Journal of Cheminformatics

SOFTWARE

Open Access

Moldina: a fast and accurate search algorithm for simultaneous docking of multiple ligands

Radek Halfar[1*], Jiří Damborský[2,3], Sérgio M. Marques[2,3*] and Jan Martinovič[1*]

Abstract

Protein-ligand docking is a computational method routinely used in many structural biology applications. It usually involves one receptor and one ligand. The docking of multiple ligands, however, can be important in several situations, such as the study of synergistic effects, substrate and product inhibition, or competitive binding. This can be a challenging and computationally demanding process. By integrating Particle Swarm Optimization into the established AutoDock Vina framework, we provided a powerful tool capable of accelerating drug discovery, and computational enzymology. Here we present Moldina (Multiple-Ligand Molecular Docking over AutoDock Vina), a new algorithm built upon AutoDock Vina. Through comprehensive testing against AutoDock Vina, the algorithm exhibited comparable accuracy in predicting ligand binding conformations while significantly reducing the computational time up to several hundred times. Moldina and the benchmark data are freely available at https://opencode.it4i.eu/permed/moldina-multiple-ligand-molecular-docking-over-autodock-vina and https://github.com/It4innovations/moldina-multiple-ligand-molecular-docking-over-autodock-vina.

Scientific Contribution

This efficient and accurate performance positions our algorithm as a valuable asset for researchers conducting fragment-based drug discovery or high-throughput virtual screening.

Keywords Multiple-ligand, Molecular docking, AutoDock Vina, Fragment-based drug design, Substrate inhibition, Competitive binding

*Correspondence:
Radek Halfar
radek.halfar@vsb.cz
Sérgio M. Marques
smarques@mail.muni.cz
Jan Martinovič
jan.martinovic@vsb.cz
[1] IT4Innovations, VSB - Technical University of Ostrava, 70800 Ostrava, Czech Republic
[2] Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic
[3] International Clinical Research Centre, St. Anne's University Hospital, 656 91 Brno, Czech Republic

Introduction

Molecular docking is an extremely valuable tool for many applications, such as drug discovery and the development of efficient biocatalysts. It uses the three-dimensional structural knowledge of biomolecules (receptors) to predict the binding mode and affinity of small molecules (ligands) onto those receptors. The main advantage of docking compared to other theoretical methods is its speed, which makes it suitable for screening extensive libraries of potential binders [1, 2]. Many search algorithms and scoring methods are available, also bound to different trade-offs between accuracy and computational cost. Due to being free, open-source, and showing fast

convergence rates, AutoDock Vina is one of the most widely used software tools for molecular docking [3–5].

The simultaneous docking of multiple ligands (hereafter simply called "multiple docking") may represent a much more challenging task than traditional docking. The existence of several ligands, which may compete for the binding site and interact with each other, increases the complexity of the problem significantly. This scenario, however, can be extremely relevant in a multitude of situations. Some examples are fragment-based drug design, studies of enzymatic mechanisms, substrate inhibition, interactions of inhibitors with the substrate, synergistic or competitive binding of different ligands, docking of explicit waters, or even the crowding effects of non-specific binders. In the field of medicine, in particular, multiple docking can be invaluable in a variety of applications. It can be instrumental in fragment-based drug design, where multiple small molecule fragments are concurrently docked to identify promising candidates for larger, more potent drug molecules [6]. It may also be used to predict synergistic drug combinations for enhanced therapeutic efficacy. Lastly, it could play a crucial role in identifying allosteric modulators, which can influence the binding behavior of primary ligands and offer novel therapeutic avenues [7–13]. However, unlike classical docking, there are not many programs available to perform multiple docking.

Until recently, for docking multiple ligands, the users had to dock them sequentially, without the possibility of performing the task simultaneously. Expectedly, this could incur biases and inaccuracies. The first approach to perform this task was introduced by Li and Li in 2010 in a protocol they termed Multiple-Ligand Simultaneous Docking (MLSD) [8]. They employed particle swarm optimization (PSO) for local search and implemented it on the AutoDock 4 platform. At present, a few sophisticated software tools are available for simulating the simultaneous molecular docking of multiple ligands. Notable among these tools is S4MPLE [9, 14], grounded in genetic algorithm optimization techniques, and AutoDock Vina since version 1.2.0 (published in 2021) [15]. This latter algorithm combines Monte-Carlo iterated search for global optimization with the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method [16] for local refinement of the ligand conformations.

The application of PSO to MLSD is described in the study of Li and Li [8]. This study systematically compares the efficacy of the Lamarckian genetic algorithm and PSO across a range of test models. Notably, this proposed method is seamlessly integrated into the AutoDock4 framework, sharing both a scoring function and a customized Solis-Wets local search algorithm. Unfortunately, the MLSD program is no longer available. Hence, we decided to develop our own PSO implementation over the docking package Autodock Vina, which we named Moldina (Multiple-Ligand Molecular Docking over AutoDock Vina).

Materials and methods

System preparation for benchmarking

The systems under study were obtained from the RCSB Protein Data Bank [17] (PDB IDs: 1s63, 5x72, 2qfo, 2xdu, 2yei, 2yej, 3hz1, 2flh, 4oxk and 5ogo). The water molecules, ions, co-crystallization molecules, and additional chains were removed with PyMOL 2.3.2 [18]. The pdb4amb module of AmberTools 16 [19] was used to remove double side chains of any residues, when existing, keeping only the most populous conformations. The hydrogen atoms were added to the protein and ligands with the *reduce* program of AmberTools 16, using dynamic optimization of their position (-build -nuclear options), and the ligands were saved in separate PDB files. With the *antechamber*module of AmberTools 16, unique atom names were assigned to the ligand atoms, for a more convenient calculation of the root-mean-square deviation (RMSD) of the docked ligands from the crystallographic poses. Using Avogadro [20], the ligands were moved randomly in space and minimized with the UFF force field [21] and the steepest descent algorithm. The input files of the ligands and receptors in PDB format were converted to the AutoDock Vina-compatible format PDBQT using MGLTools [22] to add the Gasteiger atomic charges [23]. The ligand molecules in the crystal structures reported above were fragmented using PyMOL. We used our chemical intuition to break the ligands into smaller moieties by removing aliphatic carbons connecting rings or larger moieties. Hydrogen atoms were added afterwards with PyMOL. In several cases (PDB IDs 2qfo, 2xdu, 2yej and 3hz1) the ligands were already too small to fragment.

