

Article

Not peer-reviewed version

Accurate Pocket Identification for Binding-Site-Agnostic Docking

[Yaroslav Balytskyi](#)*, Inna Hubenko, Alina Balytska, [Christopher Kelly](#)

Posted Date: 24 July 2025

doi: 10.20944/preprints202502.0352.v2

Keywords: Protein–ligand interactions; Blind docking; Binding pocket prediction; Soft mask segmentation; Deep residual network; Convolutional Neural Network



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Accurate Pocket Identification for Binding-Site-Agnostic Docking

Yaroslav Balytskyi ^{1,*}, Inna Hubenko ², Alina Balytska ² and Christopher V. Kelly ^{1,3}

¹ Department of Physics and Astronomy, Wayne State University, Detroit, MI, 48201, USA

² Cherkasy Medical Academy, Cherkasy, 18000, Ukraine

³ Center for Integrative Metabolic and Endocrine Research, School of Medicine, Wayne State University, Detroit, MI, 48201, USA

* Correspondence: hr6998@wayne.edu

Abstract

Accurate identification of druggable pockets and their features is essential for structure-based drug design and effective downstream docking. Here, we present RAPID-Net, a deep learning-based algorithm designed for the accurate prediction of binding pockets and seamless integration with docking pipelines. On the PoseBusters benchmark, RAPID-Net-guided AutoDock Vina achieves 54.9% of Top-1 poses with RMSD < 2 Å and satisfying the PoseBusters chemical-validity criterion, compared to 49.1% for DiffBindFR. On the most challenging time split of PoseBusters aiming to assess generalization ability (structures submitted after September 30, 2021), RAPID-Net-guided AutoDock Vina achieves 53.1% of Top-1 poses with RMSD < 2 Å and PB-valid, versus 59.5% for AlphaFold 3. Notably, in 92.2% of cases, RAPID-Net-guided Vina samples at least one pose with RMSD < 2 Å (regardless of its rank), indicating that pose ranking, rather than sampling, is the primary accuracy bottleneck. The lightweight inference, scalability, and competitive accuracy of RAPID-Net position it as a viable option for large-scale virtual screening campaigns. Across diverse benchmark datasets, RAPID-Net outperforms other pocket prediction tools, including PURESNet and Kalasanty, in both docking accuracy and pocket–ligand intersection rates. Furthermore, we demonstrate the potential of RAPID-Net to accelerate the development of novel therapeutics by highlighting its performance on pharmacologically relevant targets. RAPID-Net accurately identifies distal functional sites, offering new opportunities for allosteric inhibitor design. In the case of the RNA-dependent RNA polymerase of SARS-CoV-2, RAPID-Net uncovers a wider array of potential binding pockets than existing predictors, which typically annotate only the orthosteric pocket and overlook secondary cavities. RAPID-Net, along with interactive demonstration notebooks, is publicly available at: github.com/BalytskyiJaroslav/RAPID-Net, offering a user-friendly and fully reproducible framework for pocket-guided docking.

Keywords: protein–ligand interactions; blind docking; binding pocket prediction; soft mask segmentation; deep residual network; convolutional neural network

1. Introduction and Problem Statement

Molecular docking is the computational problem of predicting the most favorable ligand binding poses in a protein–ligand complex, given the experimentally determined or computer-simulated protein structure and the initial conformation of the ligand [1]. It plays a key role in drug development, typically following target identification when the biological molecule responsible for a disease is identified and validated for therapeutic intervention [2]. Once the target is identified, molecular docking is used for structure-based drug discovery and development [3,4] to identify and optimize potential lead compounds [5,6]. The docking results can guide virtual screening workflows, enabling the selection of promising candidates from vast chemical libraries [7,8]. This process helps to determine whether a compound has potential for drug development [9]. By significantly improving the speed

and efficiency of early drug discovery, molecular docking reduces the time and costs associated with the development of new therapeutics [10].

Typical protein-ligand docking pipelines rely on users specifying the binding pocket, and docking programs such as AutoDock Vina [11,12], GOLD [13], and Glide [14] use grids to confine their search to known or hypothesized protein's interaction sites. However, in the absence of such information – in “blind” or binding-site-agnostic settings – protein-ligand docking becomes significantly more difficult [15], as the docking algorithm must scan the entire protein surface, dramatically increasing the computational complexity. Traditional blind docking methods often rely on extensive sampling to explore potential binding sites across the whole protein, but this approach is computationally and time-intensive, making it impractical for large-scale virtual screening tasks [16,17]. With the advent of advanced protein structure prediction methods [18]—including AlphaFold [19], ColabFold [20], OpenFold [21], and RosettaFold [22]—the number of protein structures generated has surged, often without any corresponding ligand information. Meanwhile, only about ~ 5% of the human coding genes currently serve as commercial drug targets, leaving a wide range of disease-related targets unexplored [23]. Despite this immense potential, most new drug research and development remain focused on a limited set of well-established targets, underscoring the urgent need for novel druggable targets [24]. As more proteins lacking known binding pocket information are considered in drug discovery [25], the demand for ligand docking approaches that do not rely on prior knowledge of the binding site has grown significantly, highlighting the necessity for binding-site-agnostic methods.

In this work, we aim to develop a high-accuracy pocket-finding algorithm (**RAPID-Net**) optimized for seamless integration with standard docking pipelines to deliver efficient guidance in binding-site-agnostic workflows. Many protein binding site predictors are benchmarked solely on the geometric properties of their predicted pockets [26], but their impact on the accuracy of downstream ligand docking is rarely assessed. While many factors affect docking performance, we show further in the text that pocket prediction quality has a decisive role in determining docking success. We evaluate RAPID-Net's performance both by docking accuracy – using the widely adopted AutoDock Vina [12] – and by pocket-ligand intersection rates, as detailed further in the text. Indeed, RAPID-Net can provide the search grid centers and dimensions for any docking engine, and Vina was chosen due to its popularity.

Despite relying on Vina's relatively simple scoring function, RAPID-Net-guided docking outperforms the state-of-the-art blind-docking tool DiffBindFR [27] by a substantial margin (54.9% vs 49.1%) on the PoseBusters benchmark [28] (see Section 7). This result underscores our central claim: precise pocket identification is a decisive driver of downstream docking success. It is important to note that, unlike DiffBindFR, which incorporates the ligand-induced protein changes, in our approach, we use our pocket predictor to guide AutoDock Vina that docks to a rigid receptor, and all evaluations are performed on holo structures. Nevertheless, the same pocket and search grid localization can boost flexible-receptor engines as well, for example, DSDPFlex [29]: “By leveraging prior knowledge or information, this innovative approach is anticipated to enhance the search process within the appropriate binding pocket.” Overall, our results indicate that accurate pocket localization currently outweighs receptor flexibility as the key bottleneck, making RAPID-Net a practical route to higher accuracy in binding-site-agnostic workflows and a solid foundation on which ligand-induced conformational refinements can later be added as a subleading effect. We refer the treatment of ligand-induced conformational changes to our future work, focusing here on maximizing the pocket prediction accuracy.

For comparison with co-folding approaches such as AlphaFold 3 (AF3) [30], we follow the PoseBench protocol [31]. This protocol uses a curated subset of 130 protein-ligand complexes from PoseBusters [28] deposited after September 30, 2021, specifically designed to evaluate generalization performance rather than memorization [31]. On this challenging benchmark, RAPID-Net + Vina achieves a Top-1 success rate of 53.1%, only six percentage points below AF3's 59.5% (see Section 7),

while requiring substantially fewer computational resources and offering significantly better scalability for virtual screening.

RAPID-Net supplies an accurate search box for subsequent lightweight rigid-receptor redocking, whereas AF3 predicts the full protein-ligand complex from the protein sequence and the ligand SMILES. The significantly lower computational demand of our approach is exemplified by the 8F4J structure [28], which AF3 was unable to process as a whole: “Another PDB entry (8F4J) was too large to inference the entire system (over 5,120 tokens), so we included only protein chains within 20 of the ligand of interest.”. In contrast, RAPID-Net guided Vina to a pose with $\text{RMSD} < 2 \text{ \AA}^1$, as illustrated in Figure 2. Thus, RAPID-Net provides near-state-of-the-art accuracy while remaining practical for large-scale virtual screening.

In terms of the underlying pocket prediction function, existing pocket predicting algorithms can be broadly classified into classical and Machine Learning (ML)-driven approaches [26]. Classical methods rely on expert-defined rules or heuristics to detect pockets, whereas ML-driven algorithms learn to extract features from protein data without explicit human instructions. In this work, we adopt the definition of Machine Learning by François Chollet as “the effort to automate intellectual tasks normally performed by humans” [32]. From this perspective, the main objective of an ML framework is to unveil a meaningful data representation that allows pocket-prediction rules to emerge automatically, rather than being manually hardcoded.

Furthermore, “Deep Learning (DL) is a specific subfield of Machine Learning: a new take on learning representations from data that puts an emphasis on learning successive layers of increasingly meaningful representations” [32]. In the context of pocket identification in our study, these layers of data representations are realized through Deep Neural Networks (DNNs). Finally, in contrast to DL, so-called “shallow” learning approaches typically employ only one or a small number of consecutive data representation layers [32].

In the following Sections, we describe our DL model for pocket identification, named **RAPID-Net** (ReLU Activated Pocket Identification for Docking). We tested RAPID-Net on various benchmarks to demonstrate its efficiency and analyzed its predictions on several therapeutically important proteins. In Section 2, we provide a brief overview of existing pocket prediction algorithms. In Section 3, we discuss the architecture and design rationale behind our model. In Section 4, we present the RAPID-Net training pipeline, highlighting its key differences from existing approaches, including the use of ReLU activation in the last layer of the network to operate on a “soft” rather than binary labels. Using the RAPID-Net model developed in Sections 3 and 4, in Section 5 we integrate it into a docking protocol that is subsequently used for docking. In Section 6, we describe the evaluation metrics used to assess the quality of the model’s predictions. Sections 7 and 8 report docking results on the PoseBusters [28] and Astex Diverse Set [33] datasets, respectively. For completeness and to provide a direct side-by-side comparison with existing pocket prediction algorithms, Section 9 evaluates the geometric characteristics of predicted pockets on the Coach420 [34] and BU48 [35] datasets. To demonstrate the ability of RAPID-Net to identify distant binding sites, Section 10 analyzes its predictions for several therapeutically important proteins with known distal sites. Finally, in Section 11, we reach our conclusions and outline the potential directions for future work.

2. Overview of the Available Pocket Prediction Algorithms

Pocket prediction methods exploit a variety of different techniques to identify potential binding sites. Geometry-based techniques like Fpocket [36], Ligsite [37] and Surfnets [38] identify cavities by analyzing the geometry of the molecular surface of a protein and usually rely on the use of a grid, gaps, spheres, or tessellation [36–43]. Energy-based methods such as PocketFinder [44] rely on the calculation of interaction energies between the protein and chemical group or probe to identify cavities [44–49].

¹ A demonstration video is available at: <https://youtu.be/EkUKmoW11pE?si=3aKBkL3ZRq8ibWqo>

Conservation-based methods make use of sequence evolutionary conservation information to find patterns in multiple sequence alignments and identify conserved key residues for ligand binding site identification [50–52].

Template-based methods rely on structural information from homologues and the assumption that structurally conserved proteins might bind ligands at a similar location [34,53–57]. Combined approaches or meta-predictors combine multiple methods, or the use of multiple types of data to infer ligand binding sites, e.g., geometric features with sequence conservation [35,58–63].

Although classical geometry-based tools such as FPocket are fast and widely applicable, they can sometimes merge individual pockets into a single “blurred” cavity, for example, as observed in key SARS-CoV-2 targets [64,65]. In Section 10, we revisit the Nsp12 protein considered in [64,65] using the RAPID-Net model we developed to demonstrate its ability to identify functionally significant sites.

3. Design Rationale behind our Model

When predicting likely binding pockets, there is a well-known trade-off between precision and recall. A recent evaluation of pocket prediction methods based on the geometrical properties of their predicted pockets [26] highlights this trade-off. Currently available ML methods such as VNEGNN [67], GrASP [68], and PURESNet [69] achieve high precision (over 90%), but they systematically predict a small number of pockets, leading to low recall. As [26] points out, generating multiple predictions, some of which may be false positives, is often more useful than potentially missing viable binding sites.

To address this shortcoming and mitigate low recall, we develop RAPID-Net as an ensemble-based model to improve the stability of prediction accuracy and coverage of potential binding sites. RAPID-Net consists of five independently trained model replicas to aggregate results and increase the reliability of its predictions. For subsequent docking, RAPID-Net returns two types of pockets:

1. *Majority-voted pockets*: Consisting of voxels predicted by at least 3 out of 5 ensemble models, for high-confidence predictions.
2. *Minority-reported pockets*: Consisting of voxels predicted by at least 1 out of 5 ensemble models, increasing the overall recall.

As demonstrated on diverse benchmark datasets in the following Sections, including PoseBusters [28], Astex Diverse Set [33], Coach420 [34], and BU48 [35], the proposed ensemble-based approach yields more stable and reliable performance compared to current ML-driven pocket predictors, both considering the majority-voted and minority-reported pockets.

For instance, consider the 8DP2 protein from the PoseBusters [28] dataset, illustrated in Figure 1. RAPID-Net predicts two majority-voted sub-pockets connected by a link. However, because no ligand binding occurs in one of these sub-pockets, it constitutes a false positive. By contrast, one of the subleading poses binds within the second predicted sub-pocket and passes all tests. This example demonstrates RAPID-Net’s ability to balance accuracy and recall, providing a comprehensive and robust pocket prediction framework for complex docking tasks. Furthermore, as shown in the following Sections, the minority-reported pockets predicted by our model are often shallow pockets corresponding to secondary binding sites, in contrast to the more prominent majority-voted pockets.

Types of outputs from pocket predictors. The type of output produced by pocket predictors can be divided into three main categories.

Kalasanty [70] and PURESNet V1² [69] both detect potential ligand-binding pockets using a voxel-based representation, where each voxel is $2\text{\AA} \times 2\text{\AA} \times 2\text{\AA}$ in size.

In contrast, PRANK [71] and DeepPocket [72] reweight the alpha-spheres identified by the rule-based FPocket [73] algorithm. As a result, these methods inherit FPocket’s initial search space and any

² PURESNet V1[69] predicts cavities where the ligand is likely to reside, whereas PURESNet V2 [66] predicts residues likely to interact with the ligand. For simplicity, we refer to PURESNet V1 as “PURESNet” throughout the text, as in this work we focus on cavities for generating an accurate search grid for subsequent docking.

associated cascading errors. Moreover, alpha-sphere-based approaches may lack sufficient granularity and can struggle to capture subtle variations among binding sites, as previously noted [64,65].

GrASP [68], IF-SitePred [74], VN-EGNN [67], and PUPResNetV2 [66] provide predictions of potential ligand interactions at the residue level. Although these predictions can reveal important binding details, they are less suitable for docking workflows that require a well-defined three-dimensional region for accurate ligand placement and orientation. By precisely defining these regions, the computational search space is reduced as docking is restricted to reasonable binding pockets. As illustrated in the bottom panel of Figure 2, residue-level predictions can be difficult to interpret when defining a reasonable search grid.

Although some studies reweight residues to guide blind docking [75], and others [25] employ voxelized pocket prediction models—such as PUPResNet [69]—in tandem with AutoDock Vina [12], our work focuses on developing a voxel-based cavity prediction model that seamlessly integrates into the docking pipeline. This voxel-based approach not only provides well-defined search regions but also facilitates a more modular and interpretable workflow compared to less structured docking strategies, as illustrated in Figure 4 and discussed further in Section 5.

Furthermore, as discussed in Section 7, improving the accuracy of pocket identification directly improves docking results. In particular, AutoDock Vina [12], guided by our pocket predictor, outperforms the state-of-the-art DiffBindFR [27] docking tool, which otherwise scans the entire protein surface in “blind” settings.

Our model architecture. Figure 3 depicts the architecture of our proposed RAPID-Net model, which is similar to U-Net [76] with encoder and decoder branches. However, we implement several notable adaptations.

First, we include residual connections *only* in the encoder part of the model, as our experiments show that they are highly beneficial there, but lead to overfitting if included in the decoder. This approach differs from Kalasanty [70], which omits residual blocks altogether [77], and PUPResNet [69], which uses them throughout the network.

Second, although the standard SE-ResNet [78] architecture typically includes attention blocks at multiple layers, studies on breast-cancer imaging [79] and Raman spectra classification [80] have shown that using a *single* attention block can be highly effective while mitigating overfitting. Our experiments similarly suggest that in inherently noisy datasets, adding too many attention modules can amplify noise and degrade performance.

As we demonstrate in the following Sections, adopting a moderate level of attention, limiting residual connections, and incorporating a modified loss function significantly improves performance compared to earlier models.

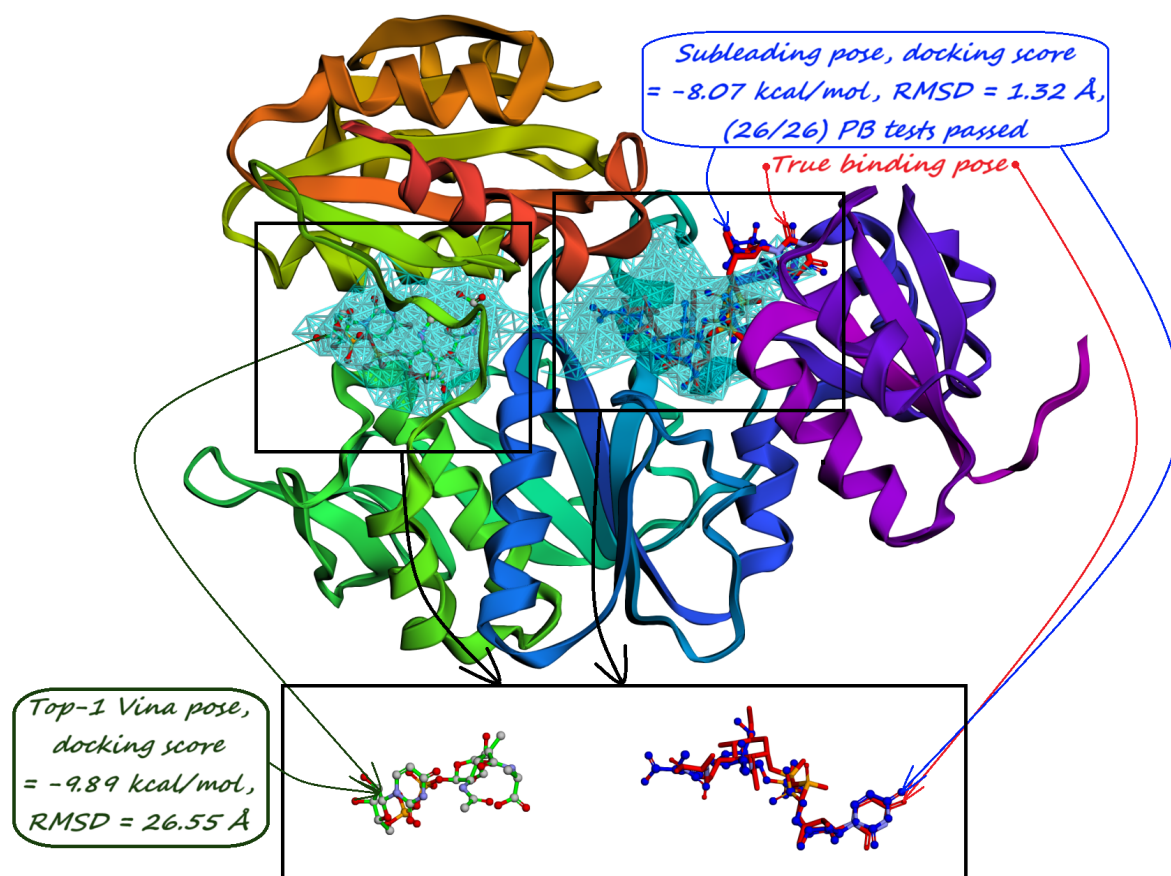


Figure 1. 8DP2 protein structure from the PoseBusters [28] dataset. The Top-1 Vina pose is docked in a predicted sub-pocket devoid of true ligand binding pose, while the subleading pose correctly occupies a second predicted sub-pocket and passes all validation tests.

