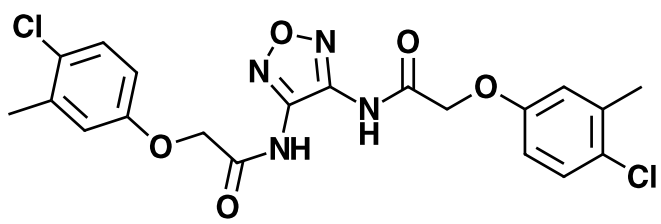
# A Pose Prediction Approach based on Ligand 3D Shape Similarity: Lessons learned in D3R challenge

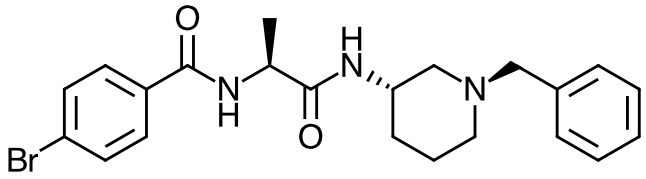
Ashutosh Kumar

Research Scientist
Structural Bioinformatics Team,
Division of Structural and Synthetic Biology,
Centre for Life Science Technologies
RIKEN, Yokohama, Japan

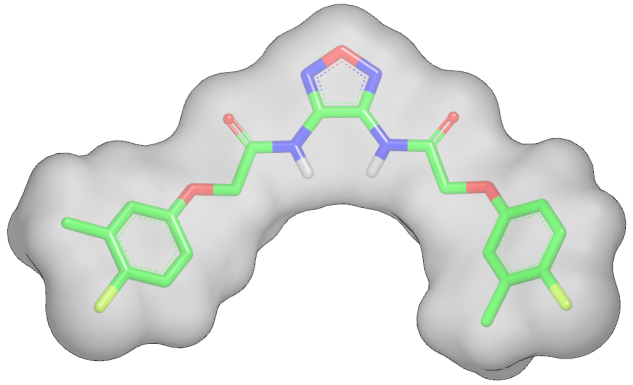
# Ligand 3D Shape Similarity

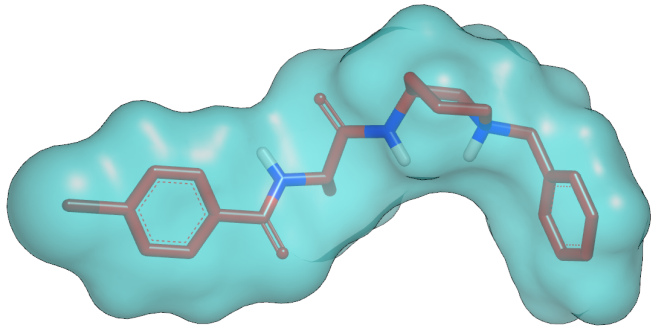


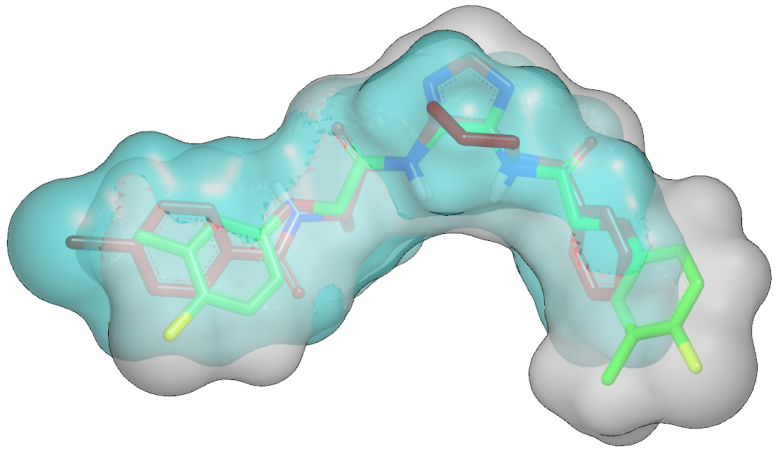

Ligand 1




Ligand 2



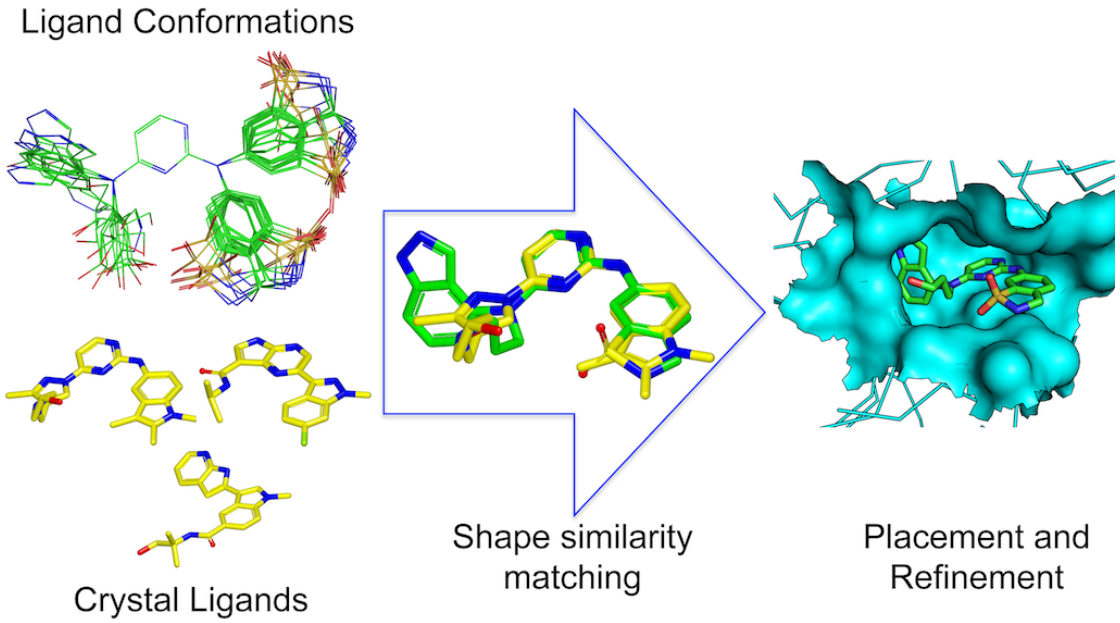


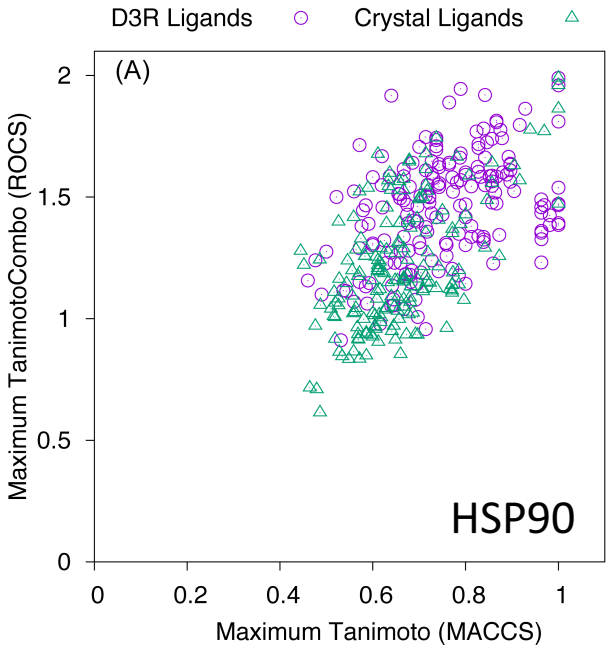


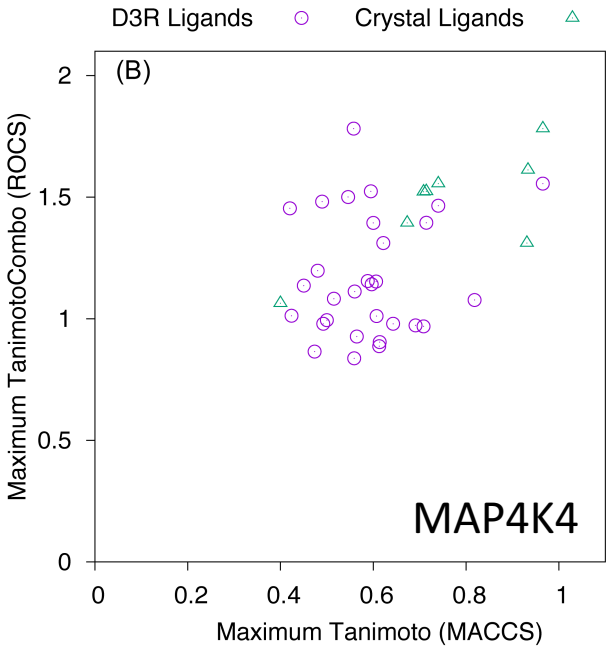
- Ligand 3D shape similarity has been routinely used as sole virtual screening approach or in combination with other structure based methods such as molecular docking.
- Other utilities in drug discovery include use as ligand scoring approach