We benchmarked the docking of single ligands, and for that, we used a set of binders of the SARS-CoV-2 Protease M^{pro} [24], available on GitHub (https://github.com/shanizev/Benchmarking-SARS-CoV-2/, accessed on February 2025). We converted the proteins and ligand files to AutoDock Vina-compatible format PDBQT using MGLTools. Some cases presented problems in the file compatibility and were discarded, resulting in a total of 182 cases successfully studied.

Docking with Moldina

Moldina is constructed upon the framework of the original AutoDock Vina 1.0, and the specification of the input parameters adheres to the same guidelines as in that program. The sole alterations pertain to the representation of the input ligands, which are now listed in a text file (with each ligand on a separate line). The parameter for specifying the name of this text file is "ligands". There is no restriction on the number of input ligands. Additionally, a new parameter, "search_alg_PSO", has been introduced. To employ the Particle Swarm Optimization (PSO), this parameter must be set to 1. Otherwise, the ligand docking process will utilize the original AutoDock Vina algorithm. The region of interest was, in all docking calculations done here, delineated by a 20×20×20 Å box centered on the geometric center of the original ligands in the corresponding crystal structure, mirroring the parameter settings of AutoDock Vina 1.2 in the benchmark (see below). Each docking calculation was repeated thirty times to assess the output variability. For measuring the docking accuracy in each run, the RMSD was calculated for the heavy atoms of the ligands by comparing the best-ranked binding poses with the crystallographic structures.

Since the AutoDock Vina scoring function does not provide details on the contributions from individual residues or molecules, we evaluated the affinity of each individual ligand with the protein by systematically removing one ligand at a time. This was done for each ternary complex obtained from every docking run with Moldina on the main benchmark set. The resulting two binary protein-ligand complexes were then evaluated by the Vina score (using the –score_only parameter) to obtain the binding energy for each ligand.

Docking with AutoDock Vina 1.2.0

AutoDock Vina 1.2 [15] was used to compare the docking of multiple ligands with Moldina. For that, the region of interest was defined by a 20×20×20 Å box centered at the geometric center of the multiple ligands in the respective crystal structure. The default Vina scoring function was used. The exhaustiveness parameter was specified as 8 (the default) or 100, and the number of CPUs varied between 1 and 16. Each docking calculation was repeated thirty times with exhaustiveness 8, or ten times with exhaustiveness 100, to assess the output variability. The benchmark of single ligands on the SARS-CoV-2 Protease M^{pro} was performed in ten replicates using exhaustiveness 8.

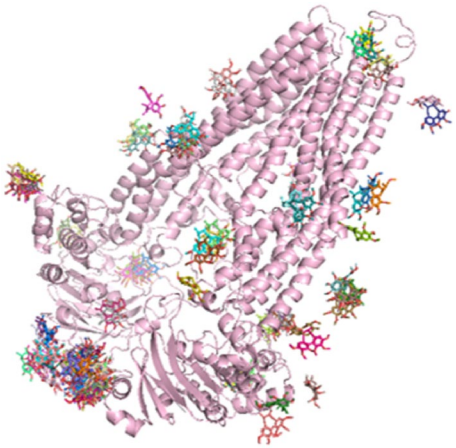

Fig. 1 An example of ligands forming an initial swarm. Ligand clusters are discernible within the coordinates identified through pre-search efforts. Within these specific spatial domains, there exists a notable likelihood of encountering a favorable site for ligand docking. Consequently, a substantial portion of the initial swarm is strategically allocated to these regions

Results and discussion

Particle swarm optimization

We improved the original optimization algorithm utilized in AutoDock Vina, which involved combining Monte Carlo with BFGS, by incorporating Particle Swarm Optimization (PSO) with a novel swarm initialization technique.

Our methodology encompasses an initial exploration of the search space (pre-search) by individually docking input ligands in each of the search space octants using PSO with randomly initialized swarms. The resulting conformations from this preliminary search are then subjected to random perturbations and combined to create a swarm for the final (global) PSO optimization. We present an instance where ligands, following this approach, form an initial swarm when interacting with a protein in Fig. 1

Following the creation of this population, we initiate the global docking process employing PSO, which spans the entire search space. Subsequently, once the PSO optimization concludes, we initiate a local optimization phase using BFGS, similar to the original algorithm. This local optimization aims to identify potentially improved local conformations. For a visual representation of the workflow involving the docking of multiple ligands, please refer to the flow chart depicted in Fig. 2. The incorporation of this optimization algorithm was executed to serve as a prospective substitute for the original optimization algorithm within AutoDock Vina 1.0 (wherein the new parameter search_alg_PSO was set to 1).

