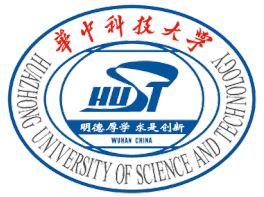
ICLR
International Conference On Learning Representations



Fast and Accurate Blind Flexible Docking

Zizhuo Zhang, Lijun Wu, Kaiyuan Gao, Jiangchao Yao, Tao Qin, Bo Han

TMLR Group, Hong Kong Baptist University
Shanghai AI Laboratory
Huazhong University of Science and Technology
CMIC, Shanghai Jiao Tong University
Microsoft Research AI4Science


Paper


Dataset


Code

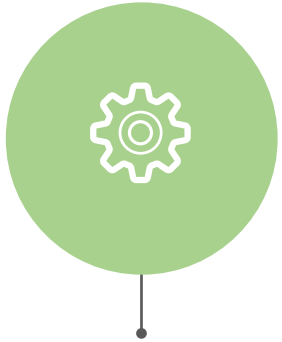
Outline

Background

Molecular Docking & Application
Rigid vs. Flexible

Related Work

Regression-based
Sampling-based



Motivation

Rigid → Flexible
High Efficiency

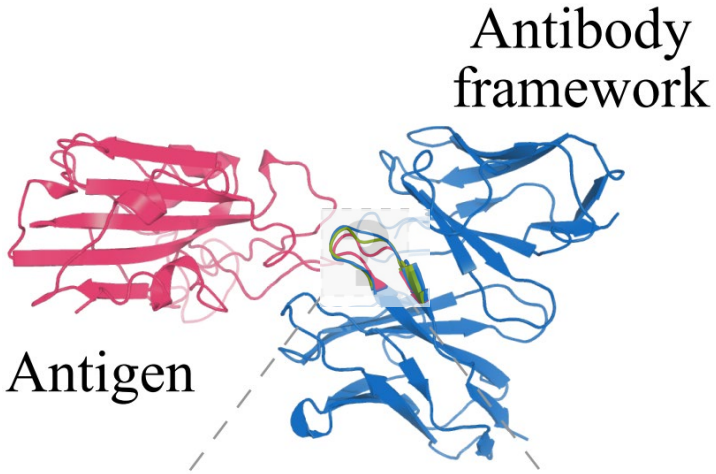
Method

Regression-based
Flexible Docking

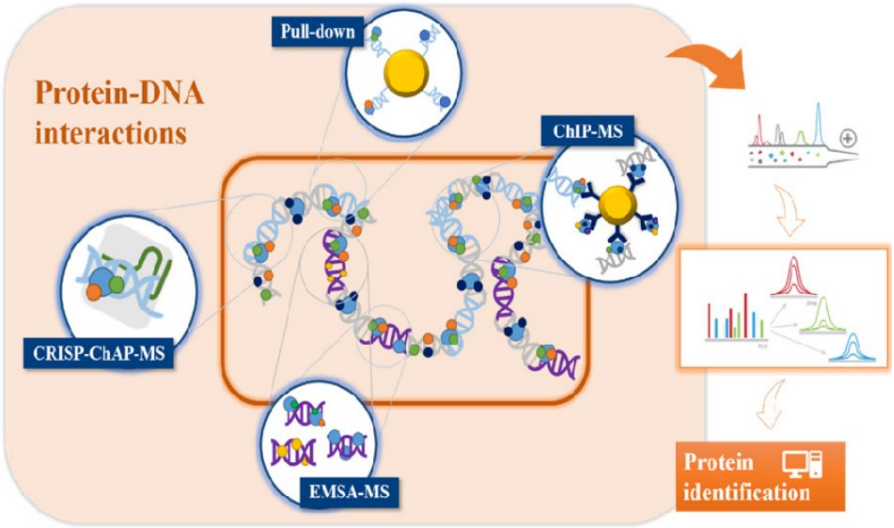
Experiments

Better Performance
Faster Computation

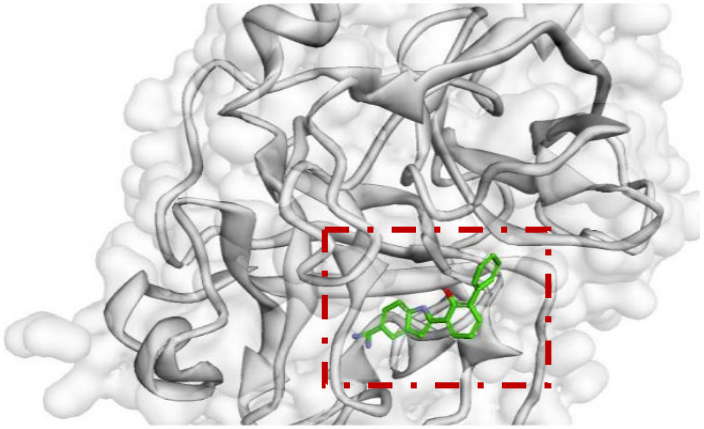
Background: Biomolecular Interaction



Antibody-Antigen



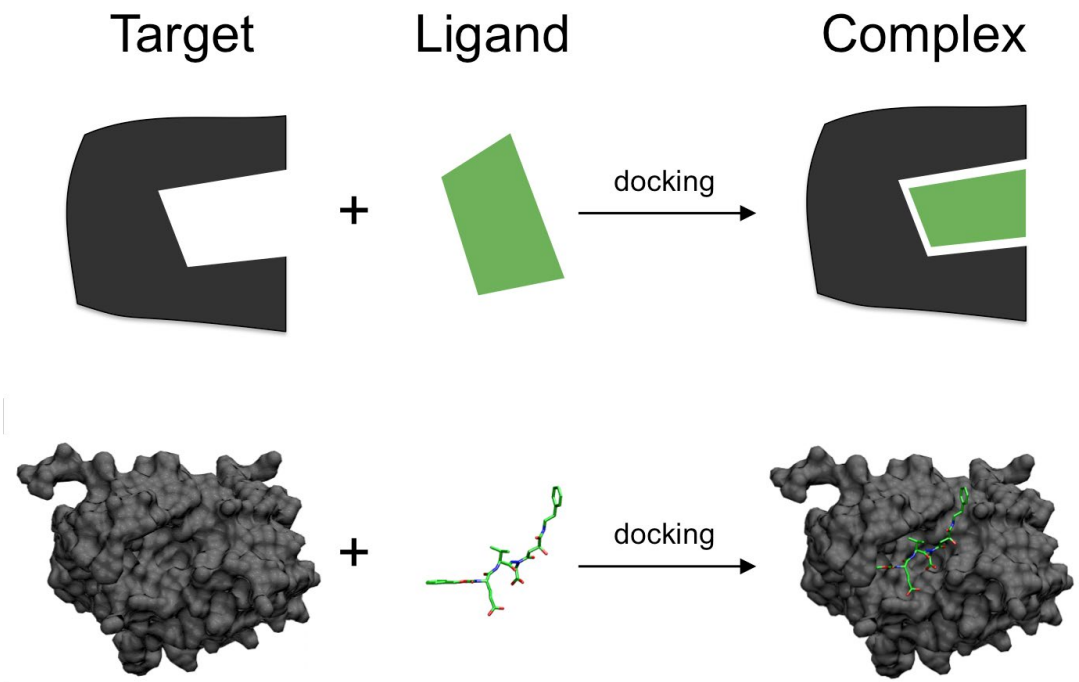
Proteins-Nucleic Acids (DNA, RNA)



Protein-Ligand (Molecule)

- Proteins, molecules, nucleic acids all have 3D structures in real-world scenarios.
- **Biomolecular interactions are the fundamental basis of life**, as they dynamically interact with each other to drive biological processes

Background: Molecular (Protein-Ligand) Docking



Input:

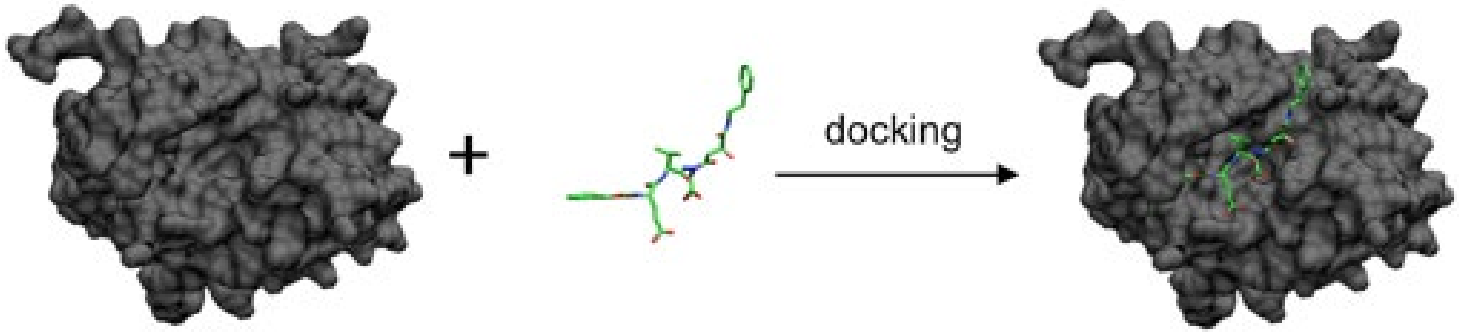
- Protein 3D structure (holo / apo protein)
- Molecule 3D structure randomly initialized by RDKit (apo ligand)

Output:

- Docked (Bound) protein-molecule complex 3D structure
  - ✓ Docked molecule 3D structure (holo ligand)
  - ✓ Docked protein 3D structure (holo protein)

- Molecular Docking, also called Protein-Ligand Docking, or Protein-Molecule Docking.
- Given a protein target, Protein-Molecule Docking aims to predict **how a molecule interacts with the protein** to have its effects.

