Swiss Institute of Bioinformatics

INTRODUCTION TO COMPUTER-AIDED DRUG DESIGN

Section 7
Biomolecular Complex Structure Prediction for Drug Design

Ute F. Röhrig, Molecular Modeling Group, SIB Lausanne

June 10, 2026, Lausanne

Unil.
Université de Lausanne
Faculté de biologie
et de médecine

1

Agenda

Session	Lecture	Practice
1	Prologue: molecular representation	
	Introduction to (computer-aided) drug design	
	Origin of 3D structures	
	Molecular recognition	Use of UCSF ChimeraX to analyze protein-ligand complexes
2	Binding free energy estimation	
	Introduction to molecular docking	Ligand-protein docking with **SwissDock**
3	Introduction to molecular (virtual) screening	Ligand-based virtual screening with **SwissSimilarity**
4	Short introduction on target prediction of small molecules	Use of **SwissTargetPrediction** to perform reverse screening.
5	Introduction to ADME, pharmacokinetics, druglikeness	Estimate physicochemical, pharmacokinetic, druglike and related properties with **SwissADME**
6	Short introduction to bioisosterism	Use of **SwissBioisostere** to perform bioisosteric design
7	Biomolecular Complex Structure Prediction	Use of UCSF ChimeraX/OpenFold to predict and analyze biomolecular complexes

2

2

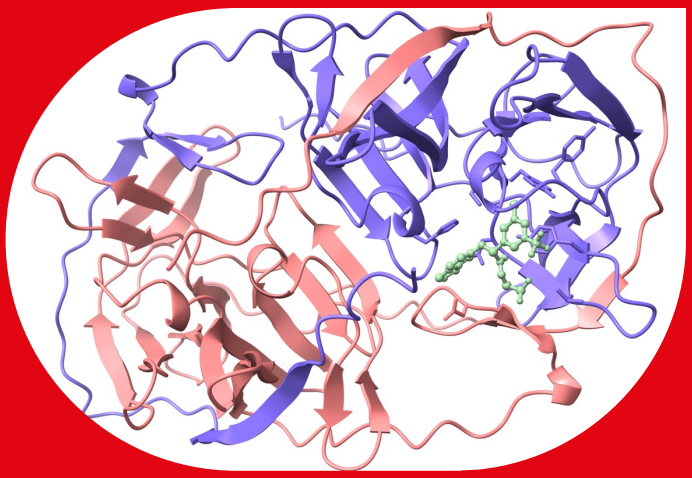
Outline of Section 7

Biomolecular Complex Structure Prediction

- Introduction
- Docking
- Co-Folding
- Exercises with ChimeraX and AlphaFold/OpenFold

Biomolecular Complex: Assembly of biomacromolecules, such as proteins, DNA, RNA, carbohydrates, lipids, co-factors, metals, ligands,…

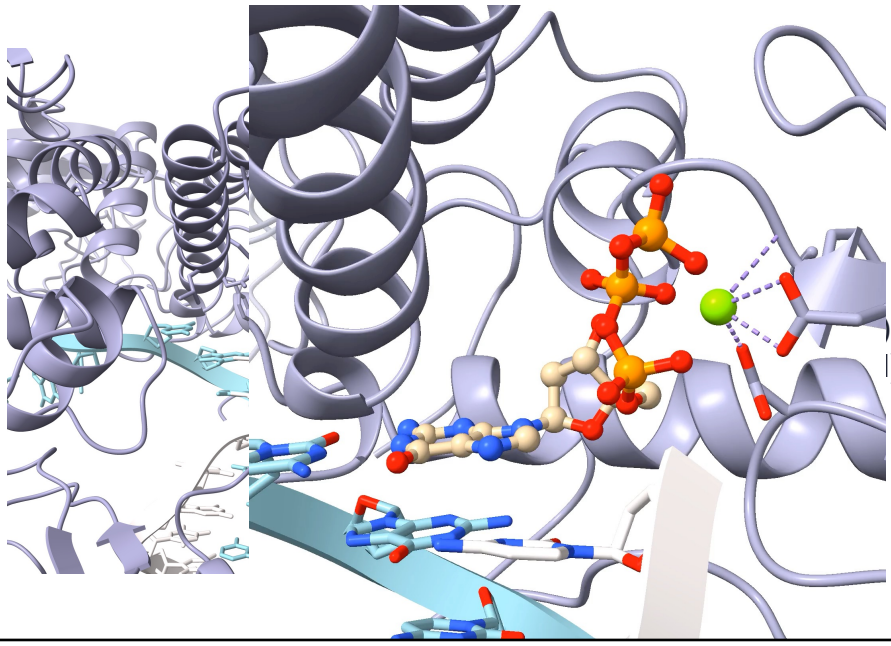
3



Biomolecular Complex Structure Prediction

4

Biomolecular Complex Structure Prediction - Docking



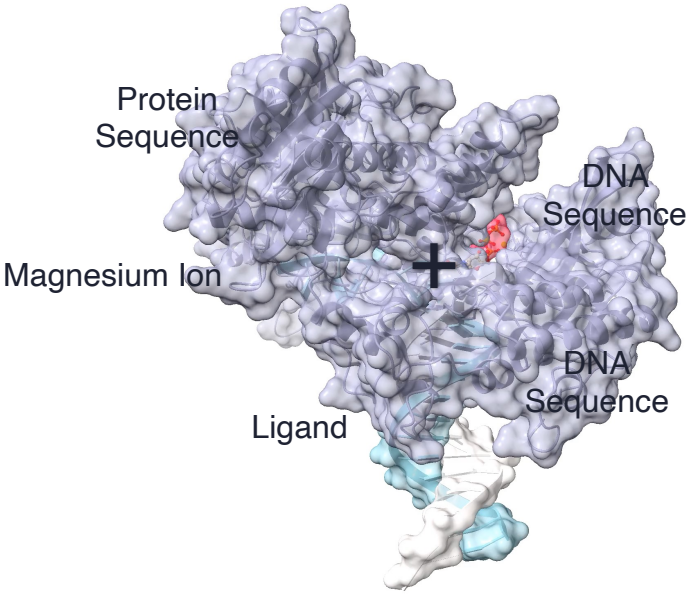
)OP(=O)(O)OCC
)NC1=O)CC3O

= ?

5

5

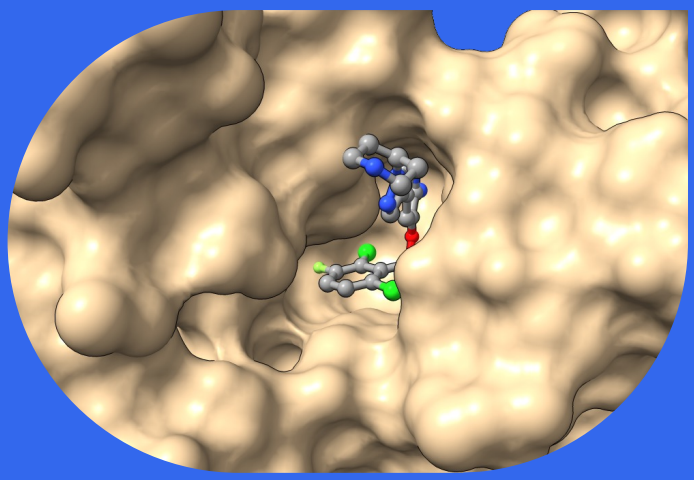
Biomolecular Complex Structure Prediction – Co-Folding



= ?

6

6

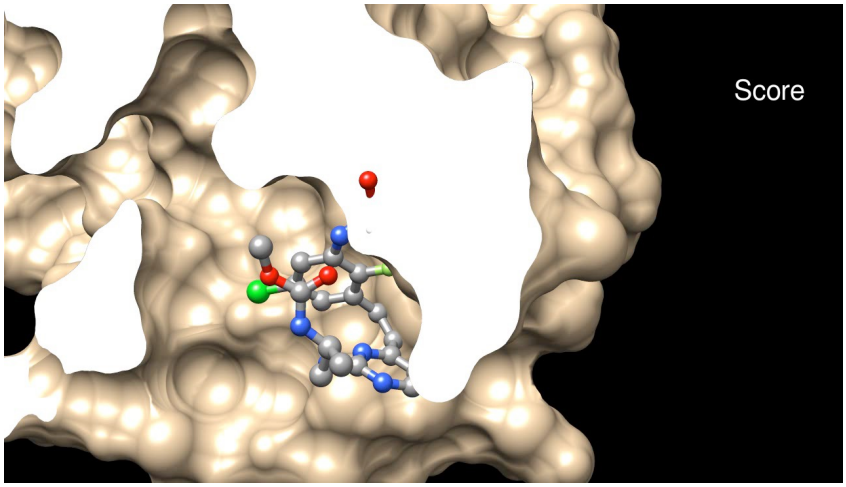


Evaluation of Docking Algorithms for Drug Design

7

Docking of a Small Molecule to a Protein Receptor

Objective: Predict the preferred binding mode of a small molecule to a receptor

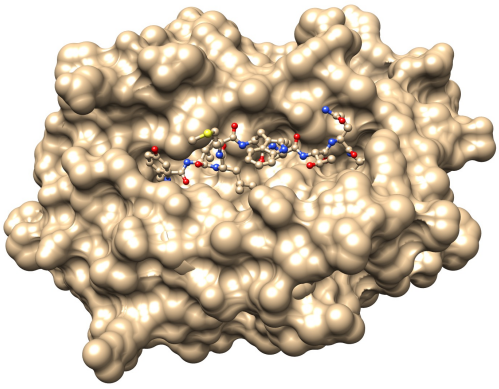


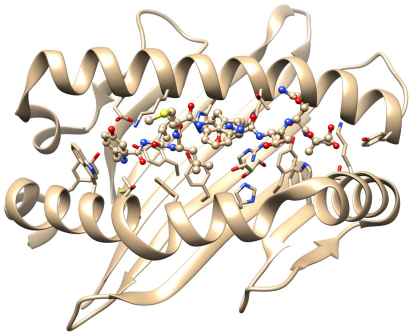
8

8

Peptide – Protein Docking

- Same or dedicated algorithms as for ligand-protein docking
- Many degrees of freedom, sampling problem
- Example: predict peptide/MHC complexes for immunology