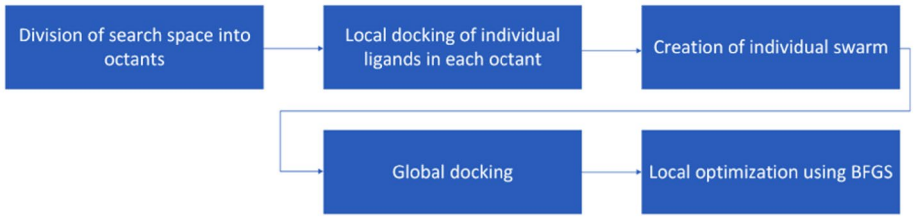

Fig. 2 Flow chart of Moldina algorithm. Demonstration of the computational procedure for simultaneously docking multiple ligands using Moldina. Initially, the search space is partitioned into octants, where each ligand undergoes individual docking. This approach facilitates the discovery of suitable conformations for each ligand. Subsequently, the identified conformations are perturbed to prevent the algorithm from being trapped in local minima and amalgamated together. Using this formed initial swarm for global docking, the optimization process employing particle swarm optimization is initiated. Following the completion of particle swarm optimization (PSO), the identified conformations undergo further refinement through local optimization using (BFGS), akin to the original AutoDock Vina

Benchmark

Our newly developed Moldina algorithm underwent testing across ten tasks, each involving the crystallographic structures of proteins complexed with two ligands. Specifically, two of these tasks were sourced from the recent AutoDock Vina 1.2 paper [15], which introduces an updated AutoDock Vina capable of accommodating multiple ligands (PDB IDs 1s63 and 5x72). The next five tasks (PDB IDs 2qfo, 2xdu, 2yei, 2yej and 3hz1) were taken from the S4MPLE publication [9], and the last three (PDB IDs 2flh, 4oxk and 5ogo) from searching the RCSB Protein Data Bank [17]. Given the algorithm's heuristic nature, each docking process underwent thirty iterations using Moldina and ten or thirty using Vina (the

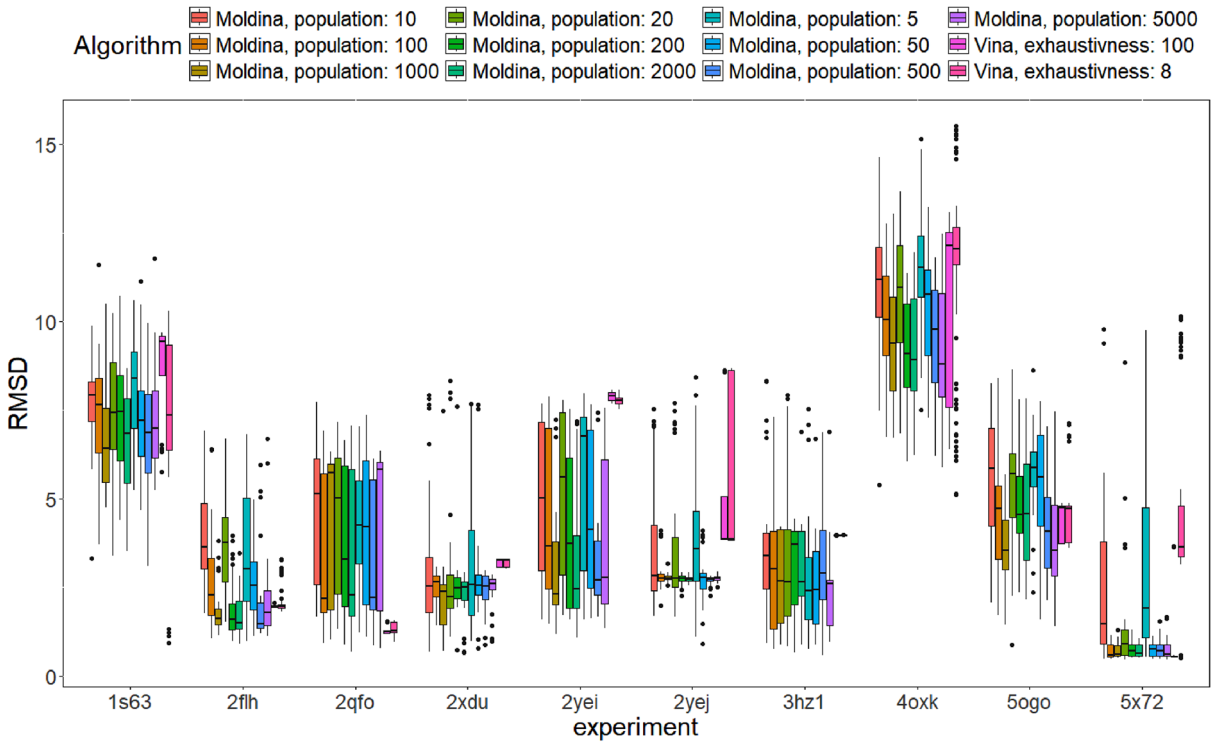

Fig. 3 Accuracy of Moldina and AutoDock Vina for the benchmark set. RMSD (in Å) distribution of the best docking binding modes achieved by Moldina (for particle swarm optimization population size from 5 to 5000) and AutoDock Vina 1.2 (exhaustiveness 8 and 100), in comparison with the crystallographic ones (PDB ID codes are listed on the X-axis). In most cases, Moldina achieved comparable or better results than AutoDock Vina 1.2, for particle swarm optimization population size equal to or greater than 50, or in a few cases above 100. However, Moldina achieved these results in a fraction of the time compared to AutoDock Vina (Table 1). The plots show the results obtained for all the replicas performed in each case, and the error bars represent the interquartile range for the multiple runs (thirty runs for nearly all the cases, and ten for AutoDock Vina with exhaustiveness 100)

smaller number of runs was due to the higher computational demands of AutoDock Vina with exhaustiveness 100) to assess the output variability. The benchmarking was performed in comparison with AutoDock Vina 1.2. Moldina was tested with several PSO populations.