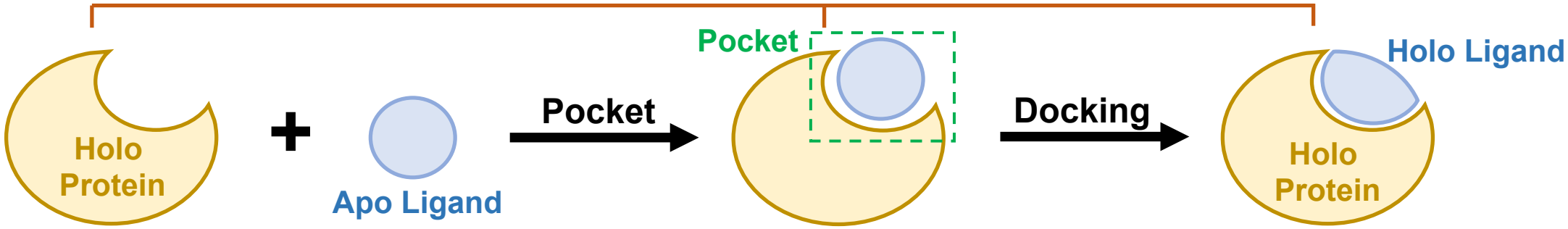
Background: Application



- **Drug Design and Discovery:** Used to predict the binding modes of small molecules with protein targets, aiding in design and discovery candidate drugs.
- **Virtual Screening:** Involves computationally screening large libraries of compounds to identify potential active compounds.
- **Protein-Ligand Interaction Studies:** Helps in understanding protein-ligand interactions, such as properties, binding affinities, others.
- Combination with other AI4Science Techniques: AlphaFold, RoseTTAFold, etc.

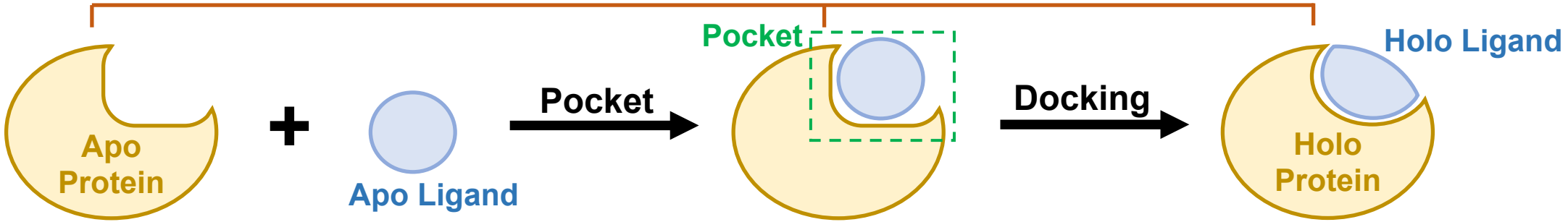
Rigid Docking vs. Flexible Docking

Rigid Docking assumes protein rigidity, i.e., **protein structure remains unchanged** during docking process



Rigid Docking: $(holo\ protein + apo\ ligand) \rightarrow holo\ ligand$

Flexible Docking relaxes the protein rigidity assumption,
i.e., **protein structure exhibits dynamic behaviors** during docking process, especially pocket region

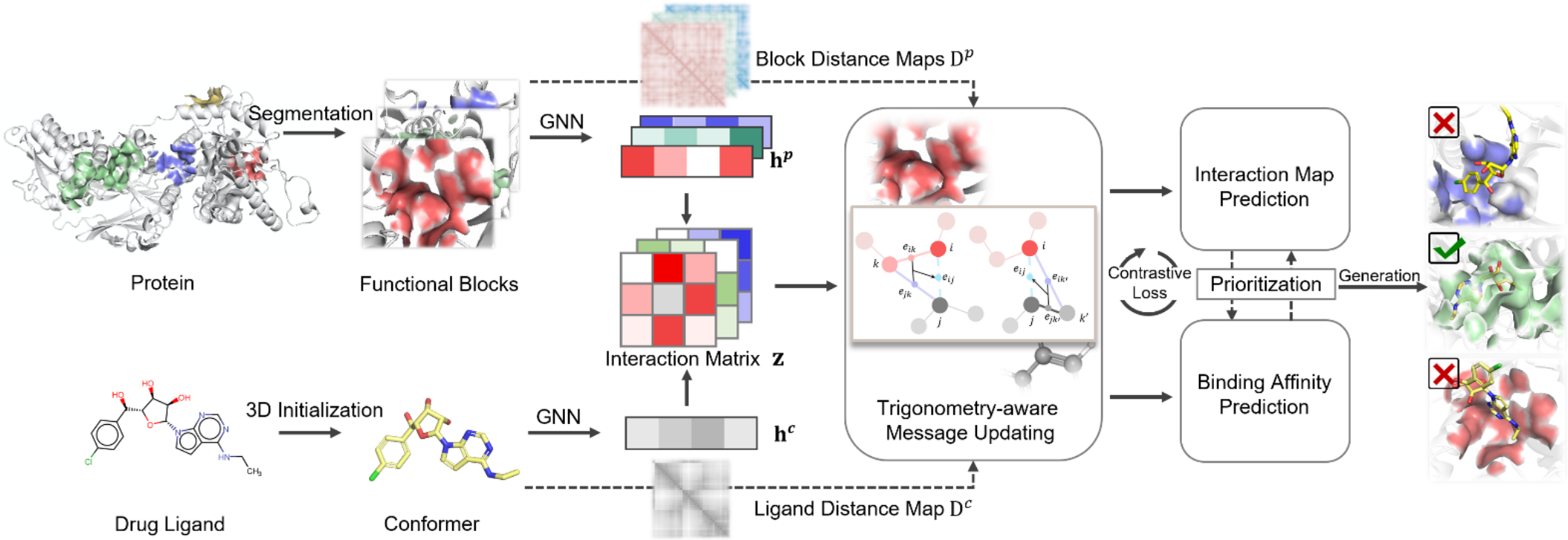


Flexible Docking: $(apo\ protein + apo\ ligand) \rightarrow (holo\ protein, holo\ ligand)$

Related Work: Rigid Molecular Docking

EquiBind (ICML 2022), TankBind (NeurIPS 2022), E3Bind (ICLR 2023):

- Regression-based Methods, directly predict binding structures.
- Encode protein (pocket) and molecule via E(3)-equivariant GNNs
- Use external pocket detection tool P2Rank to pre-identify pocket candidates (TankBind, E3Bind)

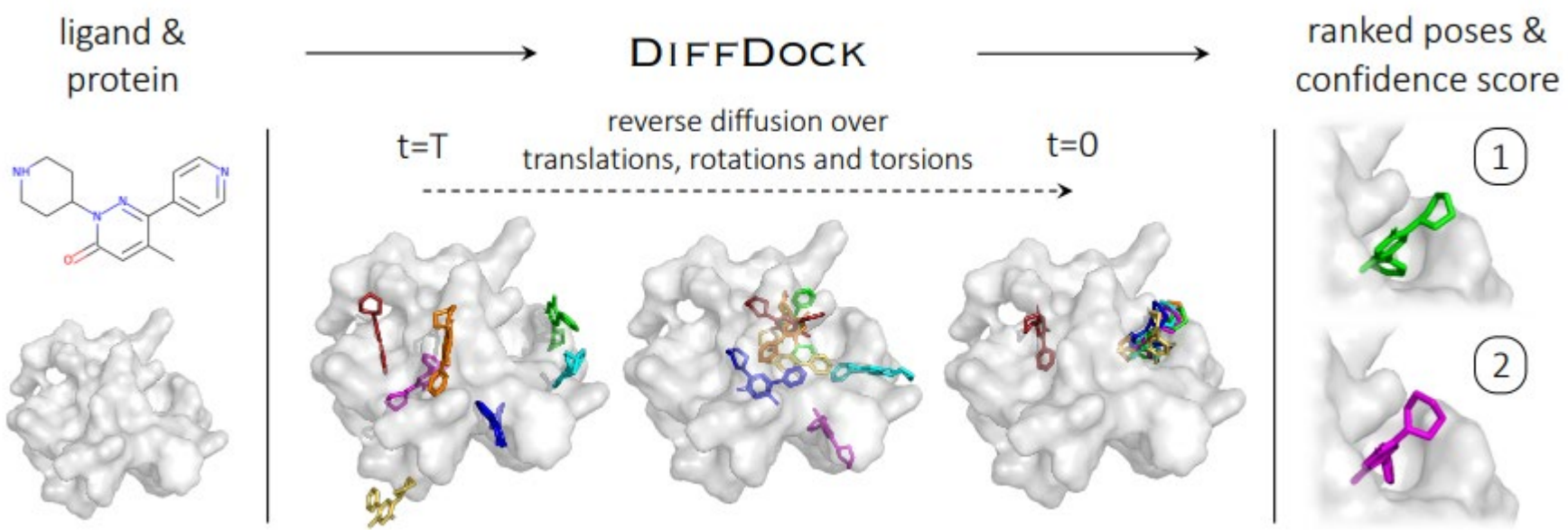


TankBind (NeurIPS 2022)

Related Work: Rigid Molecular Docking

DiffDcok (ICLR 2023), DiffDock-L (ICLR 2024):

- Sampling-based Methods, use diffusion model to generate multiple structures during inference and rank them using a confidence model.
- Randomly sampled initial structures are denoised via a reverse diffusion over translational, rotational, and torsional degrees of freedom $(m + 6)$.