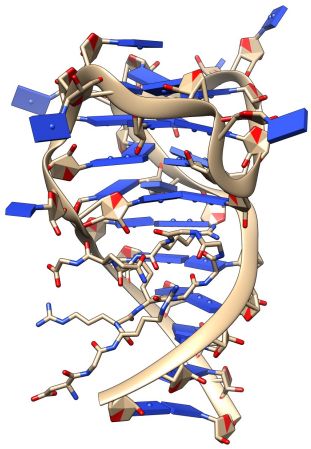


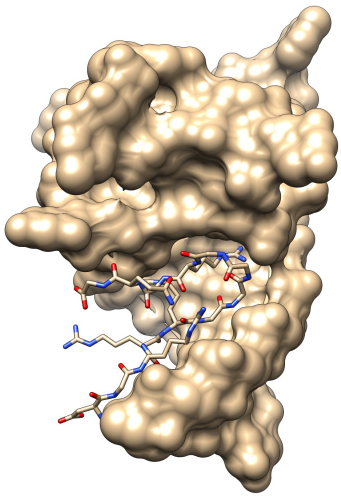
9

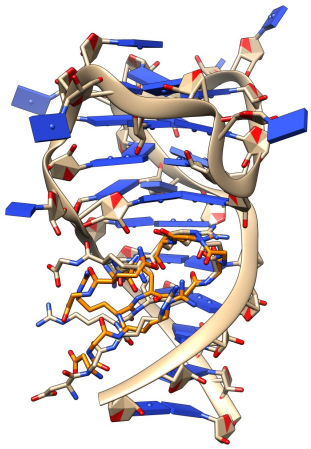
9

Peptide – DNA/RNA Docking

- Same algorithms as for ligand-protein docking (if scoring function applicable)
- Targets: DNA/RNA with specific tertiary structures





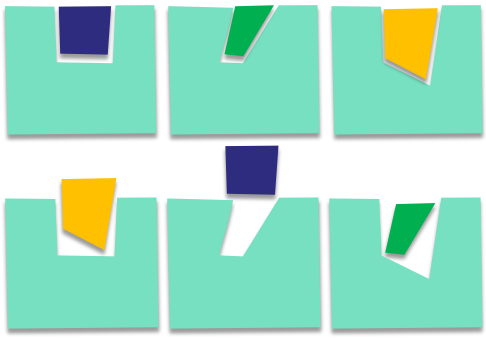


10

10

Docking - Definitions

- Re-docking:
  docking of a ligand into a structure of the macromolecule in complex with that same ligand; perfect induced fit; used for code evaluation
- Cross-docking:
  docking of a ligand into a structure of the macromolecule without or with a different ligand; induced fit absent or different
- Success:
  capacity to predict a binding mode close to the native binding mode if known (re-docking, benchmark)



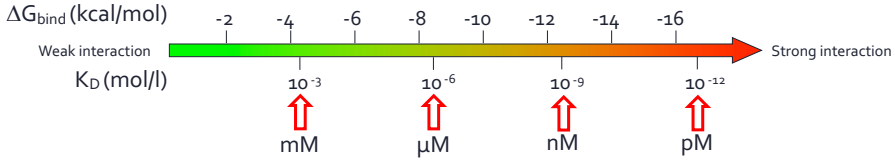
Calculation of Binding Affinities (Binding Free Energies)

$$K_D = \frac{[A]\,[B]}{[AB]}$$

K_D : dissociation constant (equilibrium constant)

$$\Delta G_{bind} = RT \ln(K_D) = \Delta H - T\Delta S$$

G: Gibbs free energy
R: universal gas constant
T: temperature
H: enthalpy
S: entropy



Docking Power Assessment

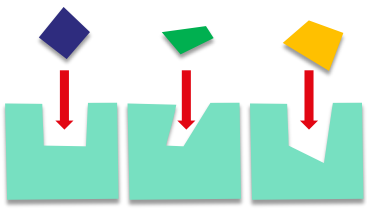
13

Docking Power

Docking power

identify the native ligand binding pose starting from a random ligand conformation

Re-Docking



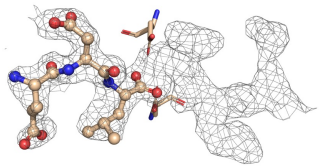
Challenge for Assessment Setup

assemble a set of diverse high-quality ligand-protein structures with reliable associated data ("benchmark set")

14 **"PDBbind Set"**: Su, M. et al., **Comparative Assessment of Scoring Functions: The CASF-2016 Update**, *J. Chem. Inf. Model.* **59**, 895 (2019)

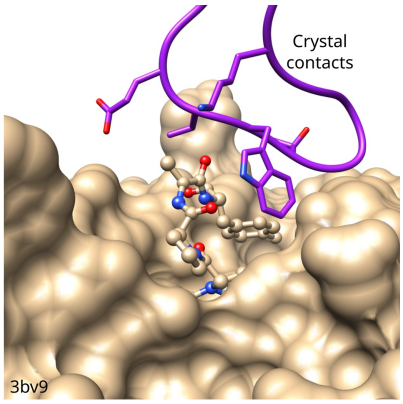
14

Issues in Structural Data

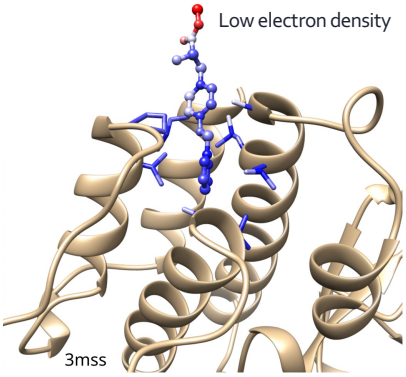


Extra electron density

1a30



3bv9

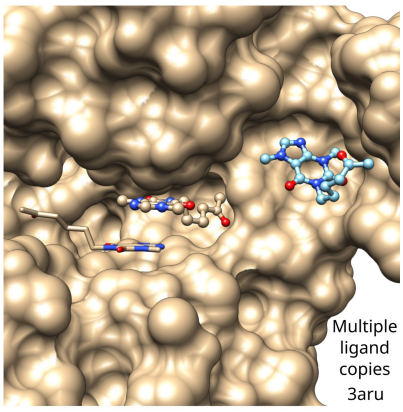


3mss

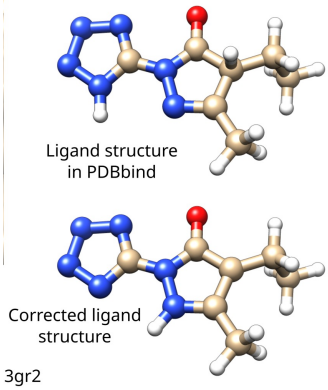
Su, M. et al., **Comparative Assessment of Scoring Functions: The CASF-2016 Update**, *J. Chem. Inf. Model.* **59**, 895 (2019)
U.F. Röhrig et al., **Attracting Cavities 2.0: Improving the Flexibility and Robustness for Small-Molecule Docking**, *J. Chem. Inf. Model.* **63**, 3925 (2023)

15

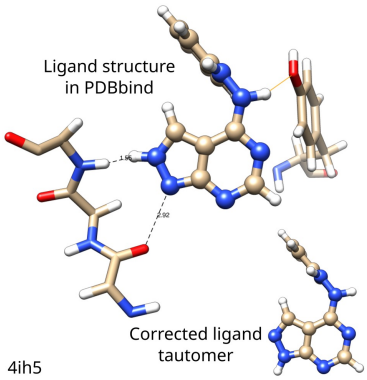
Issues in Structure Preparation



3aru



3gr2



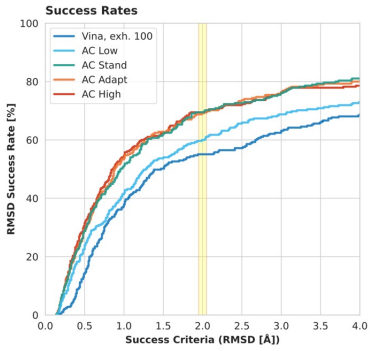
4ih5

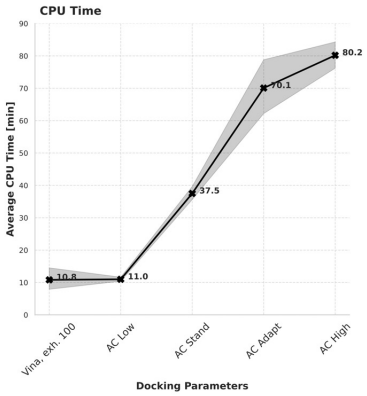
Su, M. et al., **Comparative Assessment of Scoring Functions: The CASF-2016 Update**, *J. Chem. Inf. Model.* **59**, 895 (2019)
U.F. Röhrig et al., **Attracting Cavities 2.0: Improving the Flexibility and Robustness for Small-Molecule Docking**, *J. Chem. Inf. Model.* **63**, 3925 (2023)

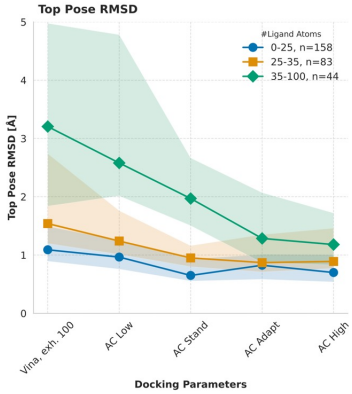
16

Re-Docking Evaluation (AC 3.0/AutoDock Vina)