The accuracy of the multiple docking was assessed by the root-mean-square deviation (RMSD) of the ligands in the docked conformations (using the best-ranked pose) in comparison with the crystallographic ones. Generally, the accuracy of Moldina improved (the mean RMSD decreased) by increasing the PSO population size, but for PSO populations higher than 50 the improvements were usually minor (Fig. 3 and Supplementary Table S1.). In the great majority of the cases, Moldina performed better than AutoDock Vina, when using a PSO population size of 50 or lower and the default AutoDock Vina parameters (exhaustiveness 8). The biding poses obtained for the complete benchmark set can be found in Fig. 4, and the corresponding results from AutoDock Vina are presented in Supplementary Figure S1. In a representative Case Study 1 (described in detail in the Supplementary Material), the complex of the heat shock protein 90 bound with two molecular fragments (PDB ID: 3hz1) [25] was reproduced by Moldina and AutoDock Vina with mean RMSD values of 2.54 and 3.98 Å, respectively. For both calculations, the docked molecules were correctly located in the respective crystallographic binding, showing only some shifts in a few functional groups (Fig. 5). Whereas for a single ligand, one could consider a docking result with an RMSD value above 3 Å as a rather poor prediction, for two simultaneously docked ligands, RMSD values of ca. 5 Å or higher may (arguably) be acceptable.

AutoDock Vina performed clearly better than Moldina only in one task (PDB ID 2qfo). Although Moldina managed to find very similar poses to the crystallographic structure (as can be observed in Fig. 4), often that was not the case in multiple other runs. The reasons for this behavior were not clear, since these ligands are not particularly large nor was the binding site very wide. In a few cases, the accuracy of Moldina improved significantly by increasing the PSO population to 1000 or more (PDB IDs 2flh, 2yei and 5ogo). For two systems (PDB IDs 1s63 and 4oxk), however, both failed to predict the crystallographic binding modes by far (mean RMSD > 7 Å and lowest RMSD > 4.6 Å). This was due to the intrinsic complexity of the systems, which contain two large ligands with a large number of rotable bonds (many degrees of freedom during the conformational search). This is a common problem for most docking software, which very often fail to predict the binding of very bulky and flexible ligands. In the case of 4oxk, the large volume of the

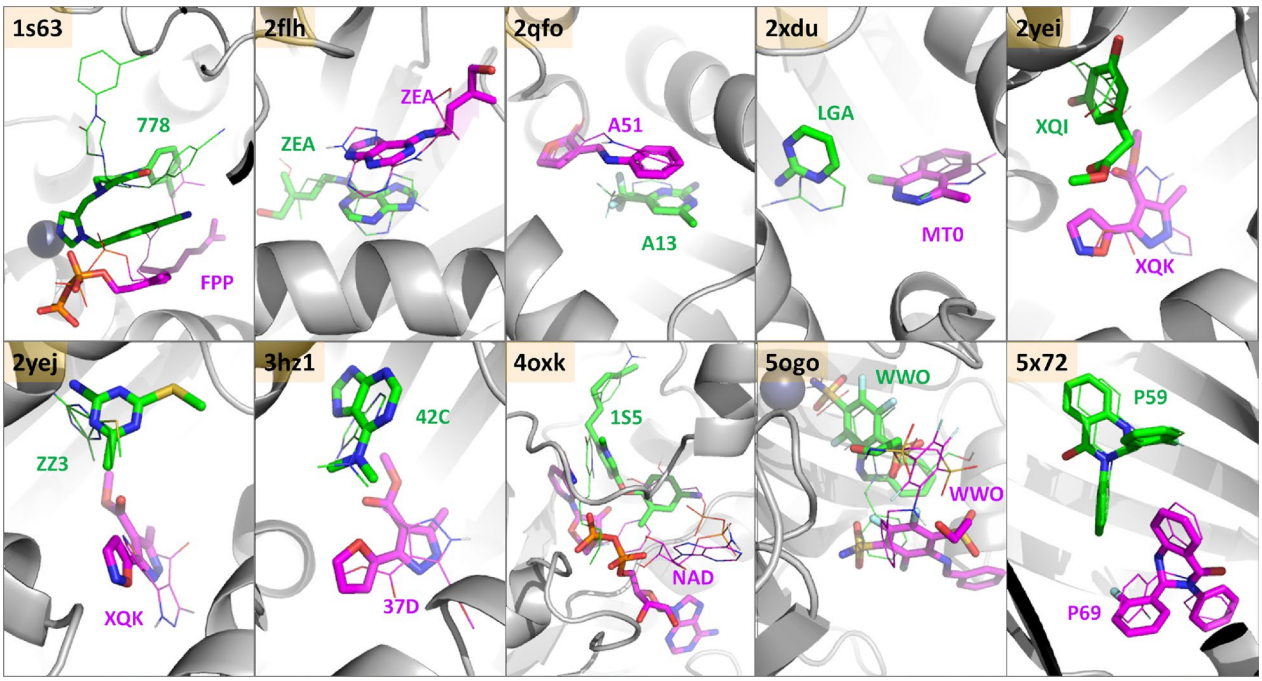

Fig. 4 Docked binding poses for the benchmark set, obtained from Moldina with particle swarm optimization population size 50. Superimposition of the best binding modes (represented as wires) with the crystallographic poses of the ligands (represented as sticks). The respective PDB ID codes are labeled on the top-left corner of each image; the two ligands are represented in green and magenta, and the respective names are shown by the labels with the same colors