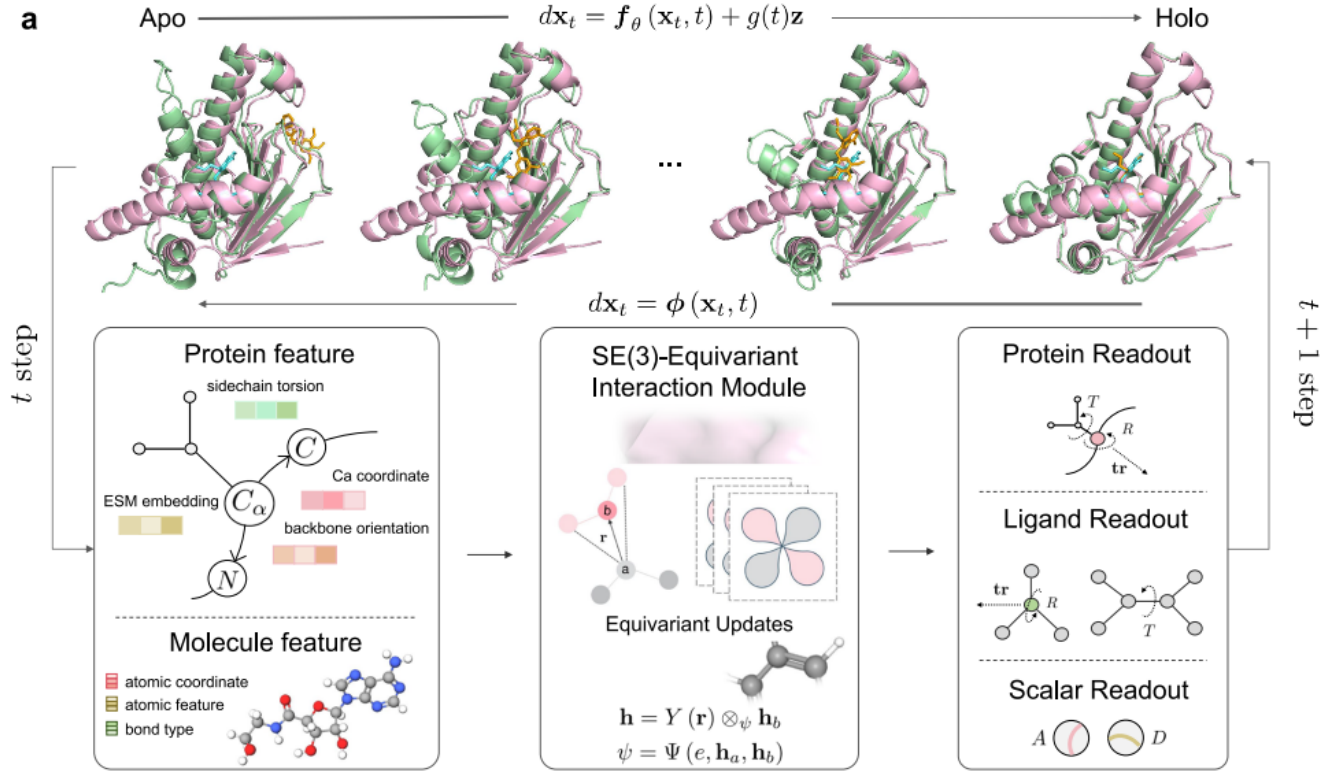


DiffDock (ICLR 2023)

Related Work: Flexible Molecular Docking

DynamicBind (Nature Communication 2024), ReDock (ICML 2024):

- Sampling-based Methods, use diffusion model to generate structures.
- Reverse diffusion to recover protein from AlphaFold2-initialized apo structures to holo structures.
- Reverse diffusion to denoise ligand from randomly initialized apo structures to holo structures.
- Use a confidence model to rank multiple generated structure.

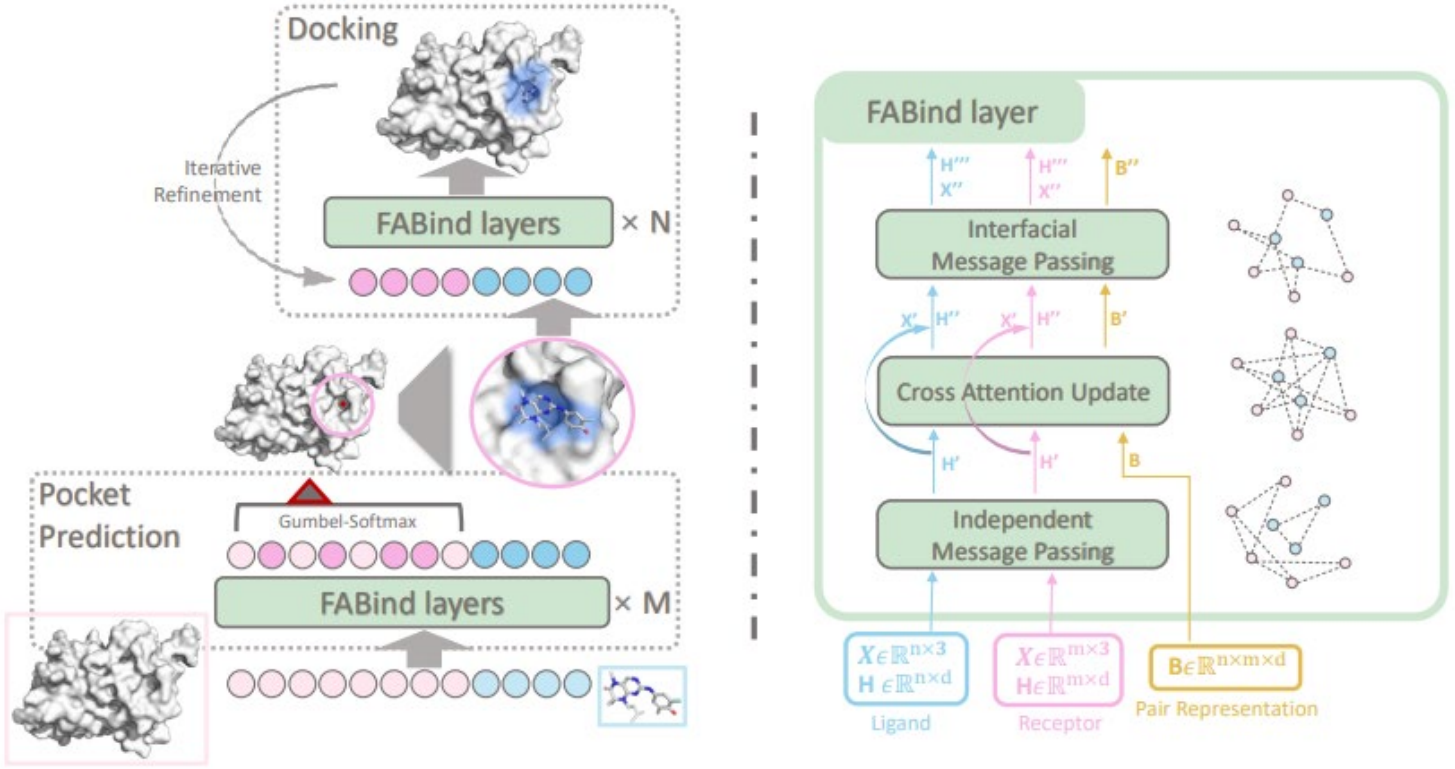


DynamicBind (Nature Communication 2024)

Related Work: FABind and FABind+

FABind (NeurIPS 2023), FABind+ (KDD 2025):

- Regression-based Methods, directly predict binding structures.
- Multi-task framework to both predict pocket residues and holo ligand structures.
- No need for external pocket detection tools.
- No need for a confidence model to rank sampled structures.