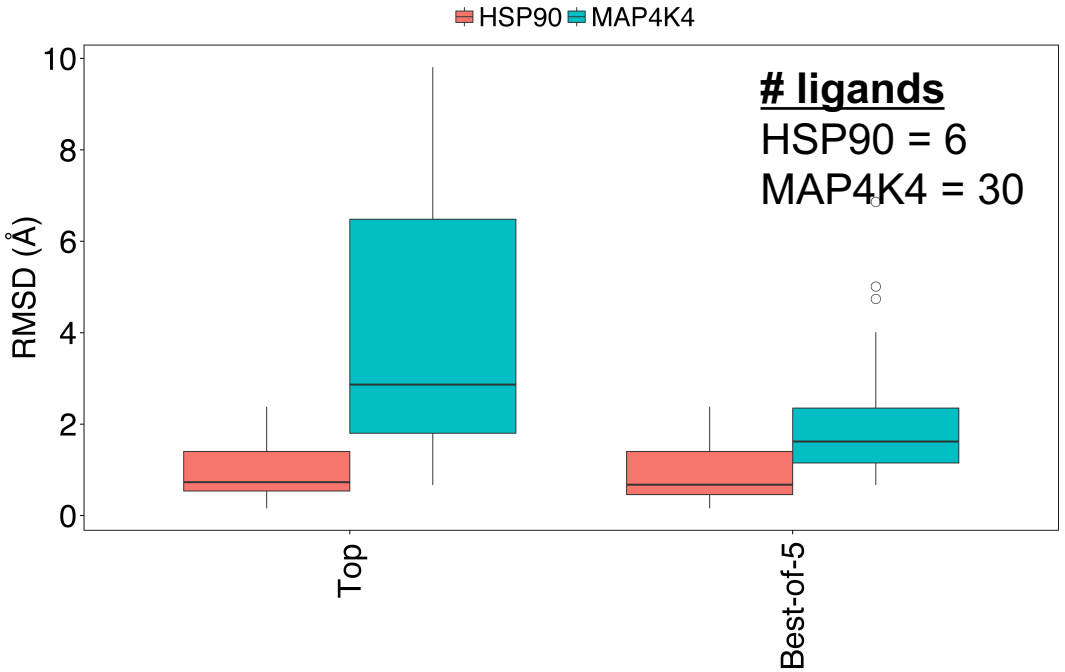
# PoPSS: Pose Prediction using Shape Similarity in D3R GC1





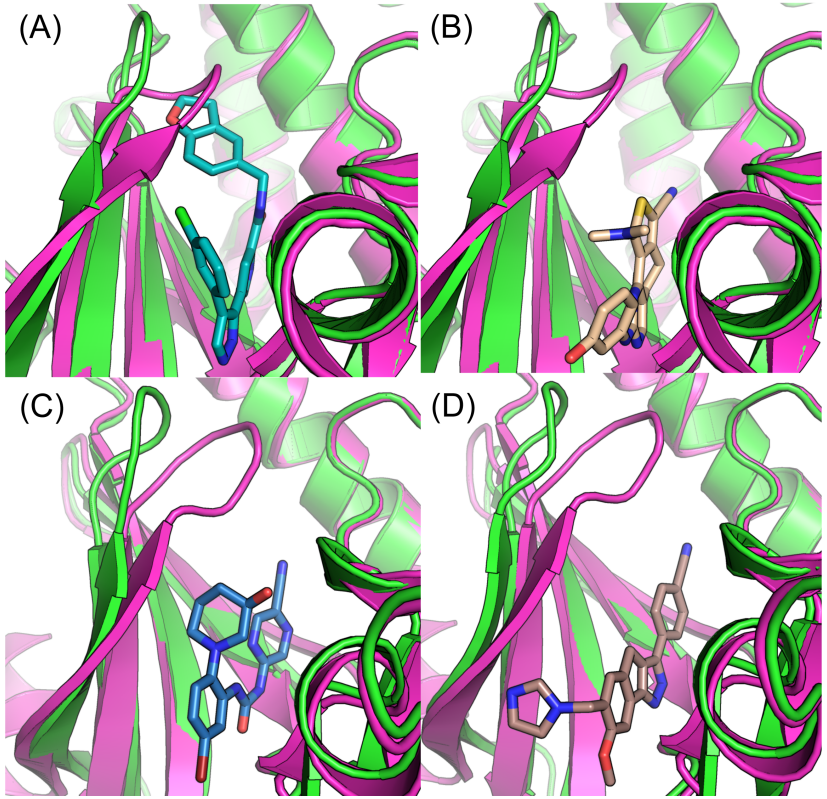



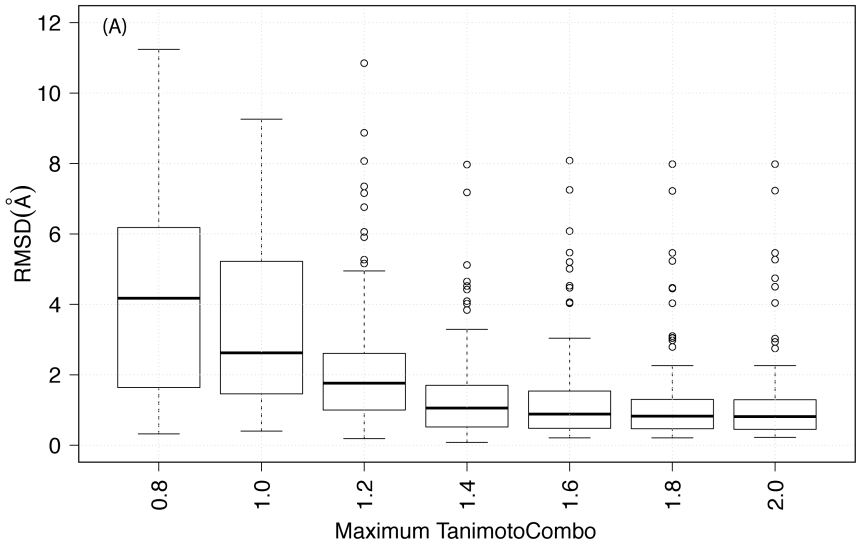

| Protein | Number of submissions | | Mean RMSD (Å) | Median RMSD (Å) | Mean RMSD Rank | Median RMSD Rank |
|---|---|---|---|---|---|---|
| HSP90 | 42 | Top | 1.02 | 0.73 | 2 | 4 |
| | | Best of 5 | 0.99 | 0.68 | 4 | 5 |
| MAP4K4 | 27 | Top | 4.04 | 2.87 | 3 | 3 |
| | | Best of 5 | 2.09 | 1.62 | 1 | 1 |

# What did we learn in D3R GC1





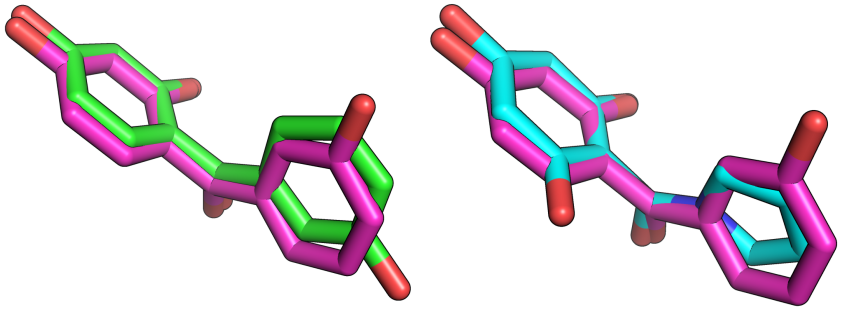

- Selection of right receptor conformation is very important in accurately predicting binding poses
- Receptor conformation suitable for one chemotype may not be suitable for others




- Available shape similarity correlates with pose prediction performance
- Poses can be reliably predicted if maximum available shape similarity is more than 1.4 TanimotoCombo

- Predicted pose scored using Chemgauss4 scoring function
- Scoring performance was very bad
- Random

| Coefficient | HSP90 | MAP4K4 |
|---|---|---|
| Pearson | 0.16 | 0.05 |
| Kendall Tau | 0.10 | 0.02 |



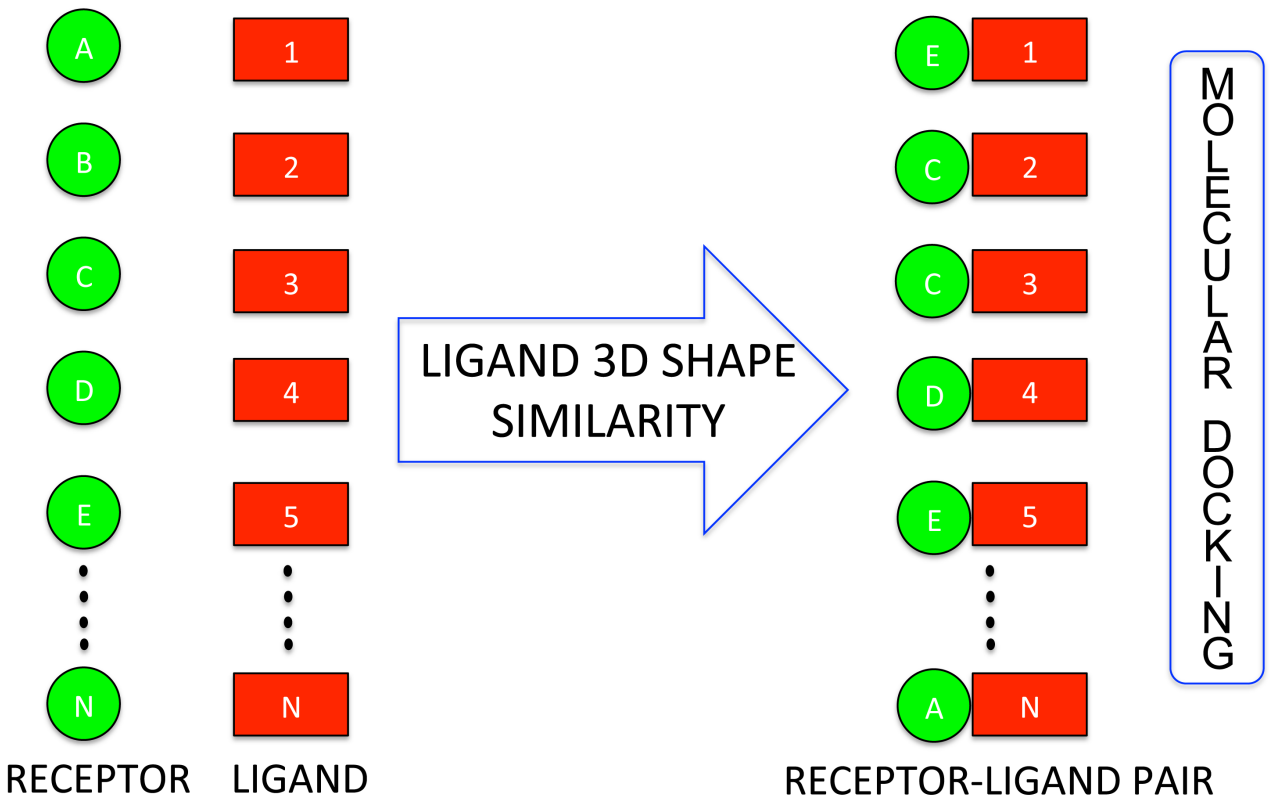
Crystal (green) v/s predicted (magenta)