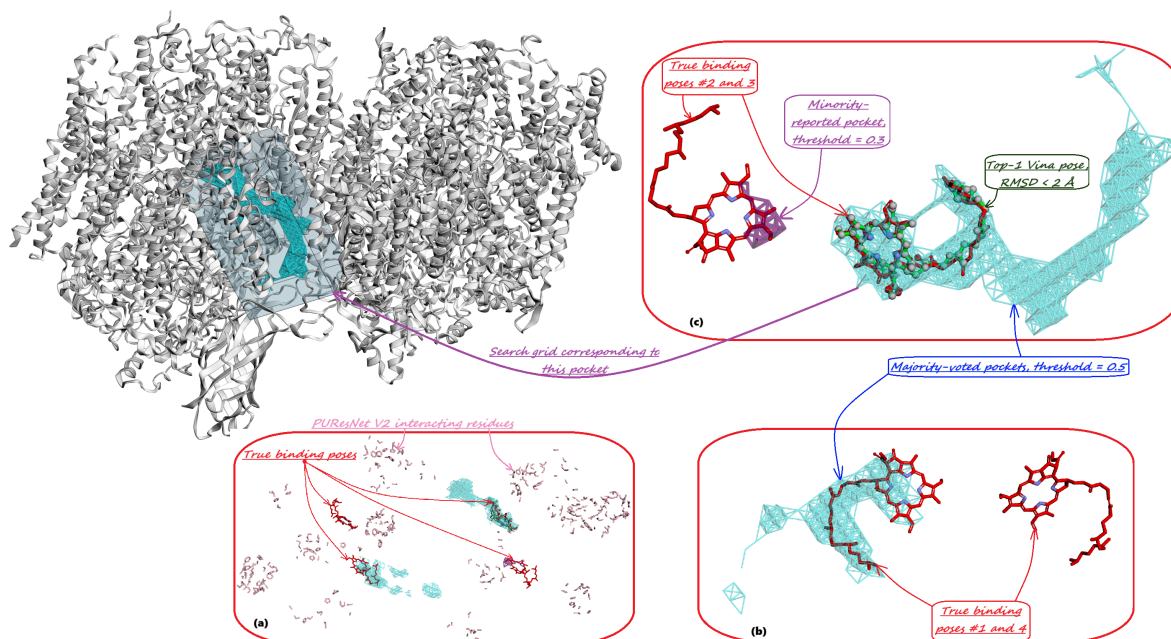


Figure 2. For the 8F4J protein structure from the PoseBusters [28] dataset, our docking approach achieves RMSD < 2 Å, while AlphaFold 3 [30] is unable to process the full structure due to input size limitations. Unlike residue-focused methods such as PURESNet V2 [66], which produce complex predictions that are difficult to interpret, our approach offers clear, plug-and-play guidance for docking grid selection, as illustrated in the bottom panel.

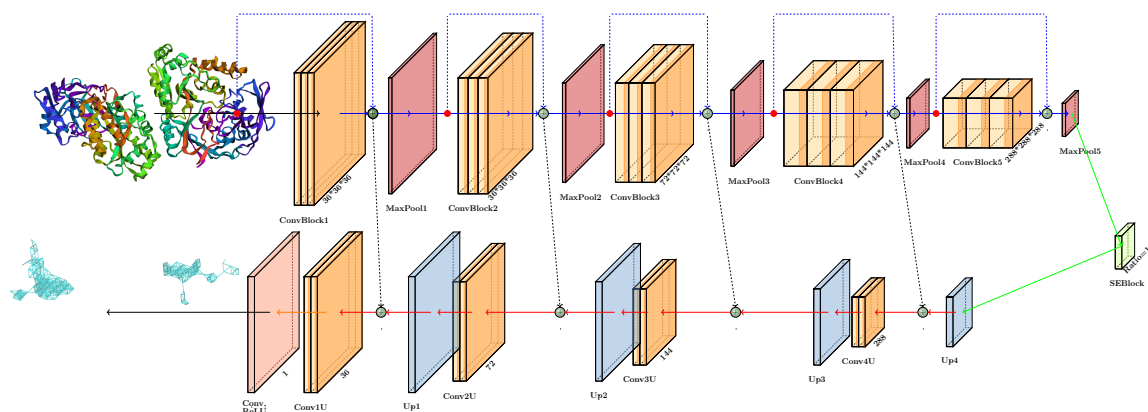


Figure 3. Schematic representation of **RAPID-Net** (ReLU-Activated Pocket Identification for Docking). Key improvements that distinguish RAPID-Net from previous approaches include a ReLU activation in the final layer, the usage of a soft dice loss function, including a single SE-attention block, and removing redundant residual connections.

4. Model Training and Inference Pipeline

We trained our model using the sc-PDB [81] dataset, which contains protein structures with annotated binding sites. In this dataset, cavities are defined using VolSite [82], which maps the pharmacophore properties of nearby protein atoms onto a three-dimensional grid. This method assumes that ligand-induced conformational changes remain relatively small [83], treating each cavity atom as a “pseudoatom” to denote an interaction point rather than a physical atom.

Following [69], we used a curated subset of sc-PDB [81] in which redundant protein structures were filtered out based on their Tanimoto index [84]. Each protein structure was then placed in a $36 \times 36 \times 36$ grid, with each voxel corresponding to a $2\text{\AA} \times 2\text{\AA} \times 2\text{\AA}$ unit cell. We extracted 18 features per voxel using the `tfbio` [85] package to provide a representation of the protein environment for the model training.

In contrast to previous studies, we apply the medical image segmentation practices [86–88] to perform *soft* segmentation of ligand-binding pockets, rather than treating it as a strict binary classification problem. In Kalasanty [70] and PURESNet [69], each voxel is considered part of a binding pocket if it contains at least one cavity pseudoatom. The number of pseudoatoms per voxel is then clipped to the range $[0, 1]$, and the model is trained with a Dice loss function as part of a binary classification task:

$$\Delta_{\text{Dice}}(V_{\text{true}}^{\text{clipped}}, V_{\text{predicted}}^{\text{clipped}}) = 1 - \frac{2|V_{\text{true}}^{\text{clipped}} \cap V_{\text{predicted}}^{\text{clipped}}|}{|V_{\text{true}}^{\text{clipped}}| + |V_{\text{predicted}}^{\text{clipped}}|}, \quad (1)$$

where $V_{\text{true}}^{\text{clipped}}$ and $V_{\text{predicted}}^{\text{clipped}}$ represent the binary masks of the true and predicted pockets, respectively, \cap indicates their intersection, and $|V_{\text{true}}^{\text{clipped}}|$ and $|V_{\text{predicted}}^{\text{clipped}}|$ denote the number of occupied voxels in the true and predicted masks, respectively.

Recently, PURESNet V2 [66] sought to improve performance by replacing the Dice loss function with a focal loss [89]. However, our extensive experiments indicate that adopting a “soft” approach [86–88] is better suited in this context. In contrast to the binary “yes/no” classification of occupied voxels, the density of pseudoatoms serves as an indicator of the proximity of a voxel to the interior or boundary of a pocket. Higher densities typically occur near polar or charged residues and ligand functional groups, whereas lower densities are often associated with hydrophobic regions.

Drawing inspiration from medical image segmentation methods [86–88], we replace the sigmoid output in the final layer with a ReLU [90] (hence the name of our model, ReLU Activated Pocket Identification for Docking, **RAPID-Net**):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \rightarrow \text{ReLU}(x) = \max(0, x) \quad (2)$$

For the model training, we also replace the conventional Dice loss in Eqn. 1 with its “soft” variant:

$$\Delta_{\text{Soft Dice, } L^2}(V_{\text{true}}, V_{\text{predicted}}) = 1 - \frac{2 \sum_{\text{Grid}} (V_{\text{true}} V_{\text{predicted}})}{\sum_{\text{Grid}} (V_{\text{true}}^2 + V_{\text{predicted}}^2)} \quad (3)$$

Several ways of implementing the “soft” Dice loss have been proposed [87,88], but we found that this simplest version, based on the L^2 norm, works best.

To mitigate class imbalance, as the number of non-interacting voxels sharply exceeds the number of interacting ones, we applied class reweighting using `scikit-learn` [91]. During inference, a voxel is classified as part of a binding pocket if the model output exceeds `threshold = 0.5`. When ensembling five models, we apply morphological closing using the `binary_closing` function from the `scipy.ndimage` package [92] to mitigate sparsity in the predicted pocket regions across model replicas.

Unlike previous approaches that use `cavity6.mol2` labels [69,70], we train our model with threshold-less pocket labels (`cavityALL.mol2`). In `cavity6.mol2`, annotations are restricted to regions within 6.5 Å of the ligand’s heavy atoms [81], potentially overlooking distal functional regions such as allosteric pockets, exosites, or flexible loops. These sites can critically influence drug binding and resistance, making them key therapeutic targets. Secondary binding sites, which often lie beyond the 6.5 Å boundary, mediate interactions with larger substrates or cofactors and help position them within the catalytic pocket, as further discussed in Section 10. By adopting threshold-less labels, our model identifies both catalytic and more distant sites, thereby improving predictions for proteins that contain such secondary binding regions.

This approach also makes the Distance Center Center (DCC) metric, which was previously used to evaluate Kalasanty [70] and PURESNet [69], less relevant. DCC defines the distance between the geometrical center of the predicted pocket and the true ligand binding pose, but “tunnels” can extend from the primary binding pocket to distal residues that lie far from the ligand but have significant therapeutic implications (see Section 10 for examples). Furthermore, the threshold-less pocket definition is less sensitive to the ligand size, thereby improving the generalizability of our model.

5. Pocket-informed Docking Protocol and its Rationale

For each predicted pocket, we define the center of the search grid as the average of the maximum and minimum x , y , and z coordinates of all pseudoatoms in the pocket:

$$\text{center}(x, y, z) = \frac{\max(x, y, z) + \min(x, y, z)}{2}, \quad (4)$$

where $\max(x, y, z)$ and $\min(x, y, z)$ are the maximum and minimum coordinates, respectively. The grid dimensions are then given by:

$$\Delta x, \Delta y, \Delta z = (\max(x, y, z) - \min(x, y, z)) + 2 \cdot \text{threshold} \quad (5)$$

Figure 4 illustrates this setup for the majority-voted pocket in the ABHD5 protein [93] with a threshold value of 2 Å.

center: -5.001999855041504, 1.998999834060669, -10.999000549316406
 bsize: 24.0, 22.0, 24.0

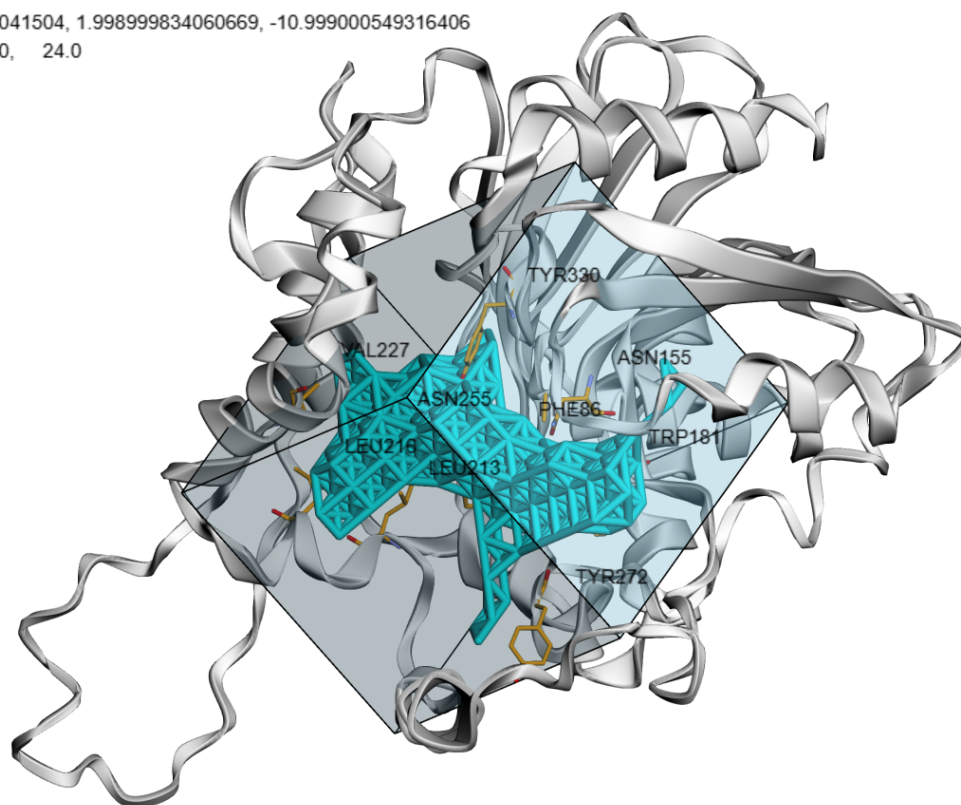


Figure 4. Majority-voted pocket predicted by RAPID-Net and its corresponding search grid with 2 Å threshold for ABHD5 protein [93].

For each majority-voted pocket, we generate search grids with thresholds of 2 Å and 5 Å. For minority-reported pockets, we use thresholds of 2 Å, 5 Å, 10 Å, and 15 Å. This strategy accounts for large ligands, such as those in the PoseBusters [28] dataset, which may extend beyond the predicted pockets, and also reduces potential inaccuracies in our model's pocket predictions.

As an example, Figure 5 shows the 8FAV protein from the PoseBusters dataset [28], where none of the predicted pockets overlap with one of the true ligand binding poses. However, the docking still succeeds because the expanded search grid thresholds encompass the true binding site. This case demonstrates how geometry-based evaluation metrics, such as those in [26], can overlook the practical success of docking, even if the pocket predictions are imperfect.

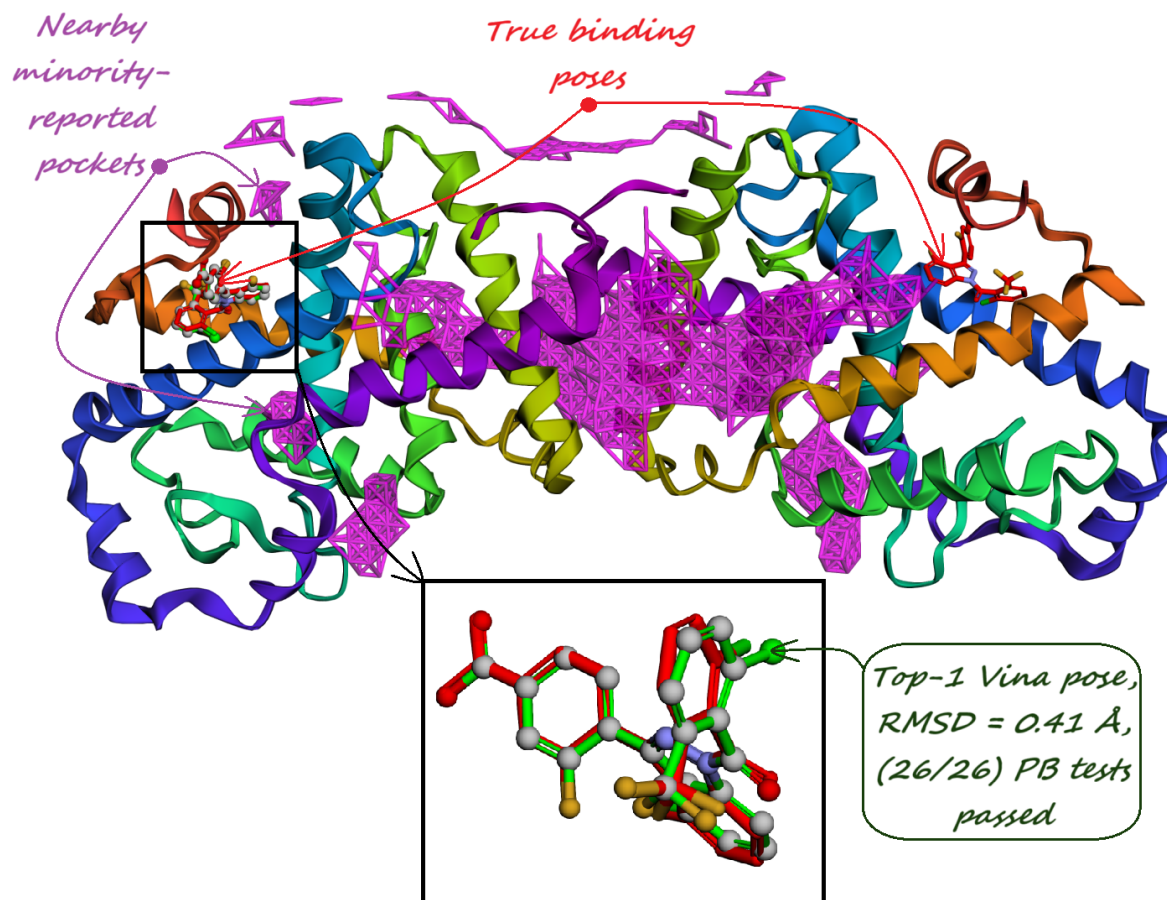


Figure 5. For the 8FAV protein structure, our model predictions are inaccurate, and none of the true ligand binding poses are within one of the predicted pockets. Nevertheless, the docking is successful because the search grids with large thresholds smooths out this drawback.

To demonstrate that the accuracy of our model is primarily a result of architectural and training improvements—not merely ensembling—we also provide single-model predictions from Kalasanty [70], PURESNet [69], and individual RAPID-Net runs. Similarly, for each predicted pocket, we generate grids with 2 Å, 5 Å, 10 Å, and 15 Å thresholds.

Finally, we perform molecular docking on each identified grid using AutoDock Vina 1.2.5 [12], with default settings (Exhaustiveness= 32, num_modes= 40), following the protocol described in [28].

6. Test Benchmarks and Evaluation Metrics

We evaluate docking performance as the percentage of Top-1 Vina poses with a root mean square deviation (RMSD) below 2 Å between the predicted and one of the true ligand poses if multiple true poses are available. RMSD is calculated as:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \|\mathbf{r}_i^{\text{true}} - \mathbf{r}_i^{\text{pred}}\|^2}, \quad (6)$$

where N is the number of ligand atoms, $\mathbf{r}_i^{\text{true}}$ and $\mathbf{r}_i^{\text{pred}}$ are positions of the i -th atom in the true and predicted poses, respectively, and $\|\cdot\|$ is the Euclidean norm. RMSD is computed using the CalcRMS function in RDKit [94].

Docking methods like AutoDock Vina [12] balance speed and accuracy [95], utilizing “fast and dirty” scoring functions. A common strategy to enhance their accuracy is to generate an ensemble of potential poses and then identify the best pose among them using more accurate scoring tools, for example [96,97]. Fast scoring functions enable the efficient exploration of numerous poses, while more

accurate methods are reserved for refinement and final pose selection. To account for this approach, we also report the sampling RMSD accuracy, defined as the percentage of cases with “at least one correct pose in the ensemble”. This metric underscores the potential gains that can be achieved through more accurate rescoring. Since our primary focus is on improving pocket identification rather than developing new rescoring functions, we defer such work to the future. Additionally, we evaluate the PB-success rate, defined as the percentage of poses with $\text{RMSD} < 2 \text{ \AA}$ that pass all chemical validity checks defined by PoseBusters [28].

Finally, similarly to [69], we compute the Pocket-Ligand Intersection (PLI) score to quantify the proportion of ligand atoms residing within the predicted pocket. Unlike [69], which computes a voxel-based intersection, we measure the average fraction of ligand heavy atoms located within 5 \AA of at least one pocket pseudoatom, yielding a more intuitive metric:

$$\text{PLI} = \frac{N_{\text{within } 5 \text{ \AA}}}{N_{\text{total}}}, \quad (7)$$

where $N_{\text{within } 5 \text{ \AA}}$ is the number of ligand heavy atoms within 5 \AA of any pocket pseudoatom, and N_{total} is the total number of heavy atoms in the ligand.

However, unlike [69], we compute this ratio for *all* protein-ligand pairs in the dataset rather than restricting it to cases where the ligand is centered within $\leq 4 \text{ \AA}$ of the pocket center. If no pockets are predicted, the PLI is set to zero. When a protein has multiple predicted pockets or multiple ligand poses, we report the maximum PLI value across all possibilities. For completeness, we also indicate how many proteins have at least one predicted pocket and how many have at least one ligand pose within the defined search grids.

Our primary dataset for evaluating docking performance is the PoseBusters [28] dataset, which consists of novel protein-ligand complexes that were not available during the development of current docking programs and thus provides a rigorous assessment of their generalization capabilities. For instance, several docking algorithms that claimed superior accuracy on CASF 2016 [95] dataset, perform drastically worse on PoseBusters [28]. Additionally, we evaluate docking accuracy on the Astex Diverse Set [33]. Although this dataset is older than PoseBusters [28], it remains a challenging benchmark for many algorithms that had previously reported high accuracy on CASF 2016 [95].

We distinguish between two docking scenarios: those that use “prior knowledge” of the binding site (which is typically unavailable in real-world applications) and truly “blind” settings that lack this information. In both the PoseBusters [28] and Astex [33] datasets, when the binding site is assumed to be known, the reference ligand coordinates are used to center a $25 \text{ \AA} \times 25 \text{ \AA} \times 25 \text{ \AA}$ bounding box. However, to evaluate RAPID-Net under truly “blind”, binding-site-agnostic conditions, we omit these reference ligands entirely.

In addition to docking accuracy, for completeness, we report the PLI rates for PoseBusters [28] and Astex [33] datasets. For direct comparison with PURESNet [69], we also evaluate PLI for the Coach420 [34] and BU48 [35] datasets, excluding any structures that appear in the training set as specified in [69]. Lastly, to highlight the advantages of our model architecture and training approach, we report docking accuracy for both single RAPID-Net models and their ensembled version. PLI is calculated for individual models, majority-voted pockets, and minority-reported pockets.

