolecules. 8. Forming new bonds— challenges and opportunities. *J Chem Inf Model* 2022;**62**:1061−77.

24. De Cesco S, Deslandes S, Therrien E, Levan D, Cueto M, Schmidt R, et al. Virtual screening and computational optimization for the discovery of covalent prolyl oligopeptidase inhibitors with activity in human cells. *J Med Chem* 2012;**55**:6306−15.
25. Abagyan R, Totrov M, Kuznetsov D. ICM—a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. *J Comput Chem* 1994;**15**: 488−506.
26. Schröder Jr, Klinger A, Oellien F, Marhöfer RJ, Duszenko M, Selzer PM. Docking-based virtual screening of covalently binding ligands: an orthogonal lead discovery approach. *J Med Chem* 2013;**56**: 1478−90.
27. Shraga A, Olshvang E, Davidzohn N, Khoshkenar P, Germain N, Shurrush K, et al. Covalent docking identifies a potent and selective MKK7 inhibitor. *Cell Chem Biol* 2019;**26**:98−108.e5.
28. Bensinger D, Stubba D, Cremer A, Kohl V, Waßmer T, Stuckert J, et al. Virtual screening identifies irreversible FMS-like tyrosine kinase 3 inhibitors with activity toward resistance-conferring mutations. *J Med Chem* 2019;**62**:2428−46.
29. Liu H, Shen C, Li H, Hou T, Yang Y. Discovery of potent covalent CRM1 inhibitors *via* a customized structure-based virtual screening pipeline and bioassays. *J Chem Inf Model* 2024;**64**:7422−31.
30. Amendola G, Ettari R, Previti S, Di Chio C, Messere A, Di Maro S, et al. Lead discovery of SARS-CoV-2 main protease inhibitors through covalent docking-based virtual screening. *J Chem Inf Model* 2021;**61**: 2062−73.
31. Scarpino A, Ferenczy GrG, Keserű GrM. Comparative evaluation of covalent docking tools. *J Chem Inf Model* 2018;**58**:1441−58.
32. Wei L, Chen Y, Liu J, Rao L, Ren Y, Xu X, et al. Cov_DOX: a method for structure prediction of covalent protein−ligand bindings. *J Med Chem* 2022;**65**:5528−38.
33. Goullieux M, Zoete V, Röhrig UF. Two-step covalent docking with attracting cavities. *J Chem Inf Model* 2023;**63**:7847−59.
34. Li H, Sze KH, Lu G, Ballester PJ. Machine-learning scoring functions for structure-based drug lead optimization. *Wiley Interdiscip Rev Comput Mol Sci* 2020;**10**:e1465.
35. Shen C, Ding J, Wang Z, Cao D, Ding X, Hou T. From machine learning to deep learning: advances in scoring functions for protein−ligand docking. *Wiley Interdiscip Rev Comput Mol Sci* 2020;**10**:e1429.
36. Moon S, Zhung W, Kim WY. Toward generalizable structure-based deep learning models for protein−ligand interaction prediction: challenges and strategies. *Wiley Interdiscip Rev Comput Mol Sci* 2024; **14**:e1705.
37. Li H, Sze KH, Lu G, Ballester PJ. Machine-learning scoring functions for structure-based virtual screening. *Wiley Interdiscip Rev Comput Mol Sci* 2021;**11**:e1478.
38. Shen C, Han X, Cai H, Chen T, Kang Y, Pan P, et al. Improving the reliability of language model-predicted structures as docking targets through geometric graph learning. *J Med Chem* 2025;**68**:1956−69.
39. Zhang X, Shen C, Zhang H, Kang Y, Hsieh CY, Hou T. Advancing ligand docking through deep learning: challenges and prospects in virtual screening. *Acc Chem Res* 2024;**57**:1500−9.
40. Corso G, Stärk H, Jing B, Barzilay R, Jaakkola T. Diffdock: diffusion steps, twists, and turns for molecular docking. *arXiv preprint arXiv: 221001776* 2022. Available from: https://doi.org/10.48550/arXiv.2210.01776.
41. Stärk H, Ganea O, Pattanaik L, Barzilay R, Jaakkola T. Equibind: geometric deep learning for drug binding structure prediction. In: *International conference on machine learning*. PMLR; 2022.
42. Zhang X, Zhang O, Shen C, Qu W, Chen S, Cao H, et al. Efficient and accurate large library ligand docking with KarmaDock. *Nat Comput Sci* 2023;**3**:789−804.
43. Lu W, Zhang J, Huang W, Zhang Z, Jia X, Wang Z, et al. DynamicBind: predicting ligand-specific protein−ligand complex structure with a deep equivariant generative model. *Nat Commun* 2024;**15**:1071.