Predicted (magenta) v/s template(cyan)

- No ligand sampling after placing the conformation
- No rewards for forming good interactions
- No penalties for bad interactions

# Cross-docking based virtual screening pipeline (CDVS)

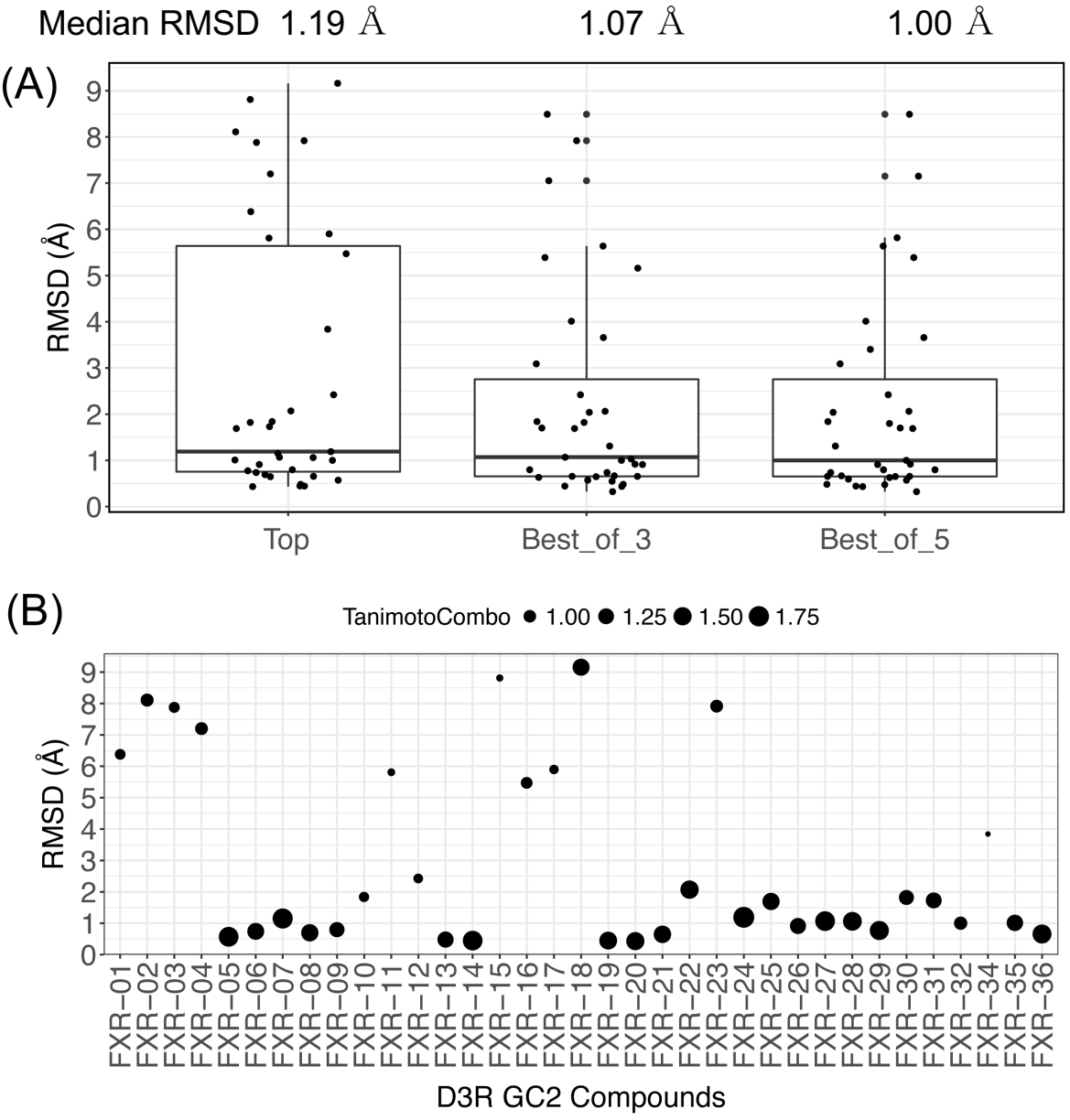


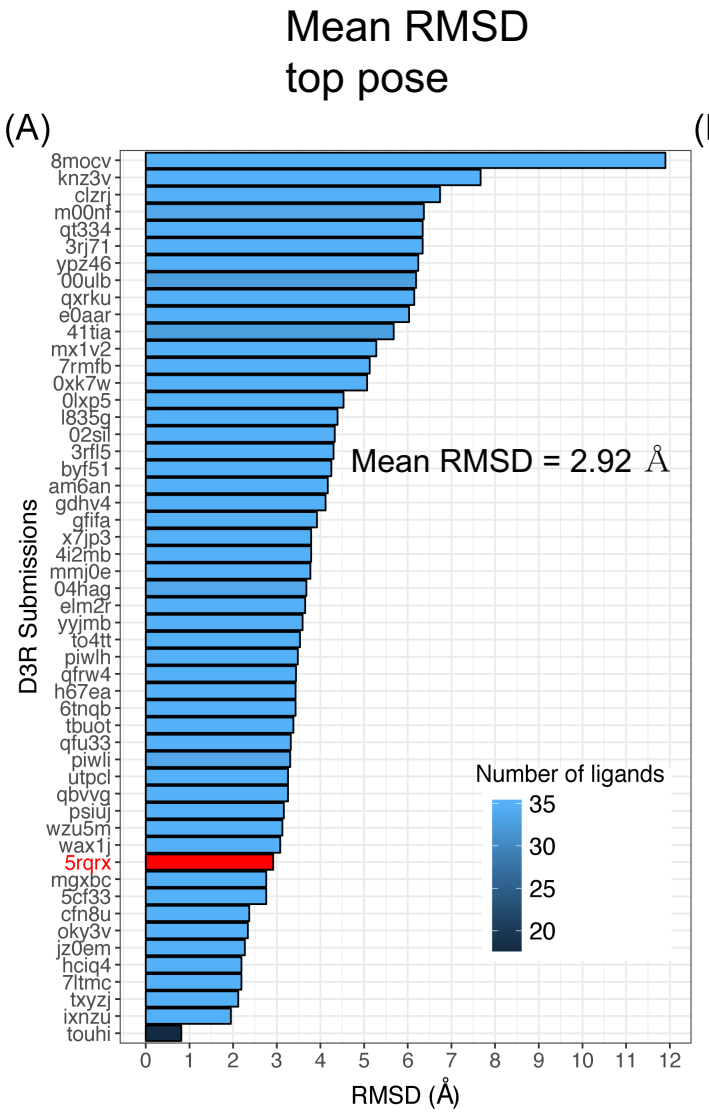


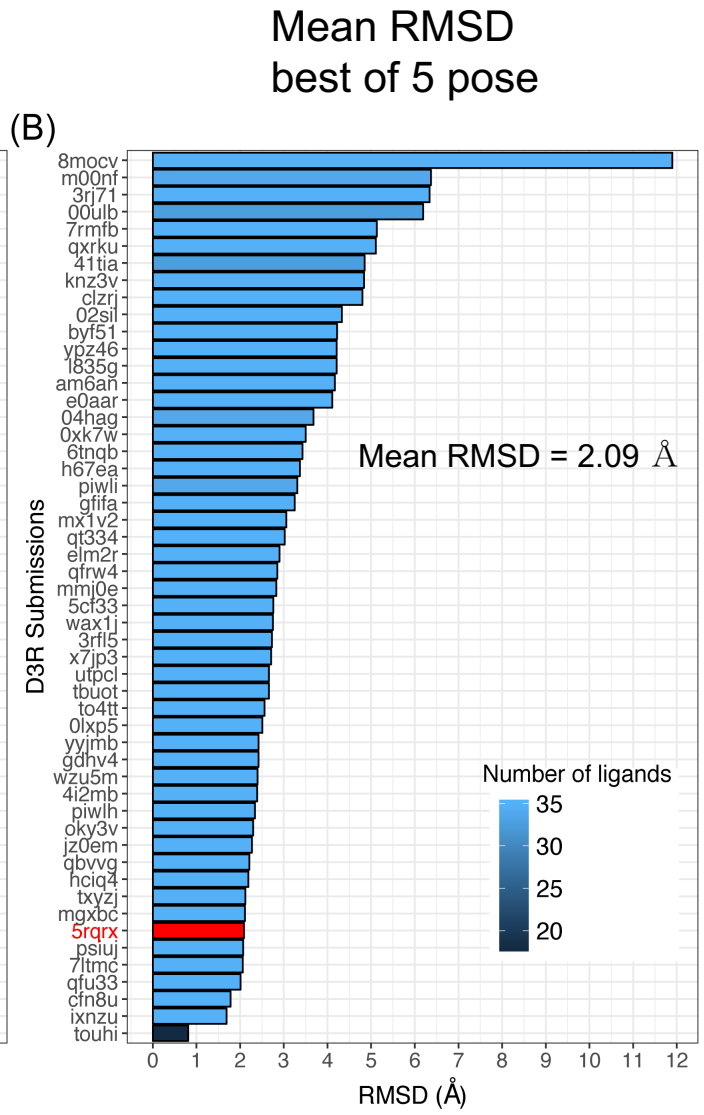

- Evaluated in D3R Grand Challenge 2
- Involved prediction of poses and potency of farnesoid X receptor (FXR) ligands.
- Stage 1: Pose prediction for 36 ligands (FXR-1-FXR-36)
- Stage 1: Affinity prediction or ranking for 102 FXR ligands
- Stage 2: Affinity prediction or ranking for 102 FXR ligands utilizing crystal information from stage 1