7. Evaluation on the PoseBusters Dataset

We obtained the following results by docking the PoseBusters [28] dataset using the aforementioned protocol. Figure 6 shows the percentage of predictions with $\text{RMSD} < 2 \text{ \AA}$ and the corresponding PB-success rates. When guided by the ensembled RAPID-Net, AutoDock Vina [12] achieved **55.8%** RMSD-correct predictions and a **54.9%** PB-success rate, outperforming DiffBindFR [27], which scored **50.2%** and **49.1%**, respectively. This result highlights the crucial impact of accurate pocket identification on the subsequent docking accuracy. Our approach uses targeted docking with the widely used AutoDock Vina [12], whereas DiffBindFR [27] scans the entire protein surface.

Furthermore, as shown in Figure 6, docking driven by individual RAPID-Net models—the ensemble components—outperforms PURESNet [69] and Kalasanty [70] in docking accuracy. Ensembling RAPID-Net models further improves docking accuracy by providing more stable results. The distribution of RMSD values for the Top-1 predicted Vina poses is shown in Figure 7. Note the issue of low recall discussed above. While the RAPID-Net ensemble predicts pockets in all 308 structures, PURESNet [69] and Kalasanty [70] predict pockets in only 278 and 263 protein structures, respectively. Since we evaluate accuracy in “blind” docking settings, cases with no predicted pockets are treated as failures by default.

Not all predicted pockets are positioned correctly. Of the 308 structures where RAPID-Net predicts at least one pocket, 307 contain at least one true ligand binding pose entirely within a search grid, meaning that docking can be successful in 307 out of 308 cases. In the remaining case, docking would inevitably fail due to an inaccurate search grid location. By contrast, PURESNet [69] and Kalasanty [70] have only 265 and 249 structures, respectively, where at least one true ligand pose is located completely within the search grid. This demonstrates that our model not only accurately identifies pockets in favorable protein structures but also provides reliable predictions across the entire dataset.

For comparison, Figure 8 presents the evaluation on a subset of novel protein-ligand complexes from the PoseBusters benchmark, as defined by [31]. Compared to the best-performing structure prediction model, AlphaFold 3, as evaluated by PoseBench, RAPID-Net + Vina achieves a Top-1 success rate of 53.1%, closely approaching AF3’s 59.5%. This result highlights RAPID-Net’s ability to achieve competitive accuracy while offering a substantially lower computational cost.

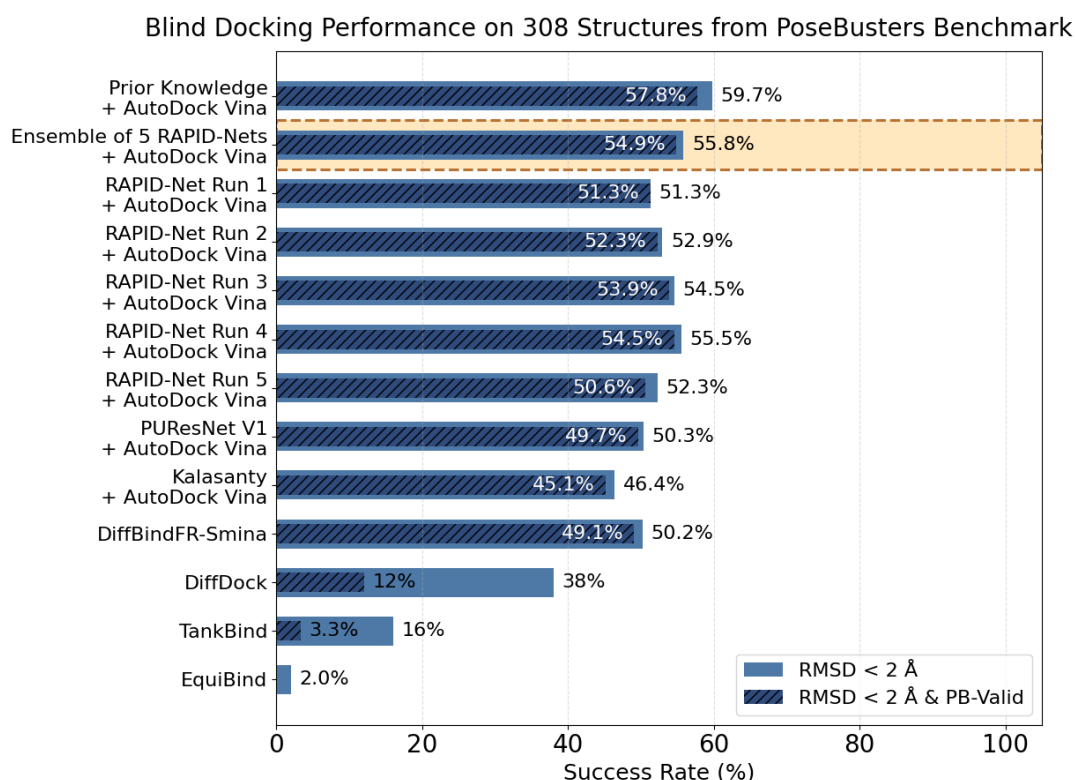


Figure 6. Comparison of Top-1 Vina accuracies when guided by different pocket identification algorithms, and a DiffBindFR-Smina [27], for PoseBusters [28] dataset.

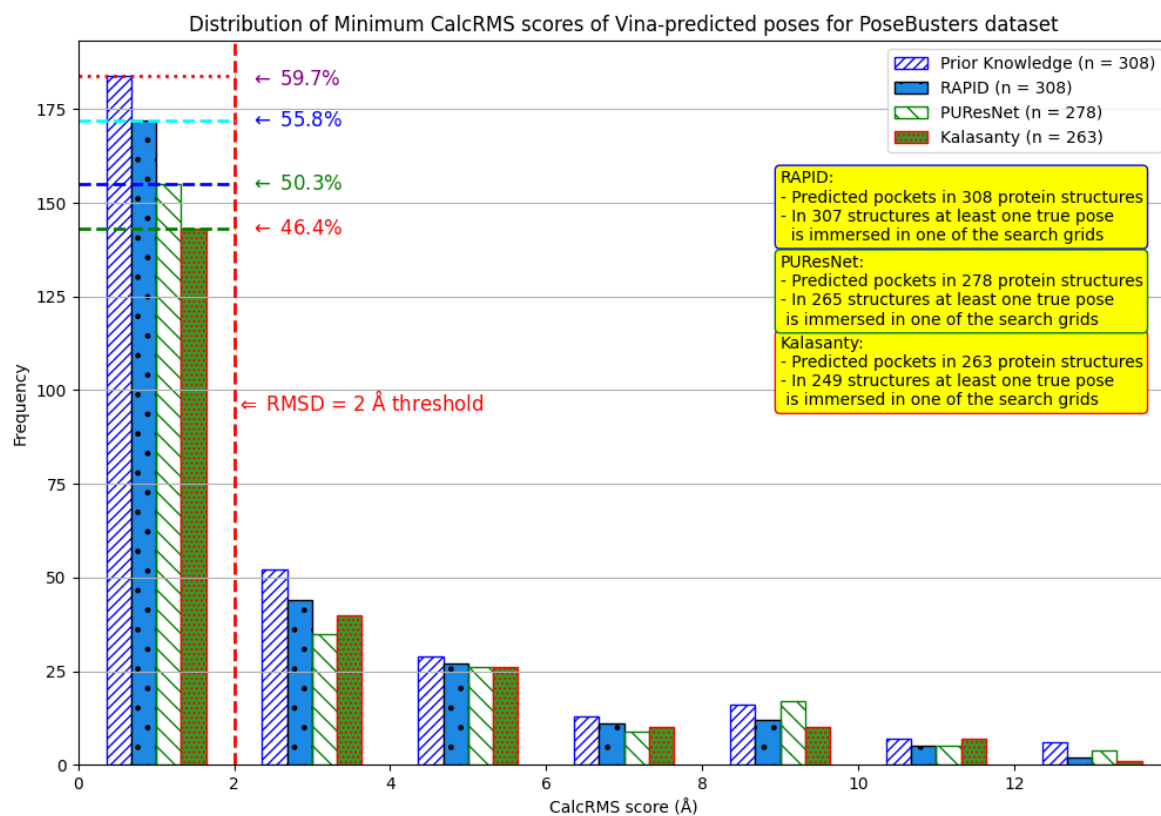


Figure 7. Distribution of RMSD values for Top-1 Vina predicted poses in the PoseBusters [28] dataset. When multiple true ligand poses are available, RMSD to the closest one is reported.

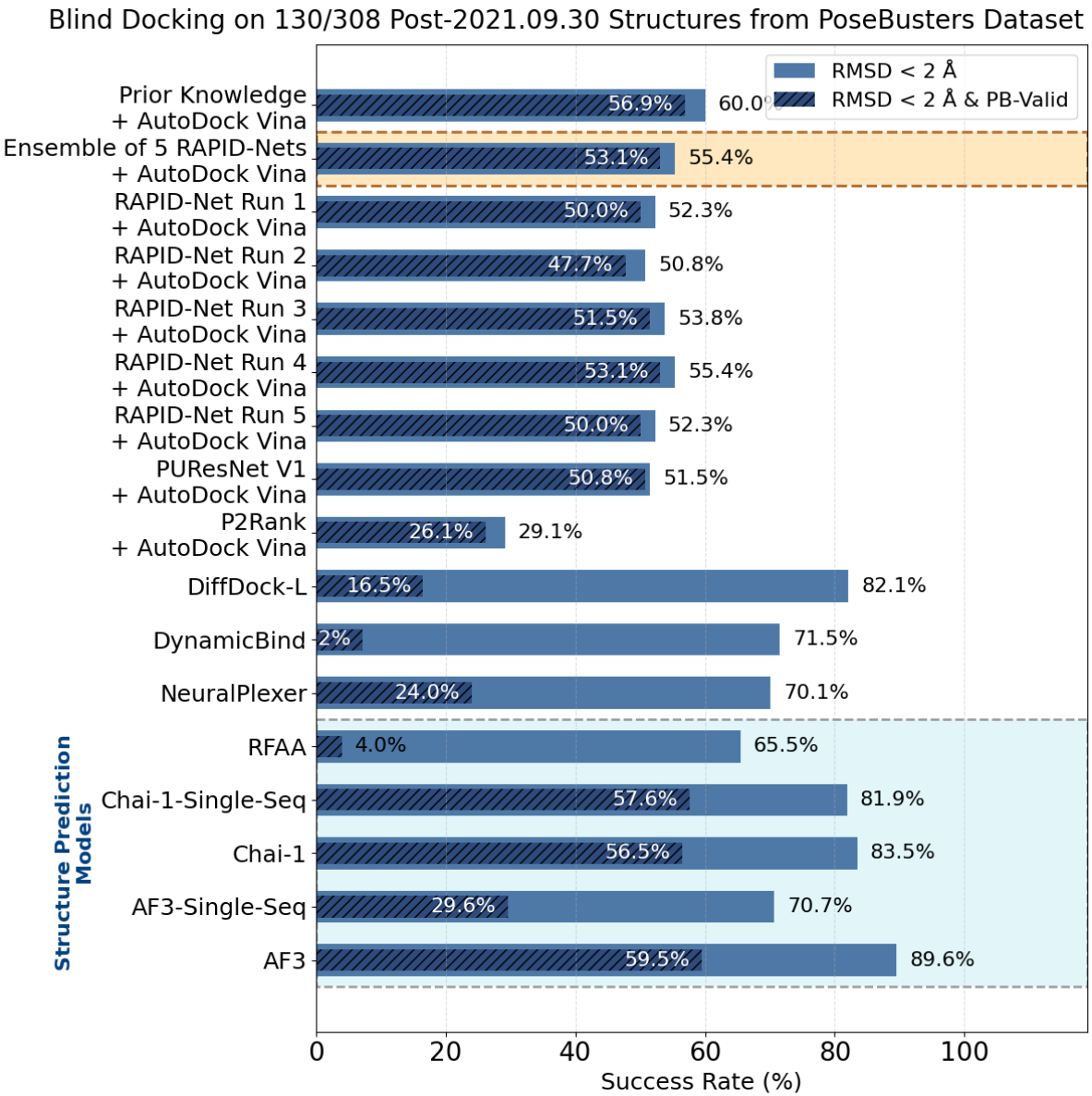


Figure 8. Performance comparison on the subset of novel PoseBusters structures as suggested in [31]. The RAPID-Net-guided AutoDock Vina achieves 53.1% compared to 59.5% of AlphaFold 3, with significantly lower computational costs.

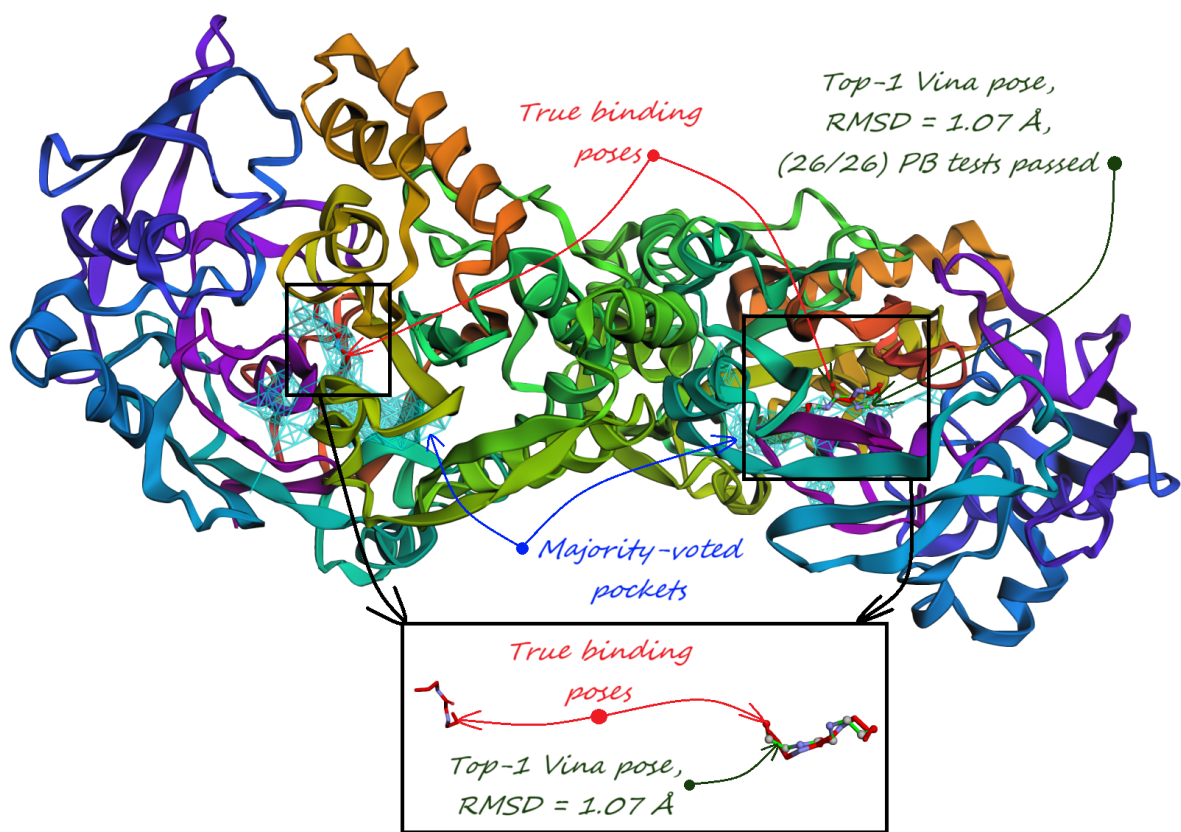


Figure 9. In 7KZ9 protein structure, RAPID-Net predicts majority-voted pockets around both true binding poses, and the Top-1 Vina pose passes all tests.

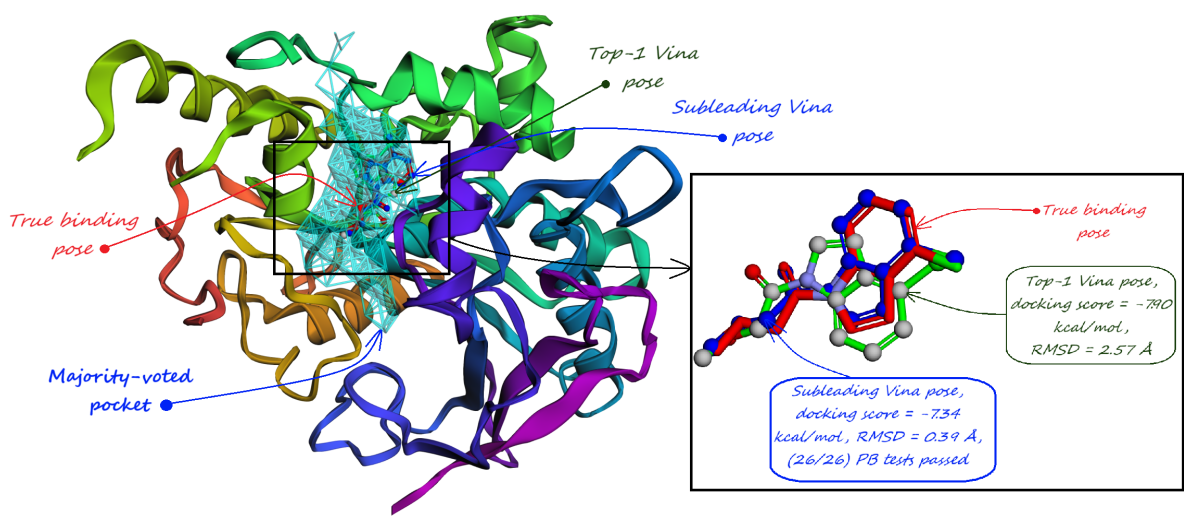


Figure 10. For the 8BTI protein structure, the Top-1 Vina pose is incorrect, but the subleading pose passes all tests.

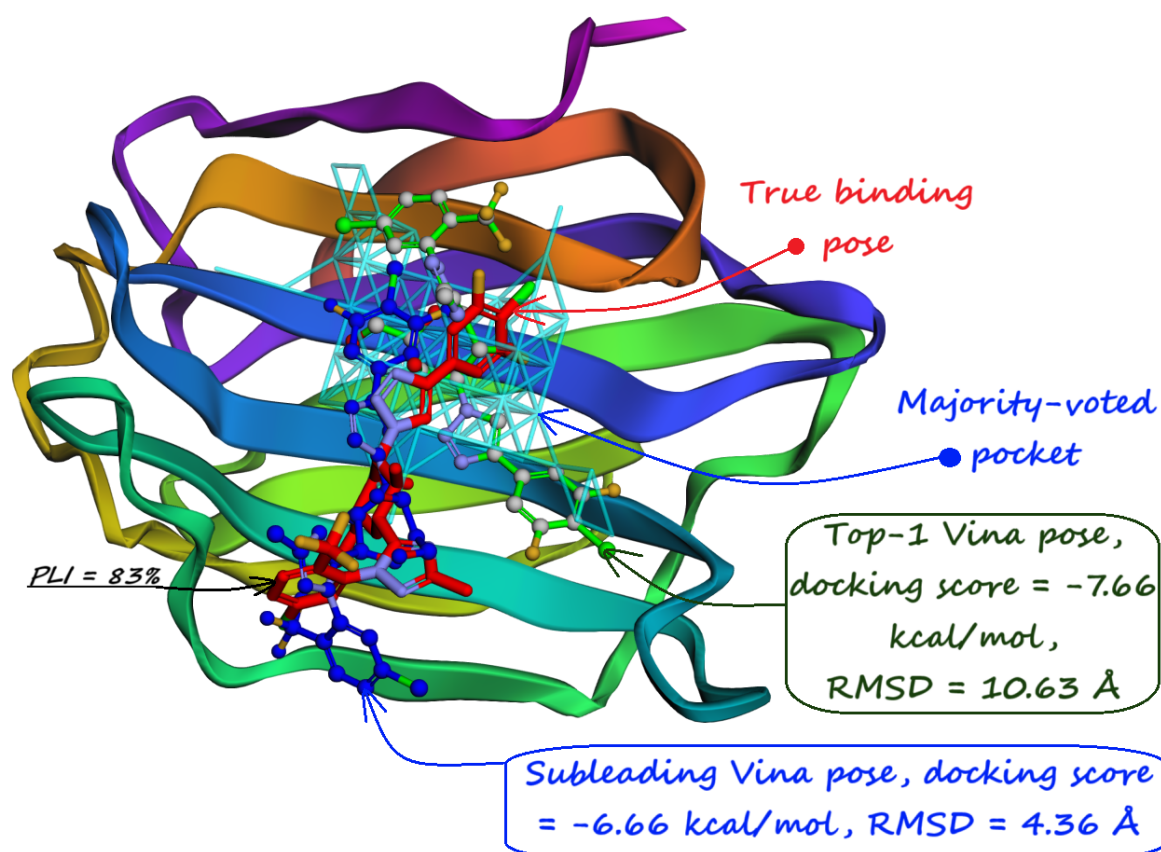


Figure 11. For the 7XFA protein structure, neither the Top-1 nor any of the subleading Vina predicted poses pass the RMSD test, despite the predicted pocket covering the ligand with PLI = 83%.