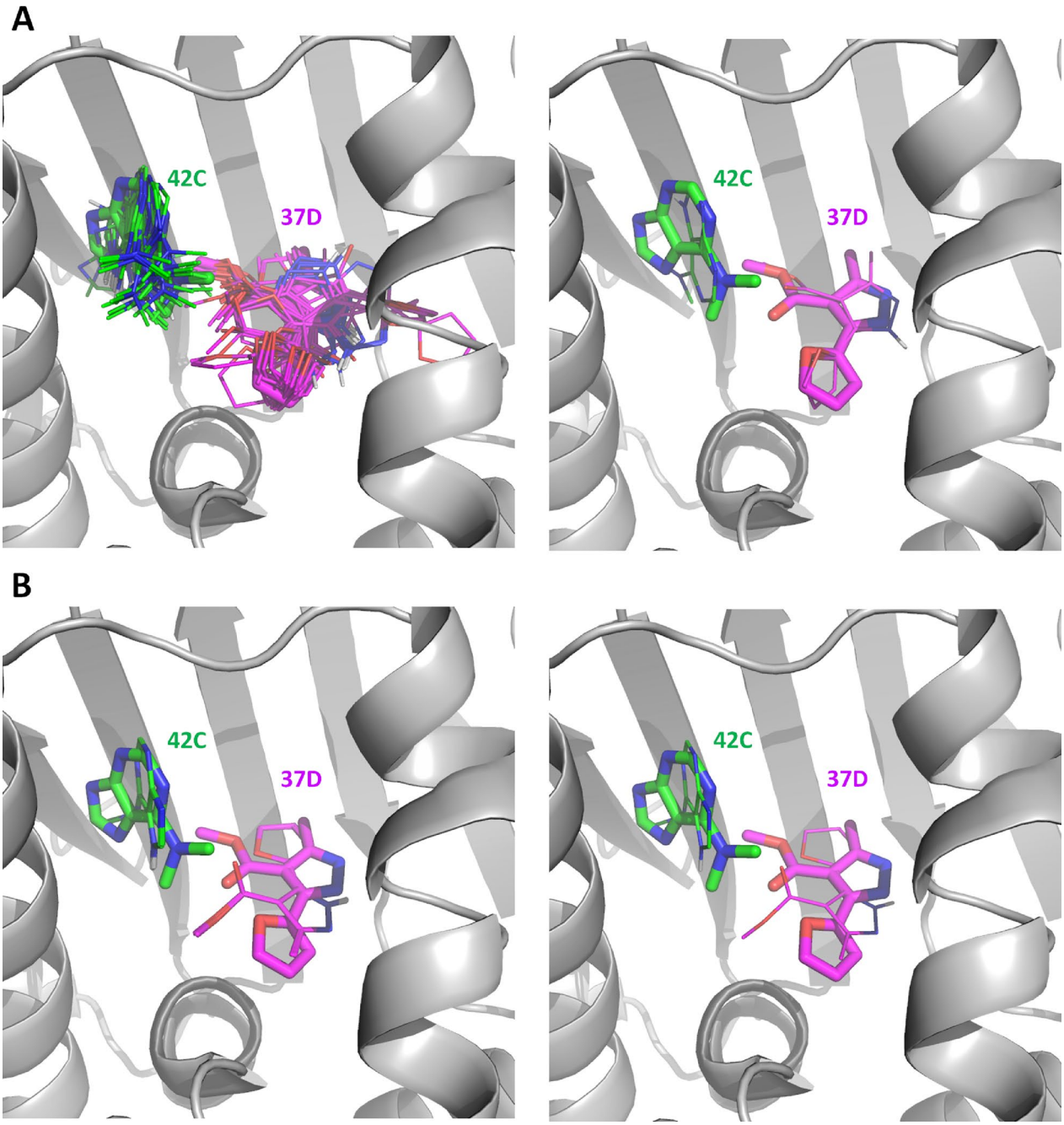

Fig. 5 Docked binding poses superimposed with the ligands in crystal structures (PDB ID: 3hz1). Calculations with Moldina with particle swarm optimization population size 50 (**A**), and AutoDock Vina with the default settings (exhaustiveness 8) (**B**). Left: superimposition of the best binding modes from all the multiple runs (for Vina, only results with 8 CPUs are displayed); right: binding mode from the run with RMSD closest to the overall mean value obtained from all the runs (2.54 Å for Moldina and 3.98 Å for AutoDock Vina). The ligands from crystal structures are represented as sticks and the docked ones as wires, 42C is shown in green and 37D in magenta. The minimum RMSD obtained over all the replicates with the same settings was 0.89 Å for Moldina and 3.96 Å for Vina

binding cavity, which is deep and extensive, aggravates this difficulty, and results in the wide range of RMSD values observed. Interestingly, Moldina performed better than Vina, and it could place the ligands in the correct side of the pocket in most of the replicates (Supplementary Fig. S2.). Regarding complex 1s63, it contains a metal ion (zinc), whose binding is often difficult to predict by docking using regular scoring functions (see Case Study

Table 1 Mean run-time (in seconds) of the docking process

	Moldina (1 CPU)										AutoDock Vina (Exhaustiveness 8)					AutoDock Vina (Exhaustiveness 100)	
	PSO population										CPU					CPU	
	5	10	20	50	100	200	500	1000	2000	5000	1	2	4	8	16	8	16
1s63	2.3	2.4	2.5	3.7	5.6	5.7	5.9	6.5	7.6	10.8	409.0	210.0	121.8	61.7	63.1	706.0	381.2
2flh	0.8	0.8	0.9	1.5	2.4	2.5	2.7	3.0	3.8	6.1	123	61.4	31.4	16.4	16.3	200.6	105.8
2qfo	0.7	0.7	0.7	1.1	1.7	1.8	1.9	2.2	2.8	4.6	51.7	26.0	48.9	7.5	23.1	81.0	43.4
2xdu	0.6	0.6	0.6	0.9	1.3	1.4	1.5	1.7	2.1	3.4	93.6	11.7	6.4	3.1	3.3	31.3	17.3
2yei	0.7	0.7	0.8	1.2	1.9	2.0	2.2	2.4	4.8	5.5	74.9	38.6	19.5	10.3	10.3	118.9	64.6
2yej	0.7	0.7	0.8	1.1	1.8	1.8	2.0	2.2	2.8	4.3	47.9	24.1	12.6	7.1	7.0	75.7	41.0
3hz1	0.7	0.7	0.8	1.2	1.9	2.0	2.1	2.4	2.9	4.6	48.4	24.7	12.9	6.9	7.1	77.3	42.3
4oxk	1.8	1.8	2.0	3.5	6.1	6.1	6.5	7.2	8.5	12.8	1312.0	722.6	404.5	214.9	225.4	2409.0	1293.8
5ogo	1.0	1.0	1.2	1.9	3.1	3.3	3.6	4.3	5.5	9.0	503.9	255.8	191.7	71.8	68.9	823.5	449.3
5x72	1.0	1.1	1.2	2.0	3.3	3.4	3.6	4.2	5.1	7.9	155.7	80.5	43.1	25.8	25.0	247.4	134.0