# Cross-docking based virtual screening pipeline (CDVS) in D3R





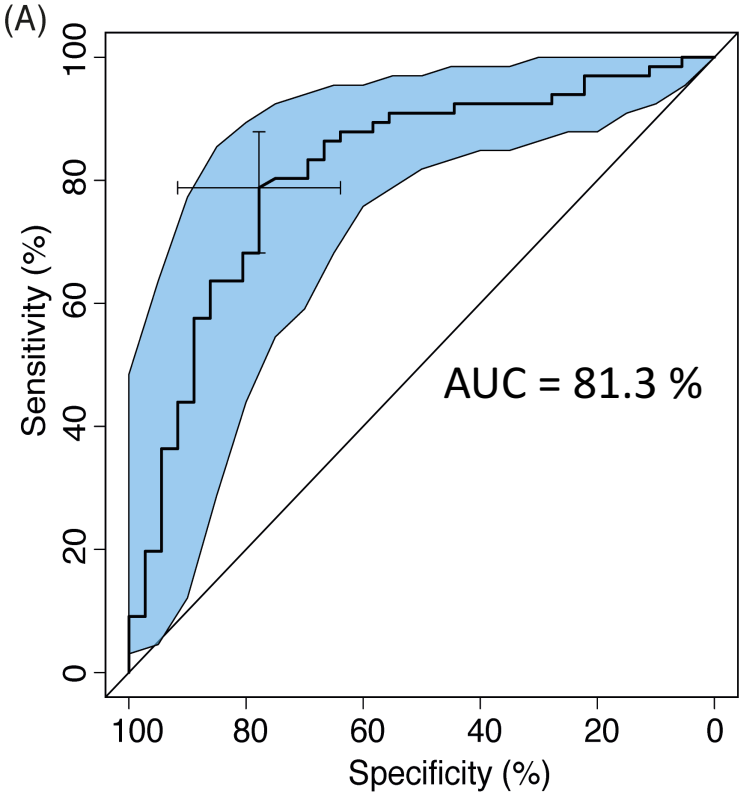



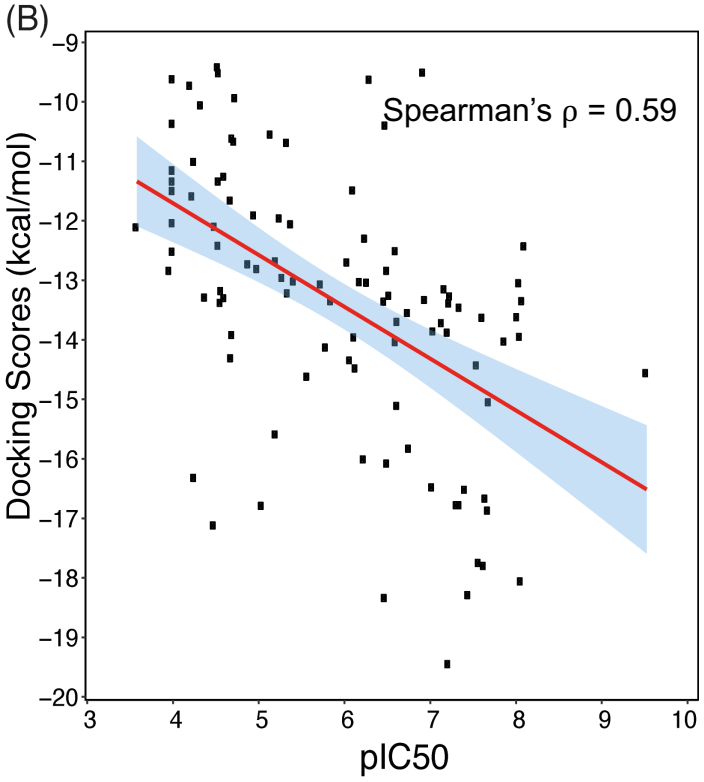



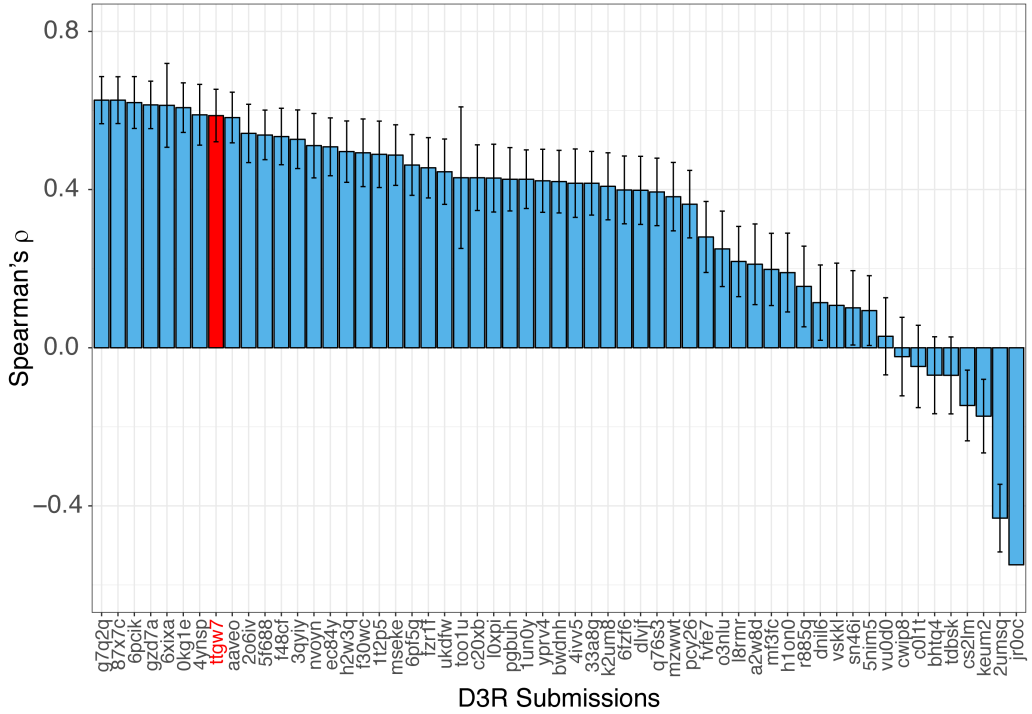
# Cross-docking based virtual screening pipeline (CDVS): Virtual screening performance

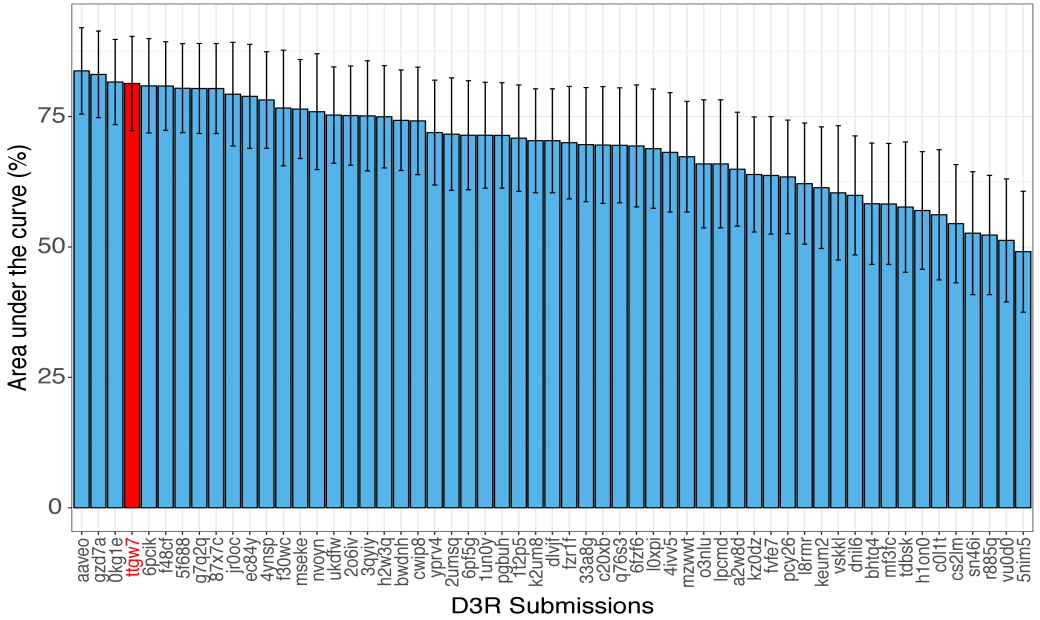


| Scores | D3R GC2 Phase 1 value | D3R GC2 Phase 1 Rank/ Participating groups |
|---|---|---|
| AUC (95 % confidence interval) (%) | 81.3 (72.3 – 90.4) | 4/57 |
| Spearman's rho (ρ) (standard error) | 0.59 (0.07) | 8/57 |
| Kendall's tau (τ) (standard error) | 0.40 (0.05) | 8/57 |