FABind (NeurIPS 2023)

Motivations

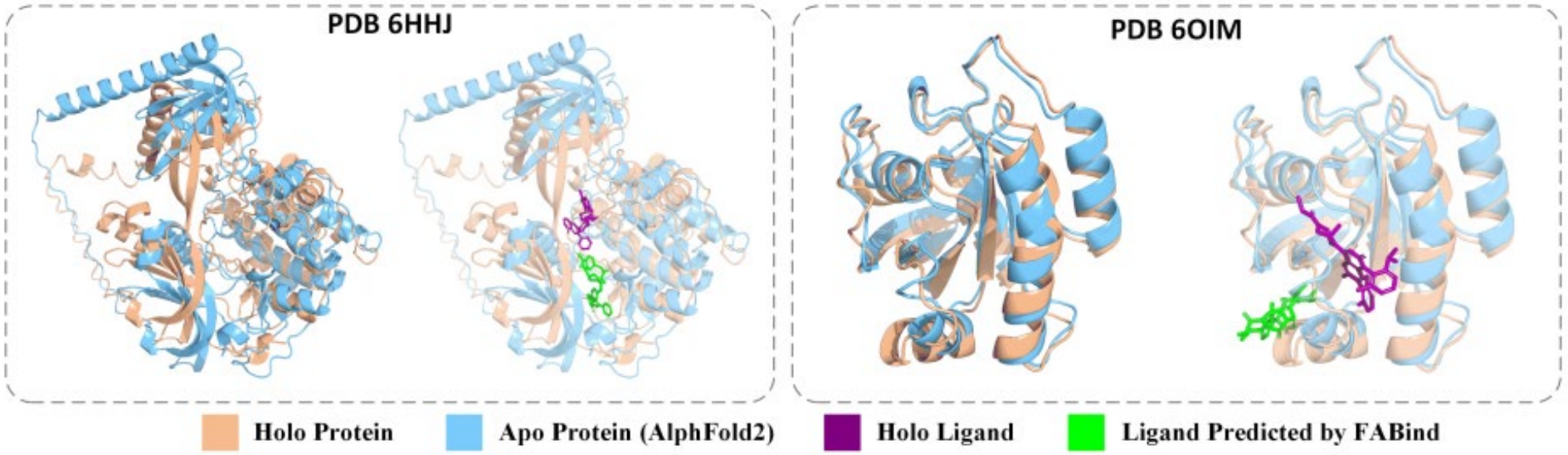
Technical Perspective: Regression-based vs. Sampling-based Docking

- Regression-based:
  - Directly predict the docked holo structures by a single forward computation.
  - Advantages: Faster computational efficiency
  - Disadvantages: Less accurate docking performance.
- Sampling-based:
  - Multi-round sampling involves multiple rounds of inference computation.
  - Train a confidence model to rank multiple sampled structures.
  - Advantages: Higher docking performance.
  - Disadvantages: Low computational efficiency.

Scenario Perspective: Rigid Docking vs. Flexible Docking

- Rigid Docking:
  - Protein rigidity assumption oversimplifies real-world protein-ligand interactions.
  - Both regression-based and sampling-based methods exist in previous work.
- Flexible Docking:
  - Flexible docking is a more realistic docking scenario considering dynamic protein behaviors.
  - Existing flexible docking methods predominantly follow the sampling-based technical line.

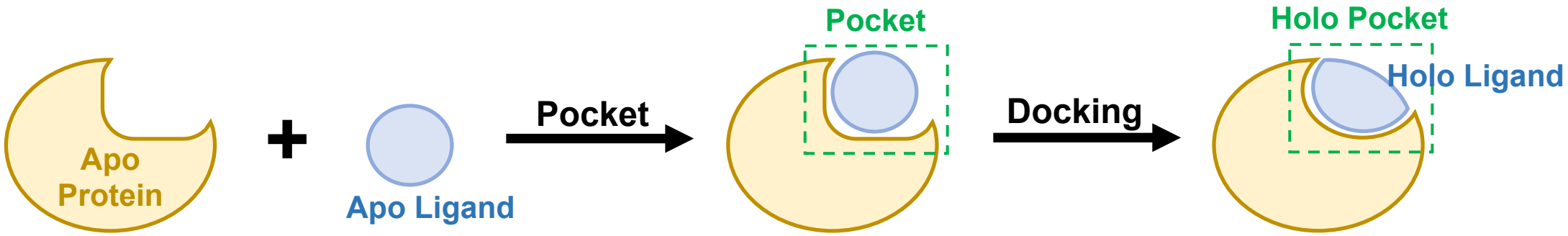
Motivations



- Existing flexible docking methods struggle with low efficiency due to sampling strategy.
- There is a research gap in exploring the potential of the regression-based paradigm for flexible docking.
- Rigid docking methods experience a noticeable performance drop when applied to flexible docking scenarios.

Can we develop a faster flexible docking method based on a regression-based paradigm?

Problem Definition: Blind Flexible Docking



Blind Flexible Docking: $(apo\ protein + apo\ ligand) \rightarrow (holo\ pocket, holo\ ligand)$

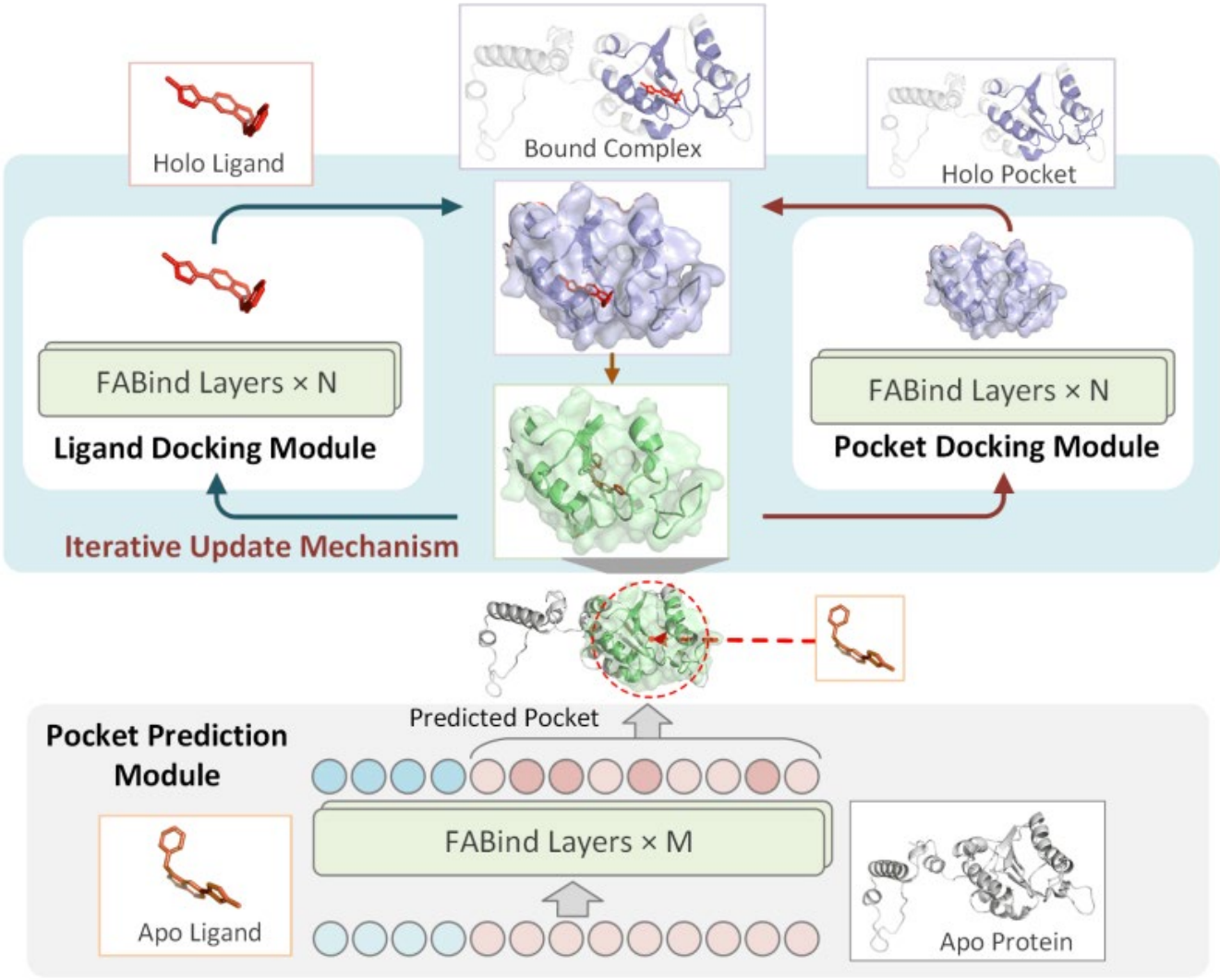
- Blind flexible docking:
  - Proteins exhibit flexibility (flexible) during docking process.
  - Binding pocket sites are unknown (blind), no prior knowledge about binding pocket sites.

- We use AlphaFold2-predicted structure to initialize apo protein structure.
- We use RDKit to randomly initialize apo ligand structure.

- Following existing methods, we use geometric heterogeneous graph to model protein-ligand or pocket-ligand complexes.