Re-docking: docking a ligand into a protein structure that was co-crystallized with *the same* ligand







- AC success rate ~ 70% (top pose RMSD <2.0 Å, including problematic structural data)
- AC better than Vina at similar CPU time, improves further with more sampling
- Big ligands require more sampling than small ones

17

U.F. Röhrig, M. Mathieu-Bugnon, V. Zoete, **Attracting Cavities 3.0: Faster and More Versatile Molecular Docking for the SwissDock Webserver**, *Bioinformatics*, ASAP

17

Screening Power Assessment

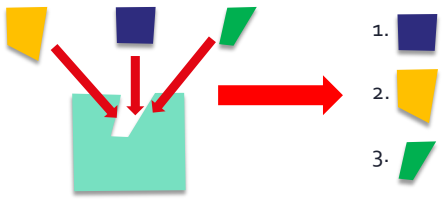
18

Screening Power

Screening power

identify the true binders to a target among a pool of molecules based on their docking-generated poses

Cross-Docking + Ranking



Challenge for Assessment Setup

assemble a set of diverse true ligands and non-binders ("decoys") for a diverse set of protein targets with reliable associated activity data

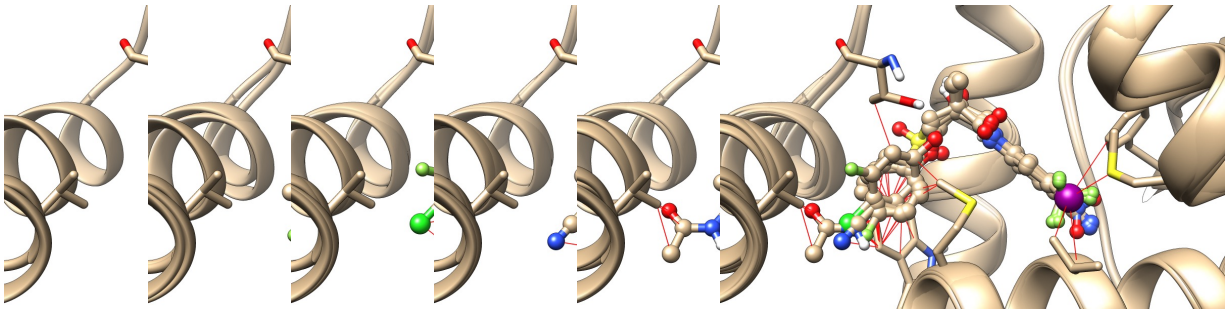
Su, M. et al., **Comparative Assessment of Scoring Functions: The CASF-2016 Update**, *J. Chem. Inf. Model.* **59**, 895 (2019)

Attracting Cavities 2.0: Cross-Docking

Cross-docking: docking a ligand into a protein structure that was co-crystallized with a *different* or *no* ligand

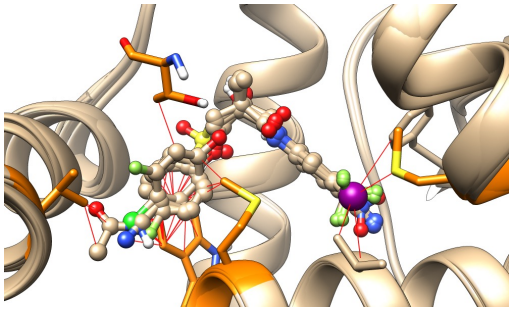
PDBbind : collection of 57 targets with 5 ligands each

Software	Success Rate	CPU Time
AC 2.0	43%	1.6 h
AutoDock Vina	33%	11 min



U.F. Röhrig, M. Goullieux, M. Bugnon, V. Zoete, *J. Chem. Inf. Model.*, **2023**, *63*, 3925-3940.

Attracting Cavities 2.0: Cross-Docking



5 selected targets from PDBbind, flexible amino acids chosen by analysis of clashes (orange)

Software	Succes rate	CPU Time
AC 2.0/fixed	25%	3.2 h
AC 2.0/flex	65%	4.3 h
AutoDock Vina	35%	12 min

U.F. Röhrig, M. Goullieux, M. Bugnon, V. Zoete, *J. Chem. Inf. Model.*, **2023**, *63*, 3925-3940.

Attracting Cavities 2.0: Ranking

Ranking: detect 5 true binders among the 285 docked ligands

Software	EF 1%	EF 5%	EF 10%
AC 2.0 + SwissParam	4.0	6.3	3.9
AutoDock Vina	8.0	6.9	3.5
Optimal Value	60	20	10

Enrichment Factor EF:

$$EF_\alpha = \frac{NTB_\alpha}{NTB_{total} \times \alpha}$$

NTB: number of true binders
α: 1%, 5%, 10%

Problem:

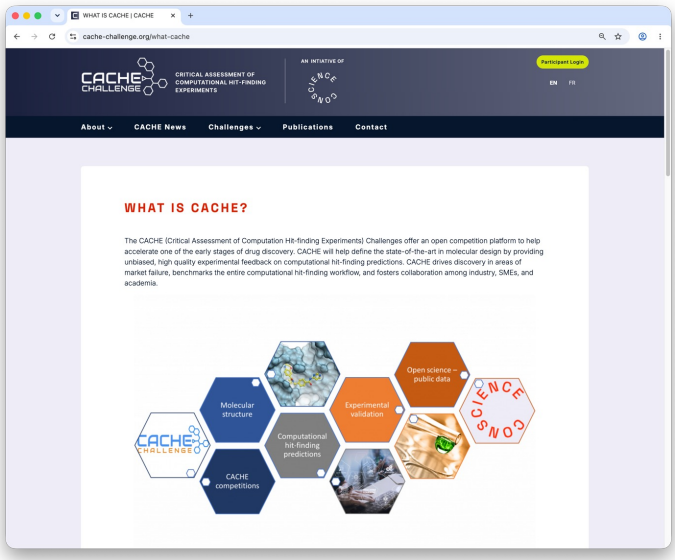
- SwissParam score was parameterized only on data from true binders, excluding decoy data
- Improvements possible
- **Predicted structures are more reliable than predicted activities**

U.F. Röhrig, M. Goullieux, M. Bugnon, V. Zoete, *J. Chem. Inf. Model.*, **2023**, *63*, 3925-3940.

Prospective Drug Design Applications

23

Blind Challenge Example: CACHE



https://cache-challenge.org/

24

24

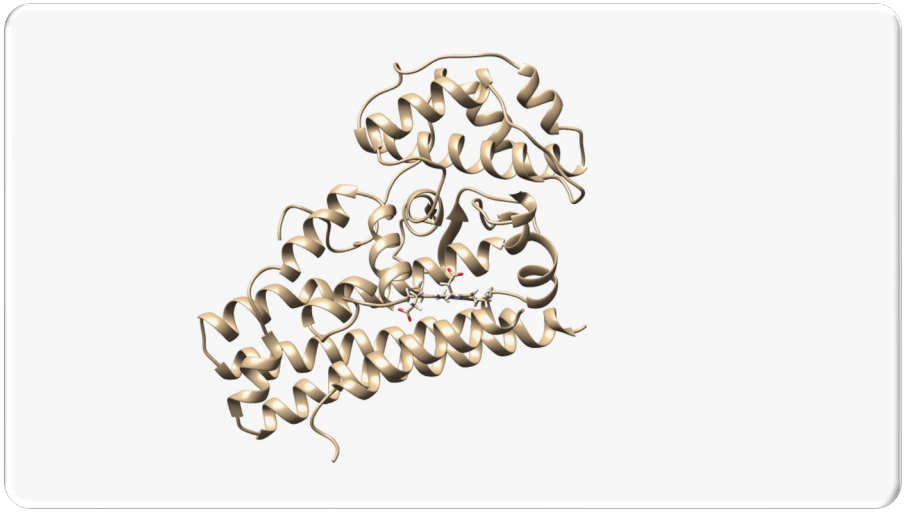
Docking-Guided Drug Design: Examples

- Shoichet Group, UCSF
- Olson Group, Scripps
- Abagyan Group, UCSD
- Gorgulla Group, St. Jude Research
- Katrich Group, USC
- Rarey Group, University of Hamburg
- Schneider Group, ETH Zurich
- Caflisch Group, UniZH

... many many more!

Docking studies without experimental validation are meaningless

Example of Fragment-Based Drug Design



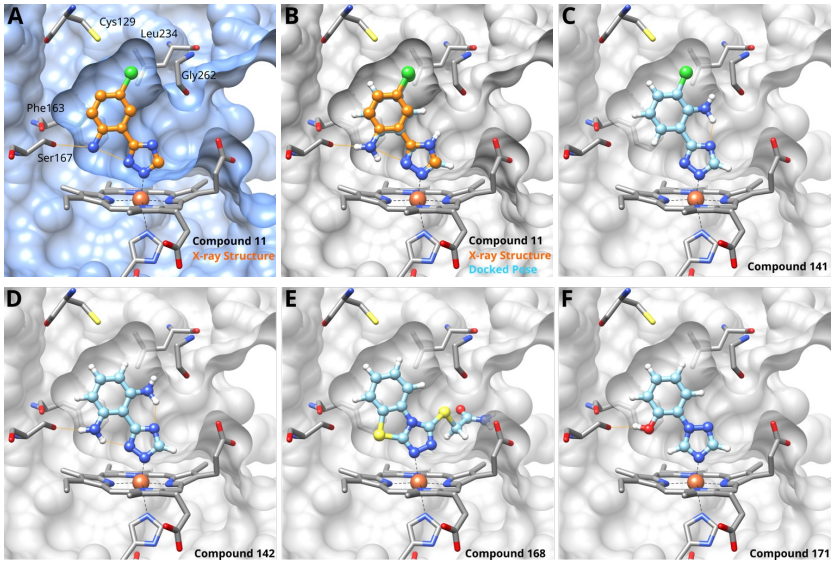
U.F. Röhrig et al., **Rational Design of 4-Aryl-1,2,3-Triazoles for Indoleamine 2,3-Dioxygenase 1 Inhibition**, *J. Med. Chem.* **55**, 5270 (2012)