Mean run-time (in seconds) of the docking process to ten different protein structures with multiple bound ligands using Moldina (1 CPU) according to different PSO population sizes, and AutoDock Vina 1.2 (exhaustiveness 8 and 100), according to the number of CPUs. Protein structures solved by protein crystallography are specified by PDB ID in the first column

2, described in detail in Supplementary Material). Conversely to these situations, when the ligands are not too flexible and the active site is narrow and compact, both docking programs showed high success rates with convergent results, like for PDB ID 5x72.

We also compared the RMSD values for the ligands separately. Generally, we did not find striking differences between the predicted poses of either ligand (Supplementary Fig. S3.). This is probably because the size of the two ligands was not remarkably different in any of the cases. The only exception is 4oxk, where NAD was more difficult to predict due to its higher number of rotable bonds, compared to ligand 1S5.

On average, Moldina achieved comparable or better results than those from AutoDock Vina. However, those results were obtained in a fraction of the time compared to AutoDock Vina, even using fewer CPU resources (Table 1). With the presented algorithm, using a single CPU, Moldina can achieve over a hundred times speedup compared to AutoDock Vina, even when leveraging AutoDock Vina with parallel computation of 16 CPUs.

The ligand binding affinities (ΔGbind) were also analyzed for Moldina and AutoDock Vina, which use the same AutoDock Vina scoring function. The results show that the affinities predicted by Moldina tended to increase (ΔGbind decreased) with the PSO population size (Supplementary Figure S4. and Supplementary Table S2.), which is consistent with the accuracy improvements. In most of the cases, the affinities were rather similar to those found by AutoDock Vina (for PSO population 50), differing by less than 2 kcal/mol. In one case (PDB ID 5x72), the mean affinity obtained by Moldina was slightly better than that obtained by AutoDock Vina. In other cases (PDB IDs 1s63, 2flh, 4oxk and 5ogo), however, the affinities were worse than obtained with AutoDock Vina, suggesting that the binding modes predicted by Moldina were not as favorable. Among these cases, one indeed corresponded to worse RMSD accuracy by Moldina (PDB IDs 2flh). However, the others showed similar (5ogo) or even better RMSD accuracy than AutoDock Vina (PDB IDs 1s63, and 4oxk).

We tried to dissect the binding interactions of each ligand with the protein. Since AutoDock Vina score does not allow assessing the contributions from the different residues or molecules, we removed one ligand at the time from the original ternary complex. In some cases, we observed a large dispersion of the binding energies for the individual ligands over the multiple runs (Supplementary Fig. S5., Table S3.). This is related, in different degrees, with the respective accuracy variability (Supplementary Fig. S3.). In some cases, the individual energies were additive (PDB IDs 2qfo, 2xdu, 3hz1, and 5x72), where the total energy of the complex is nearly the same as the sum of the parts. This suggests either an independent binding or a weak cooperative effect between the two ligands. However, in most cases those energies are not additive, being the total energy higher (less negative) than the sum of the energies for the individual ligands. This suggests a negative cooperativity between the ligands, which may be due to steric clashes between them, competition for the binding site, or disruption of favorable interactions by the presence of the second ligand. In most

of the cases, the ligands have quite similar affinities. For the remaining ones (2xdu, 2yej, 3hz1 and 4oxk), one of the ligands interacts with the protein more strongly than the other. This is mainly due to the different nature of the ligands, binding location, and their sizes.

We conducted sequential docking of the ligands within the same systems described above, and compared its accuracy to that of simultaneous docking as implemented in Moldina. Based on the RMSD distributions and mean values (Supplementary Fig. S6. and Table S4.), we observed no significant difference in accuracy for several cases (PDB IDs: 1s63, 4oxk and 5x72), while in one instance (2qfo) the results were slightly better. However, in most of the cases the performance was worse (PDB IDs: 2flh, 2xdu, 3hz1 and 5ogo) or much worse (2yei and 2yej). These findings demonstrate that simultaneous docking can offer advantages over the conventional sequential approach for multiple ligands, thereby supporting our initial hypothesis and the motivation behind the development of Moldina.

As mentioned previously, the docking of multiple small ligands is a powerful technique used in fragment-based drug design [6]. We tried to assess if Moldina can reproduce the binding of molecular fragments for the reconstruction of larger ligands, and for that, we docked simultaneously the fragments from the original ligands in the previous PDB structures. In several cases (PDB IDs 2qfo, 2xdu, 2yej and 3hz1), the original ligands were too small to be fragmented, and here the results are the same as in the multiple docking discussed above. The RMSD analysis indicated that the docked poses of individual fragments did not accurately reproduce their original positions within the larger ligands. The overall mean RMSD values ranged from 4.3 Å (considering only cases involving fragmented ligands) to 10.3 Å (Supplementary Fig. S7. and Table S5.). The high variability in RMSD values further suggests significant pose diversity across different runs and poor convergence of the results. This is not surprising, considering that the binding of small ligands is often not specific and the docking scoring functions are not sufficiently accurate to distinguish small differences in binding energies. Nonetheless, relatively low RMSDs were observed in some cases (e.g., 4oxk showed lower RMSD for the fragments than for the intact ligands). Interestingly, the mean total binding affinities obtained from docking individual fragments (Supplementary Fig. S8. and Table S6) were consistently better (e.g., for PDB ID: 2yei) or significantly better (PDB IDs: 1s63, 2flh, 4oxk, 5ogo, and 5x72) than those obtained for the corresponding full ligands (compare Supplementary Tables S2. and S6.). Such increase may be attributed to different factors, such as the fact that fragments contain additional hydrogen atoms, some of which may form hydrogen bonds with the protein, thereby enhancing the interaction with the protein and among each other. On the other hand, docking the multiple fragments with Moldina may have identified more favorable binding modes than those present in the original ligands, suggesting new positions that could potentially lead to stronger binders if the fragments were reassembled. Therefore, we believe that these results support the usefulness of Moldina for docking simultaneously multiple molecular fragments.