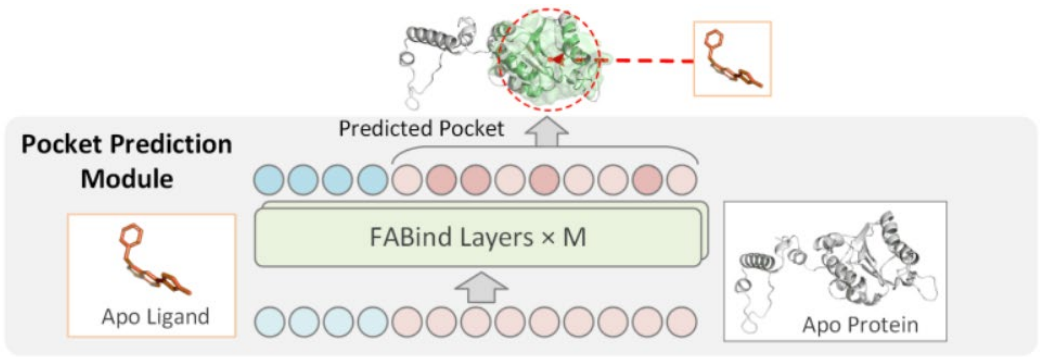
FABFlex: Fast and Accurate Blind Flexible Docking

- Regression-based multi-task framework to correspond to three subtasks:
  - ✓ Pocket prediction module:
    - identify binding pocket residues.
  - ✓ Ligand docking module:
    - predict holo ligand structure.
  - ✓ Pocket docking module:
    - predict holo pocket structure.
- Iterative update mechanism
  - Exchange predicted structures between two docking modules to further promote coordinate refinement.
- End-to-end



Overview of proposed FABFlex

Pocket Prediction Module



Input
- Protein-ligand heterogeneous graph

Output:
- Residues to form the pocket region.

- Pocket Prediction Module
  - Residue-level binary classification

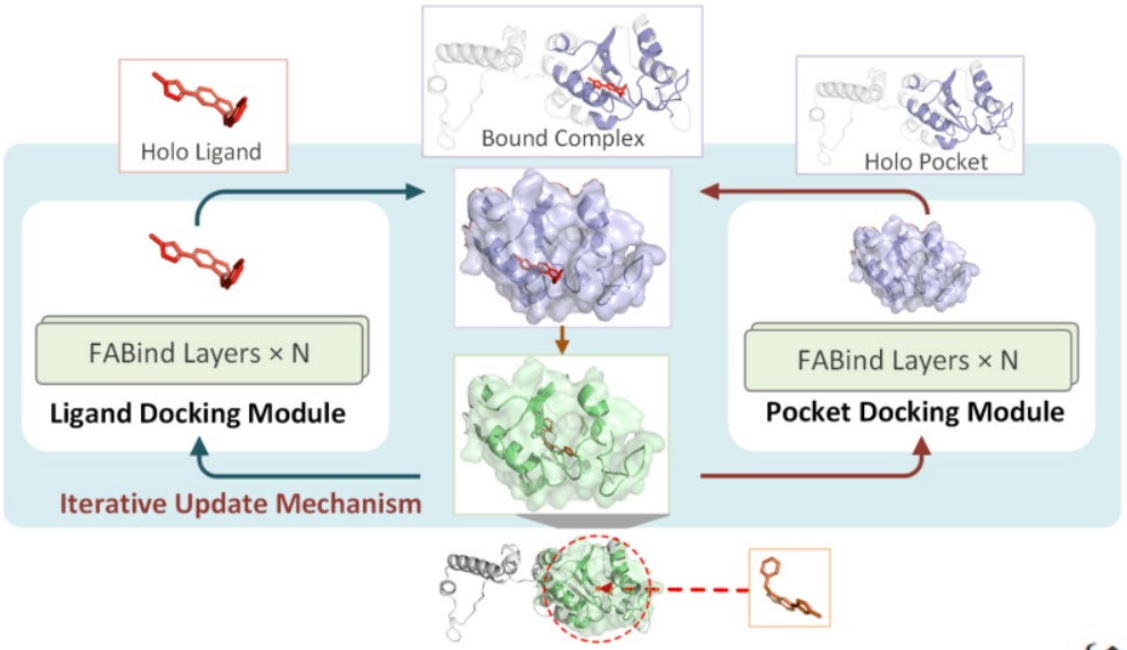
$$\{\hat{y}_j\}_{1\leq j\leq n^p} = \mathcal{M}_S(\mathcal{G}, \{\mathbf{h}_i, \mathbf{x}_i\}_{1\leq i\leq n^l}, \{\mathbf{h}_j, \mathbf{x}_j\}_{1\leq j\leq n^p}),$$
$$\{v_j\}_{1\leq j\leq \hat{n}^{p*}} = \{\hat{y}_j \odot v_j\}_{1\leq j\leq n^p} \in \hat{\mathcal{V}}^{p*},$$

  - Predict the pocket center.
  - Move ligand to pocket center for subsequent docking modules.

$$\mathcal{L}_{\text{pocket\_cls}} = \frac{1}{N}\sum_{i=1}^{N}\frac{p_i}{q_i}\{-\sum_{j=1}^{p_i}[y_j \log(\hat{y}_j) + (1-y_j)\log(1-\hat{y}_j)]\}, \qquad \mathcal{L}_{\text{pocket\_center}} = \frac{1}{N}\sum_{i=1}^{N}\text{HuberLoss}(center_i, \widehat{center}_i),$$

$$\mathcal{L}_{\text{pocket\_pred}} = \alpha_1^{\text{cls}}\mathcal{L}_{\text{pocket\_cls}} + \alpha_1^{\text{center}}\mathcal{L}_{\text{pocket\_center}}.$$

Ligand Docking Module and Pocket Docking Module



Input

- Pocket-ligand heterogeneous graph

Output:

- Predicted coordinates of holo ligand.
- Predicted coordinates of holo pocket.

- Ligand Docking Module
  - Predict holo ligand structure
- Pocket Docking Module
  - Predict holo pocket structure

$$\{\hat{\mathbf{x}}_i\}_{1\le i\le n^l} = \mathcal{M}_L(\hat{\mathcal{G}}^*, \{\mathbf{h}_i, \mathbf{x}_i\}_{1\le i\le n^l}, \{\mathbf{h}_j, \mathbf{x}_j\}_{1\le j\le \hat{n}^{p*}}),$$

$$\{\hat{\mathbf{x}}_j\}_{1\le j\le \hat{n}^{p*}} = \mathcal{M}_P(\hat{\mathcal{G}}^*, \{\mathbf{h}_i, \mathbf{x}_i\}_{1\le i\le n^l}, \{\mathbf{h}_j, \mathbf{x}_j\}_{1\le j\le \hat{n}^{p*}}),$$

$$\mathcal{L}_{\text{ligand\_coord}} = \frac{1}{N}\sum_{i=1}^{N}\text{HuberLoss}(\tilde{\mathbf{x}}_i^l, \hat{\mathbf{x}}_i^l), \quad \mathcal{L}_{\text{pocket\_coord}} = \frac{1}{N}\sum_{i=1}^{N}\text{HuberLoss}(\tilde{\mathbf{x}}_i^p, \hat{\mathbf{x}}_i^p),$$

$$\mathcal{L}_{\text{dis\_map}} = \frac{1}{N}\sum_{i=1}^{N}\text{MSELoss}(\tilde{\mathbf{D}}_i, \hat{\mathbf{D}}_i),$$

- Iterative Update Mechanism to promote prediction exchange for coordinate refinement.

Experimental Settings

- **Dataset:** PDBBind v2020
  - Protein-ligand complexes recorded before 2019 as training set and validation set.
  - Protein-ligand complexes recorded after 2019 as test set.
  - Align AlphaFold2-predicted protein structures with groud-truth structures.

PDBbind	# Complexes	avg. residues in proteins	max. residues in proteins	avg. heavy atoms in ligands	max. atoms in ligands
Train	12,807	310.59	1,290	31.30	149
Validate	734	312.14	1,025	32.52	177
Test	303	280.92	1,098	35.77	147

- **Evaluation Metrics:** RMSD (Root-Mean-Square Deviation)
  - Ligand RMSD
  - Pocket RMSD

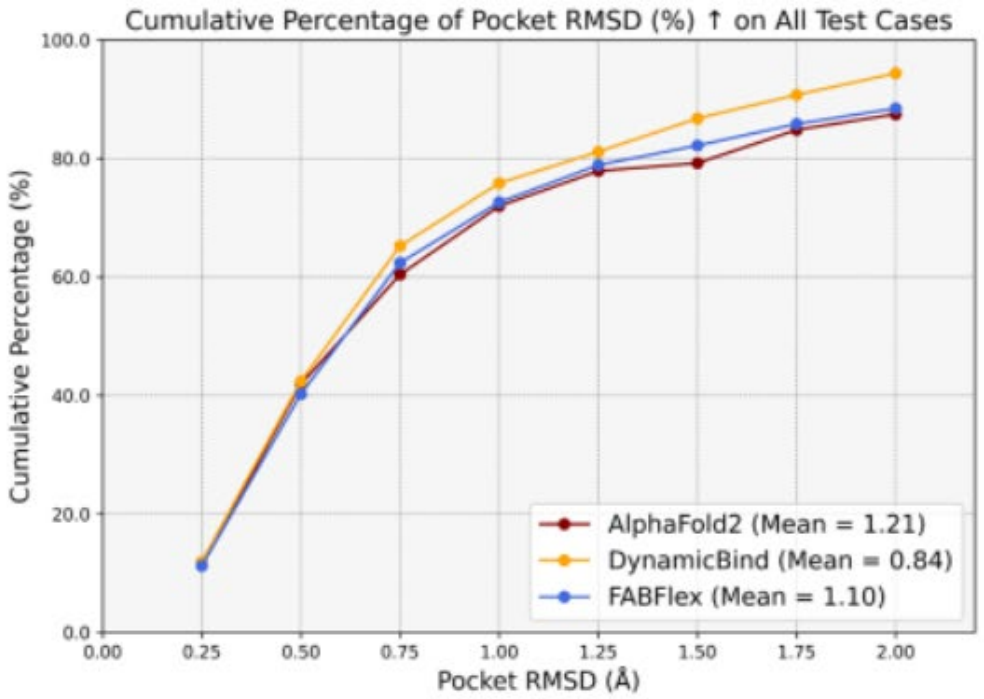
$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - x_i')^2 + (y_i - y_i')^2 + (z_i - z_i')^2}$$