For the 8BTI protein structure illustrated in Figure 10, the Top-1 Vina pose fails, but one of the subleading poses passes all tests. In contrast, for the 7XFA protein structure shown in Fig. 11, neither the Top-1 nor any of the subleading Vina poses satisfy the RMSD test, with the closest pose in the ensemble having $\text{RMSD} = 4.36 \text{ \AA}$. This occurs even though the majority-voted pocket predicted by our model covers most of the true ligand binding pose with $\text{PLI} = 83\%$.

Similarly to the 8BTI example in Figure 10, many protein structures have the following pattern: the **Top-1** AutoDock Vina pose fails the $\text{RMSD} < 2 \text{ \AA}$ criterion, yet a lower-ranked pose meets it. To illustrate this behavior, we report the *sampling accuracy*, defined as the fraction of cases in which *at least one* pose in the ensemble achieves $\text{RMSD} < 2 \text{ \AA}$. Under this metric, as shown in Figure 12, AutoDock Vina with the prior knowledge of the search box [12] has a success rate of 93.8%.

In comparison, AlphaFold 3, when run with pocket information available, achieves a **Top-1** accuracy of 93.2% with $\text{RMSD} < 2 \text{ \AA}$ [30]. We did not compute AlphaFold 3's sampling accuracy and therefore make *no claim* of overall superiority. This comparison merely illustrates that the current limitations in docking accuracy stem more from pose *ranking* than from conformational sampling.

This difference widens in fully “blind” docking settings: AutoDock Vina, guided by a RAPID-Net, achieves the 92.2% of $\text{RMSD} < 2 \text{ \AA}$ *sampling accuracy*, whereas AlphaFold 3 achieves 80.5% of $\text{RMSD} < 2 \text{ \AA}$ **Top-1** accuracy. These results suggest that improved re-ranking algorithms could further boost pocket-guided docking workflows. Furthermore, RAPID-Net performs better than PURESNet [69] and Kalasanty [70] in guiding AutoDock Vina [12], as can be observed in Figure 12.

To better illustrate this pattern, we consider four representative cases in which the **Top-1** AutoDock Vina pose fails to reach the $\text{RMSD} < 2 \text{ \AA}$ threshold, while a pose with a much lower ranking succeeds.

Figure 13 illustrates the 7PRI protein structure, which contains two true ligand binding poses—one largely covered by the majority-voted pocket predicted by RAPID-Net, and the other by a minority-

reported pocket. The Top-1 Vina pose, which fails the $\text{RMSD} < 2 \text{ \AA}$ test, is docked in the minority-reported pocket. A subleading pose in the same pocket passes the test, underscoring the importance of minority-reported pockets in occasionally yielding correct predictions.

Figure 14 shows the 7P1F protein structure, where both true ligand binding poses lie largely in the majority-voted pockets predicted by RAPID-Net. Although the Top-1 Vina is docked in one of these pockets and fails the $\text{RMSD} < 2 \text{ \AA}$ test, a subleading pose lands in another majority-voted pocket, closer to the second true ligand pose, and passes the test.

Figure 15 illustrates the 7NUT protein with two true ligand binding poses, both largely within the majority-voted pockets. Both the Top-1 Vina pose, which failed the $\text{RMSD} < 2 \text{ \AA}$ test, and the subleading pose, which passed the test, are located in the same majority-voted pocket near one of the true binding poses.

Finally, in the 7A9E protein structure shown in Figure 16, RAPID-Net predicts a single majority-voted pocket. The Top-1 Vina pose is docked in the wrong part of that pocket, while a subleading pose successfully passes the $\text{RMSD} < 2 \text{ \AA}$ test by aligning with one of the true binding poses in another part of the pocket.

These results illustrate that the AutoDock Vina [12] sampling mechanism is robust and reliably generates near-native binding poses within the ensemble. However, its main limitation is its ranking power: although it reliably generates poses, it struggles to distinguish correct poses from incorrect ones in its ranking.

Finally, Figure 17 shows the distribution of PLI values for PURESNet [69], Kalasanty [70], and RAPID-Net for this dataset. All five replicas of RAPID-Net outperform both PURESNet [69] and Kalasanty [70]. Majority voting stabilizes the PLI predictions, and including minority-reported pockets further improves the PLI.

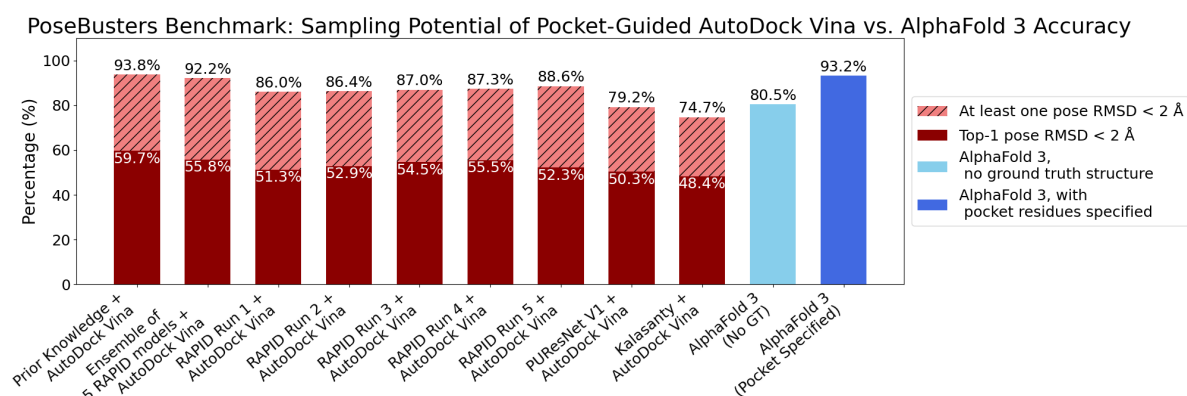


Figure 12. Solid bars represent Top-1 accuracies. Dashed segments indicate the accuracy achievable when at least one pose in the ensemble is correct—potentially attainable with a more accurate reweighting tool.

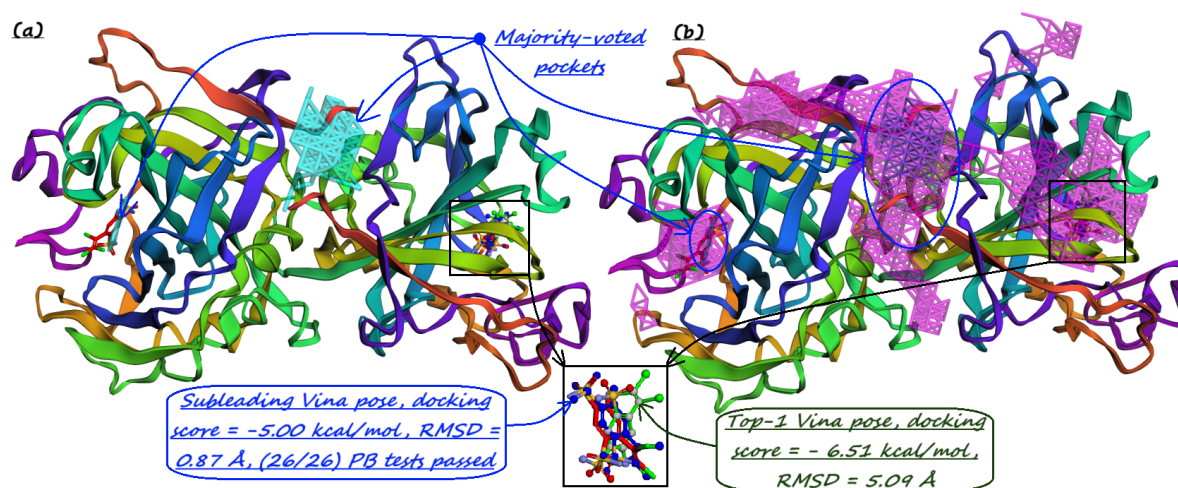


Figure 13. For the 7PRI protein structure, the Top-1 Vina pose fails while a subleading one succeeds. In (a), the majority-voted pockets are shown in cyan, and one of the true ligand binding poses largely overlaps with one of them. In (b), minimally-reported pockets are displayed as purple sticks.

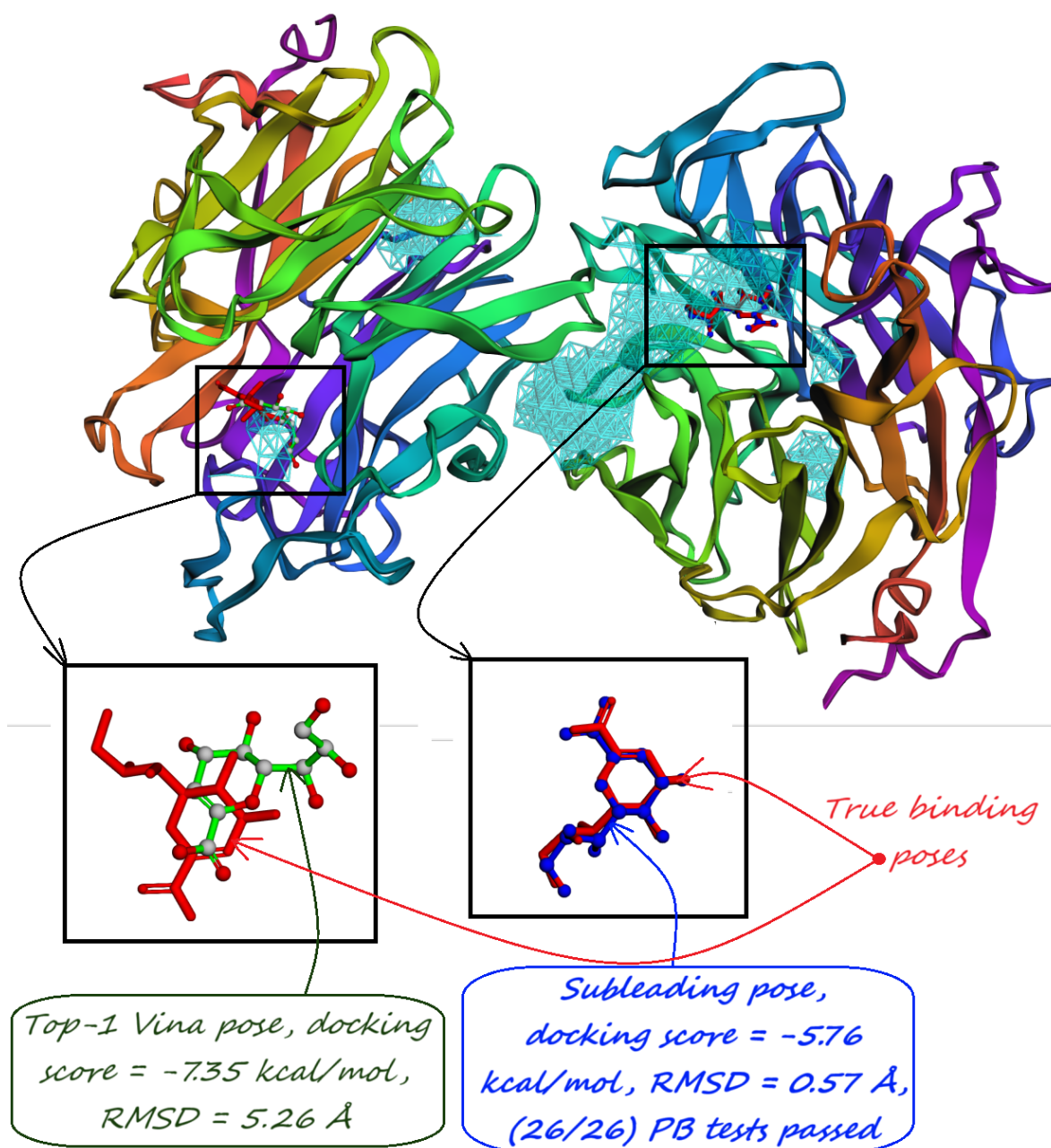


Figure 14. In the 7P1F protein structure, two true ligand binding poses exist, both largely within the majority-voted pockets predicted by RAPID-Net. The Top-1 Vina pose is within one of these pockets but is inaccurate, failing the $\text{RMSD} < 2 \text{ \AA}$ test. However, a subleading pose occupies a different pocket and passes all tests.

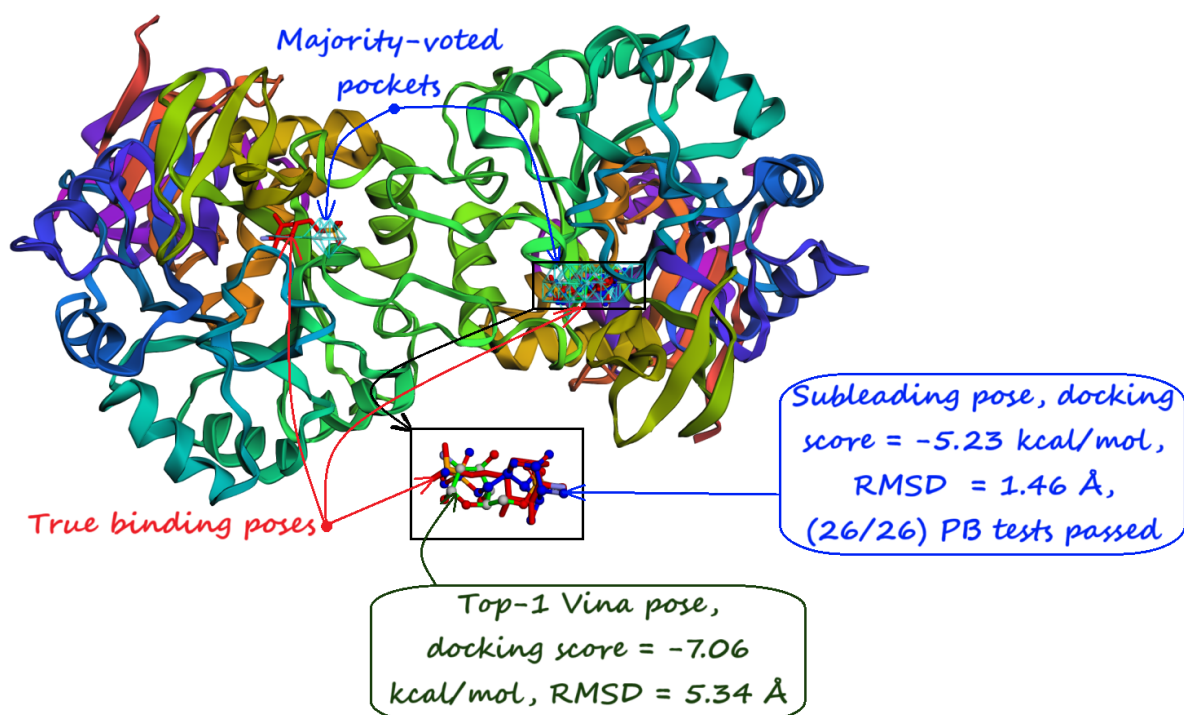


Figure 15. For the 7NUT protein structure, there are two true ligand binding poses, both largely within majority-voted pockets predicted by RAPID-Net. The Top-1 Vina pose failing the $\text{RMSD} < 2 \text{ \AA}$ test and a subleading pose passing the test are in the same majority-voted pocket.

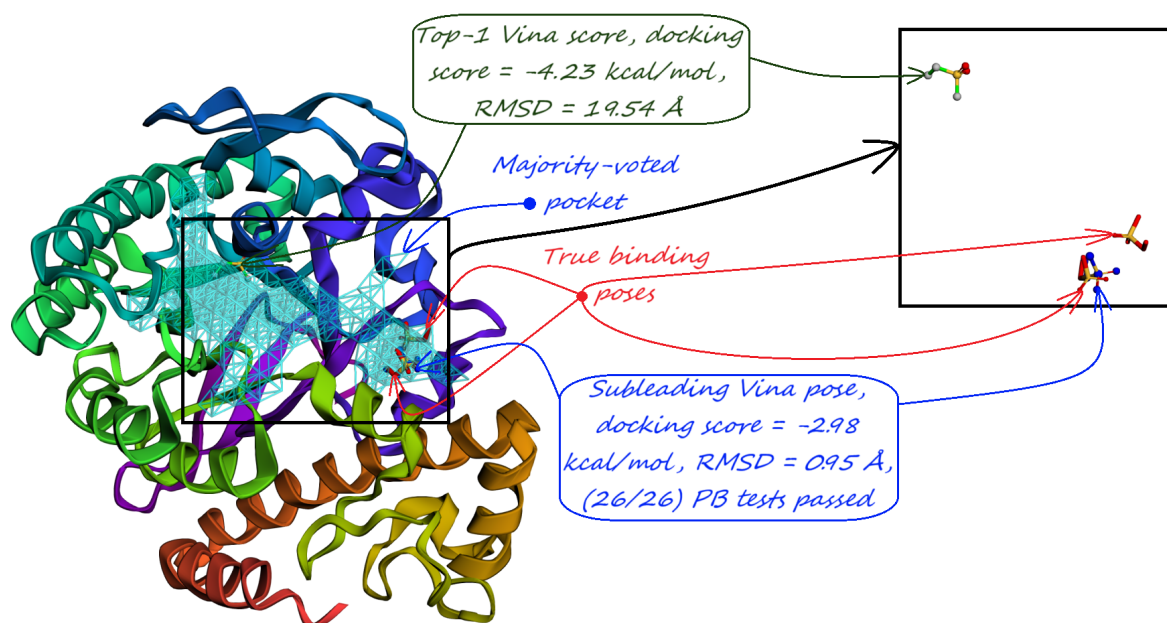


Figure 16. In the 7A9E protein structure, RAPID-Net predicts a single majority-voted pocket. The Top-1 Vina pose is located in an incorrect part of that pocket, whereas a subleading pose that passes the $\text{RMSD} < 2 \text{ \AA}$ test—aligning with one of the true ligand binding poses—is found in another part of the same pocket.

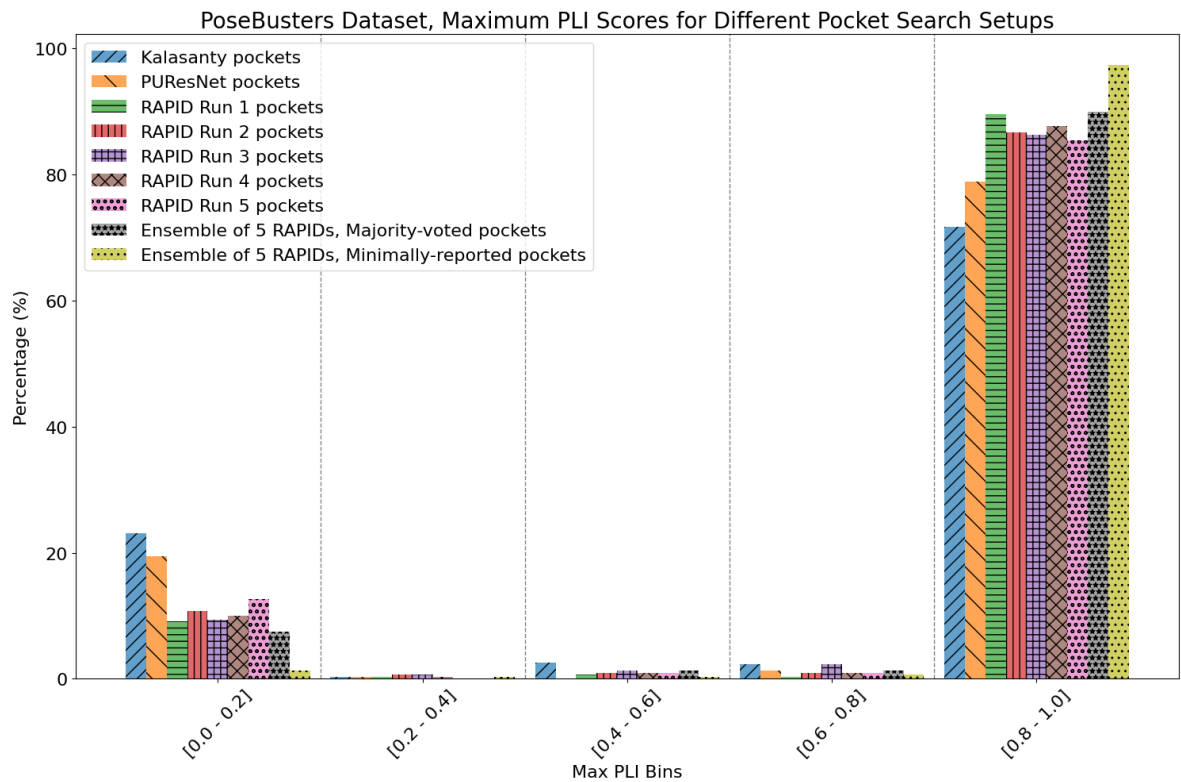


Figure 17. Distribution of the maximum PLI scores corresponding to the pockets predicted by RAPID-Net, PURESNet [69], and Kalasanty [70] for the PoseBusters dataset [28].