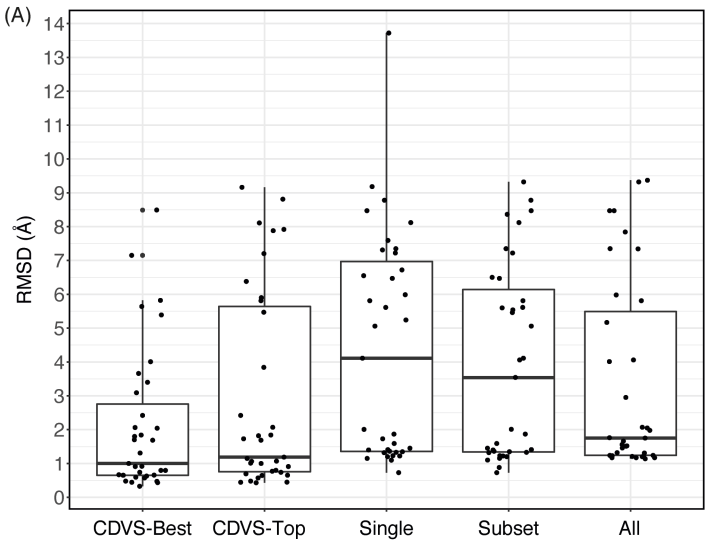



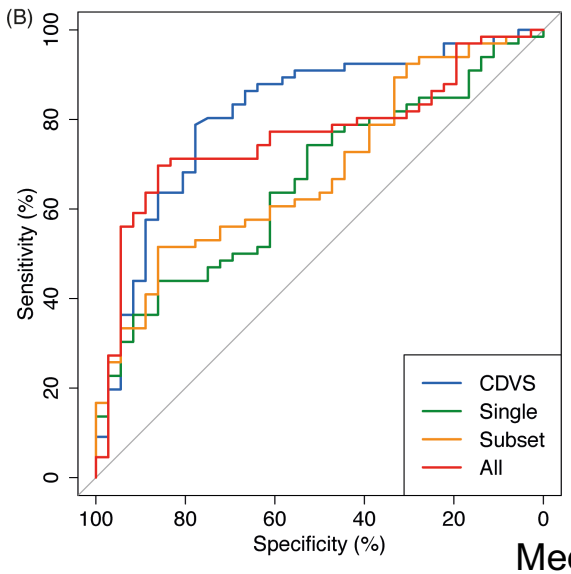



# What did we learn in D3R Grand Challenge 2

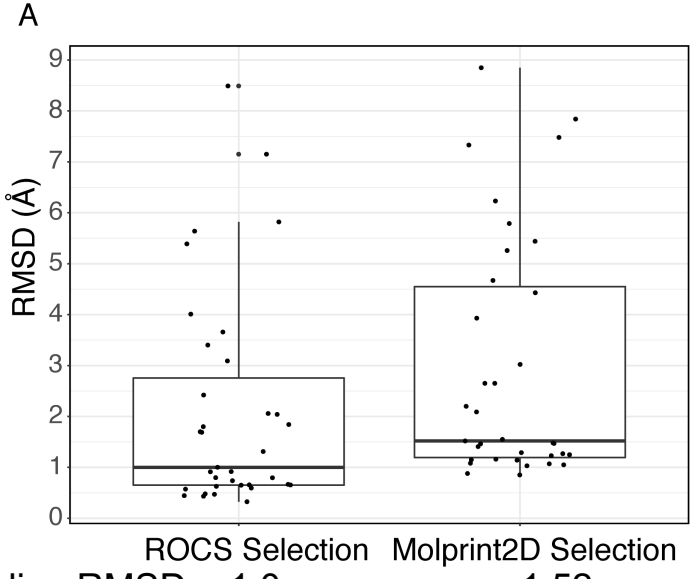


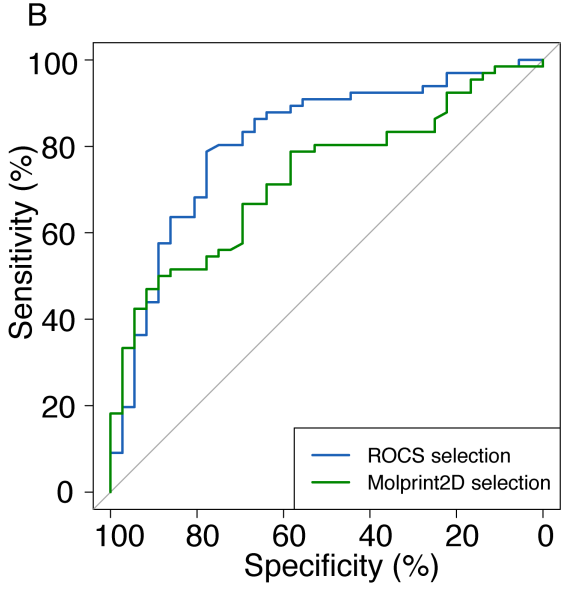
## Comparison with multiple receptor docking methods




## Receptor selection method

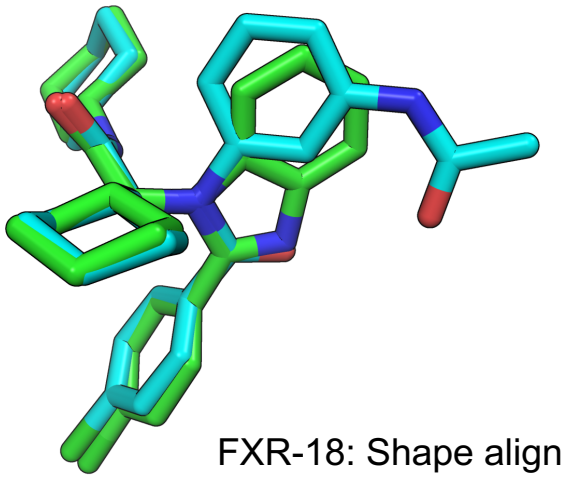


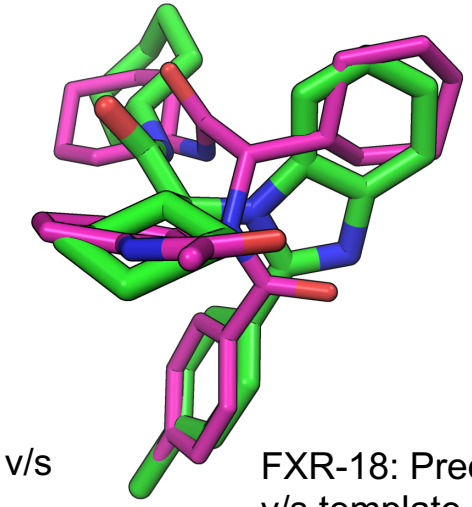

Median RMSD = 1.0 1.52
Mean RMSD = 2.09 2.90

AUC ROCS sel. = 81.3 %
AUC Molprint2D sel. = 72.8 %

| Method | Median RMSD (Å) | AUC (%) |
|---|---|---|
| CDVS | 1.19 | 81.3 |
| Single-receptor docking | 4.11 | 66.1 |
| Multiple-receptor docking (subset of receptors) | 3.54 | 68.4 |
| Multiple-receptor docking (all receptors) | 1.75 | 76.3 |



FXR-18: Shape align (cyan) v/s template (green)



FXR-18: Predicted (magenta) v/s template (green)

# D3R Grand Challenge 3

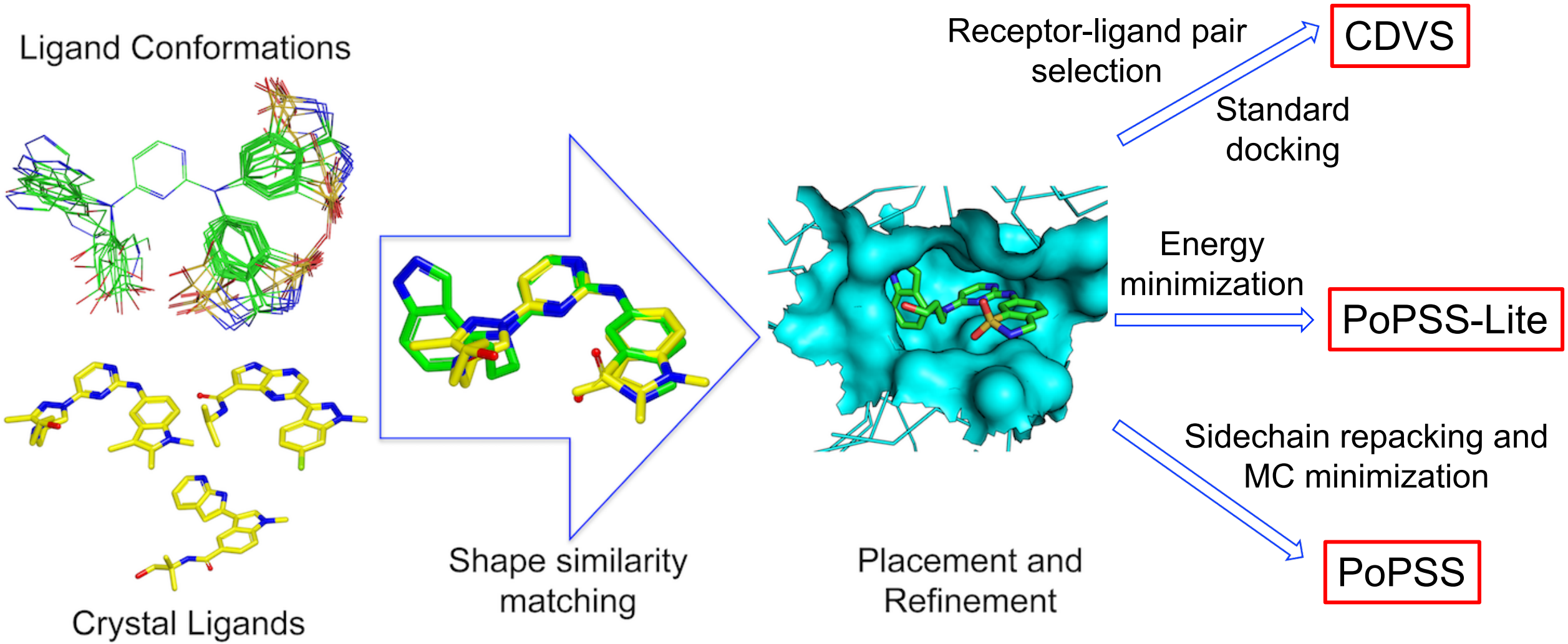


- Mega challenge
- Five subchallenge
  - Cathepsin S: Subchallenge 1, Pose prediction, ranking, FEP
  - VEGFR2
  - JAK2 SC2
  - P38-α