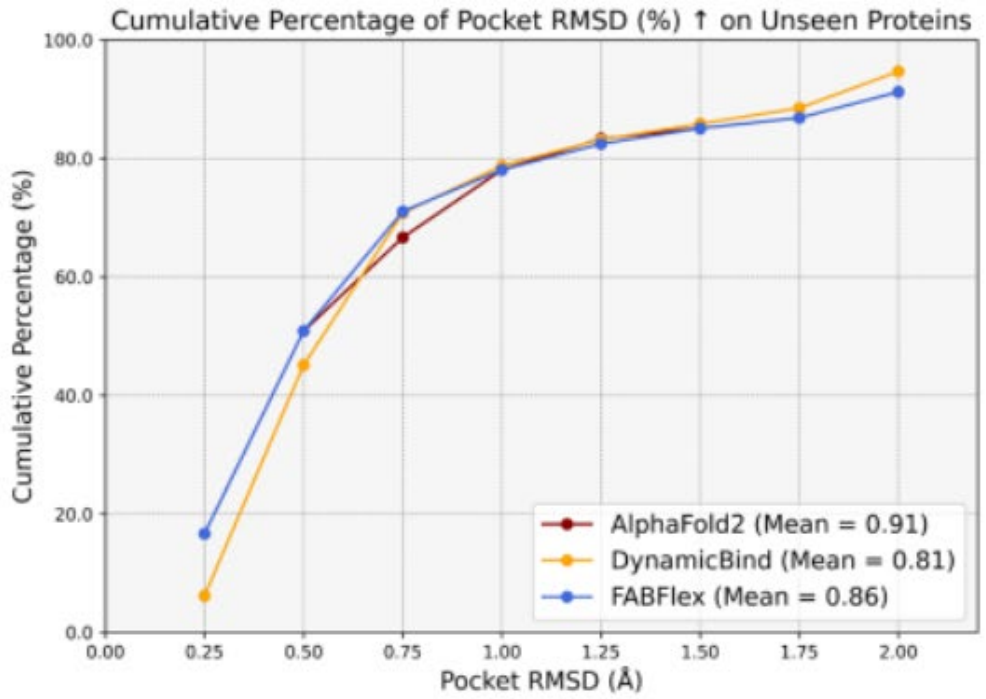
Experimental Results: Ligand RMSD

Method	Ligand RMSD												Average Runtime (s)
	On All Cases						On Unseen Protein Receptors						
	Percentiles ↓				% Below ↑		Percentiles ↓				% Below ↑		
	25%	50%	75%	Mean	< 2Å	< 5Å	25%	50%	75%	Mean	< 2Å	< 5Å	
Traditional Docking Software													
Vina	4.79	7.14	9.21	7.14	6.67	27.33	5.27	7.06	8.84	7.15	6.25	23.21	205*
Glide	2.84	5.77	8.04	5.81	14.66	40.60	2.38	5.01	**7.17**	**5.21**	21.36	49.51	1405*
Gnina	2.58	5.17	8.42	5.76	19.32	48.47	2.03	4.96	7.35	5.33	24.55	50.91	146
Deep Learning-based Rigid Docking Methods													
TankBind	2.82	4.53	7.79	7.79	8.91	54.46	2.88	4.45	7.53	7.60	4.39	58.77	0.87
FABind	2.19	3.73	8.39	6.63	22.11	60.73	2.73	4.83	9.35	7.15	8.77	50.88	0.12
FABind+	1.58	**2.79**	6.69	5.63	35.64	66.01	1.93	**3.13**	8.59	6.76	27.19	57.89	0.16
DiffDock	1.82	3.92	6.83	6.07	29.04	60.73	1.97	4.82	8.03	7.41	26.32	51.75	82.83
DiffDock-L	1.55	3.22	6.86	5.99	36.75	62.58	1.86	3.16	9.09	7.14	29.82	**61.40**	58.72
Deep Learning-based Flexible Docking Methods													
DynamicBind	1.57	3.16	7.14	6.19	33.00	64.69	2.23	4.02	10.23	8.27	20.18	54.39	102.12
FABFlex	**1.40**	2.96	**6.16**	**5.44**	**40.59**	**68.32**	**1.81**	3.51	8.03	7.17	**32.46**	59.65	0.49

- FABFlex consistently outperforms both traditional docking software and contemporary deep learning-based methods almost across all metrics.
- FABFlex effectively generalizes to new and unseen proteins.
- FABFlex is approximately 208 times faster than DynamicBind.