Docking for Hit-to-Lead and Lead Optimization

Example IDO1

- Determined X-ray structure of a scaffold (**A**)
- Docking yielded good results for co-crystallized ligand (**B**)
- Small/moderate modifications of the scaffold (**C-F**)

- Docking-guided design of new compounds
- Small, nanomolar, highly efficient inhibitors



U.F. Röhrig et al. *J. Enz. Inh. Med. Chem.* **37**, 1773 (2022)
U.F. Röhrig et al. *J. Med. Chem.* **62**, 8784 (2019)

27

Summary: Docking Predictions

Structure Prediction

- Success rates reported for re-docking (perfect induced fit): ~50% to 90%
- Strong dependence on
  - Quality of structural data
  - Structure preparation
  - Size/flexibility of the ligands
  - Experience with a docking code
- Expected success rate for real applications (cross-docking): ~20-30% less than in re-docking

Affinity Prediction

- Estimation of ΔG_{bind} very challenging from static docking poses

Real Utility: Prospective Drug Design Applications

28

Docking + Deep-Learning

29

DL Structure Models as Docking Targets

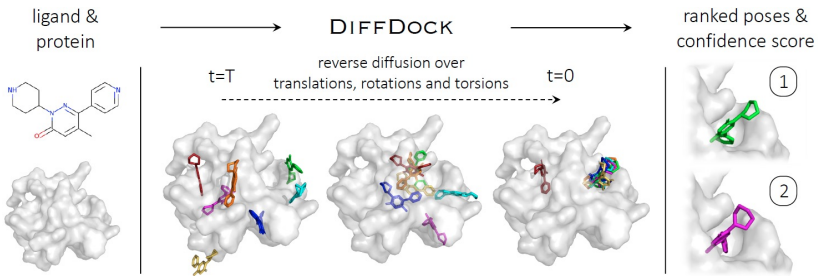
- **How accurately can one predict drug binding modes using AlphaFold models?** M. Karelina, J.J. Noh, R.O. Dror, *eLife* 12:RP89386,2023
- **Are Deep Learning Structural Models Sufficiently Accurate for Virtual Screening? Application of Docking Algorithms to AlphaFold2 Predicted Structures**, A.M. Díaz-Rovira, H. Martín, T. Beuming, L. Díaz, V. Guallar, S.S. Ray, *J. Chem. Inf. Model.* 2023, 63, 1668
- **AlphaFold2 structures guide prospective ligand discovery**, J. Lyu et al., *Science* 2024

Careful evaluation and refinement before large-scale docking

30

30

Deep-Learning Docking Algorithms



Other examples: EquiBind, TankBind

Corso, G., Stark, H., Jing, B. ,Barzilay, R., Jaakkola, T. **DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking** *ICLR* (2023).

Deep-Learning Docking Evaluation

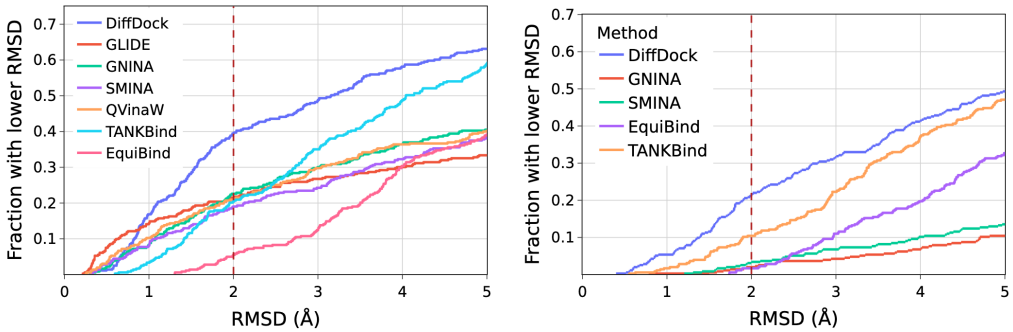
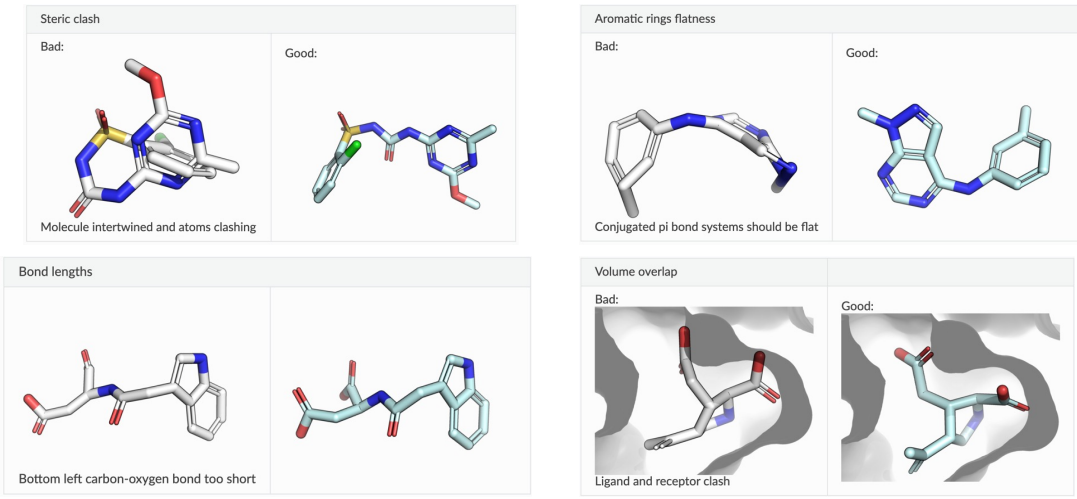


Figure 6: Cumulative density histogram of the methods' RMSD: left on holo crystal structures, right on apo ESMFold structures.

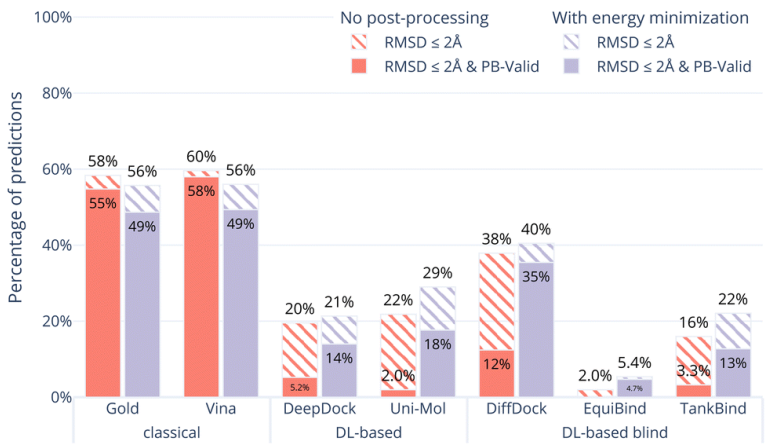
Physical Validity of Predicted Structures: PoseBusters



Buttenschoen, Morris, Deane, **PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences,** *Chem. Sci.*, **15**, 3130 (2024)
https://github.com/maabuu/posebusters

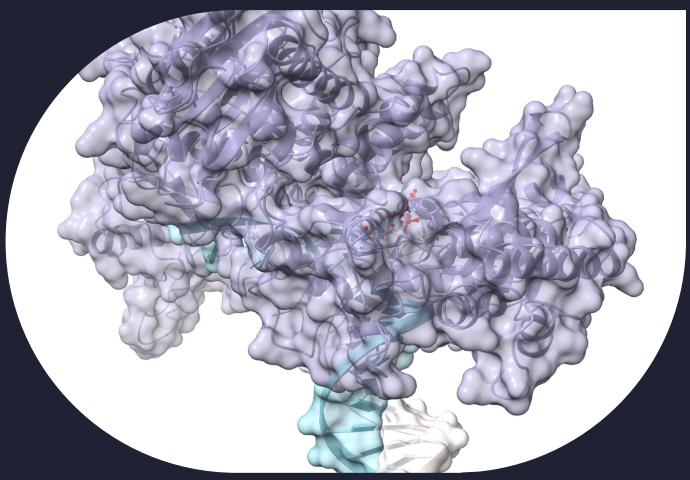
33

Physical Validity of Predicted Structures: PoseBusters



Buttenschoen, Morris, Deane, **PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences,** *Chem. Sci.*, **15**, 3130 (2024)
https://github.com/maabuu/posebusters