    Subchallenge 2, Ranking, selectivity
  - JAK2 SC3: Subchallenge 3, Ranking, activity cliff
  - TIE2: Subchallenge 4, Ranking, activity cliff
  - ABL1: Subchallenge 5, Effect of mutation on binding affinities
- Subchallenge 1
  - Cathepsin S
  - Pose prediction, ranking and FEP
  - Dataset: 26 crystal structures, 136 compound affinities
  - Pose prediction of 24 compounds and affinity prediction or ranking of 136 compounds

# Pose Prediction methods used in D3R Grand Challenge 3

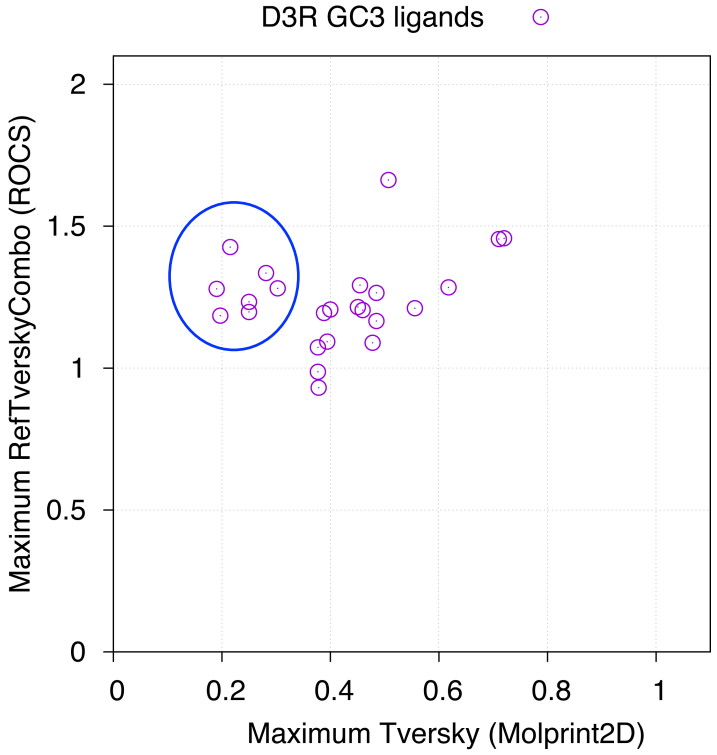


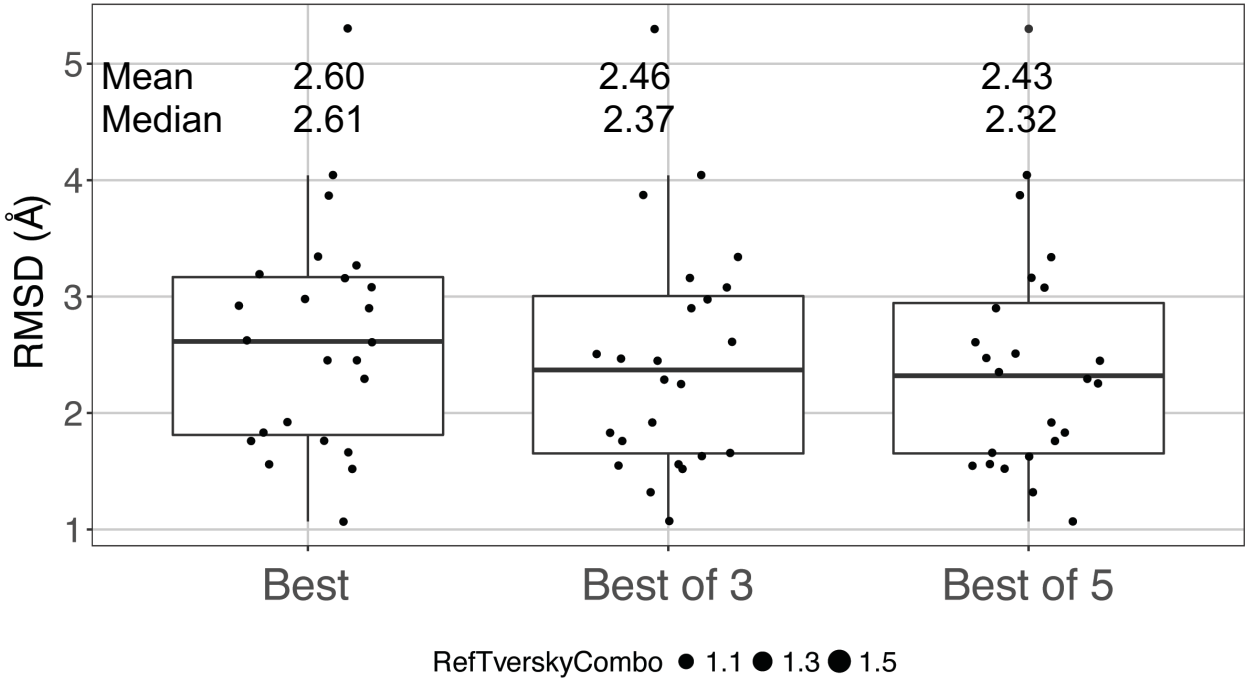


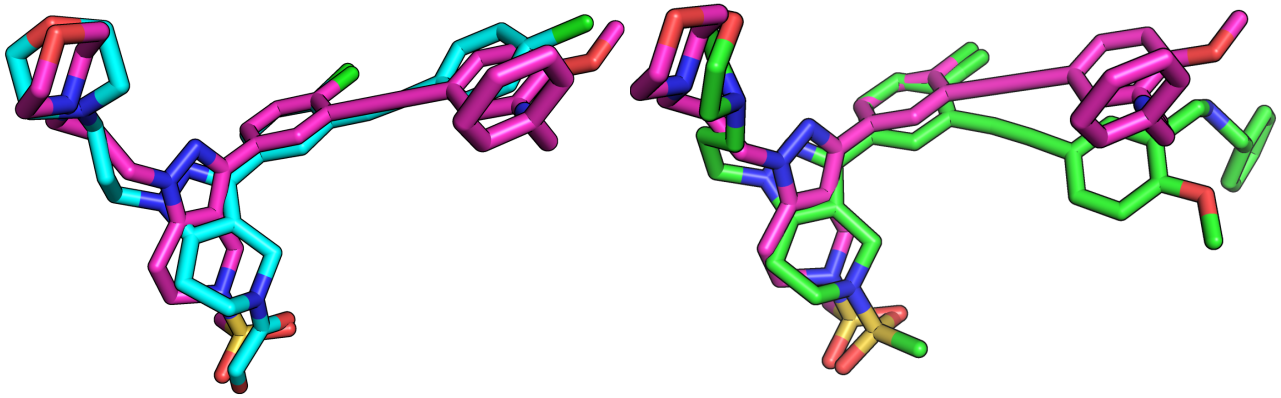
# Pose prediction performance (PoPSS)



- A maximum of 1000 conformations per ligand
- RefTverskyCombo coefficient
- Rest same as D3R GC1