Experimental Results: Pocket RMSD

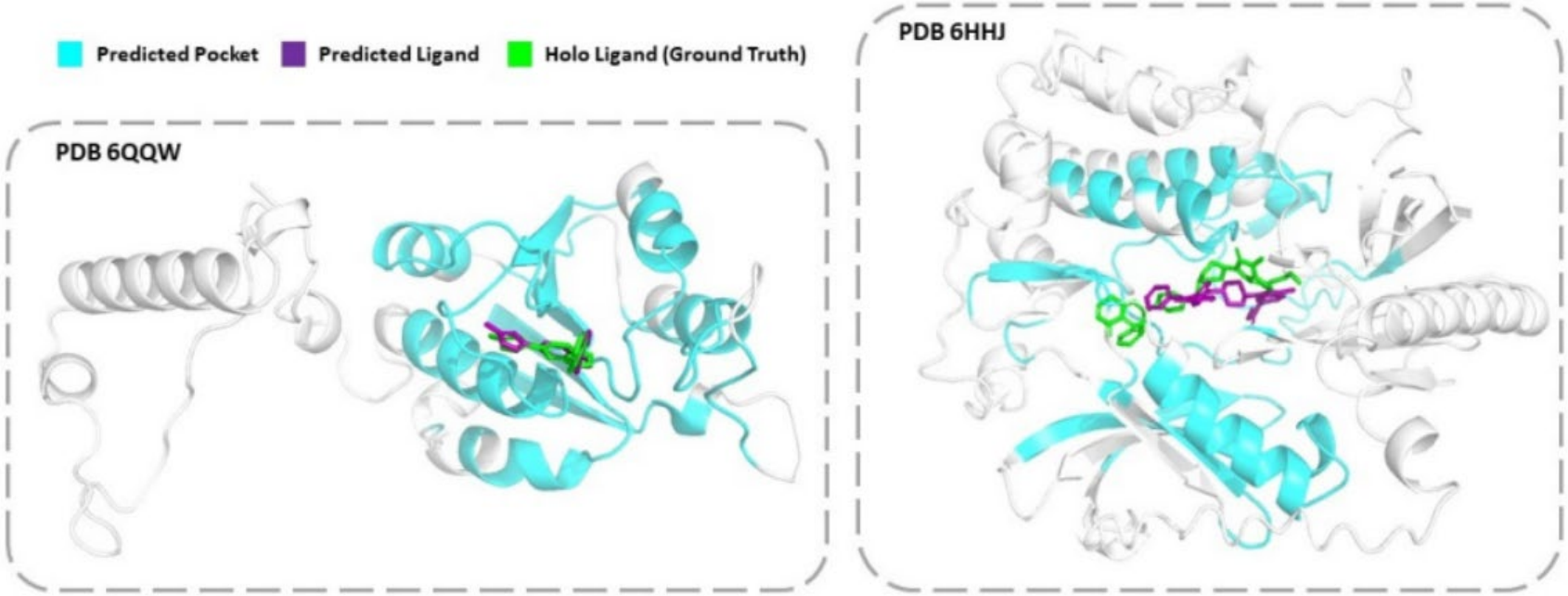




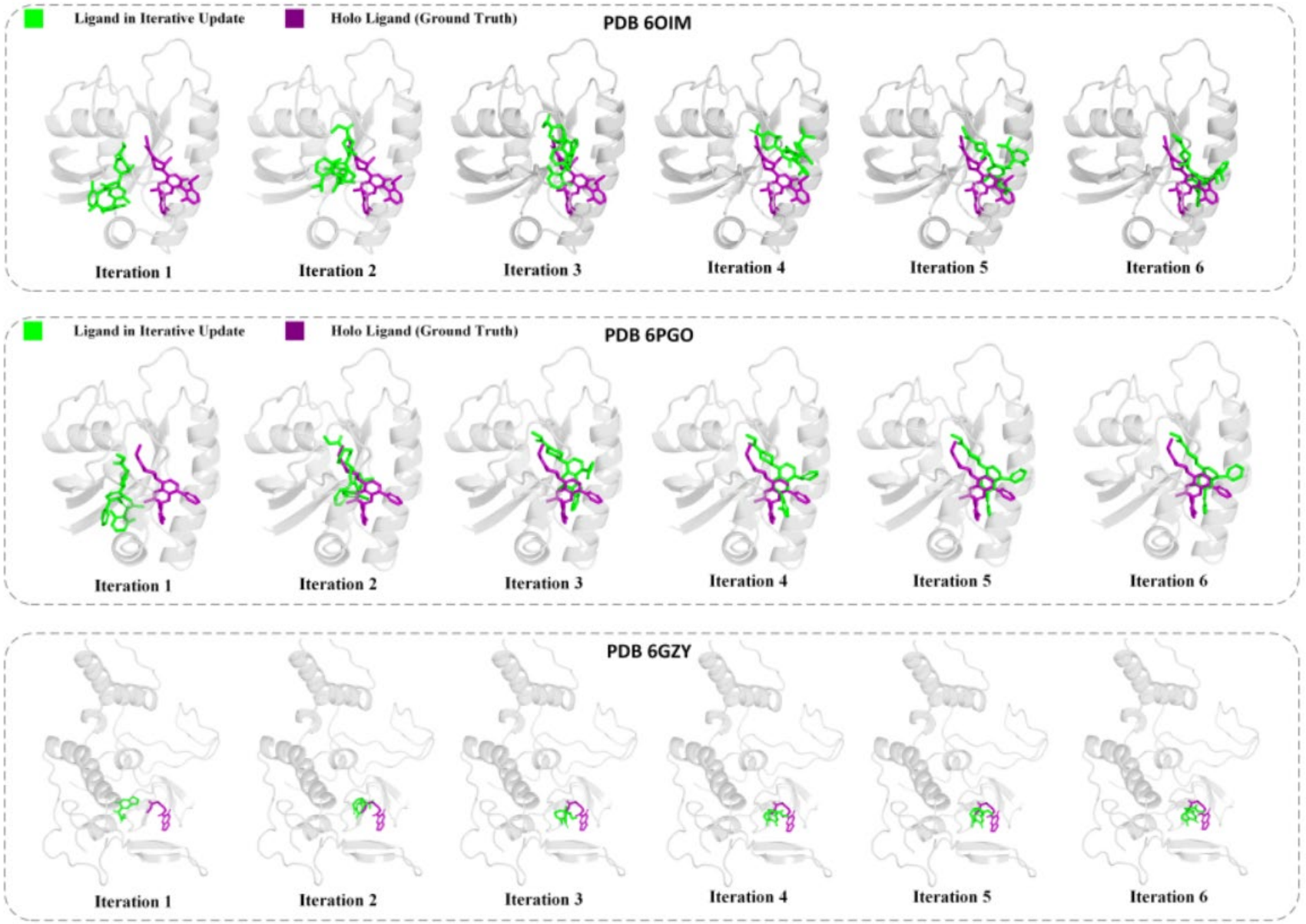
- FABFlex has a positive effect on refining pocket poses from the initial AlphaFold2 structures.
- FABFlex can improve pocket structures for those unseen proteins.

Experimental Results: Pocket Prediction

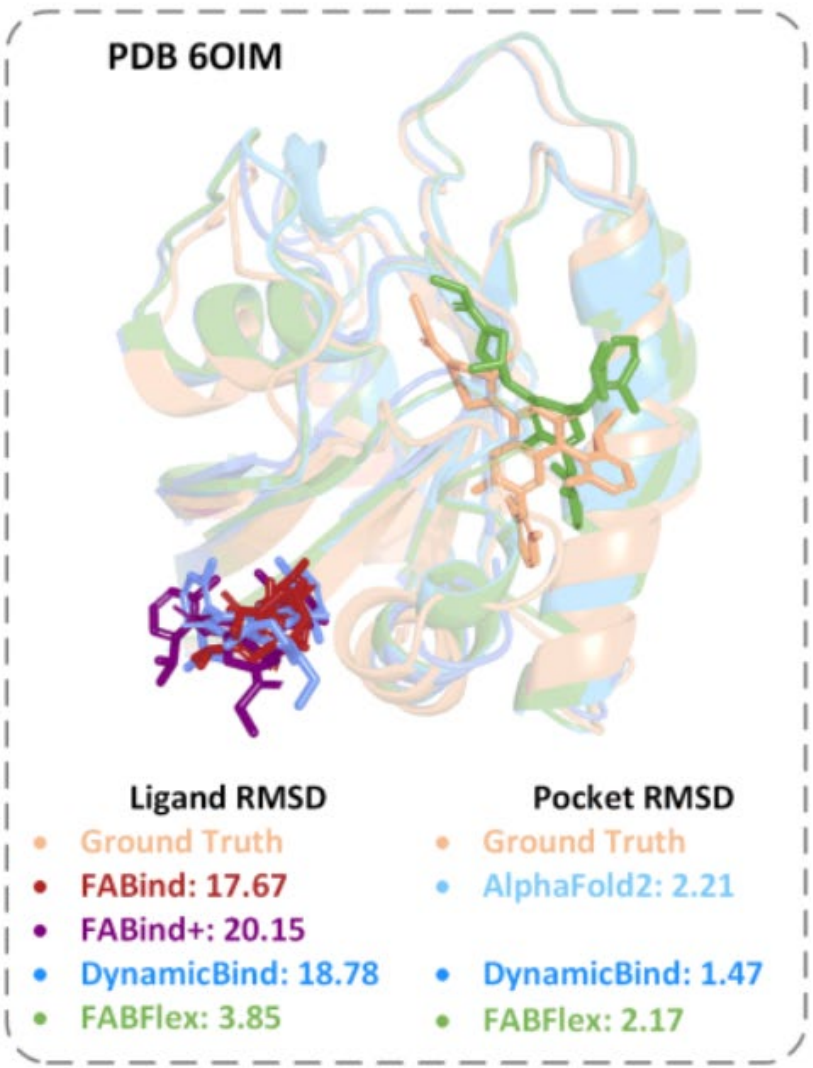
Method	Given Apo Proteins				Given Holo Proteins			
	CLS ACC % ↑	Pocket Center (Å)			CLS ACC % ↑	Pocket Center (Å)		
		MAE ↓	RMSE ↓	EucDist ↓		MAE ↓	RMSE ↓	EucDist ↓
P2Rank	-	4.04	5.69	7.85	-	4.11	6.17	8.05
FABind	73.32	**3.06**	5.30	**6.18**	73.54	**3.08**	5.35	**6.18**
FABind+	**87.18**	3.49	5.09	6.94	**87.52**	3.50	5.13	6.69
FABFlex	87.08	3.29	**4.83**	6.59	87.06	3.31	**4.89**	6.63