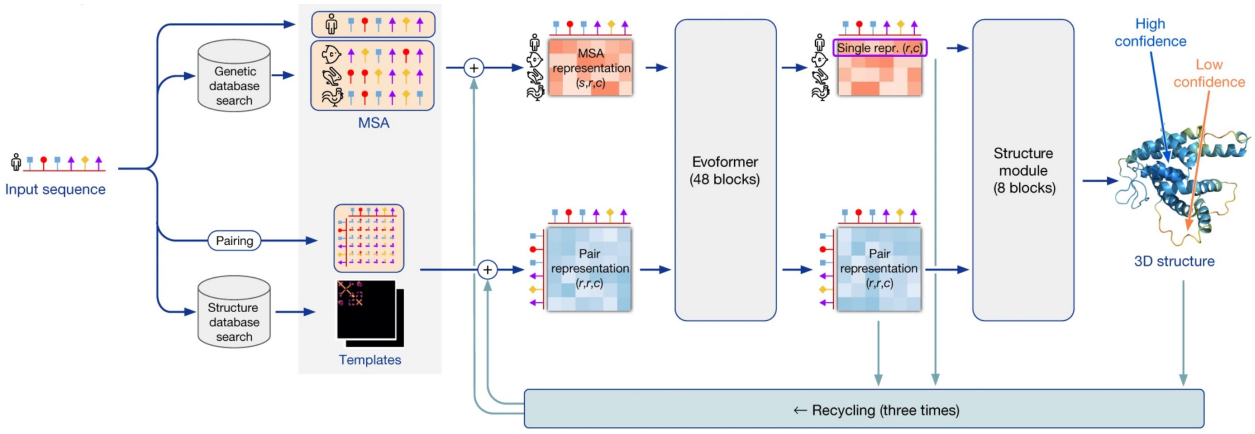
34



Structure Prediction by Co-Folding

35

AlphaFold 2: Protein Structure Prediction



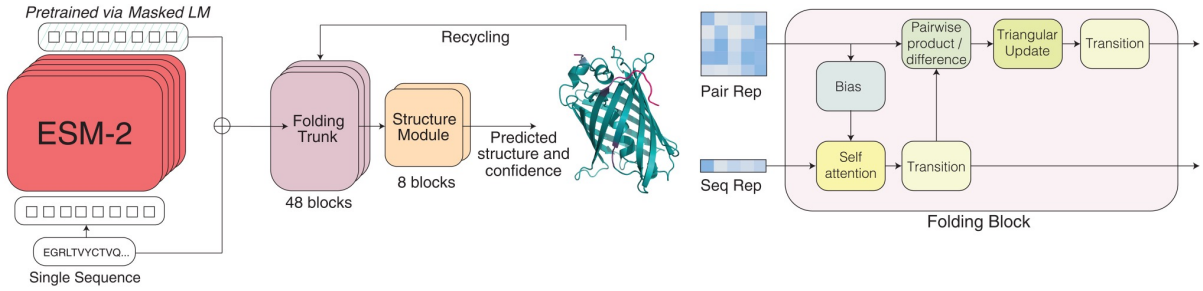
- Google DeepMind
- Multiple sequence alignment (MSA): co-evolving amino acids should be in contact
- MSA calculation on CPU, modeling on GPU
- AlphaFold Database: 214 million predicted structures

36

Jumper, J., Evans, R., Pritzel, A. *et al.* **Highly accurate protein structure prediction with AlphaFold**. *Nature* **596**, 583–589 (2021).

36

ESMFold

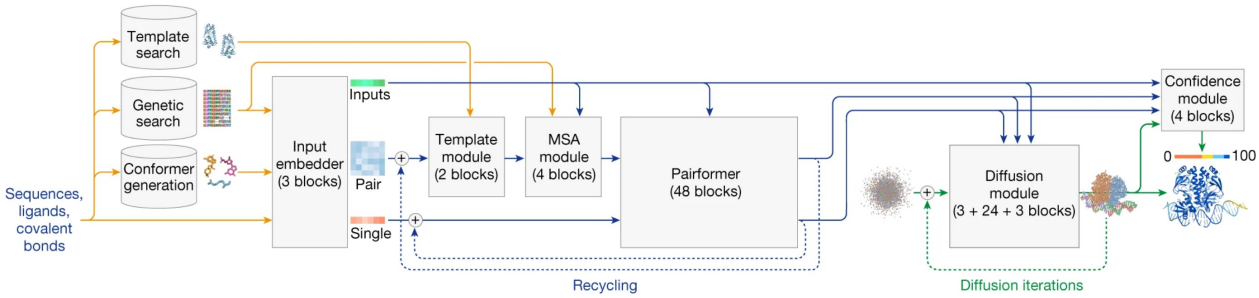


- ESM: Evolutionary Scale Modeling (Meta)
- Protein Language Model, trained on sequences and on AlphaFold predicted structures
- Faster than AlphaFold because does not require MSA (optional)
- ESM Metagenomic Atlas (772 million predicted structures)

Lin, Z., et al., **Evolutionary-scale prediction of atomic-level protein structure with a language model**, *Science*, 379, 1123 (2023)

AlphaFold 3: Biomolecular Structure Prediction

Can predict structures of multiple proteins, DNA, RNA, post-translational modifications, co-factors, and ligands



[38]Abramson, J., Adler, J., Dunger, J. *et al.* **Accurate structure prediction of biomolecular interactions with AlphaFold 3**. *Nature* **630**, 493–500 (2024).

Co-Folding Algorithms

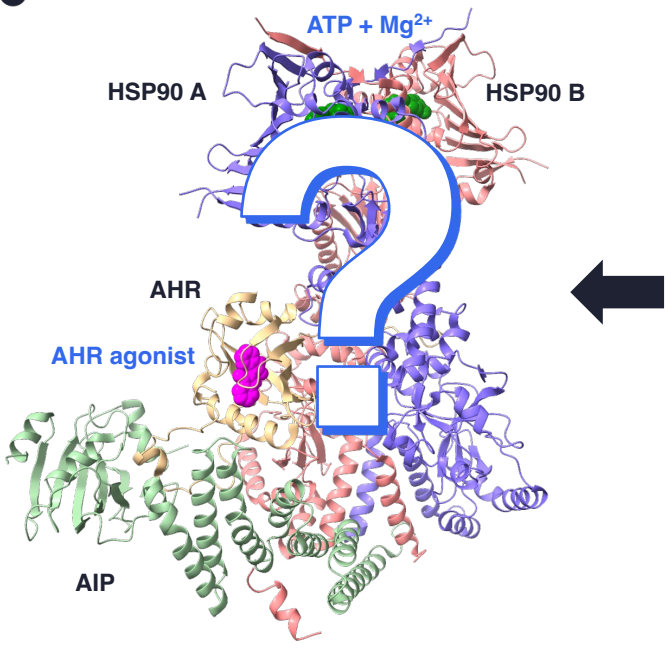
- AlphaFold 3
  - Protenix
  - Chai-1
  - Boltz-1, Boltz-2, BoltzGen
  - OpenFold
- Umol
- RoseTTaFold All-Atom
- ESM Fold 2

Different serves/websites/notebooks to test/use these algorithms

Abramson, J., Adler, J., Dunger, J. et al. **Accurate structure prediction of biomolecular interactions with AlphaFold 3**. *Nature* (2024)
Chen, X., et al. **Protenix—advancing structure prediction through a comprehensive AlphaFold3 reproduction**. bioRxiv (2025)
J. Boitreaud, et al. **Chai-1: Decoding the molecular interactions of life**. BioRxiv (2024)
J. Wohlwend, et al. **Boltz-1 democratizing biomolecular interaction modeling**. *BioRxiv* (2025)
Ahdritz, G. *et al.* **OpenFold: retraining AlphaFold2 yields new insights into its learning mechanisms and capacity for generalization**. *Nat Methods* (2024)
Bryant, P. et al,. **Structure prediction of protein-ligand complexes from sequence information with Umol**. *Nat Commun* (2024)
Krishna, R., et al., **Generalized biomolecular modeling and design with RoseTTAFold All-Atom**. *Science*, (2024)
S. Candido *et al.*, **Language Modeling Materializes a World Model of Protein Biology,** bioRxiv 2026.06.03.729735

39

Structure Prediction

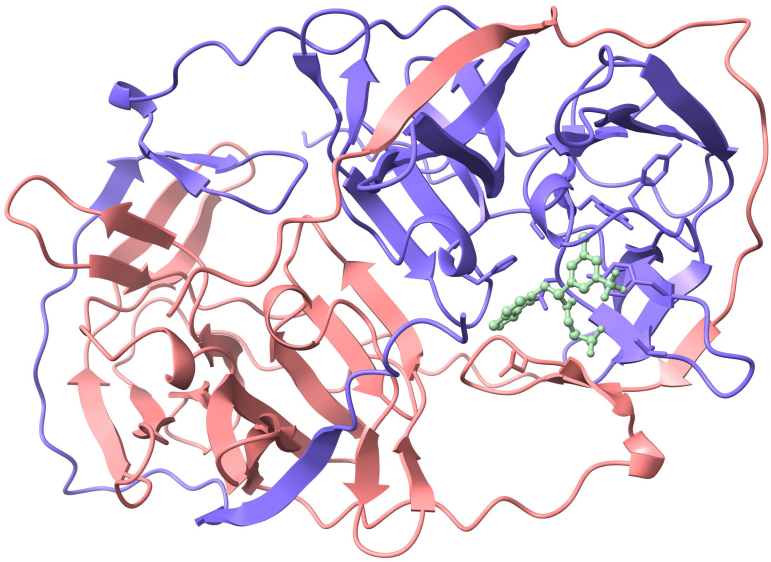



```
{
  "name": "AHR Cytosolic Complex",
  "modelSeeds": [375202786, 131558723],
  "sequences": [
    {"protein": {
        "id": [ "A", "B" ],
        "sequence": "MPEEVHHGEEEVETF…"
      }
    },
    {"protein": {
        "id": [ "C" ],
        "sequence": "MADIIARLREDGIQKRVIQEGR…"
      }
    },
    {"protein": {
        "id": [ "D" ],
        "sequence": "MNSSSANITYASR…"
      }
    },
    {"ligand": {
       "id": ["G", "H"],
       "ccdCodes": ["ATP"]
     }
    },
    {"ligand": {
       "id": ["I", "J"],
       "ccdCodes": ["MG"]
     }
    },
    {"ligand": {
        "id": ["K"],
        "smiles": "c1ccc2c(c1)/C(=C/3\\\\C(=O)c4ccccc4N3)/C(=O)N2"
      }
    }
  ]
}
```


2 * HSP90
AHR
AIP
2 * ATP
2 * Mg^{2+}
AHR agonist

40

Structure Prediction Confidence



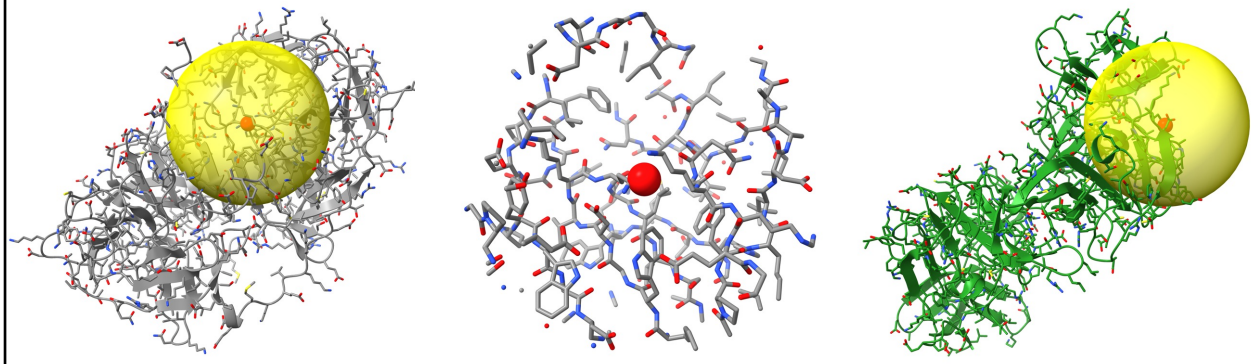
Example

Zika virus NS2B-NS3 protease with inhibitor: 2 protein chains, 1 ligand

Structure Prediction Confidence Measures: pLDDT

LDDT (Local Distance Difference Test, [0, 100]):

Comparison of two structures. A value of 100 means perfect agreement.