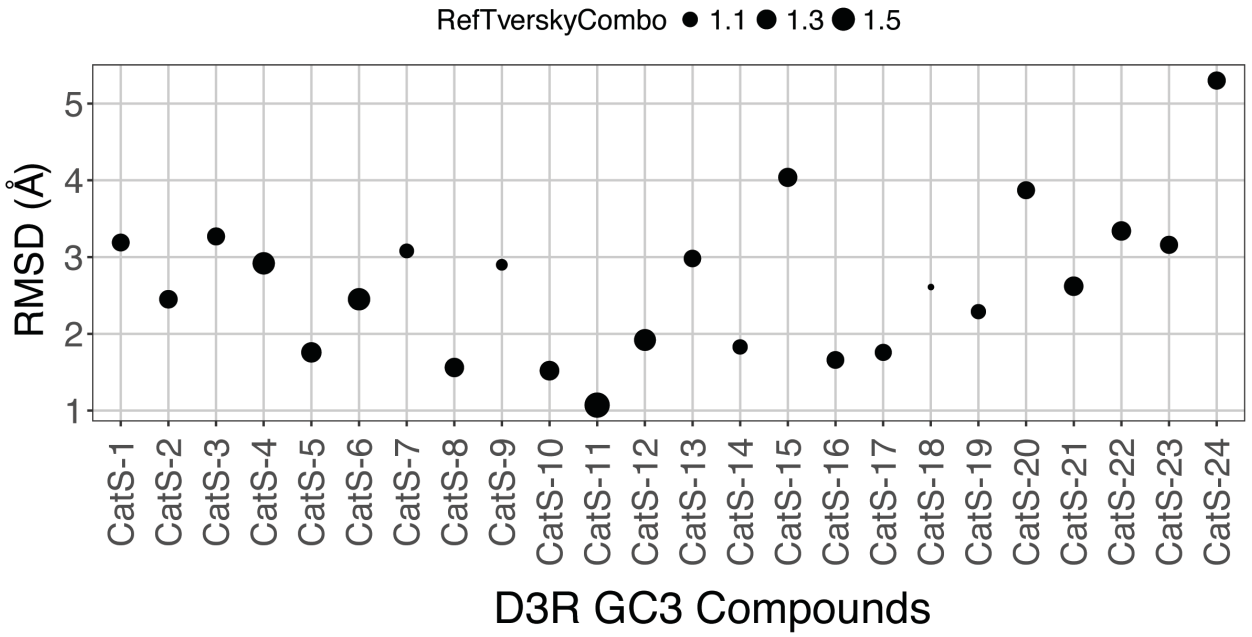


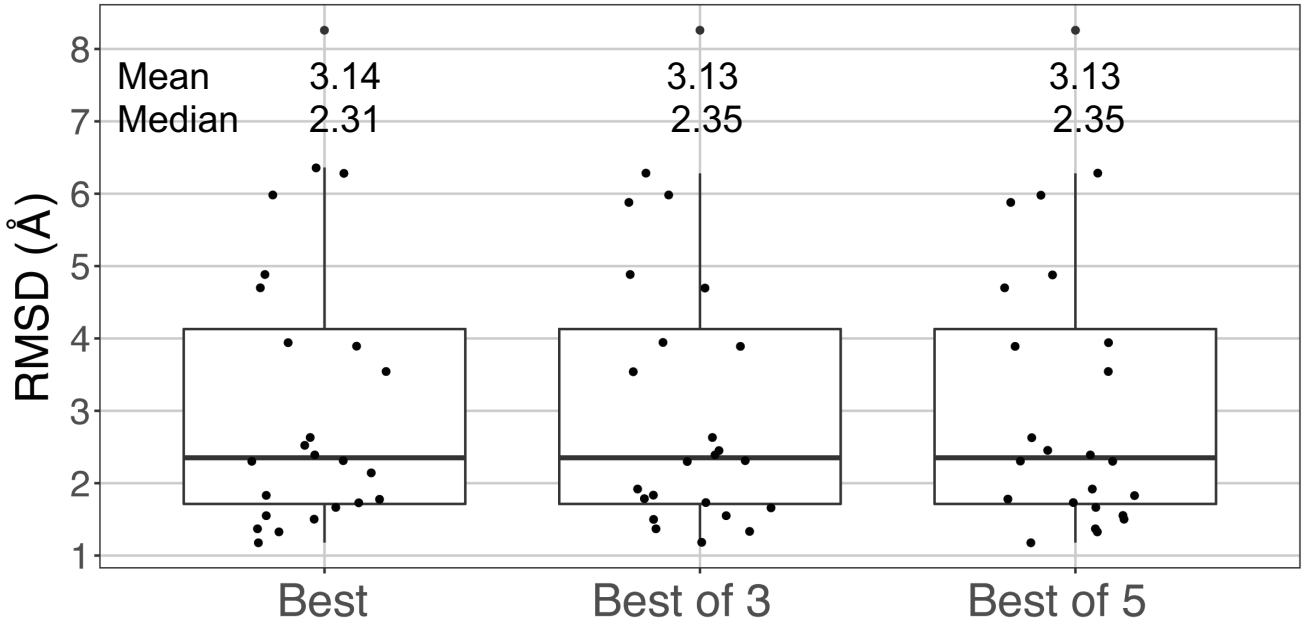
CatS-24 predicted (magenta) v/s template (cyan)

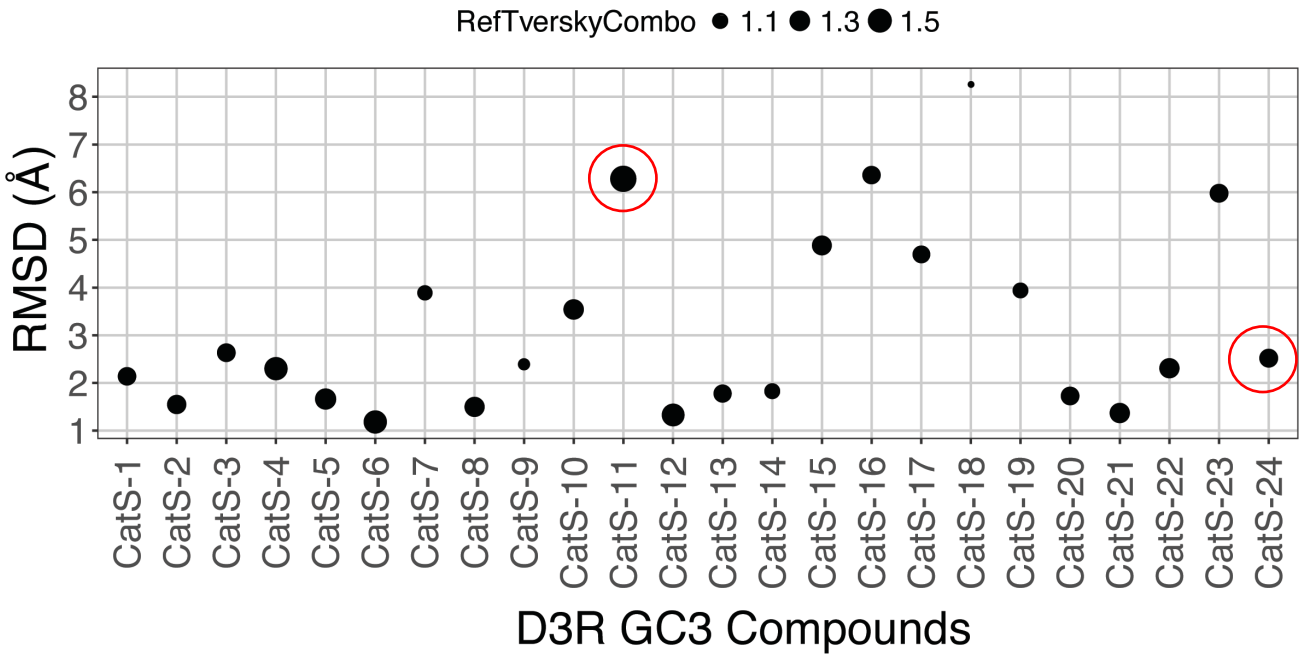
CatS-24 predicted (magenta) v/s crystal (green)

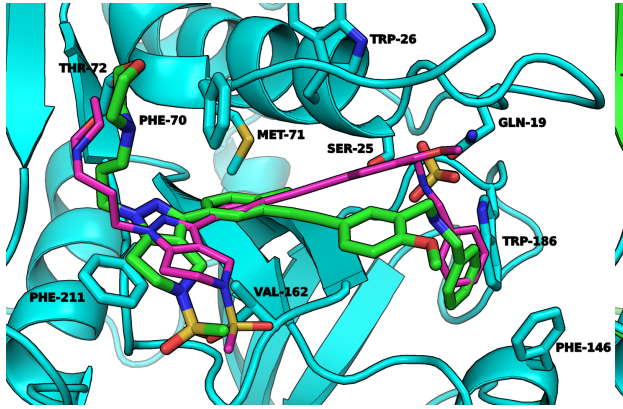


# Pose prediction performance (CDVS)

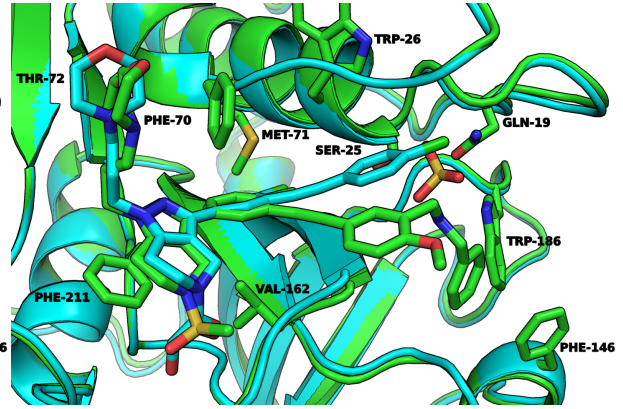




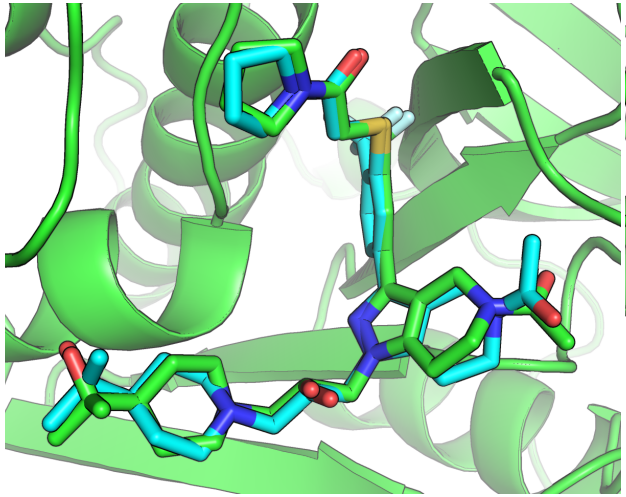



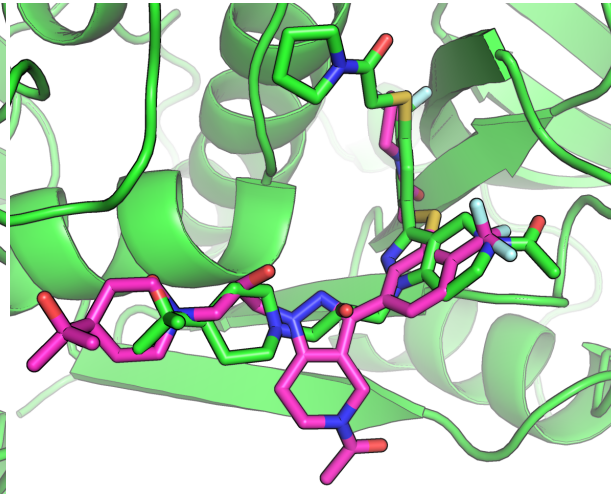

CatS-24 predicted (magenta) v/s crystal (green)




CatS-24 template (cyan) v/s crystal (green)