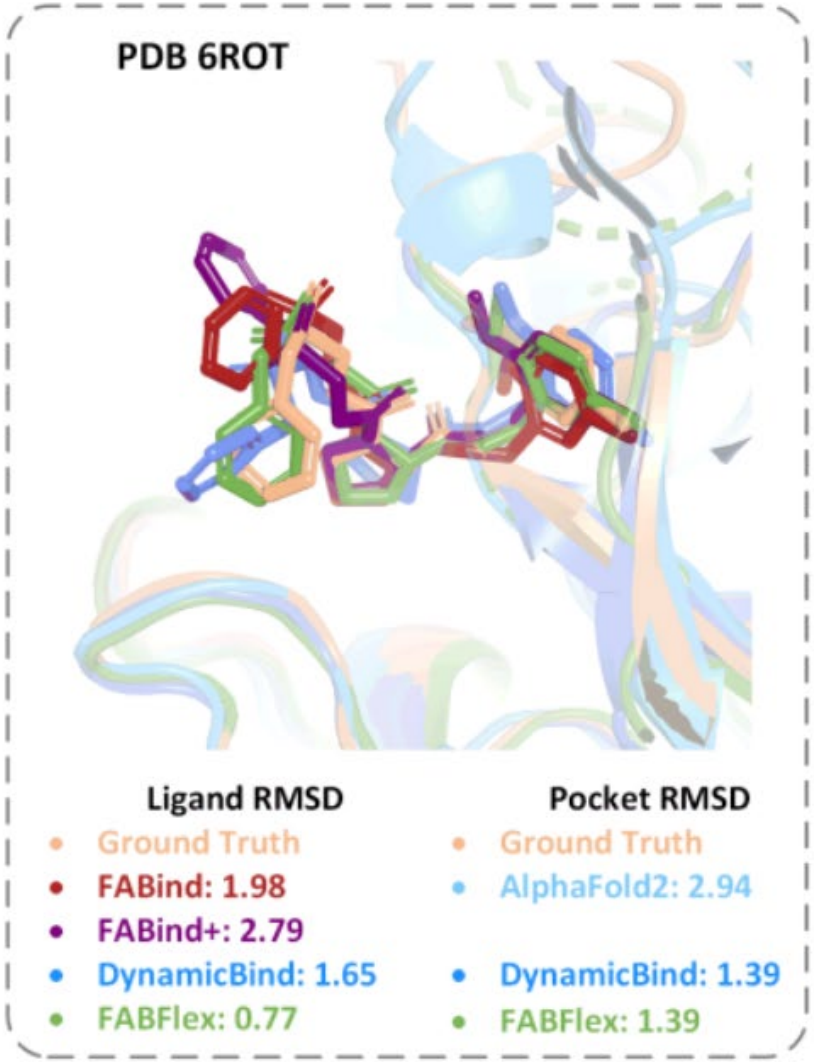


Experimental Results: Iterative Update Mechanism



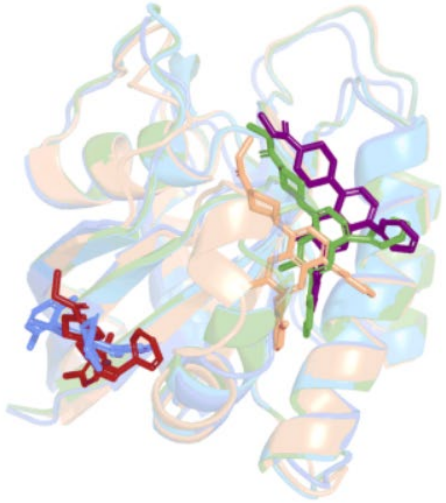
Experimental Results: Case Studies





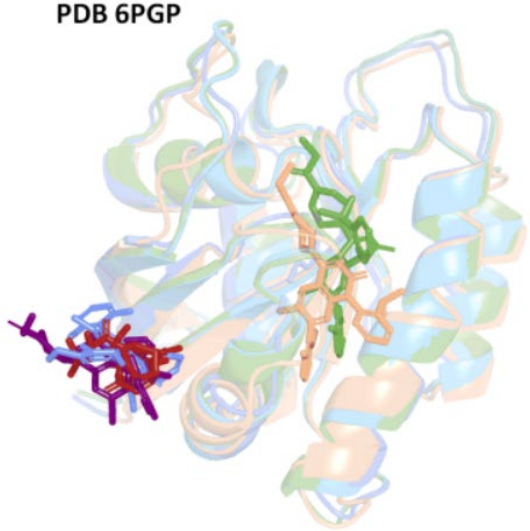
Experimental Results: Case Studies

PDB 6PGO



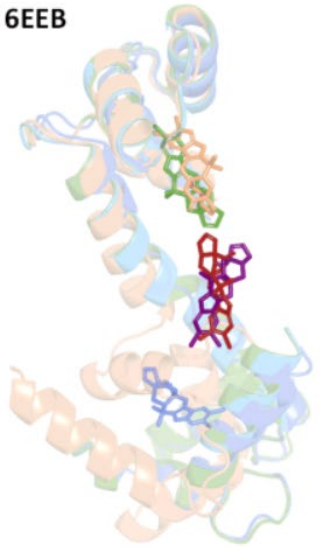
Ligand RMSD	Pocket RMSD
Ground Truth	Ground Truth
FABind: 15.66	AlphaFold2: 1.73
FABind+: 4.34	
DynamicBind: 16.69	DynamicBind: 1.52
FABFlex: 2.99	FABFlex: 1.70

PDB 6PGP



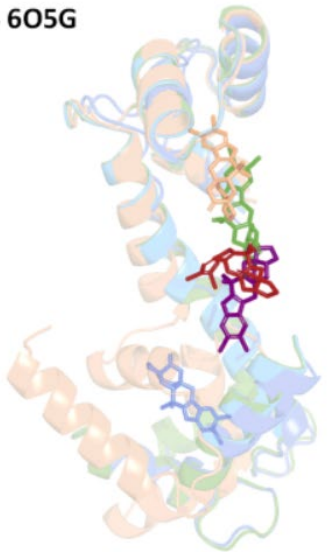
Ligand RMSD	Pocket RMSD
Ground Truth	Ground Truth
FABind: 16.66	AlphaFold2: 1.80
FABind+: 18.29	
DynamicBind: 17.56	DynamicBind: 1.23
FABFlex: 4.25	FABFlex: 1.76

PDB 6EEB



Ligand RMSD	Pocket RMSD
Ground Truth	Ground Truth
FABind: 15.18	AlphaFold2: 3.38
FABind+: 15.59	
DynamicBind: 29.28	DynamicBind: 3.34
FABFlex: 3.63	FABFlex: 3.37

PDB 6O5G



Ligand RMSD	Pocket RMSD
Ground Truth	Ground Truth
FABind: 11.95	AlphaFold2: 3.36
FABind+: 15.52	
DynamicBind: 29.12	DynamicBind: 3.34
FABFlex: 6.52	FABFlex: 3.35

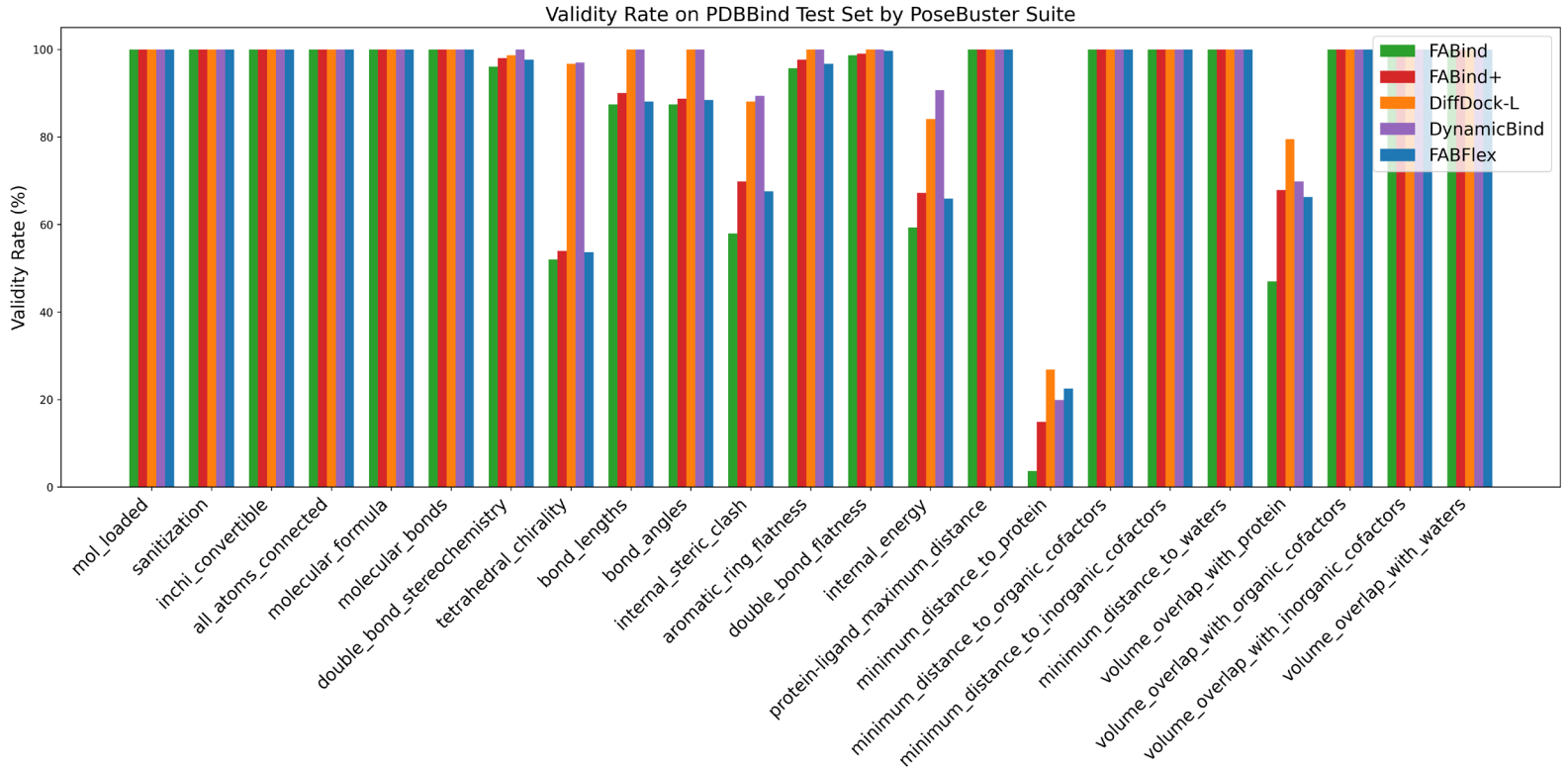
Experimental Results: Validity of Predicted Structure

Clash score

Methods	Vina	Glide	Gnina	TankBind	FABind	FABind+	DiffDock	DynamicBind	FABFlex
clash score ↓	0.02	0.08	0.05	0.41	0.51	0.45	0.33	0.27	0.37

PoseBuster

Methods	Ligand RMSD < 2Å (%)	PB-valid & Ligand RMSD < 2Å (%)
FABind	22.52	1.32
FABind+	34.11	8.94
DiffDock-L	37.42	16.89
DynamicBind	30.79	14.57
FABFlex	39.40	13.91



Reference

[1] Equibind: Geometric deep learning for drug binding structure prediction, ICML, 2022.
[2] Tankbind: Trigonometry-aware neural networks for drug-protein binding structure prediction, NeurIPS, 2022.
[3] E3Bind: An End-to-End Equivariant Network for Protein-Ligand Docking, ICLR, 2023.
[4] Fabind: Fast and accurate protein-ligand binding, NeurIPS, 2023.
[5] DynamicBind: predicting ligand-specific protein-ligand complex structure with a deep equivariant generative model, Nature Communications, 2024.
[6] Fabind+: Enhancing molecular docking through improved pocket prediction and pose generation, KDD, 2025.
[7] PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences, Chemical Science, 2024.

ICLR
International Conference On Learning Representations

Thank you!

If you have any questions, feel free to contact with us!

Email: cszzzhang@comp.hkbu.edu.hk

Paper

Dataset

Code