V. Mariani, M. Biasini, A. Barbato, T. Schwede, **lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests**, *Bioinformatics* 29, 2722 (2013)

Structure Prediction Confidence Measures: pLDDT

LDDT (Local Distance Difference Test, [0, 100]):

Comparison of two structures. A value of 100 means perfect agreement. LDDT computes the fraction of atomic distance differences inside an inclusion radius of 15 Å below a threshold *d*. The reported value is the average over 4 fractions computed with thresholds [0.5, 1.0, 2.0, 4.0] Å.

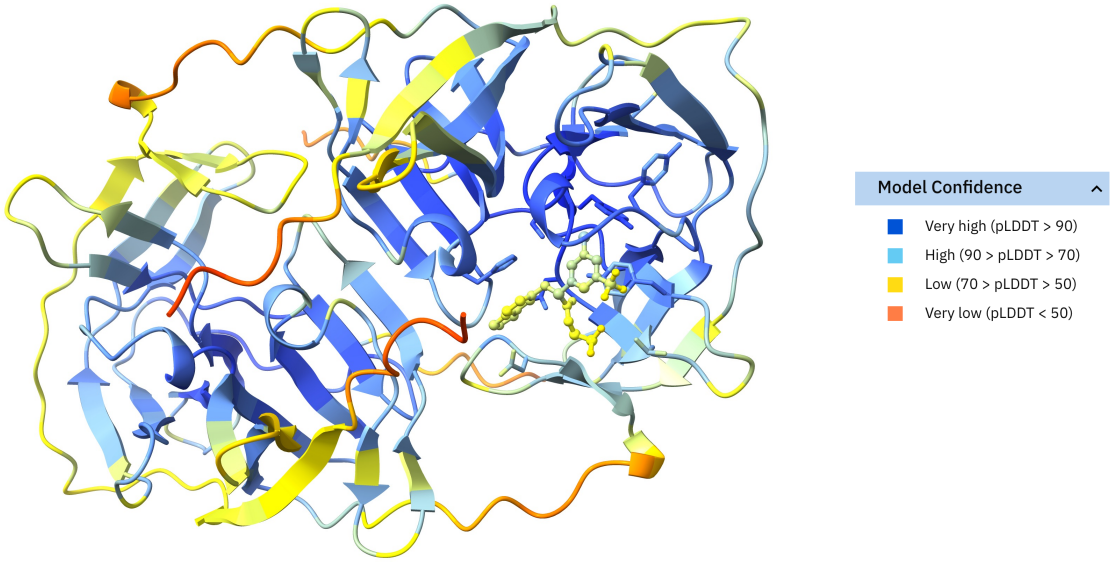
pLDDT (predicted Local Distance Difference Test, [0, 100]):

Reflects local confidence for each residue. Stored in the B-factor field of the mmCIF files. pLDDT is used to color-code the confidence in each atom of the model:

- Very high (pLDDT > 90)
- High (90 > pLDDT > 70)
- Low (70 > pLDDT > 50)
- Very low (pLDDT < 50)

V. Mariani, M. Biasini, A. Barbato, T. Schwede, **lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests**, *Bioinformatics* 29, 2722 (2013)

Structure Prediction Confidence: Example



Zika virus NS2B-NS3 protease with inhibitor: 2 protein chains, 1 ligand

Structure Prediction Confidence Measures: TM Scores

Template Modelling (TM) Score (0,1]: Comparison of two structures; measures the accuracy of the global structure of a protein and is relatively insensitive to localized inaccuracies; 1 indicates a perfect match, >0.5 assumes roughly the same fold, while <0.2 corresponds to random unrelated folds.

predicted Template Modelling (pTM) Score (0,1]: predicted TM score for a superposition between the predicted structure and the *hypothetical* true structure. A pTM score above 0.5 means the overall predicted fold for the complex might be similar to the true structure. A TM score below 0.5 means the predicted structure is likely wrong.

interface predicted Template Modelling (ipTM) Score (0,1]: measures the quality of the predicted interaction between subunits. ipTM reflects confidence in the predicted relative positions of subunits in a complex. An ipTM score > 0.8 indicates a confident, high-quality prediction; ipTM 0.6-0.8 suggests the prediction is uncertain; and ipTM < 0.6 suggests the complex may be incorrect.

45 Zhang and Skolnick, **Scoring function for automated assessment of protein structure template quality,** *Proteins*. **57**, 702 (2004)

45

Structure Prediction Confidence Measures: Ranking Score

Ranking Score (-100, 1.5):

AlphaFold 3 samples five predictions per random seed. The model with the highest ranking score is returned by default.

ranking_score = 0.8 × *ipTM* + 0.2 × *pTM* + 0.5 × *fraction_disordered* − 100 × *has_clash*

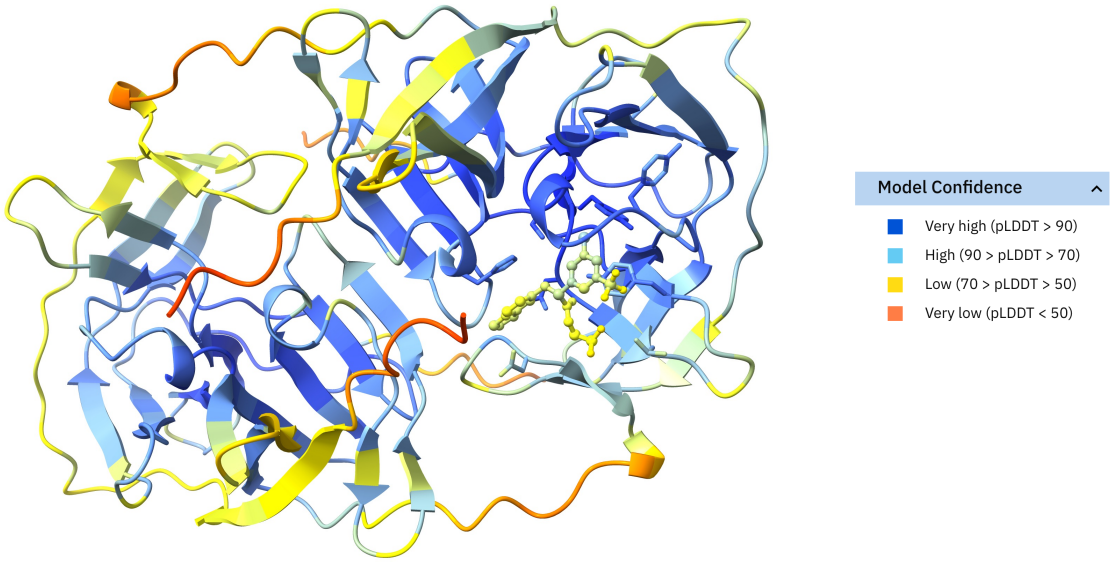
fraction_disordered: A scalar in the range 0-1 that indicates what fraction of the prediction structure is disordered, as measured by accessible surface area

has_clash: A Boolean, i.e. a yes/no value, indicating if the structure has a significant number of clashing atoms (more than 50% of a chain, or a chain with more than 100 clashing atoms).