44. Shen C, Hu Y, Wang Z, Zhang X, Pang J, Wang G, et al. Beware of the generic machine learning-based scoring functions in structure-based virtual screening. *Briefings Bioinf* 2021;**22**:bbaa070.
45. Volkov M, Turk JA, Drizard N, Martin N, Hoffmann B, Gaston-Mathé Y, et al. On the frustration to predict binding affinities from protein−ligand structures with deep neural networks. *J Med Chem* 2022;**65**:7946−58.
46. Buttenschoen M, Morris GM, Deane CM. PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chem Sci* 2024;**15**:3130−9.
47. Gu S, Shen C, Zhang X, Sun H, Cai H, Luo H, et al. Benchmarking AI-powered docking methods from the perspective of virtual screening. *Nat Mach Intell* 2025;**7**:509−20.
48. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;**630**:493−500.
49. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;**596**:583−9.
50. Cai H, Shen C, Jian T, Zhang X, Chen T, Han X, et al. CarsiDock: a deep learning paradigm for accurate protein−ligand docking and screening based on large-scale pre-training. *Chem Sci* 2024;**15**: 1449−71.
51. Shen C, Zhang X, Deng Y, Gao J, Wang D, Xu L, et al. Boosting protein−ligand binding pose prediction and virtual screening based on residue−atom distance likelihood potential and graph transformer. *J Med Chem* 2022;**65**:10691−706.
52. Liu DC, Nocedal J. On the limited memory BFGS method for large scale optimization. *Math Program* 1989;**45**:503−28.
53. Shen C, Zhang X, Hsieh CY, Deng Y, Wang D, Xu L, et al. A generalized protein−ligand scoring framework with balanced scoring, docking, ranking and screening powers. *Chem Sci* 2023;**14**:8129−46.
54. Su M, Yang Q, Du Y, Feng G, Liu Z, Li Y, et al. Comparative assessment of scoring functions: the CASF-2016 update. *J Chem Inf Model* 2018;**59**:895−913.
55. Shen C, Song J, Hsieh CY, Cao D, Kang Y, Ye W, et al. DrugFlow: an AI-driven one-stop platform for innovative drug discovery. *J Chem Inf Model* 2024;**64**:5381−91.
56. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The protein data bank. *Nucleic Acids Res* 2000;**28**:235−42.
57. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* 2013;**27**:221−34.
58. Jorgensen WL, Maxwell DS, Tirado Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 1996;**118**: 11225−36.
59. Olsson MH, Søndergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pK_a predictions. *J Chem Theor Comput* 2011;**7**:525−37.
60. Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchimaya M. Epik: a software program for pK_a prediction and protonation state generation for drug-like molecules. *J Comput Aided Mol Des* 2007;**21**:681−91.
61. Guo XK, Zhang Y. CovBinderInPDB: a structure-based covalent binder database. *J Chem Inf Model* 2022;**62**:6057−68.
62. Du H, Gao J, Weng G, Ding J, Chai X, Pang J, et al. CovalentInDB: a comprehensive database facilitating the discovery of covalent inhibitors. *Nucleic Acids Res* 2021;**49**:D1122−9.
63. Du H, Zhang X, Wu Z, Zhang O, Gu S, Wang M, et al. CovalentInDB 2.0: an updated comprehensive database for structure-based and ligand-based covalent inhibitor design and screening. *Nucleic Acids Res* 2025;**53**:D1322−7.
64. Meli R, Biggin PC. spyrmsd: symmetry-corrected RMSD calculations in Python. *J Cheminf* 2020;**12**:49.
65. Arafet K, Serrano Aparicio N, Lodola A, Mulholland AJ, González FV, Świderek K, et al. Mechanism of inhibition of SARS-CoV-2 M pro by N3 peptidyl Michael acceptor explained by QM/MM simulations and design of new derivatives with tunable chemical reactivity. *Chem Sci* 2021;**12**:1433−44.

66. Dos Santos AM, Oliveira ARS, da Costa CH, Kenny PW, Montanari CA, Varela JdJG, et al. Assessment of reversibility for covalent cysteine protease inhibitors using quantum mechanics/molecular mechanics free energy surfaces. *J Chem Inf Model* 2022;**62**: 4083−94.

67. Giudetti G, Polyakov I, Grigorenko BL, Faraji S, Nemukhin AV, Krylov AI. How reproducible are QM/MM simulations? Lessons from computational studies of the covalent inhibition of the SARS-CoV-2 main protease by carmofur. *J Chem Theor Comput* 2022;**18**:5056−67.