The dependence of the docking time on the number of atoms in the ligands and the number of rotatable bonds was also investigated (see Supplementary Figures S9., and S10.). The results indicate a linear dependence of the docking time on the number of atoms in the ligands. This dependence was not confirmed for the number of rotatable bonds. However, these results can be affected by the size of the protein. For more accurate results, it would be necessary to perform an analysis on a larger number of cases. Overall, our results revealed that the outcomes achieved by the newly proposed algorithm closely resemble those of AutoDock Vina or were improved. Importantly, Moldina shortened the running times by two orders of magnitude compared to AutoDock Vina.

Finally, we assessed the ability of Moldina to dock single molecules and compared it with AutoDock Vina (both using the default parameters mentioned above). We evaluated 182 complexes of different compounds bound to the coronavirus CoV-2 Main Protease (M^{pro}) [24]. Generally, the results showed that Moldina and AutoDock Vina performed with similar accuracy (as per mean RMSD values with respect to the crystallographic poses of the ligands; Supplementary Table S7). However, Moldina outperformed Vina more frequently, with 63 cases where its mean RMSD was at least 1.0 Å lower than that from Vina, in contrast with only 34 cases where Vina achieved a better RMSD by 1.0 Å or more. These results demonstrate the ability of Moldina to dock single ligands in addition to the multiple docking, which was intended to be its main purpose.

Conclusions

Here, we developed Moldina, an algorithm and software tool that uses the underlying Autodock Vina method and optimizes it for the simultaneous docking of multiple ligands, thus enhancing its capabilities in molecular docking applications. This algorithm replaces the optimization algorithm of AutoDock Vina with PSO and introduces a novel swarm initialization method, enabling more consistent outcomes while utilizing a smaller initial population. The benchmarking of Moldina against AutoDock Vina 1.2 demonstrates the efficacy of our approach.

The presented algorithm achieves better or comparable accuracy in the binding conformations but also exhibits a remarkable reduction of the computational time up to several hundred folds. We also found that simultaneous docking is advantageous in comparison to the conventional sequential approach for multiple ligands, showcasing its usefulness. Overall, the efficiency enhancements position our algorithm as a valuable tool for researchers and practitioners seeking to expedite multiple docking predictions for different purposes, such as fragment-based drug discovery, high-throughput virtual screening of cooperative binders, studying competitive inhibitors, substrate and product inhibition, and many more.

Abbreviations

MLSD	Multiple-ligand simultaneous docking
PSO	Particle swarm optimization
BFGS	Broyden–Fletcher–Goldfarb–Shanno method

ΔGbind Docking binding affinitySupplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13321-025-01005-4.

Supplementary material 1.

Acknowledgements
The authors would like to express gratitude to the National Supercomputing Centre IT4Innovation for providing computational resources for this project, and colleagues from Loschmidt Laboratories for providing domain expertise for developing the presented algorithm.

Author contributions
R.H. developed the software, analyzed and visualized the results; S.M.M. prepared the data, analyzed, visualized, and interpreted the docking results, and wrote the case studies; R.H. and S.M.M. wrote the manuscript; J.D. and J.M. supervised the research project and edited the manuscript. All authors contributed to the project methodology and reviewed the manuscript.

Funding
This work has been supported by the ELIXIR-CZ project (nr. LM2023055), part of the international ELIXIR infrastructure, and the e-INFRA CZ project (ID:90254) supported by the Ministry of Education, Youth and Sports of the Czech Republic. The work was further supported by the project from the National Institute for Neurology Research (nr. LX22NPO5107 MEYS) financed by the European Union - Next Generation EU, and the Technology Agency of the Czech Republic (nr. Permed-TN01000013). This project was also supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreements Nos. 857560 and 101136607.

Data availability
Project name: Moldina - Multiple Ligand Molecular Docking over AutoDock Vina. Project home page: https://opencode.it4i.eu/permed/moldina-multiple-ligand-molecular-docking-over-autodock-vina and https://github.com/It4innovations/moldina-multiple-ligand-molecular-docking-over-autodock-vina. Operating system(s): Linux. Programming language: C++.

Declarations

Competing interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received: 28 November 2024 Accepted: 30 March 2025
Published online: 28 April 2025