To highlight the necessity of model ensembling to provide stable pocket predictions, Table 1 summarizes the level of pocket and ligand overlap and lists:

- The number of protein structures with at least one predicted pocket (shown in the “Nonzero” column). Since no protein in all our datasets is a true negative, failure to predict any pocket for a given structure is automatically an incorrect prediction.
- The number of structures with viable search grids encompassing at least one true ligand binding pose (“Within 15 Å” column), where the docking may succeed.
- The average Pocket-Ligand Intersection (PLI) value for all protein structures.

Table 1. Comparison of pocket predictions across different methods for PoseBusters [28] dataset consisting of 308 protein structures.

Method	Nonzero	Within 15 Å	PLI
Kalasanty	263	249	74.34%
PURESNet	278	265	79.51%
RAPID Run 1	307	298	90.17%
RAPID Run 2	305	289	87.90%
RAPID Run 3	308	296	88.71%
RAPID Run 4	308	300	88.57%
RAPID Run 5	308	292	86.32%
RAPID ensemble	308	307	91.44%/98.09%

For the RAPID-Net model, individual runs show variability in pocket prediction quality. For example, Run 5 predicts pockets for all 308 protein structures but with lower accuracy, resulting in the lowest average PLI of 86.32% among the five runs. In contrast, Run 1 fails to predict pockets for one structure out of 308 but achieves the highest average PLI of 90.17% among five runs. Combining predictions via majority voting among these RAPID-Net runs increases the average PLI to 91.44%. Further incorporating minority-reported pockets boosts the average PLI even higher, to 98.09%. In

the final ensemble of pockets predicted by RAPID-Net, all 308 protein structures have at least one predicted pocket and only one structure lacks a search grid covering at least one true ligand binding pose. These observations highlight the importance of combining pocket predictions from multiple runs to achieve reliable results for subsequent docking.

In the next Section 8, we perform docking on the Astex Diverse Set [33] following a similar procedure as described for the PoseBusters [28] dataset.

8. Evaluation on the Astex Diverse Set

The docking results of AutoDock Vina [12] guided by different pocket-finding algorithms for the Astex Diverse Set [33], consisting of 85 protein structures, are shown in Figure 18. The corresponding RMSD distribution of the Top-1 predicted Vina poses is presented in Figure 19, and the associated PLI scores are shown in Figure 20.

Several striking features emerge from these graphs. First, although each individual Run of our model outperforms the docking accuracy achieved when guiding AutoDock Vina [12] using PUPresNet [69] or Kalasanty [70], ensembling these individual Runs does not improve the overall docking accuracy. This happens even though the ensembled version of RAPID-Net achieves a higher ligand-pocket intersection, as shown in Figure 20 and Table 2.

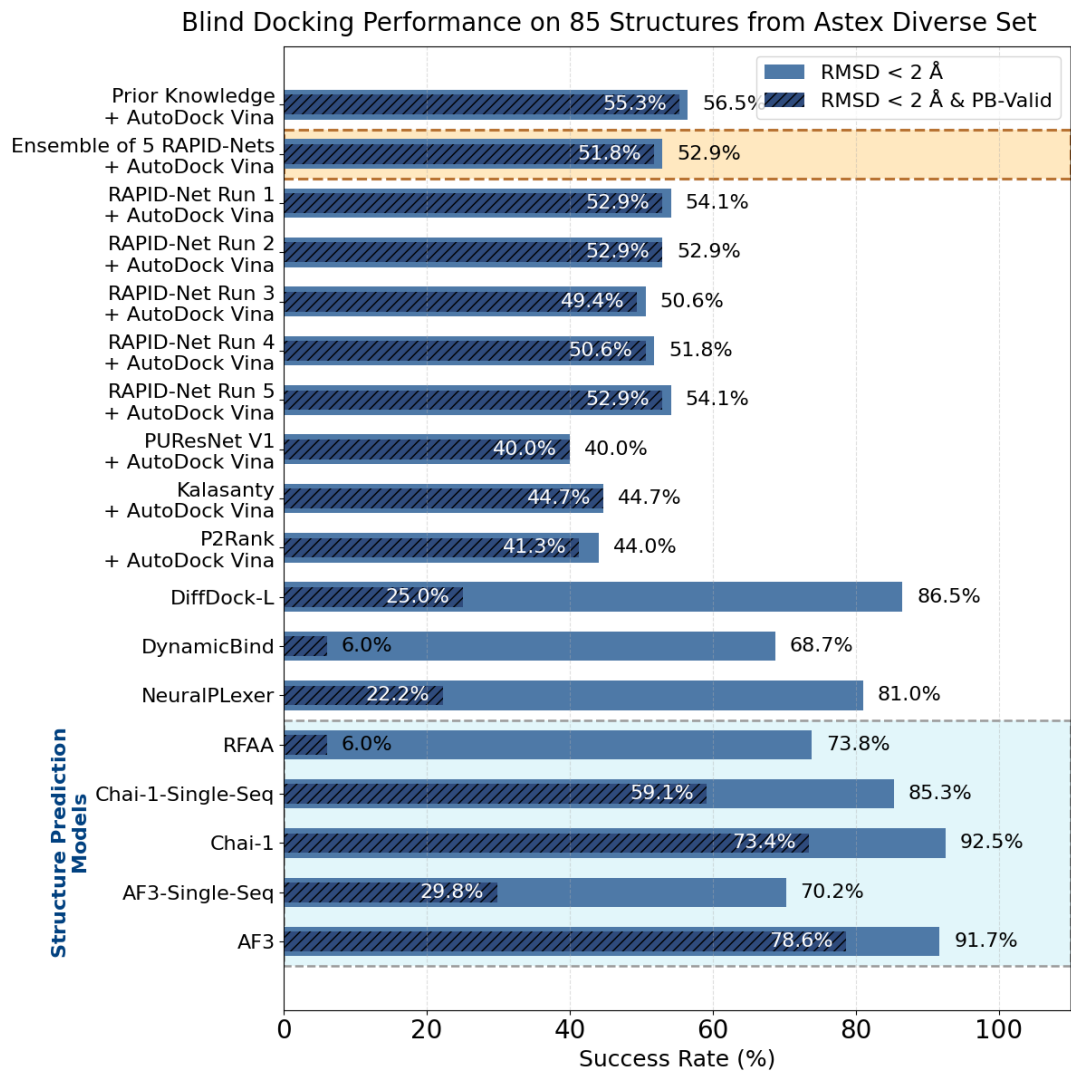


Figure 18. Comparison of docking accuracies for the Astex Diverse Set [33] when AutoDock Vina [12] is guided by different pocket prediction algorithms.

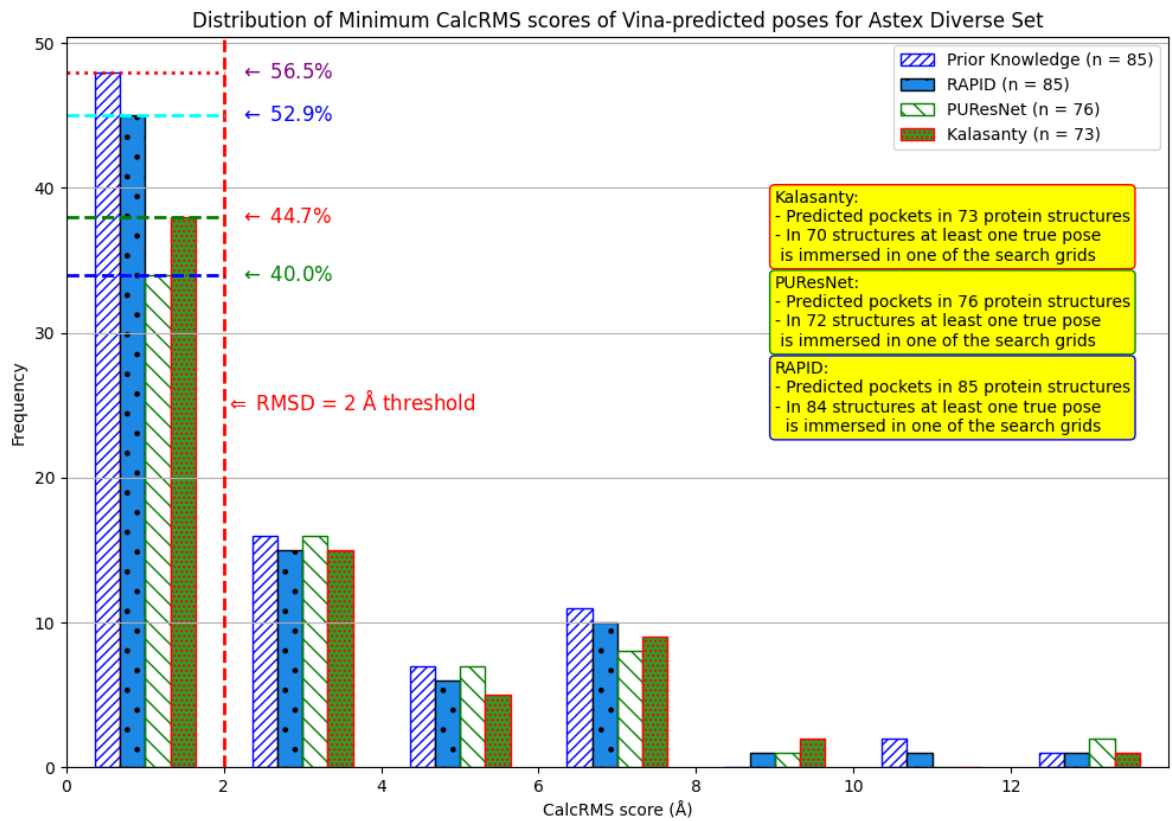


Figure 19. Distribution of RMSD values for Top-1 Vina predicted poses in the Astex Diverse Set [28]. When multiple true ligand poses are available, the RMSD to the closest one is considered.

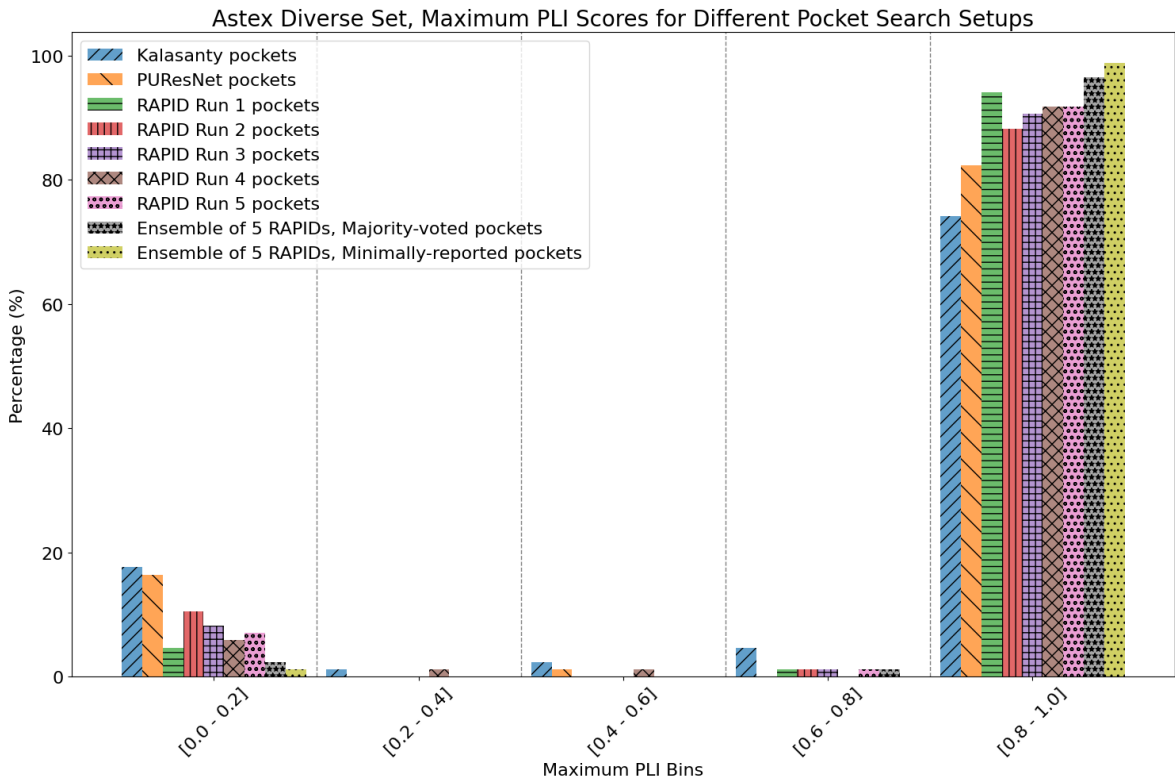


Figure 20. Distribution of the maximum PLI scores corresponding to the pockets predicted by RAPID-Net, PURESNet [69], and Kalasanty [70] for the Astex Diverse Set [33].

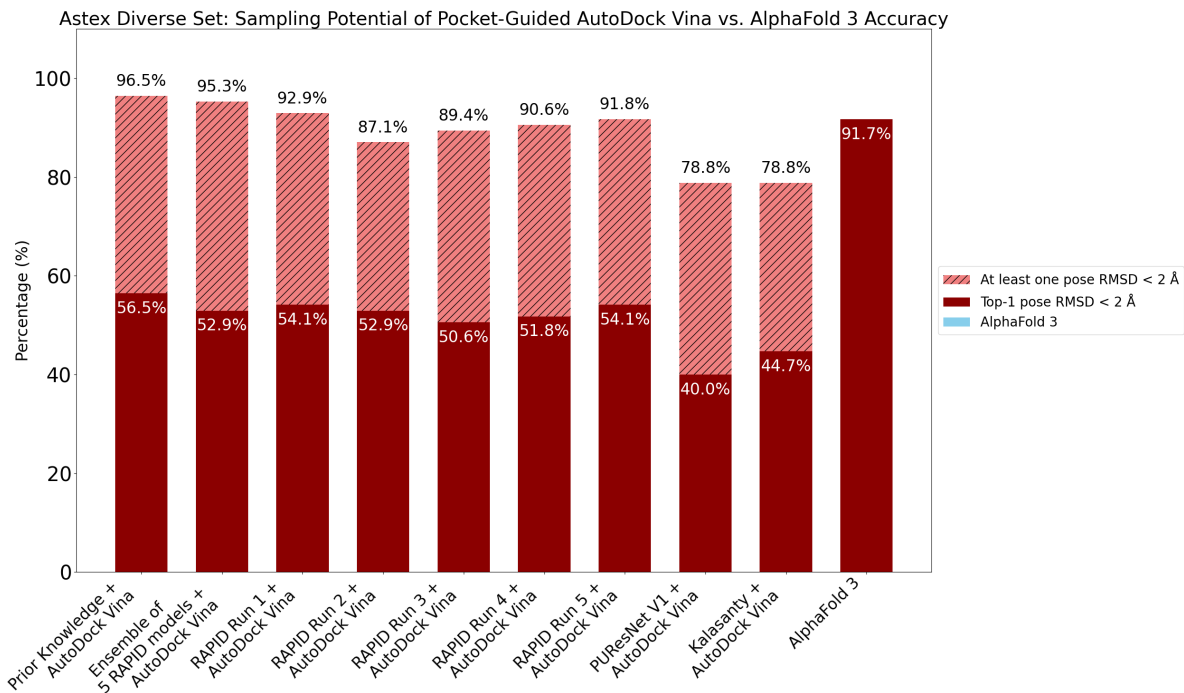


Figure 21. Docking accuracy of AutoDock Vina [12] for the Astex Diverse Set [33] when guided by different pocket prediction algorithms, comparing Top-1 and ensemble accuracy.

Furthermore, unlike the PoseBusters [28] dataset, when AutoDock Vina [12] is guided by Kalasanty [70] on the Astex Diverse Set [33], it achieves better docking accuracy than when guided by PURESNet [69], even though PURESNet [69] has more structures with at least one true ligand binding pose completely within the search grid (72 vs. 70, respectively).

Table 2. Comparison of pocket predictions across different methods for Astex Diverse Set [33] consisting of 85 protein structures.

Method	Nonzero	Within 15 Å	PLI
Kalasanty	73	70	78.72%
PURESNet	76	72	82.36%
RAPID Run 1	85	83	94.83%
RAPID Run 2	85	79	89.00%
RAPID Run 3	85	78	91.09%
RAPID Run 4	85	83	92.61%
RAPID Run 5	85	81	92.47%
RAPID ensemble	85	84	97.15%/98.82%

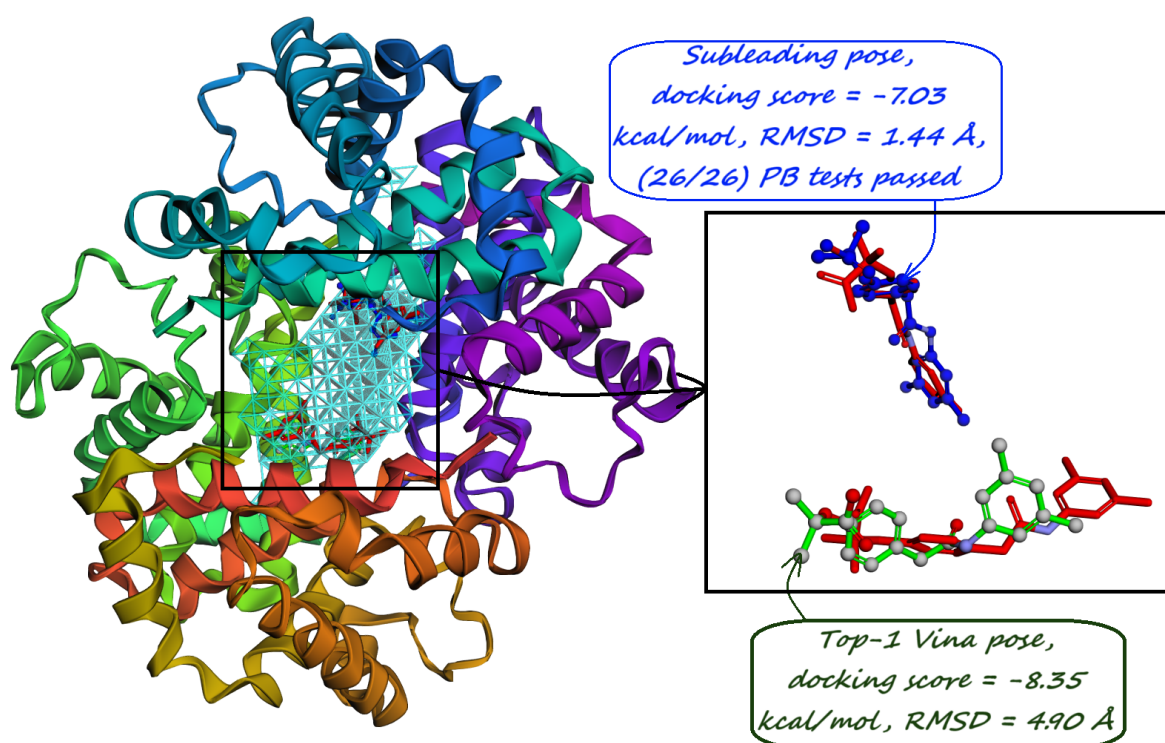


Figure 22. In the 1G9V protein structure from Astex Diverse Set [33], two true ligand binding poses are observed within the same majority-voted pocket predicted by RAPID-Net. While the Top-1 Vina pose fails the RMSD test, the subleading one passes all validation tests.

Nevertheless, when considering sampling accuracy—achieved either by the Top-1 pose or any subleading pose, as illustrated in Figure 21—the ensembled version of RAPID-Net outperforms each of its individual Runs: 95.3% versus 92.9%, 87.1%, 89.4%, 90.6%, and 91.8%, respectively. Additionally, PURESNet [69] and Kalasanty [70] both achieve the same sampling accuracy of 78.8%. A comparison of Tables 1 and 2 alongside Figs. 12 and 21 shows that RAPID-Net achieves a higher PLI rate on the Astex Diverse Set [33] than on the PoseBusters [28] dataset. However, the Top-1 docking accuracy is higher for PoseBusters [28], while the “at least one correct pose in the ensemble” accuracy is better for the Astex Diverse Set [33]. These results further emphasize and illustrate that, in addition to accurate pocket identification, another major bottleneck in the docking process is the accurate reranking of the generated poses.

In the next Section, we compare our model against other available pocket predictors in terms of PLI metrics using their original test datasets to enable a direct side-by-side comparison.