46

46

Structure Prediction Confidence: Example



Zika virus NS2B-NS3 protease with inhibitor: 2 protein chains, 1 ligand

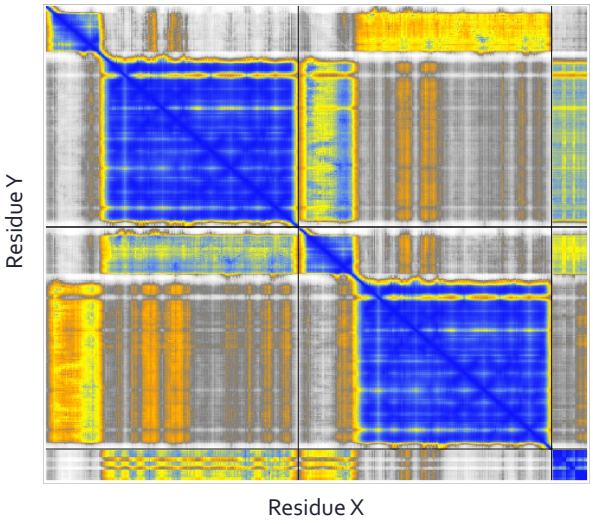
pTM 0.59; ipTM 0.69; fraction disordered 0.2, has_clash 0.0; ranking score 0.77

47

Structure Prediction Confidence Measures: PAE

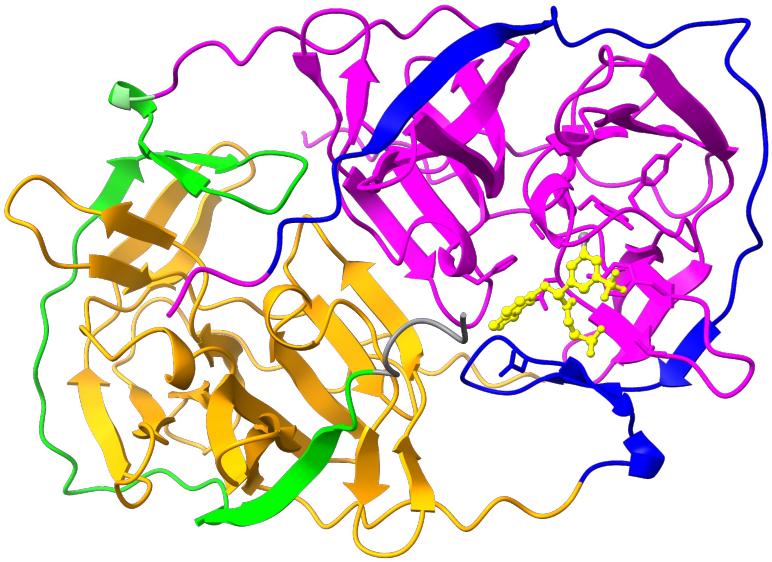
Predicted Aligned Error (PAE):

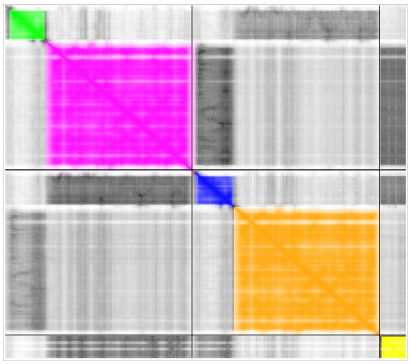
- Measure of how confident the model is in the *relative position* of *two residues*
- PAE [Å] is the expected positional error at residue X, if the predicted and true structures were aligned on residue Y
- Effectively a measure of how confident the model is that the domains/chains are well packed and that the relative placement of the domains/chains is correct
- Blue: small error (< 5 Å)
- Yellow/orange: large error (5–15 Å)
- Grey/white: very large error (> 20 Å)



48

Structure Prediction Confidence: Example



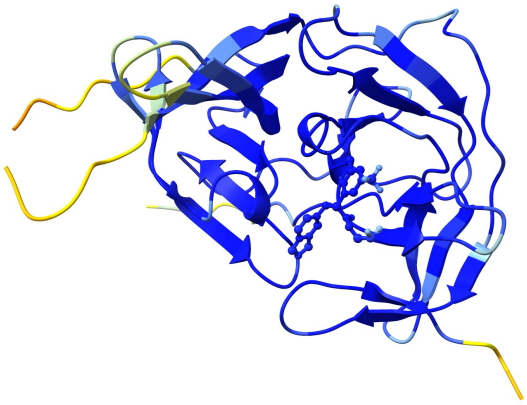


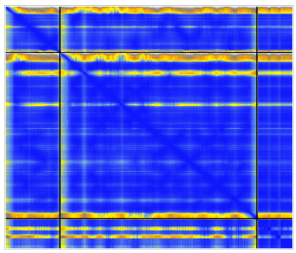
49

Zika virus NS2B-NS3 protease with inhibitor: 2 protein chains, 1 ligand
pTM 0.59; ipTM 0.69; fraction disordered 0.2, has_clash 0.0; ranking score 0.77

49

Structure Prediction Confidence: Example

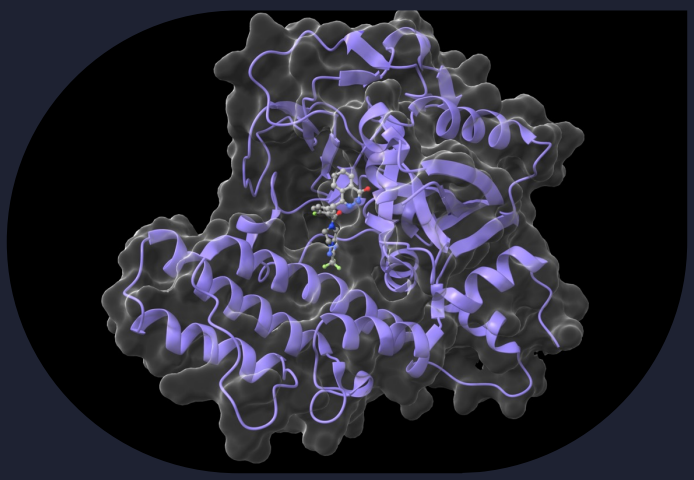




50

Zika virus NS2B-NS3 protease with inhibitor (PDB ID 9rm7): 1 cleaved protein chain, 1 ligand
pTM 0.94; ipTM 0.92; fraction disordered 0.24, has_clash 0.0; ranking score 1.05

50



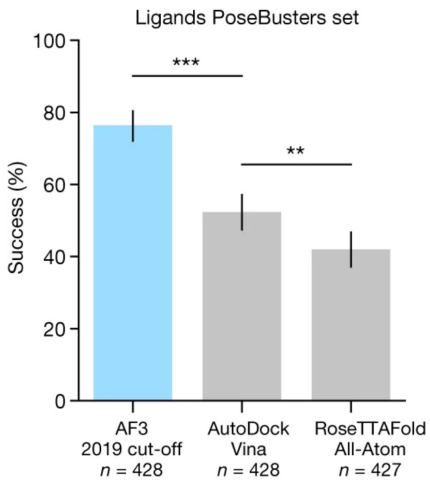
Evaluation of Co-Folding Methods for Drug Design

51

Co-Folding Performance Evaluation

"Time split": using data published after the model training deadline for model evaluation

Problem: Many new structures in the PDB are very similar to already known structures ("data leakage")



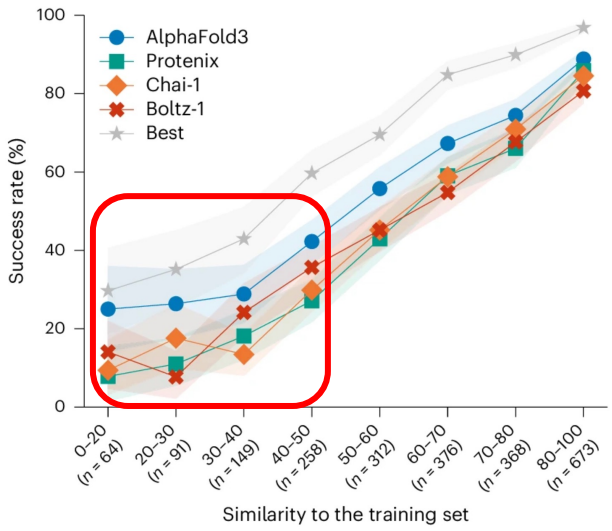
52 Abramson, J., Adler, J., Dunger, J. et al. **Accurate structure prediction of biomolecular interactions with AlphaFold 3**. *Nature* (2024)

52

Co-Folding Performance Evaluation

Similarity to data in training set:
Active site similarity * ligand similarity

Area of interest
for drug design



Škrinjar, P., Eberhardt, J., Studer, G. *et al.* **Evaluating generalization in protein–ligand cofolding methods**. *Nat Struct Mol Biol* (2026)

Runs N' Poses Benchmark Set

- 2,600 PDB protein-ligand complexes published after model training cutoff date
- Similarity = active site similarity * ligand similarity
- Open access:
  - Annotations
  - Predictions
  - Input files
  - Analysis scripts

Škrinjar, P., Eberhardt, J., Studer, G. *et al.* **Evaluating generalization in protein–ligand cofolding methods**. *Nat Struct Mol Biol* (2026)

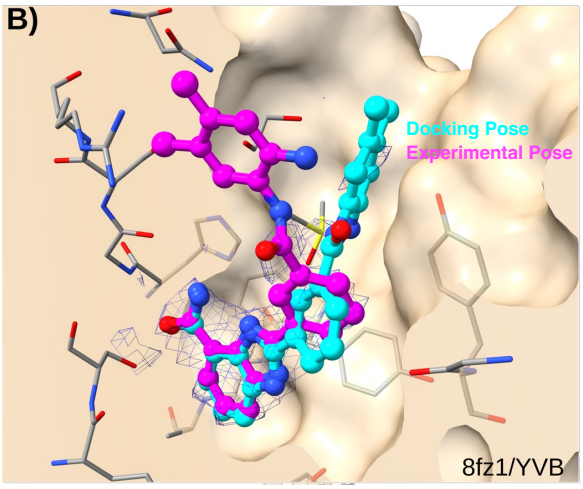
Runs N' Poses Benchmark Set: Evaluation for Drug Design

Problems :

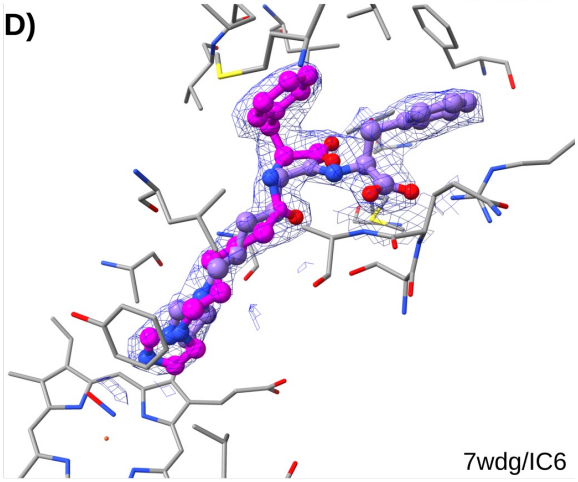
- 19% of ligands are not well-defined by electron density
- 24% of ligands are phosphates
- Many ligands are co-factors, not synthetic molecules
- 26% of complexes belong to the top 10 structural clusters