References

1. Torres PHM, Sodero ACR, Jofily P, Silva-Jr FP (2019) Key topics in molecular docking for drug design. Int J Mol Sci 20(18):4574. https://doi.org/10.3390/ijms20184574
2. Agu PC, Afiukwa CA, Orji OU, Ezeh EM, Ofoke IH, Ogbu CO, Ugwuja EI, Aja PM (2023) Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management. Sci Reports 13(1):13398. https://doi.org/10.1038/s41598-023-40160-2
3. Gaillard T (2018) Evaluation of autodock and autodock vina on the casf-2013 benchmark. J Chem Inf Model 58(8):1697–1706. https://doi.org/10.1021/acs.jcim.8b00312
4. Nguyen NT, Nguyen TH, Pham TNH, Huy NT, Bay MV, Pham MQ, Nam PC, Vu VV, Ngo ST (2019) Autodock vina adopts more accurate binding poses but autodock4 forms better binding affinity. J Chem Inf Model 60(1):204–211. https://doi.org/10.1021/acs.jcim.9b00778
5. Tang S, Chen R, Lin M, Lin Q, Zhu Y, Ding J, Hu H, Ling M, Wu J (2022) Accelerating autodock vina with gpus. Molecules 27(9):3041. https://doi.org/10.3390/molecules27093041
6. Xu W, Kang C (2025) Fragment-based drug design: From then until now, and toward the future. J Med Chem 68(5):5000–5004. https://doi.org/10.1021/acs.jmedchem.5c00424
7. Li H, Liu A, Zhao Z, Xu Y, Lin J, Jou D, Li C (2011) Fragment-based drug design and drug repositioning using multiple ligand simultaneous docking (mlsd): Identifying celecoxib and template compounds as novel inhibitors of signal transducer and activator of transcription 3 (stat3). J Med Chem 54(15):5592–5596. https://doi.org/10.1021/jm101330h
8. Li H, Li C (2010) Multiple ligand simultaneous docking: Orchestrated dancing of ligands in binding sites of protein. J Comput Chem 31(10):2014–2022. https://doi.org/10.1002/jcc.21486
9. Hoffer L, Horvath D (2012) S4mple – sampler for multiple protein–ligand entities: Simultaneous docking of several entities. J Chem Inf Model 53(1):88–102. https://doi.org/10.1021/ci300495r
10. Li H, Xiao H, Lin L, Jou D, Kumari V, Lin J, Li C (2014) Drug design targeting protein–protein interactions (ppis) using multiple ligand simultaneous docking (mlsd) and drug repositioning: Discovery of raloxifene and bazedoxifene as novel inhibitors of il-6/gp130 interface. J Med Chem 57(3):632–641. https://doi.org/10.1021/jm401144z
11. Raghavendra S, Aditya Rao SJ, Kumar V, Ramesh CK (2015) Multiple ligand simultaneous docking (mlsd): A novel approach to study the effect of inhibitors on substrate binding to ppo. Comput Biol Chem 59:81–86. https://doi.org/10.1016/j.compbiolchem.2015.09.008
12. Gupta A, Chauhan SS, Gaur AS, Parthasarathi R (2022) Computational screening for investigating the synergistic regulatory potential of drugs and phytochemicals in combination with 2-deoxy-d-glucose against sars-cov-2. Struct Chem 33(6):2179–2193. https://doi.org/10.1007/s11224-022-02049-0
13. Li H, Komori A, Li M, Chen X, Yang AWH, Sun X, Liu Y, Hung A, Zhao X, Zhou L (2023) Multi-ligand molecular docking, simulation, free energy calculations and wavelet analysis of the synergistic effects between natural compounds baicalein and cubebin for the inhibition of the main protease of sars-cov-2. J Mol Liq 374:121253. https://doi.org/10.1016/j.molliq.2023.121253
14. Hoffer L, Chira C, Marcou G, Varnek A, Horvath D (2015) S4mple—sampler for multiple protein-ligand entities: Methodology and rigid-site docking benchmarking. Molecules 20(5):8997–9028. https://doi.org/10.3390/molecules20058997
15. Eberhardt J, Santos-Martins D, Tillack AF, Forli S (2021) Autodock vina 1.2.0: New docking methods, expanded force field, and python bindings. J Chem Inf Model 61(8):3891–3898. https://doi.org/10.1021/acs.jcim.1c00203
16. Fletcher R (1988) Practical methods of optimization. https://api.semanticscholar.org/CorpusID:123487779
17. Berman HM (2000) The protein data bank. Nucleic Acids Res 28(1):235–242. https://doi.org/10.1093/nar/28.1.235

18. Schrödinger L (2019) The PyMOL Molecular Graphics System, Version 2.3.2. https://pymol.org/
19. Case DA, Betz RM, Cerutti DS, Cheatham TE, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Homeyer N, Izadi S, Janowski P, Kaus J, Kovalenko A, Lee TS, LeGrand S, Li P, Lin C, Luchko T, Luo R, Madej B, Mermelstein D, Merz KM, Monard G, Nguyen H, Nguyen HT, Omelyan I, Onufriev A, Roe DR, Roitberg A, Sagui C, Simmerling CL, Botello-Smith WM, Swails J, Walker RC, Wang J, Wolf RM, Wu X, Xiao L, Kollman PA (2016) AMBER 2016. University of California, San Francisco, University of California, San Francisco. http://ambermd.org/
20. Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR (2012) Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J Cheminform 4(1):1–17
21. Rappé AK, Casewit C, Colwell KS, Goddard WA, Skiff WM (1992) Uff, a full periodic table force field for molecular mechanics and molecular dynamics simulations. J Am Chem Soc 114:10024–10035
22. Sanner M (1999) Python: a programming language for software integration and development. J Mol Graph Model 17(1):57–61
23. Gasteiger J, Marsili M (1978) A new model for calculating atomic charges in molecules. Tetrahedron Lett 19(34):3181–3184. https://doi.org/10.1016/S0040-4039(01)94977-9
24. Zev S, Raz K, Schwartz R, Tarabeh R, Gupta PK, Major DT (2021) Benchmarking the ability of common docking programs to correctly reproduce and score binding modes in sars-cov-2 protease mpro. J Chem Inf Model 61(6):2957–2966. https://doi.org/10.1021/acs.jcim.1c00263
25. Barker JJ, Barker O, Courtney SM, Gardiner M, Hesterkamp T, Ichihara O, Mather O, Montalbetti CAGN, Müller A, Varasi M, Whittaker M, Yarnold CJ (2010) Discovery of a novel hsp90 inhibitor by fragment linking. ChemMedChem 5(10):1697–1700. https://doi.org/10.1002/cmdc.201000219

Publisher's Note