9. Evaluation on Coach420 and BU48 Datasets

For direct comparison with Kalasanty [70] and PURESNet [69], we evaluate the corresponding PLI rates on the Coach420 [34] and BU48 [35] datasets. Following [69], we exclude protein structures present in the training sets, resulting in 298 and 62 protein-ligand structures for Coach420 [34] and BU48 [35], respectively.

Table 3. Comparison of pocket predictions across different methods for Coach420 [34] dataset consisting of 298 protein structures.

Method	Nonzero	Within 15 Å	PLI
Kalasanty	273	248	76.41%
PUResNet	280	259	78.37%
RAPID Run 1	296	285	85.68%
RAPID Run 2	298	279	86.48%
RAPID Run 3	293	264	76.78%
RAPID Run 4	298	288	91.39%
RAPID Run 5	298	275	80.59%
RAPID ensemble	298	297	86.75%/95.49%

Table 4. Comparison of pocket predictions across different methods for BU48 [35] dataset consisting of 62 protein structures.

Method	Nonzero	Within 15 Å	PLI
Kalasanty	54	53	77.80%
PUResNet	42	39	46.90%
RAPID Run 1	61	60	86.52%
RAPID Run 2	62	62	97.57%
RAPID Run 3	61	60	86.68%
RAPID Run 4	60	58	84.47%
RAPID Run 5	62	43	46.97%
RAPID ensemble	62	62	95.79%/99.82%

Unlike the original PUResNet [69] paper where PLI values were evaluated only on protein structures where the distance-center-to-center (DCC) between predicted pocket centers and ligand centers was $\leq 4\text{\AA}$, we report the results for all protein structures, over the whole dataset. This comparison seems more appropriate because a rough initial approximation to set up the search grid is often sufficient to achieve successful docking, as illustrated in Figure 5. Furthermore, as discussed in the next Section, the RAPID-Net model sometimes predicts meaningful “tunnels” or “bridges” to distant binding sites that indirectly influence ligand binding. In such cases, the DCC metric becomes irrelevant since it does not reflect the functional importance of these remote interactions.

Table 3 and Figure 23 summarize the results for the Coach420 dataset [34]. Except for Run 3, every RAPID-Net run surpasses PUResNet [69] and Kalasanty [70] in terms of an average PLI score. Furthermore, when comparing the number of protein structures in which at least one true ligand binding pose is entirely within the search grid, up to the largest grid with a threshold of 15 Å, all RAPID-Net Runs outperform both PUResNet [69] and Kalasanty [70]. Similarly, Table 4 and Figure 24 present the results for the BU48 dataset [35]. Here, all RAPID-Net Runs, except Run 5, outperform both Kalasanty and PUResNet in terms of an average PLI and the number of protein structures where at least one true ligand binding pose is completely contained within the search grids.

Different RAPID-Net Runs have varying performances across the four test datasets, reflecting the inherent element of randomness in the model training. For example, on the PoseBusters [28] dataset, Run 2 has the fewest protein structures with viable predicted search grids, as shown in Table 1. In contrast, Run 3 produces the fewest viable grids for the Astex Diverse Set [33] and Coach420 [34], as shown in Tables 2 and 3, while Run 5 shows the lowest PLI for BU48 [35], as shown in Table 4. Notably, despite having lower coverage on Coach420 [34] and BU48 [35], Run 5 achieves the highest “at least one correct pose in the ensemble” rate on PoseBusters [28], as shown in Figure 12.

These results further emphasize the importance of ensembling model predictions to capture all potential binding sites and account for variability in model training. By combining five RAPID-Net models, we mitigate performance defects of individual model Runs, yielding more robust and reliable results. At the same time, as can be observed from all the data that we presented, across all four test

datasets, RAPID-Net consistently exhibits stronger generalization ability than both PURESNet [69] and Kalasanty [70].

In addition to docking accuracy and the coverage of true ligand binding poses by predicted pockets, it is crucial to identify relevant distant sites that indirectly influence ligand binding, which is the topic of the next Section.

10. Identification of Allosteric Sites, Exosites, and Flexible Regions for Drug Design

As noted earlier, a common problem with Deep Learning-based pocket prediction methods is that they may miss secondary pockets due to the scarcity of labeled training data, while classical methods may produce “blurry” pockets [65]. To demonstrate the ability of RAPID-Net to accurately detect both primary and secondary pockets, we focus on the therapeutically relevant case of the Nsp12 protein, an RNA-dependent RNA polymerase (RdRp), which plays a central role in viral RNA replication and is a key target for the development of antiviral drugs against SARS-CoV-2.

RNA-dependent RNA Polymerase (RCSB PDB: 7BV2). As illustrated in Figure 25, RAPID-Net successfully identifies the orthosteric remdesivir binding site [64] – containing residues 448, 545, 553, 555, 618, 759, 760, and 761 – as its “majority-voted” pocket, consistent with predictions from other pocket prediction methods, PURESNet V1 [69], Kalasanty [70], and PURESNet V2 [66]. In addition to the main binding site, computational studies have proposed allosteric [98] and secondary pockets [99] that can be used for drug development. Unlike other pocket predictors, RAPID-Net also highlights putative allosteric and secondary pockets reported in the literature [98,99], demonstrating its ability to recover both primary and secondary binding sites.

Additionally, to illustrate RAPID-Net’s ability to identify remote sites of therapeutic interest, we consider four proteins where such sites are well documented.

Thrombin (RCSB PDB: 1DWC). The RCSB PDB entry 1DWC [100] is the crystal structure of human α -thrombin in complex with the inhibitor MD-805. Thrombin plays a pivotal role in the coagulation cascade and is therefore a key target in treating acute coronary syndromes [101]. As shown in Figure 26, Kalasanty [70] and PURESNet [69] predict binding pockets largely surrounding the ligand. By contrast, our model, RAPID-Net, detects an additional bulge extending towards residues 71, 73, 75, 76, and 77. These residues belong to the anion-binding Exosite I, which interacts with negatively charged substrates and cofactors such as fibrinogen, thrombomodulin, and COOH-terminal peptide of hirudin [102–106]. Our model’s prediction, a “tunnel” connecting the active site to the Exosite I in Figure 26, suggests possible long-range interactions, consistent with previous studies of long-range allosteric communication in thrombin [106].

Human Carbonic Anhydrase I (RCSB PDB: 1AZM) and II (RCSB PDB: 3HS4). Human carbonic anhydrase I (hCAI) in complex with a sulfonamide drug (PDB: 1AZM) [107] and hCAII in complex with acetazolamide (AZM) (PDB: 3HS4) [108] are key targets for glaucoma treatment [109] and diuretic therapy [110], with broader potential for treating obesity, cancer, and Alzheimer’s disease [110]. In both isoforms, His64 acts as the primary proton shuttle [111–113].

As shown in Figure ??, while PURESNet V1 [69] and Kalasanty [70] predict the primary ligand binding region in 1AZM, our model additionally predicts a hook-shaped bulge wrapping around His64, suggesting an allosteric interaction.

In hCAII (3HS4), a histidine cluster (His3, His4, His10, His15, His17, His64) mediates proton exchange between the active site and the environment [110]. Notably, 3HS4 contains two additional AZM molecules in shallow surface pockets—one previously identified [114,115] and another in a novel binding site [108].

As shown in Figure ??, RAPID-Net identifies two majority-voted pockets: one around the catalytic site and another one near the histidine cluster. As shown in Figure ??, the minority-reported pockets predicted by our model have $PLI = 92\%$ with these secondary binding poses in shallow pockets.

This highlights the ability of our model to predict all binding sites in the protein, both primary and secondary.

In comparison, for this protein structure, Kalasanty [70] predicts only the main binding site, PURESNet V1 [69] detects none, and PURESNet V2 [66] identifies only the interacting residues within the main binding site.

Ls-AChBP (RCSB PDB: 2ZJV). RCSB PDB ID 2ZJV [116] represents *Lymnaea stagnalis* acetylcholine-binding protein (AChBP) bound to neonicotinoid clothianidin. AChBP serves as a model for nicotinic acetylcholine receptors (nAChRs) and their allosteric transitions [117], providing insights into therapies for neurological disorders such as Alzheimer's disease, schizophrenia, depression, attention deficit hyperactivity disorder, and tobacco addiction [118–120].

For the 2ZJV protein structure, neither Kalasanty [70] nor PURESNet V1 [69] predict any pockets. However, the majority-voted pockets from our model and the interacting residues of PURESNet V2 [66] illustrated in Figure 28 are predicted around all five true ligand binding positions.

Furthermore, our model's pockets include distinct bridges toward residues Thr155-Asp160 in the flexible F-loop, suggesting indirect ligand interactions. This finding aligns with the results from [116], who highlighted that induced-fit movements of loop regions, including the F-loop, are essential for ligand recognition by neonicotinoids.

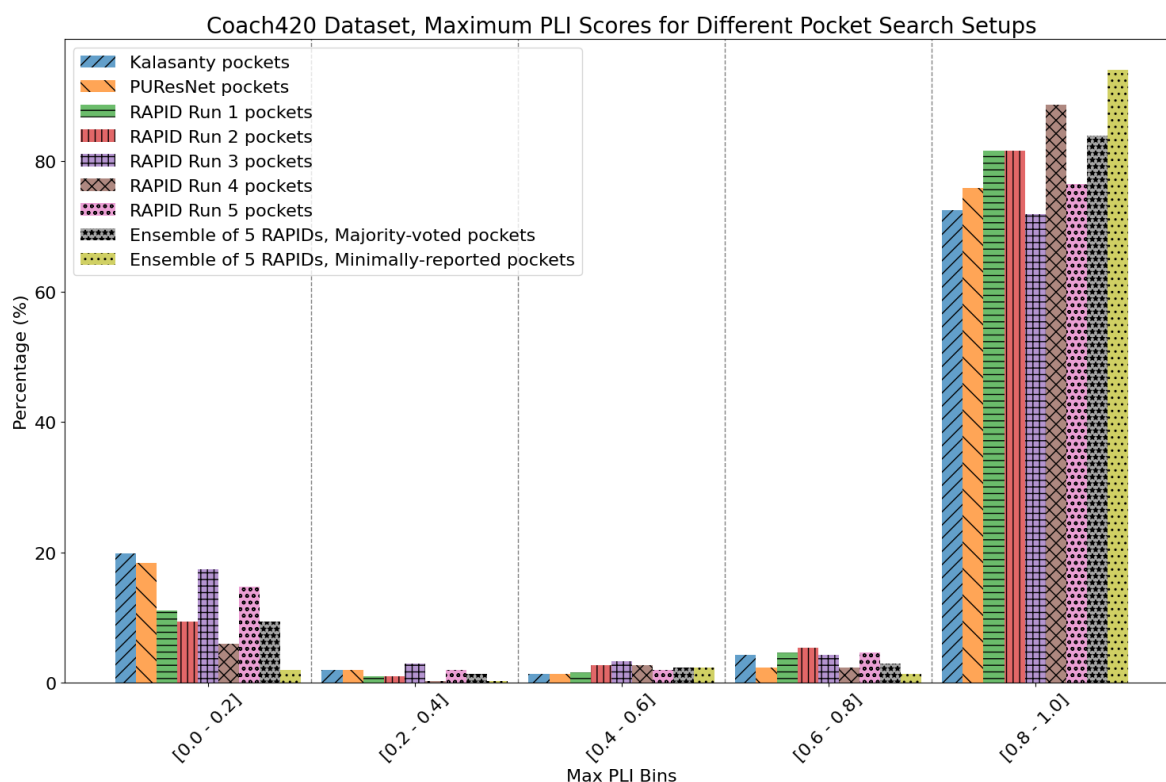


Figure 23. Distribution of the maximum PLI scores corresponding to the pockets predicted by RAPID-Net, PURESNet [69], and Kalasanty [70] for the Coach420 [34] dataset.

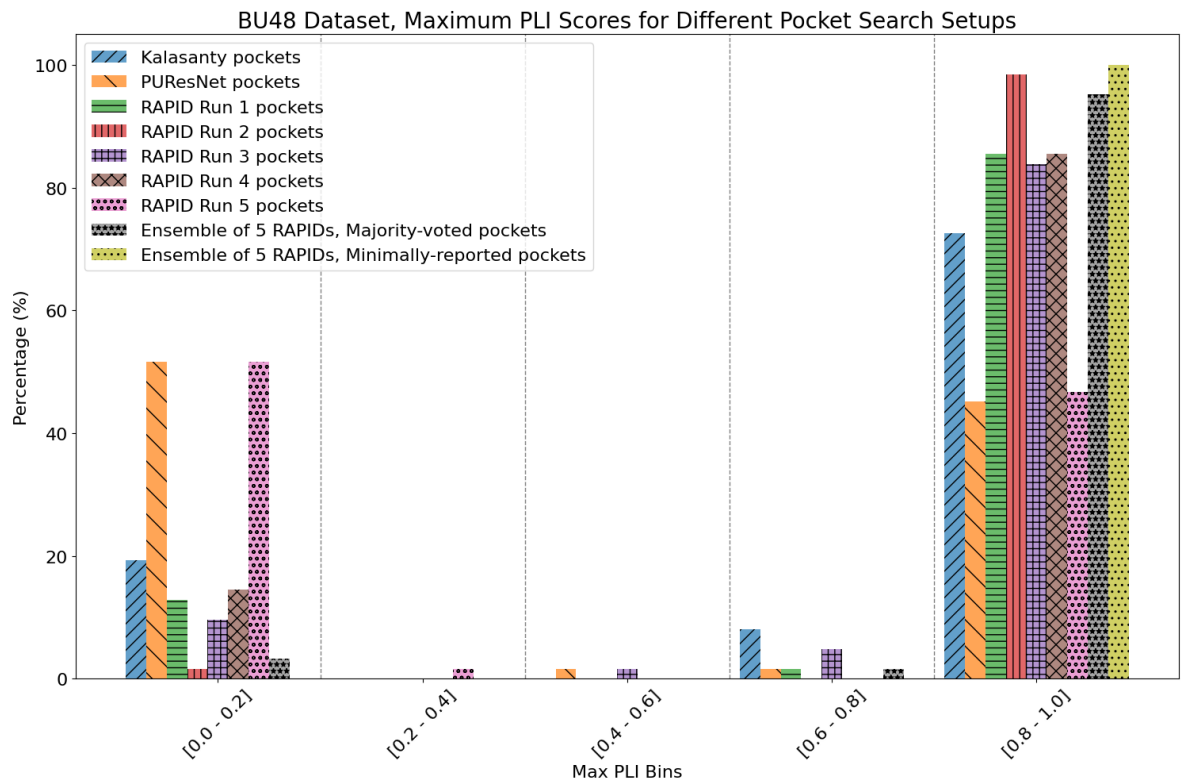


Figure 24. Distribution of the maximum PLI scores corresponding to the pockets predicted by RAPID-Net, PURESNet [69], and Kalasanty [70] for the BU48 [35] dataset.

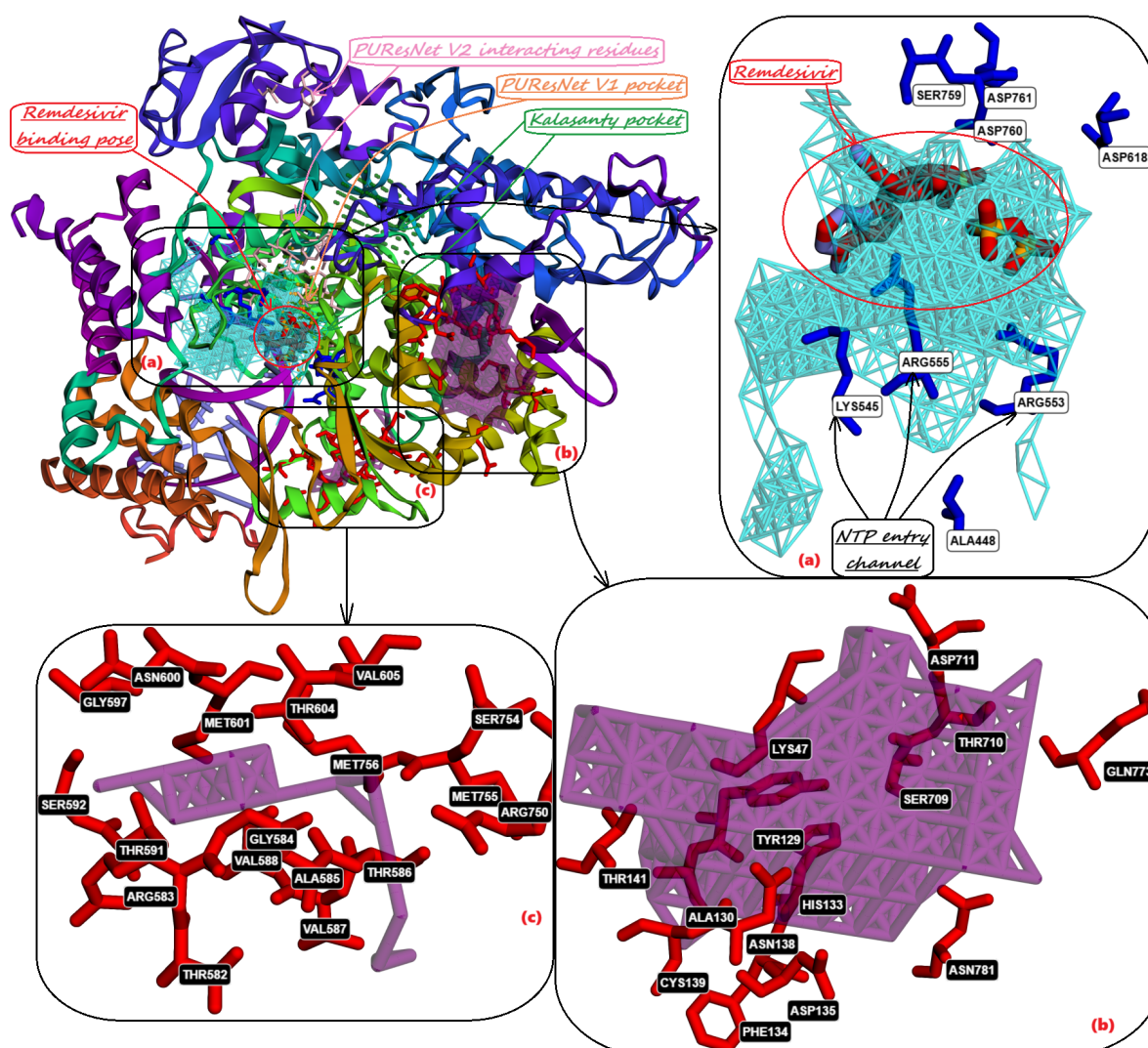


Figure 25. For the Nsp12 protein [64], the pockets predicted by Kalasanty [70] and PURESNet [69] – shown as green and orange dots, respectively – and the likely interacting residues predicted by PURESNet V2 [66] correspond to the orthosteric remdesivir binding site [64]. In contrast, RAPID-Net not only identifies the orthosteric site – corresponding to the majority-voted pocket, shown as cyan sticks – but also successfully detects putative secondary and allosteric pockets reported in previous studies [98,99], which correspond to the minimally-reported pockets predicted by our model, shown as purple sticks.