U.F. Röhrig, M. Mathieu-Bugnon, V. Zoete, **Comparative Assessment of the Utility of Co-Folding and Docking for Small-Molecule Drug Design**, bioRxiv (2025)

PDB Structures: no "Ground Truth" but Noisy Data

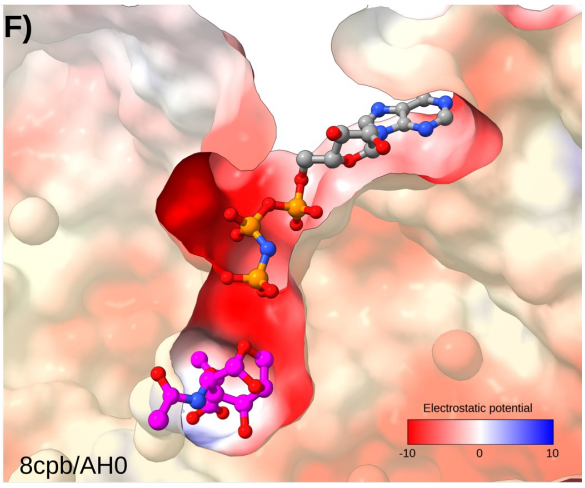


Ambiguous data



Mutually excluding data

PDB Structures: no "Ground Truth" but Noisy Data



Unphysical data
Binding site has strongly negative electrostatic potential
Co-factor carries a negative charge of -4
Ligand carries a negative charge of -1

Missing cations!

Runs N' Poses Benchmark Set: Evaluation for Drug Design

Problems :

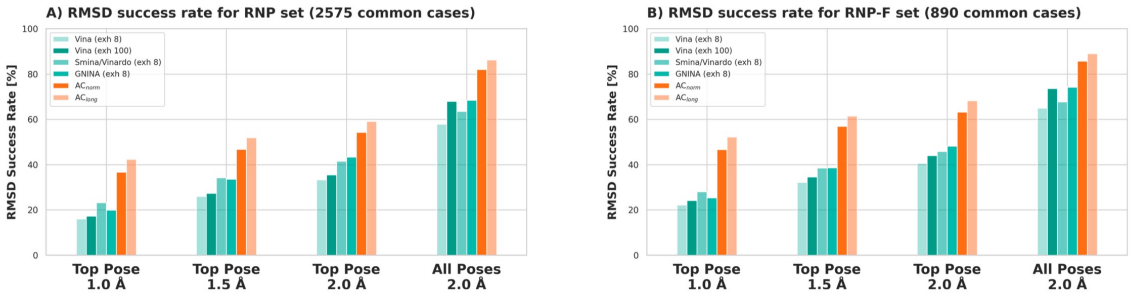
- 19% of ligands are not well-defined by electron density
- 24% of ligands are phosphates
- Many ligands are co-factors, not synthetic molecules
- 26% of complexes belong to the top 10 structural clusters

Filtering (RNP-F, 859 complexes)

- Ligand well defined by electron density
- Maximally 100 similar ligands in the training set (removes many cofactors and phosphates)
- Maximally 5 complexes of the same structural cluster

U.F. Röhrig, M. Mathieu-Bugnon, V. Zoete, **Comparative Assessment of the Utility of Co-Folding and Docking for Small-Molecule Drug Design**, bioRxiv (2025)

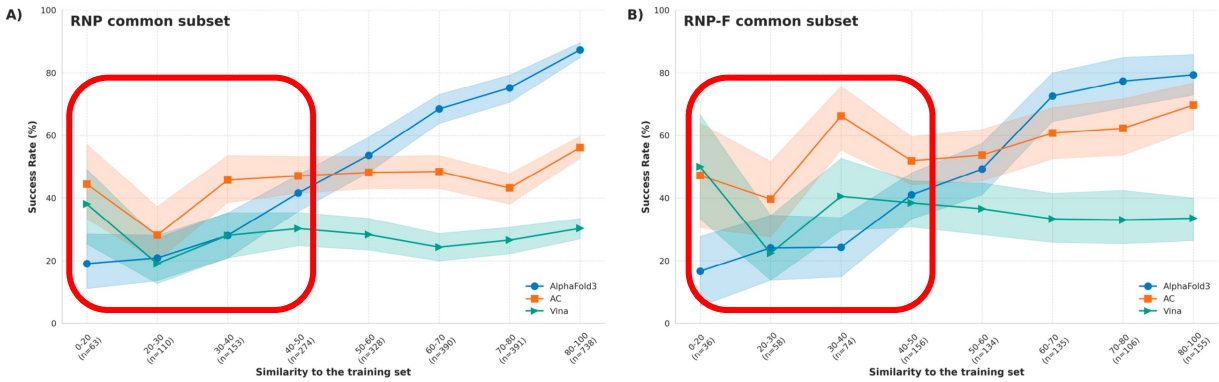
Re-Docking Results (AC 3.0/AutoDock Vina)



- AC outperforms all Vina-based methods (but requires more CPU time)
- Success rates are higher for the filtered set
- Overall success rates are low (problematic data and large ligands)

U.F. Röhrig, M. Mathieu-Bugnon, V. Zoete, **Comparative Assessment of the Utility of Co-Folding and Docking for Small-Molecule Drug Design**, bioRxiv (2025)

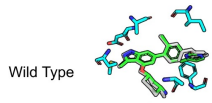
Comparison of Co-Folding et Re-Docking



A better comparison would be Co-Folding vs. Cross-Docking!

U.F. Röhrig, M. Mathieu-Bugnon, V. Zoete, **Comparative Assessment of the Utility of Co-Folding and Docking for Small-Molecule Drug Design**, bioRxiv (2025)

Physical Validity of Predicted Interactions



Grey: experimental binding mode
Green: AlphaFold 3 predicted structure

Contact residues mutated to alanine

Contact residues mutated to phenylalanine

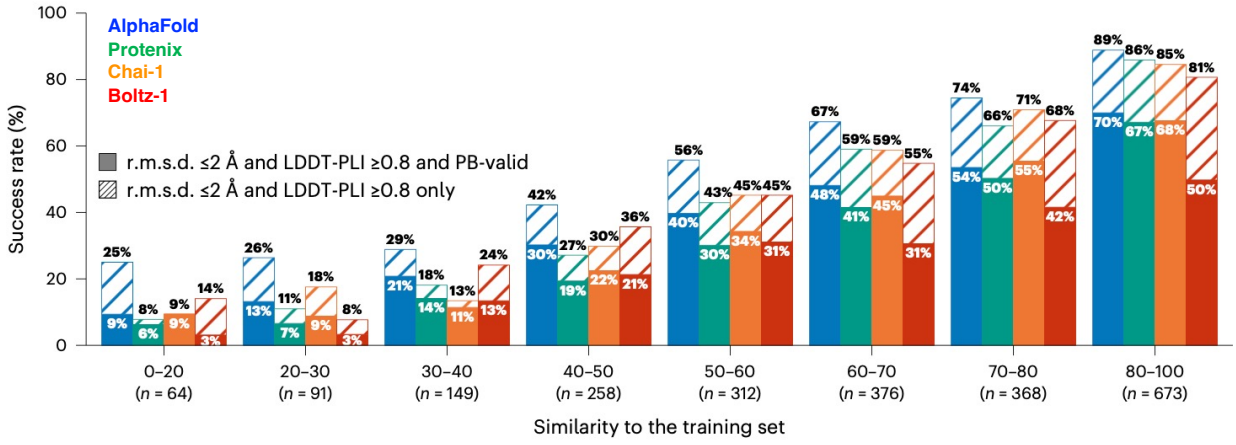
Contact residues mutated to most dissimilar amino acid

61

Masters, M.R., Mahmoud, A.H. & Lill, M.A. **Investigating whether deep learning models for co-folding learn the physics of protein-ligand interactions.** *Nat Commun* **16**, 8854 (2025)

61

Physical Validity of Predicted Structures: PoseBusters



62

Škrinjar, P., Eberhardt, J., Studer, G. *et al.* **Evaluating generalization in protein–ligand cofolding methods**. *Nat Struct Mol Biol* (2026)

62

Structure-Prediction Terminology

- **Avoid "physics-based" versus "deep learning"!**
- Docking scoring functions are not purely physics-based but to a greater or lesser extent fitted to reproduce experimental data (machine-learning elements)
- Also avoid "docking" versus "deep learning" (there are DL docking methods)
- Better terminology: mechanistic vs. data-driven (or end-to-end learned)
- Difference:
  - Traditional methods **deduce** the best binding mode by applying known rules of physics
  - DL methods **induce** rules about binding from observed examples, without being told what those rules are

Mechanistic vs. Data-Driven Docking

Aspect	Mechanistic (Traditional)	Data-Driven (Deep Learning)
Knowledge source	Physical laws & empirical parameters	Statistical patterns in structural data
Scoring function	Pre-specified (force-field, empirical, knowledge-based)	Learned end-to-end from training data
Pose generation	Explicit conformational search	Direct prediction or generative sampling (diffusion, GNN)
Search & scoring	Separate steps	Unified in one model
Interpretability	High (energy terms)	Low (black box)
Protein flexibility	Limited (rigid receptor typical)	Implicitly captured via training
Failures	Rare or difficult to describe interactions (transition metals)	Out-of-distribution data

Summary: Structure Prediction by Co-Folding

- Powerful tools for structure prediction; very rapidly evolving field
- Works well for predicting complexes of compounds similar to those encountered during training
- Can fail on new complexes , e.g. new binding sites and new ligands
- Does not learn the physical principles of interactions
- Does not adequately account for the effect of active site mutations
- Should be used in combination with other mechanistic methods for drug design

65

66