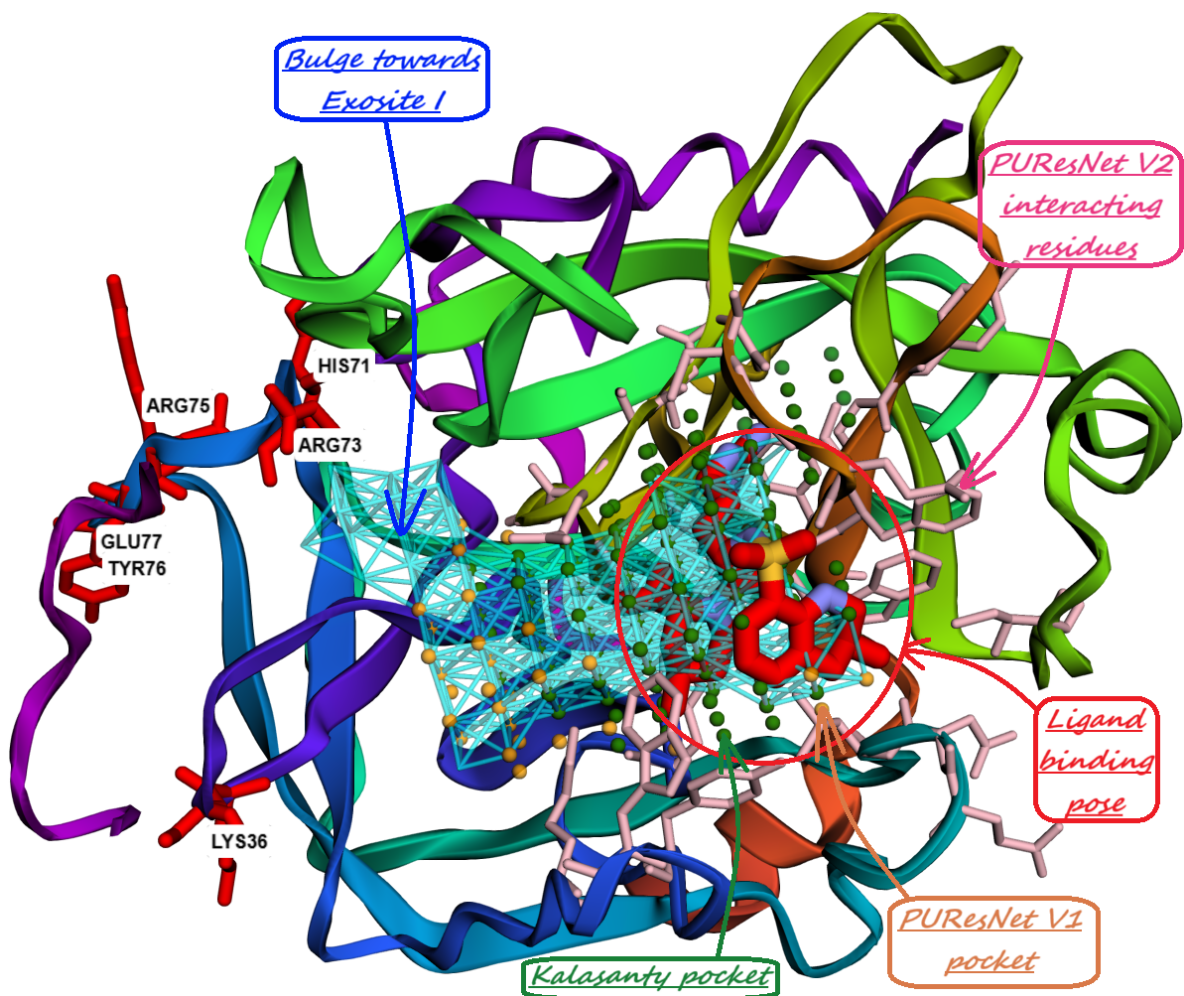


Figure 26. Thrombin (RCSB PDB: 1DWC). Unlike the pockets predicted by Kalasanty [70] and PUResNet [69] –shown as green and orange dots, respectively – the majority-voted pocket predicted by our model, shown as cyan sticks, has a distinct bulge extending toward residues associated with Exosite I.

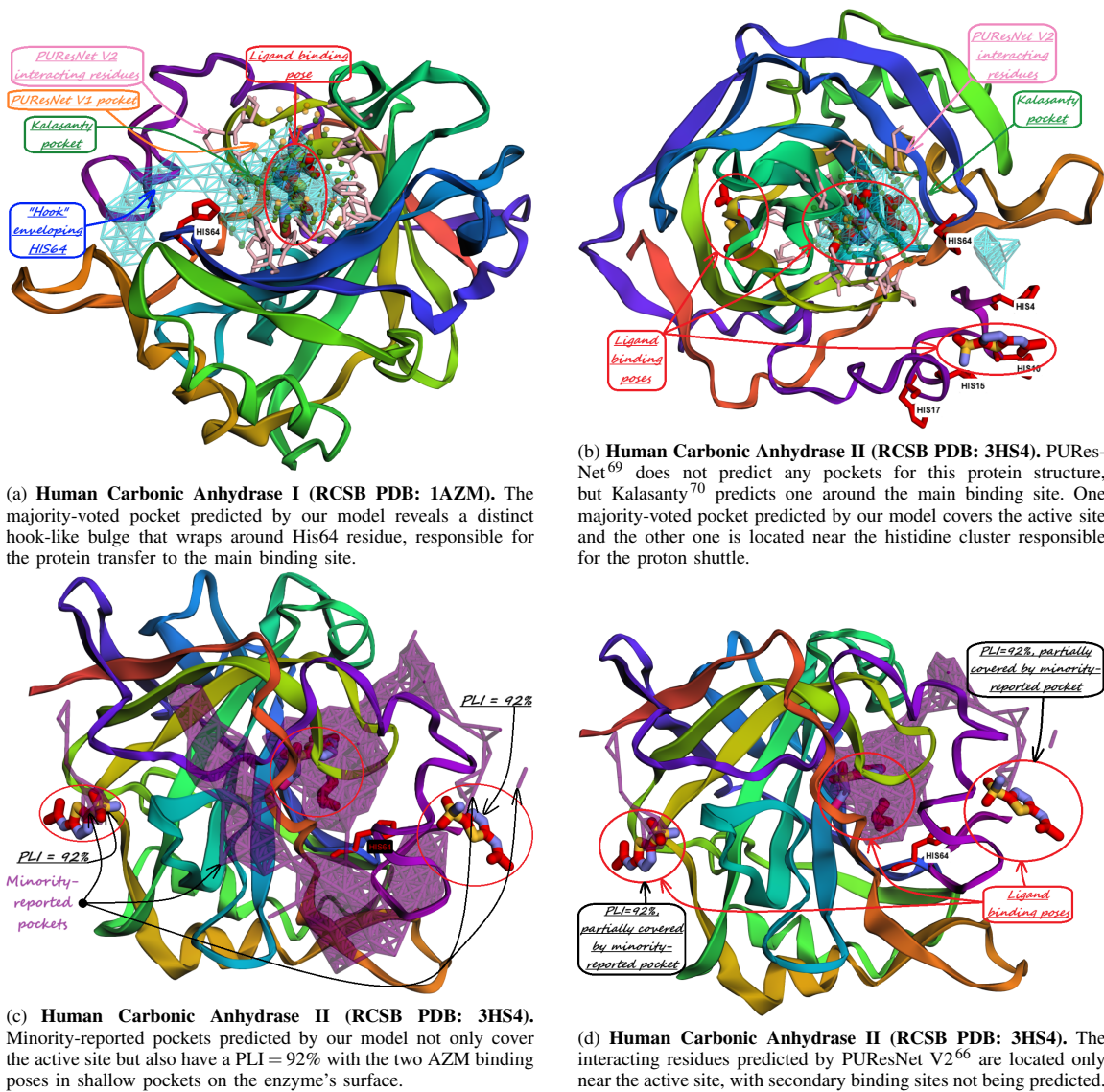


Figure 27. Majority-voted pockets predicted by RAPID-Net are represented by cyan sticks, while the minimally-reported pockets are shown by purple sticks. The interacting residues predicted by PURESNet V2 [66] are shown in yellow.

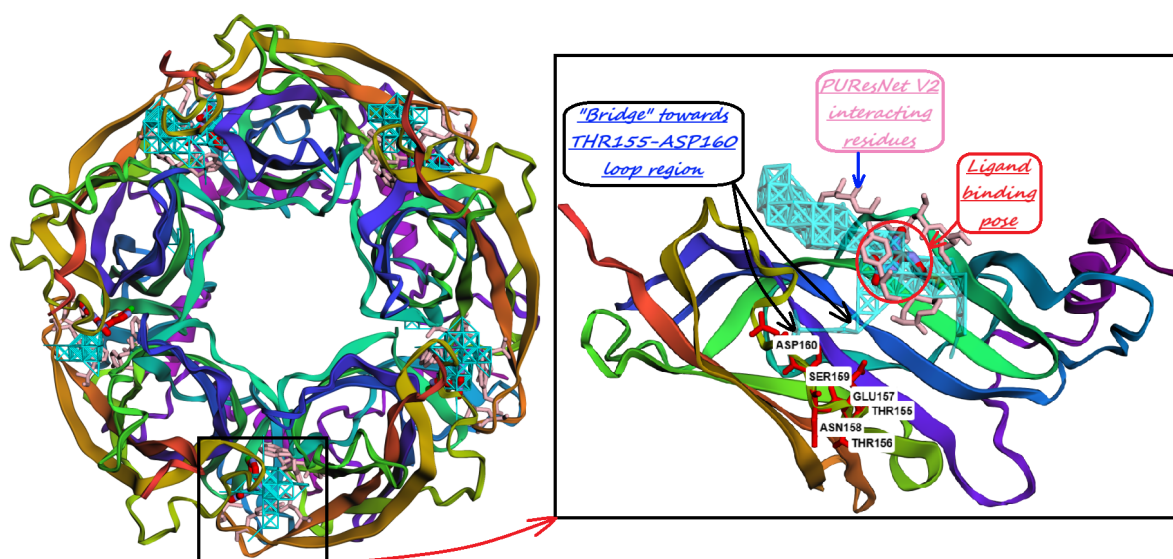


Figure 28. Ls-AChBP (RCSB PDB: 2ZJV). The majority-voted pocket predicted by our model has a distinct bridge towards the Thr155-Asp160 region belonging to the F loop. Kalasanty [70] and PURESNet V1 [69] do not predict any pockets for this protein structure, while PURESNet V2 [66] predicts interacting residues only in the immediate vicinity of the binding pose.

These examples demonstrate our model's capability to identify regions influencing ligand binding—even without direct ligand contact. By revealing distal regions such as Exosite I in thrombin, His64 in hCA, and remote loop segments in Ls-AChBP, our approach provides key structural insights that can drive allosteric inhibitor design and broader therapeutic innovation.

11. Discussion and Conclusions

As the number of protein structures without known binding sites continues to grow, performing binding-site-agnostic (or “blind”) docking has become crucial for structure-guided drug design. Successful blind docking relies heavily on accurately identifying the search grid where the ligand is likely to bind. However, most existing pocket prediction tools operate irrespectively of docking pipelines and are evaluated using metrics that do not directly correlate with docking success. To address this gap, we developed RAPID-Net, an ML-based pocket identification tool specifically designed for seamless integration with docking pipelines. We tested RAPID-Net's effectiveness in guiding blind docking using AutoDock Vina 1.2.5 [12], but our approach can be easily adapted to any docking software that requires a well-defined search grid.

When guided by RAPID-Net, AutoDock Vina [12] outperforms DiffBindFR [27] by over 5% in blind docking accuracy on the PoseBusters [28] dataset, highlighting the direct relationship between improved pocket identification and increased docking accuracy. Furthermore, RAPID-Net provides precise and compact search grids, enabling the docking of ligands to large proteins such as 8F4J from the PoseBusters [28] dataset, which tools like AlphaFold 3 [30] cannot process in whole, highlighting the critical importance of accurate and focused search areas for cost-effective docking.

We found that another major factor limiting docking accuracy is the reranking of generated poses. In addition to precise pocket identification, developing improved reranking tools could further enhance the overall performance of our combined scheme by more effectively selecting the most favorable poses from the generated ensemble.

We attribute the success of RAPID-Net to the following key changes and innovations we introduced:

- **Soft Labeling and ReLU Activation:** Unlike conventional binary segmentation tasks, we applied soft labels and ReLU activation in the output layer, drawing inspiration from medical image

segmentation techniques. This approach enables the model to more effectively differentiate between the internal regions of pockets and their boundaries.

- **Attention Mechanism:** We integrated a single attention block within the encoder-decoder bottleneck, enhancing the model's ability to focus on relevant features while reducing the risk of overfit in inherently noisy data environments.
- **Simplified Architecture:** By eliminating excessive residual connections, we streamlined the model architecture, resulting in improved performance. This simplification was effective given the noisy nature of the dataset.

Furthermore, RAPID-Net demonstrates the ability to identify “bridges” to distal sites further than 6.5 Å from the primary pocket that are often critical for therapeutic intervention. This makes RAPID-Net a promising tool for targeting remote therapeutic sites and broadening the scope of structure-based drug design.

Additionally, beyond improved docking accuracy, in our future work, we plan to integrate RAPID-Net into de novo pocket-conditioned drug design pipelines by guiding the generation of novel compounds based on the shape features of the predicted pockets. Although our current study does not directly exploit these features, using the predicted region solely as a search box, as illustrated in Figure 4, they may prove highly valuable for generative models [121–125].

Usage of Our Model and Reproducibility of Results

All code and relevant data (proteins, pockets, poses) are publicly available on GitHub³. To reproduce our results, interested readers should download these files and follow the instructions provided in our repository.

To apply our trained model to predict pockets for other proteins and perform subsequent docking with AutoDock Vina [12], we provide an accompanying notebook with step-by-step guidance in the same repository. Additionally, our pocket predictor can be integrated with other docking tools, for example, by using spherical search grids. Users may also refine the ensemble of poses generated by AutoDock Vina [12] guided by our pocket predictor by any preferred reweighting method, leveraging our pocket predictions to improve docking accuracy.

Author Contributions: YB designed the neural network, conducted model training and docking experiments, and drafted the manuscript. IH and AB contributed to data analysis and provided medical expertise. CVK supervised the research and contributed to data analysis. All authors contributed to the proofreading and approved the final version of the manuscript.

Acknowledgments: Research reported in this publication was supported by the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health under award number R01DK076629. We are grateful to Richard J. Barber and the Barber Integrative Metabolic Research Program for financial support. We also thank James Granneman and Kelly McNear for valuable discussions.

References

1. Kukul, A.; et al. *Molecular modeling of proteins*; Vol. 443, Springer, 2008.
2. An, J.; Totrov, M.; Abagyan, R. Comprehensive identification of “druggable” protein ligand binding sites. *Genome Informatics* **2004**, *15*, 31–41.
3. De Ruyck, J.; Brysbaert, G.; Blossey, R.; Lensink, M.F. Molecular docking as a popular tool in drug design, an in silico travel. *Advances and Applications in Bioinformatics and Chemistry* **2016**, pp. 1–11.
4. Rezaei, M.A.; Li, Y.; Wu, D.; Li, X.; Li, C. Deep learning in drug design: protein-ligand binding affinity prediction. *IEEE/ACM transactions on computational biology and bioinformatics* **2020**, *19*, 407–417.
5. Vamathevan, J.; Clark, D.; Czodrowski, P.; Dunham, I.; Ferran, E.; Lee, G.; Li, B.; Madabhushi, A.; Shah, P.; Spitzer, M.; et al. Applications of machine learning in drug discovery and development. *Nature reviews Drug discovery* **2019**, *18*, 463–477.

³ <https://github.com/BalytskyiJaroslaw/RAPID-Net>

6. Saikia, S.; Bordoloi, M. Molecular docking: challenges, advances and its use in drug discovery perspective. *Current drug targets* **2019**, *20*, 501–521.
7. Zoete, V.; Grosdidier, A.; Michielin, O. Docking, virtual high throughput screening and in silico fragment-based drug design. *Journal of cellular and molecular medicine* **2009**, *13*, 238–248.
8. Lionta, E.; Spyrou, G.; K Vassilatis, D.; Cournia, Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Current topics in medicinal chemistry* **2014**, *14*, 1923–1938.
9. Patrick, G.L. *An introduction to medicinal chemistry*; Oxford university press, 2023.
10. Hernández-Santoyo, A.; Tenorio-Barajas, A.Y.; Altuzar, V.; Vivanco-Cid, H.; Mendoza-Barrera, C. Protein-protein and protein-ligand docking. *Protein engineering-technology and application* **2013**, pp. 63–81.
11. Trott, O.; Olson, A.J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry* **2010**, *31*, 455–461.
12. Eberhardt, J.; Santos-Martins, D.; Tillack, A.F.; Forli, S. AutoDock Vina 1.2. 0: New docking methods, expanded force field, and python bindings. *Journal of chemical information and modeling* **2021**, *61*, 3891–3898.
13. Verdonk, M.L.; Cole, J.C.; Hartshorn, M.J.; Murray, C.W.; Taylor, R.D. Improved protein–ligand docking using GOLD. *Proteins: Structure, Function, and Bioinformatics* **2003**, *52*, 609–623.
14. Friesner, R.A.; Banks, J.L.; Murphy, R.B.; Halgren, T.A.; Klicic, J.J.; Mainz, D.T.; Repasky, M.P.; Knoll, E.H.; Shelley, M.; Perry, J.K.; et al. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry* **2004**, *47*, 1739–1749.
15. Grasso, G.; Di Gregorio, A.; Mavkov, B.; Piga, D.; Labate, G.F.D.; Danani, A.; Deriu, M.A. Fragmented blind docking: a novel protein–ligand binding prediction protocol. *Journal of Biomolecular Structure and Dynamics* **2022**, *40*, 13472–13481.
16. Hassan, N.M.; Alhossary, A.A.; Mu, Y.; Kwoh, C.K. Protein-ligand blind docking using QuickVina-W with inter-process spatio-temporal integration. *Scientific reports* **2017**, *7*, 15451.
17. Koes, D.R.; Baumgartner, M.P.; Camacho, C.J. Lessons learned in empirical scoring with smina from the CSAR 2011 benchmarking exercise. *Journal of chemical information and modeling* **2013**, *53*, 1893–1904.
18. Shen, Y. Predicting protein structure from single sequences. *Nature Computational Science* **2022**, *2*, 775–776.
19. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *nature* **2021**, *596*, 583–589.
20. Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: making protein folding accessible to all. *Nature methods* **2022**, *19*, 679–682.
21. Ahdriz, G.; Bouatta, N.; Floristean, C.; Kadyan, S.; Xia, Q.; Gerecke, W.; O'Donnell, T.J.; Berenberg, D.; Fisk, I.; Zanichelli, N.; et al. OpenFold: Retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization. *Nature Methods* **2024**, pp. 1–11.
22. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373*, 871–876.
23. Overington, J.P.; Al-Lazikani, B.; Hopkins, A.L. How many drug targets are there? *Nature reviews Drug discovery* **2006**, *5*, 993–996.
24. Santos, R.; Ursu, O.; Gaulton, A.; Bento, A.P.; Donadi, R.S.; Bologa, C.G.; Karlsson, A.; Al-Lazikani, B.; Hersey, A.; Oprea, T.I.; et al. A comprehensive map of molecular drug targets. *Nature reviews Drug discovery* **2017**, *16*, 19–34.
25. Huang, Y.; Zhang, H.; Jiang, S.; Yue, D.; Lin, X.; Zhang, J.; Gao, Y.Q. DSDP: a blind docking strategy accelerated by GPUs. *Journal of Chemical Information and Modeling* **2023**, *63*, 4355–4363.
26. Utgés, J.S.; Barton, G.J. Comparative evaluation of methods for the prediction of protein–ligand binding sites. *Journal of Cheminformatics* **2024**, *16*, 126.
27. Zhu, J.; Gu, Z.; Pei, J.; Lai, L. DiffBindFR: an SE(3) equivariant network for flexible protein-ligand docking. *Chemical Science* **2024**, *15*, 7926–7942. Supplementary information available: PDF (1982K), <https://doi.org/10.1039/D3SC06803J>.
28. Buttenschoen, M.; Morris, G.M.; Deane, C.M. PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chemical Science* **2024**, *15*, 3130–3139. <https://doi.org/10.1039/D3SC05678A>.
29. Dong, C.; Huang, Y.P.; Lin, X.; Zhang, H.; Gao, Y.Q. DSDPFlex: Flexible-Receptor Docking with GPU Acceleration. *Journal of Chemical Information and Modeling* **2024**, *64*, 5252–5265. <https://doi.org/10.1021/acs.jcim.4c01715>.

30. Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Jumper, J.M. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2024**, pp. 1–3.
31. Morehead, A.; Giri, N.; Liu, J.; Neupane, P.; Cheng, J. Deep Learning for Protein-Ligand Docking: Are We There Yet? In Proceedings of the ICML AI4Science Workshop, 2024. selected as a spotlight presentation.
32. Chollet, F. *Deep learning with Python*; Simon and Schuster, 2021.
33. Hartshorn, M.J.; Verdonk, M.L.; Chessari, G.; Brewerton, S.C.; Mooij, W.T.; Mortenson, P.N.; Murray, C.W. Diverse, high-quality test set for the validation of protein- ligand docking performance. *Journal of Medicinal Chemistry* **2007**, *50*, 726–741.
34. Roy, A.; Yang, J.; Zhang, Y. COFACTOR: An Accurate Comparative Algorithm for Structure-Based Protein Function Annotation. *Nucleic Acids Res.* **2012**, *40*, W471–W477. <https://doi.org/10.1093/nar/gks372>.
35. Huang, B.; Schroeder, M. LIGSITE csc: predicting ligand binding sites using the Connolly surface and degree of conservation. *BMC Structural Biology* **2006**, *6*, 1–11. <https://doi.org/10.1186/1472-6807-6-19>.
36. Le Guilloux, V.; Schmidtke, P.; Tufféry, P. Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinformatics* **2009**, *10*, 168. <https://doi.org/10.1186/1471-2105-10-168>.
37. Hendlich, M.; Rippmann, F.; Barnickel, G. LIGSITE: Automatic and Efficient Detection of Potential Small Molecule-Binding Sites in Proteins. *J. Mol. Graph. Model.* **1997**, *15*, 359–363, 389. [https://doi.org/10.1016/S1093-3263\(98\)00002-3](https://doi.org/10.1016/S1093-3263(98)00002-3).
38. Laskowski, R.A. SURFNET: A Program for Visualizing Molecular Surfaces, Cavities, and Intermolecular Interactions. *J. Mol. Graph.* **1995**, *13*, 323–330, 307–308. [https://doi.org/10.1016/0263-7855\(95\)00073-9](https://doi.org/10.1016/0263-7855(95)00073-9).
39. Levitt, D.G.; Banaszak, L.J. POCKET: A Computer Graphics Method for Identifying and Displaying Protein Cavities and Their Surrounding Amino Acids. *J. Mol. Graph.* **1992**, *10*, 229–234. [https://doi.org/10.1016/0263-7855\(92\)80074-N](https://doi.org/10.1016/0263-7855(92)80074-N).
40. Kleywegt, G.J.; Jones, T.A. Detection, Delineation, Measurement and Display of Cavities in Macromolecular Structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1994**, *50*, 178–185. <https://doi.org/10.1107/S0907444993011355>.
41. Liang, J.; Edelsbrunner, H.; Woodward, C. Anatomy of Protein Pockets and Cavities: Measurement of Binding Site Geometry and Implications for Ligand Design. *Protein Sci.* **1998**, *7*, 1884–1897. <https://doi.org/10.1002/pro.5560070905>.
42. Brady, G.P.J.; Stouten, P.F. Fast Prediction and Visualization of Protein Binding Pockets with PASS. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 383–401. <https://doi.org/10.1023/A:1008193205570>.
43. Weisel, M.; Proschak, E.; Schneider, G. PocketPicker: Analysis of Ligand Binding-Sites with Shape Descriptors. *Chem. Cent. J.* **2007**, *1*, 7. <https://doi.org/10.1186/1752-153X-1-7>.
44. An, J.; Totrov, M.; Abagyan, R. Pocketome via Comprehensive Identification and Classification of Ligand Binding Envelopes. *Mol. Cell. Proteomics* **2005**, *4*, 752–761. <https://doi.org/10.1074/mcp.M400031-MCP200>.
45. Goodford, P.J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *J. Med. Chem.* **1985**, *28*, 849–857. <https://doi.org/10.1021/jm00145a002>.
46. An, J.; Totrov, M.; Abagyan, R. Comprehensive Identification of “Druggable” Protein Ligand Binding Sites. *Genome Inform.* **2004**, *15*, 31–41.
47. Laurie, A.T.R.; Jackson, R.M. Q-SiteFinder: An Energy-Based Method for the Prediction of Protein–Ligand Binding Sites. *Bioinformatics* **2005**, *21*, 1908–1916. <https://doi.org/10.1093/bioinformatics/bti315>.
48. Ghersi, D.; Sanchez, R. EasyMIFS and SiteHound: A Toolkit for the Identification of Ligand-Binding Sites in Protein Structures. *Bioinformatics* **2009**, *25*, 3185–3186. <https://doi.org/10.1093/bioinformatics/btp558>.
49. Ngan, C.H.; Bohnuud, T.; Mottarella, S.E.; Beglov, D.; Villar, E.A.; Hall, D.R.; Vajda, S. FTSite: High Accuracy Detection of Ligand Binding Sites on Unbound Protein Structures. *Bioinformatics* **2012**, *28*, 286–287. <https://doi.org/10.1093/bioinformatics/btr649>.
50. Armon, A.; Graur, D.; Ben-Tal, N. ConSurf: An Algorithmic Tool for the Identification of Functional Regions in Proteins by Surface Mapping of Phylogenetic Information. *J. Mol. Biol.* **2001**, *307*, 447–463. <https://doi.org/10.1006/jmbi.2001.4414>.
51. Pupko, T.; Bell, R.E.; Mayrose, I.; Glaser, F.; Ben-Tal, N. Rate4Site: An Algorithmic Tool for the Identification of Functional Regions in Proteins by Surface Mapping of Evolutionary Determinants Within Their Homologues. *Bioinformatics* **2002**, *18*, S71–S77. https://doi.org/10.1093/bioinformatics/18.suppl_1.S71.
52. Xie, Z.R.; Hwang, M.J. Ligand-Binding Site Prediction Using Ligand-Interacting and Binding Site-Enriched Protein Triangles. *Bioinformatics* **2012**, *28*, 1579–1585. <https://doi.org/10.1093/bioinformatics/bts180>.

53. Zvelebil, M.J.; Barton, G.J.; Taylor, W.R.; Sternberg, M.J.E. Prediction of Protein Secondary Structure and Active Sites Using the Alignment of Homologous Sequences. *J. Mol. Biol.* **1987**, *195*, 957–961. [https://doi.org/10.1016/0022-2836\(87\)90595-2](https://doi.org/10.1016/0022-2836(87)90595-2).
54. Wass, M.N.; Kelley, L.A.; Sternberg, M.J.E. 3DLigandSite: Predicting Ligand-Binding Sites Using Similar Structures. *Nucleic Acids Res.* **2010**, *38*, W469–W473. <https://doi.org/10.1093/nar/gkq406>.
55. Yang, J.; Roy, A.; Zhang, Y. Protein–Ligand Binding Site Recognition Using Complementary Binding-Specific Substructure Comparison and Sequence Profile Alignment. *Bioinformatics* **2013**, *29*, 2588–2595. <https://doi.org/10.1093/bioinformatics/btt447>.
56. Lee, H.S.; Im, W. Ligand Binding Site Detection by Local Structure Alignment and Its Performance Complementarity. *J. Chem. Inf. Model.* **2013**, *53*, 2462–2470. <https://doi.org/10.1021/ci400206h>.
57. Brylinski, M.; Feinstein, W.P. eFindSite: Improved Prediction of Ligand Binding Sites in Protein Models Using Meta-Threading, Machine Learning and Auxiliary Ligands. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 551–567. <https://doi.org/10.1007/s10822-013-9660-5>.
58. Glaser, F.; Pupko, T.; Paz, I.; Bell, R.E.; Bechor-Shental, D.; Martz, E.; Ben-Tal, N. A Method for Localizing Ligand Binding Pockets in Protein Structures. *Proteins* **2006**, *62*, 479–488. <https://doi.org/10.1002/prot.20769>.
59. Halgren, T.A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377–389. <https://doi.org/10.1021/ci800324m>.
60. Capra, J.A.; Laskowski, R.A.; Thornton, J.M.; Singh, M.; Funkhouser, T.A. Predicting Protein Ligand Binding Sites by Combining Evolutionary Sequence Conservation and 3D Structure. *PLoS Comput. Biol.* **2009**, *5*, e1000585. <https://doi.org/10.1371/journal.pcbi.1000585>.
61. Huang, B. MetaPocket: A Meta Approach to Improve Protein Ligand Binding Site Prediction. *OMICS* **2009**, *13*, 325–330. <https://doi.org/10.1089/omi.2009.0045>.
62. Bray, T.; Marsden, B.D.; Blundell, T.L. SitesIdentify: A Protein Functional Site Prediction Tool. *BMC Bioinformatics* **2009**, *10*, 379. <https://doi.org/10.1186/1471-2105-10-379>.
63. Brylinski, M.; Skolnick, J. FINDSITE: A Threading-Based Approach to Ligand Homology Modeling. *PLoS Comput. Biol.* **2009**, *5*, e1000405. <https://doi.org/10.1371/journal.pcbi.1000405>.
64. Yin, W.; Mao, C.; Luan, X.; Shen, D.D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **2020**, *368*, 1499–1504.
65. Gervasoni, S.; Vistoli, G.; Talarico, C.; Manelfi, C.; Beccari, A.R.; Studer, G.; Tauriello, G.; Waterhouse, A.M.; Schwede, T.; Pedretti, A. A comprehensive mapping of the druggable cavities within the SARS-CoV-2 therapeutically relevant proteins by combining pocket and docking searches as implemented in pockets 2.0. *International journal of molecular sciences* **2020**, *21*, 5152.
66. Jeevan, K.; Palistha, S.; Tayara, H.; Chong, K.T. PURESNetV2.0: a deep learning model leveraging sparse representation for improved ligand binding site prediction. *Journal of Cheminformatics* **2024**, *16*, 1–16.
67. Sestak, F.; Schneckenteiter, L.; Brandstetter, J.; Hochreiter, S.; Mayr, A.; Klambauer, G. VN-EGNN: E (3)-Equivariant Graph Neural Networks with Virtual Nodes Enhance Protein Binding Site Identification. *arXiv preprint arXiv:2404.07194* **2024**.
68. Smith, Z.; Strobel, M.; Vani, B.P.; Tiwary, P. Graph attention site prediction (grasp): Identifying druggable binding sites using graph neural networks with attention. *Journal of chemical information and modeling* **2024**, *64*, 2637–2644.
69. Kandel, J.; Tayara, H.; Chong, K.T. PURESNet: prediction of protein-ligand binding sites using deep residual neural network. *Journal of cheminformatics* **2021**, *13*, 1–14.
70. Stepniewska-Dziubinska, M.M.; Zielenkiewicz, P.; Siedlecki, P. Improving detection of protein-ligand binding sites with 3D segmentation. *Scientific Reports* **2020**, *10*, 5035.
71. Krivák, R.; Hoksza, D. Improving protein-ligand binding site prediction accuracy by classification of inner pocket points using local features. *Journal of cheminformatics* **2015**, *7*, 1–13.
72. Aggarwal, R.; Gupta, A.; Chelur, V.; Jawahar, C.; Priyakumar, U.D. DeepPocket: ligand binding site detection and segmentation using 3D convolutional neural networks. *Journal of Chemical Information and Modeling* **2021**, *62*, 5069–5079.
73. Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: an open source platform for ligand pocket detection. *BMC bioinformatics* **2009**, *10*, 1–11.

74. Carbery, A.; Buttenschoen, M.; Skyner, R.; von Delft, F.; Deane, C.M. Learnt representations of proteins can be used for accurate prediction of small molecule binding sites on experimentally determined and predicted protein structures. *Journal of Cheminformatics* **2024**, *16*, 32.
75. Shen, A.; Yuan, M.; Ma, Y.; Du, J.; Wang, M. PGBind: pocket-guided explicit attention learning for protein–ligand docking. *Briefings in Bioinformatics* **2024**, *25*, bbae455.
76. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. *arXiv preprint arXiv:1505.04597* **2015**.
77. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
78. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 7132–7141.
79. Jiang, Y.; Chen, L.; Zhang, H.; Xiao, X. Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. *PloS One* **2019**, *14*, e0214587.
80. Balytskyi, Y.; Kalashnyk, N.; Hubenko, I.; Balytska, A.; McNear, K. Enhancing Open-World Bacterial Raman Spectra Identification by Feature Regularization for Improved Resilience against Unknown Classes. *Chemical & Biomedical Imaging* **2024**, *2*, 442–452.
81. Desaphy, J.; Bret, G.; Rognan, D.; Kellenberger, E. sc-PDB: a 3D-database of ligandable binding sites—10 years on. *Nucleic Acids Research* **2015**, *43*, D399–D404.
82. Desaphy, J.; Azdimousa, K.; Kellenberger, E.; Rognan, D. Comparison and druggability prediction of protein–ligand binding sites from pharmacophore-annotated cavity shapes. *Journal of Chemical Information and Modeling* **2012**, *52*, 2287–2299.
83. Kellenberger, E.; Schalon, C.; Rognan, D. How to measure the similarity between protein ligand-binding sites? *Current Computer-Aided Drug Design* **2008**, *4*, 209.
84. Bajusz, D.; Rácz, A.; Héberger, K. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *Journal of Cheminformatics* **2015**, *7*, 1–13.
85. García-García, J.; Others. tfbio: Molecular features featurization library for TensorFlow. <https://github.com/gnina/tfbio>, 2018. Accessed: YYYY-MM-DD.
86. Gros, C.; Lemay, A.; Cohen-Adad, J. SoftSeg: Advantages of soft versus binary training for image segmentation. *Medical Image Analysis* **2021**, *71*, 102038.
87. Wang, Z.; Popordanoska, T.; Bertels, J.; Lemmens, R.; Blaschko, M.B. Dice semimetric losses: Optimizing the dice score with soft labels. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Cham, 2023; pp. 475–485.
88. Milletari, F.; Navab, N.; Ahmadi, S.A. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV). IEEE, 2016, pp. 565–571.
89. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. In Proceedings of the Proceedings of the IEEE International Conference on Computer Vision (ICCV), October 2017, pp. 2980–2988.
90. Nair, V.; Hinton, G.E. Rectified Linear Units Improve Restricted Boltzmann Machines. In Proceedings of the Proceedings of the 27th International Conference on Machine Learning (ICML), 2010, pp. 807–814.
91. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
92. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **2020**, *17*, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>.
93. Varadi, M.; Anyango, S.; Deshpande, M.; et al.. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research* **2022**, *50*, D439–D444. <https://doi.org/10.1093/nar/gkab1061>.
94. Landrum, G. Rdkit documentation. *Release* **2013**, *1*, 4.
95. Su, M.; Yang, Q.; Du, Y.; Feng, G.; Liu, Z.; Li, Y.; Wang, R. Comparative assessment of scoring functions: the CASF-2016 update. *Journal of chemical information and modeling* **2018**, *59*, 895–913.
96. Li, H.; Han, L.; Liu, W.; Wang, R. DELPHI Scorer: A Machine-Learning-Based Scoring Function for Protein-Ligand Interactions. *Journal of Chemical Information and Modeling* **2017**, *57*, 47–55. <https://doi.org/10.1021/acs.jcim.6b00389>.

97. Shen, C.; Zhang, X.; Deng, Y.; Gao, J.; Wang, D.; Xu, L.; Kang, Y. Boosting protein–ligand binding pose prediction and virtual screening based on residue–atom distance likelihood potential and graph transformer. *Journal of Medicinal Chemistry* **2022**, *65*, 10691–10706. <https://doi.org/10.1021/acs.jmedchem.2c00520>.
98. Faisal, S.; Badshah, S.L.; Kubra, B.; Sharaf, M.; Emwas, A.H.; Jaremko, M.; Abdalla, M. Computational study of SARS-CoV-2 rna dependent rna polymerase allosteric site inhibition. *Molecules* **2021**, *27*, 223.
99. Parvez, M.S.A.; Karim, M.A.; Hasan, M.; Jaman, J.; Karim, Z.; Tahsin, T.; Hasan, M.N.; Hosen, M.J. Prediction of potential inhibitors for RNA-dependent RNA polymerase of SARS-CoV-2 using comprehensive drug repurposing and molecular docking approach. *International journal of biological macromolecules* **2020**, *163*, 1787–1797.
100. Banner, D.W.; Hadváry, P. Crystallographic analysis at 3.0-Å resolution of the binding to human thrombin of four active site-directed inhibitors. *Journal of Biological Chemistry* **1991**, *266*, 20085–20093.
101. Weitz, J.I.; Buller, H.R. Direct thrombin inhibitors in acute coronary syndromes: present and future. *Circulation* **2002**, *105*, 1004–1011. <https://doi.org/10.1161/hc0802.105443>.
102. Crawley, J.T.B.; Zanardelli, S.; Chion, C.K.N.K.; Lane, D.A. The central role of thrombin in hemostasis. *Journal of Thrombosis and Haemostasis* **2007**, *5*, 95–101. <https://doi.org/10.1111/j.1538-7836.2006.02248.x>.
103. Adams, T.E.; Huntington, J.A. Thrombin-cofactor interactions: structural insights into regulatory mechanisms. *Arteriosclerosis, thrombosis, and vascular biology* **2006**, *26*, 1738–1745.
104. Bock, P.; Panizzi, P.; Verhamme, I. Exosites in the substrate specificity of blood coagulation reactions. *Journal of Thrombosis and Haemostasis* **2007**, *5*, 81–94.
105. Troisi, R.; Balasco, N.; Autiero, I.; Vitagliano, L.; Sica, F. Exosite binding in thrombin: a global structural/dynamic overview of complexes with aptamers and other ligands. *International Journal of Molecular Sciences* **2021**, *22*, 10803.
106. Petrera, N.S.; Stafford, A.R.; Leslie, B.A.; Kretz, C.A.; Fredenburgh, J.C.; Weitz, J.I. Long range communication between exosites 1 and 2 modulates thrombin function. *Journal of Biological Chemistry* **2009**, *284*, 25620–25629.
107. Chakravarty, S.; Kannan, K. Drug-protein interactions: refined structures of three sulfonamide drug complexes of human carbonic anhydrase I enzyme. *Journal of molecular biology* **1994**, *243*, 298–309.
108. Sippel, K.H.; Robbins, A.H.; Domsic, J.; Genis, C.; Agbandje-McKenna, M.; McKenna, R. High-resolution structure of human carbonic anhydrase II complexed with acetazolamide reveals insights into inhibitor drug design. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **2009**, *65*, 992–995.
109. Pfeiffer, N. Innovative glaucoma therapy: Glaucoma therapy with topical carbonic anhydrase inhibitors. *Der Ophthalmologe: Zeitschrift der Deutschen Ophthalmologischen Gesellschaft* **2001**, *98*, 953–960. <https://doi.org/10.1007/s003470170039>.
110. Supuran, C.T. Carbonic anhydrases: Novel therapeutic applications for inhibitors and activators. *Nature Reviews Drug Discovery* **2008**, *7*, 168–181. <https://doi.org/10.1038/nrd2467>.
111. Silverman, D.N.; Lindskog, S. The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. *Accounts of Chemical Research* **1988**, *21*, 30–36.
112. Nair, S.K.; Christianson, D.W. Unexpected pH-dependent conformation of His-64, the proton shuttle of carbonic anhydrase II. *Journal of the American Chemical Society* **1991**, *113*, 9455–9458.
113. Lomelino, C.L.; Andring, J.T.; McKenna, R. Crystallography and its impact on carbonic anhydrase research. *International journal of medicinal chemistry* **2018**, *2018*, 9419521.
114. Jude, K.M.; Banerjee, A.L.; Haldar, M.K.; Manokaran, S.; Roy, B.; Mallik, S.; Srivastava, D.; Christianson, D.W. Ultrahigh resolution crystal structures of human carbonic anhydrases I and II complexed with “two-prong” inhibitors reveal the molecular basis of high affinity. *Journal of the American Chemical Society* **2006**, *128*, 3011–3018.
115. Srivastava, D.; Jude, K.M.; Banerjee, A.L.; Haldar, M.; Manokaran, S.; Kooren, J.; Mallik, S.; Christianson, D.W. Structural analysis of charge discrimination in the binding of inhibitors to human carbonic anhydrases I and II. *Journal of the American Chemical Society* **2007**, *129*, 5528–5537.
116. Ihara, M.; Okajima, T.; Yamashita, A.; Oda, T.; Hirata, K.; Nishiwaki, H.; Morimoto, T.; Akamatsu, M.; Ashikawa, Y.; Kuroda, S.; et al. Crystal structures of *Lymnaea stagnalis* AChBP in complex with neonicotinoid insecticides imidacloprid and clothianidin. *Invertebrate Neuroscience* **2008**, *8*, 71–81. <https://doi.org/10.1007/s10158-008-0069-3>.
117. Taly, A.; Corringier, P.J.; Guedin, D.; Lestage, P.; Changeux, J.P. Nicotinic receptors: allosteric transitions and therapeutic targets in the nervous system. *Nature reviews Drug discovery* **2009**, *8*, 733–750.

118. Arneric, S.P.; Holladay, M.; Williams, M. Neuronal nicotinic receptors: a perspective on two decades of drug discovery research. *Biochemical pharmacology* **2007**, *74*, 1092–1101.
119. Levin, E.D.; Rezvani, A.H. Nicotinic interactions with antipsychotic drugs, models of schizophrenia and impacts on cognitive function. *Biochemical Pharmacology* **2007**, *74*, 1182–1191.
120. Romanelli, M.N.; Gratteri, P.; Guandalini, L.; Martini, E.; Bonaccini, C.; Gualtieri, F. Central nicotinic receptors: structure, function, ligands, and therapeutic potential. *ChemMedChem: Chemistry Enabling Drug Discovery* **2007**, *2*, 746–767.
121. Cremer, J.; Le, T.; Noé, F.; Clevert, D.A.; Schütt, K.T. PILOT: equivariant diffusion for pocket-conditioned de novo ligand generation with multi-objective guidance via importance sampling. *Chemical Science* **2024**, *15*, 14954–14967.
122. Qiao, A.; Xie, J.; Huang, W.; Zhang, H.; Rao, J.; Zheng, S.; Yang, Y.; Wang, Z.; Li, G.B.; Lei, J. A 3D pocket-aware and affinity-guided diffusion model for lead optimization. *arXiv preprint arXiv:2504.21065* **2025**.
123. Alakhdar, A.; Poczos, B.; Washburn, N. Diffusion models in de novo drug design. *Journal of Chemical Information and Modeling* **2024**, *64*, 7238–7256.
124. Zhang, O.; Zhang, J.; Jin, J.; Zhang, X.; Hu, R.; Shen, C.; Cao, H.; Du, H.; Kang, Y.; Deng, Y.; et al. ResGen is a pocket-aware 3D molecular generation model based on parallel multiscale modelling. *Nature Machine Intelligence* **2023**, *5*, 1020–1030.
125. Huang, L.; Xu, T.; Yu, Y.; Zhao, P.; Chen, X.; Han, J.; Xie, Z.; Li, H.; Zhong, W.; Wong, K.C.; et al. A dual diffusion model enables 3D molecule generation and lead optimization based on target pockets. *Nature Communications* **2024**, *15*, 2657.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.