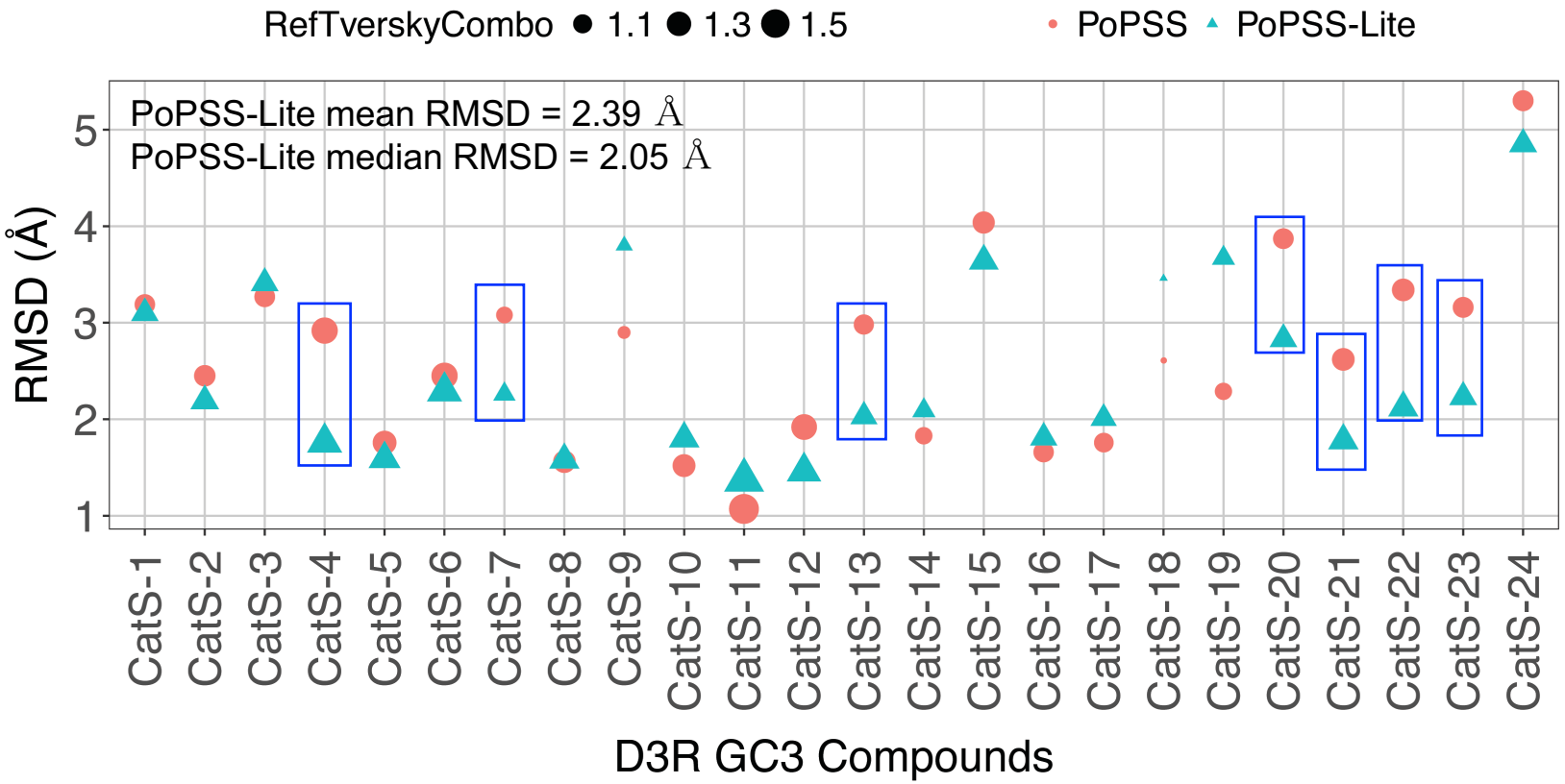


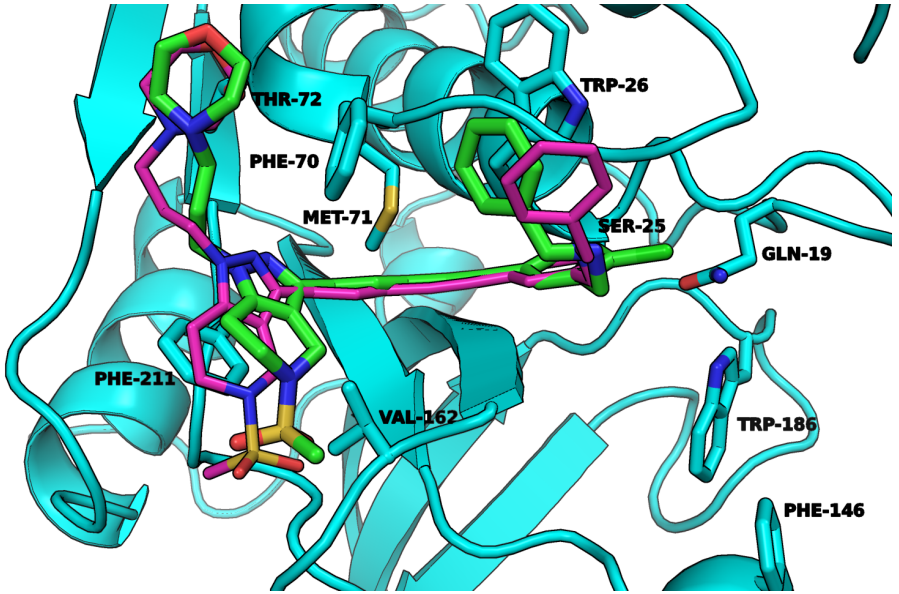
CatS-11 aligned (cyan) v/s crystal (green)



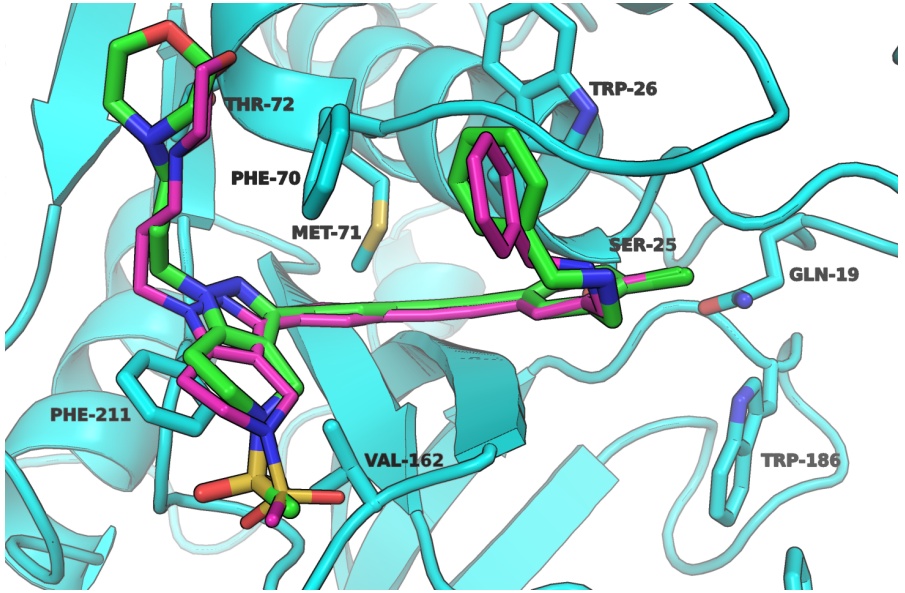
CatS-11 predicted (magenta) v/s crystal (green)

# Pose prediction performance (PoPSS-Lite)




PoPSS predicted CatS-21 (magenta)

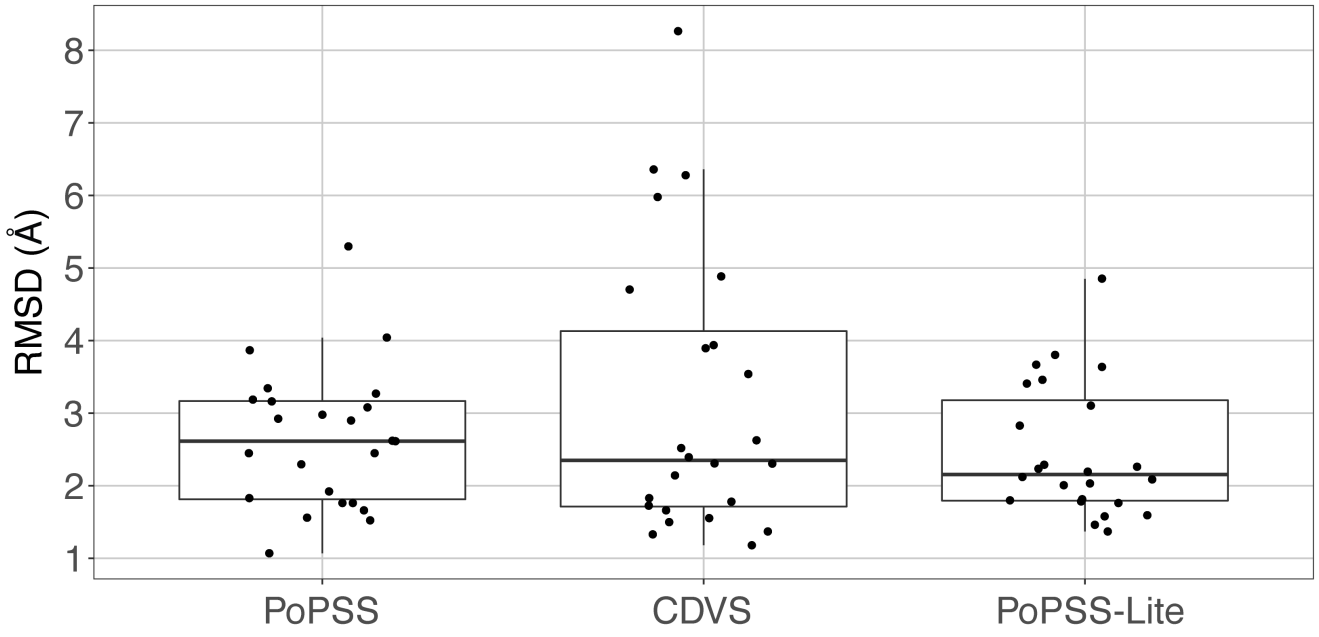



PoPSS-Lite predicted CatS-21 (magenta)

| Improved/Deteriotated | Number |
|---|---|
| Improved | 14/24 |
| Deteriorated | 10/24 |
| Improved by 0.5 Å RMSD | 7/24 |
| Deteriorated by 0.5 RMSD | 3/24 |

# Comparison of three methods in D3R GC3



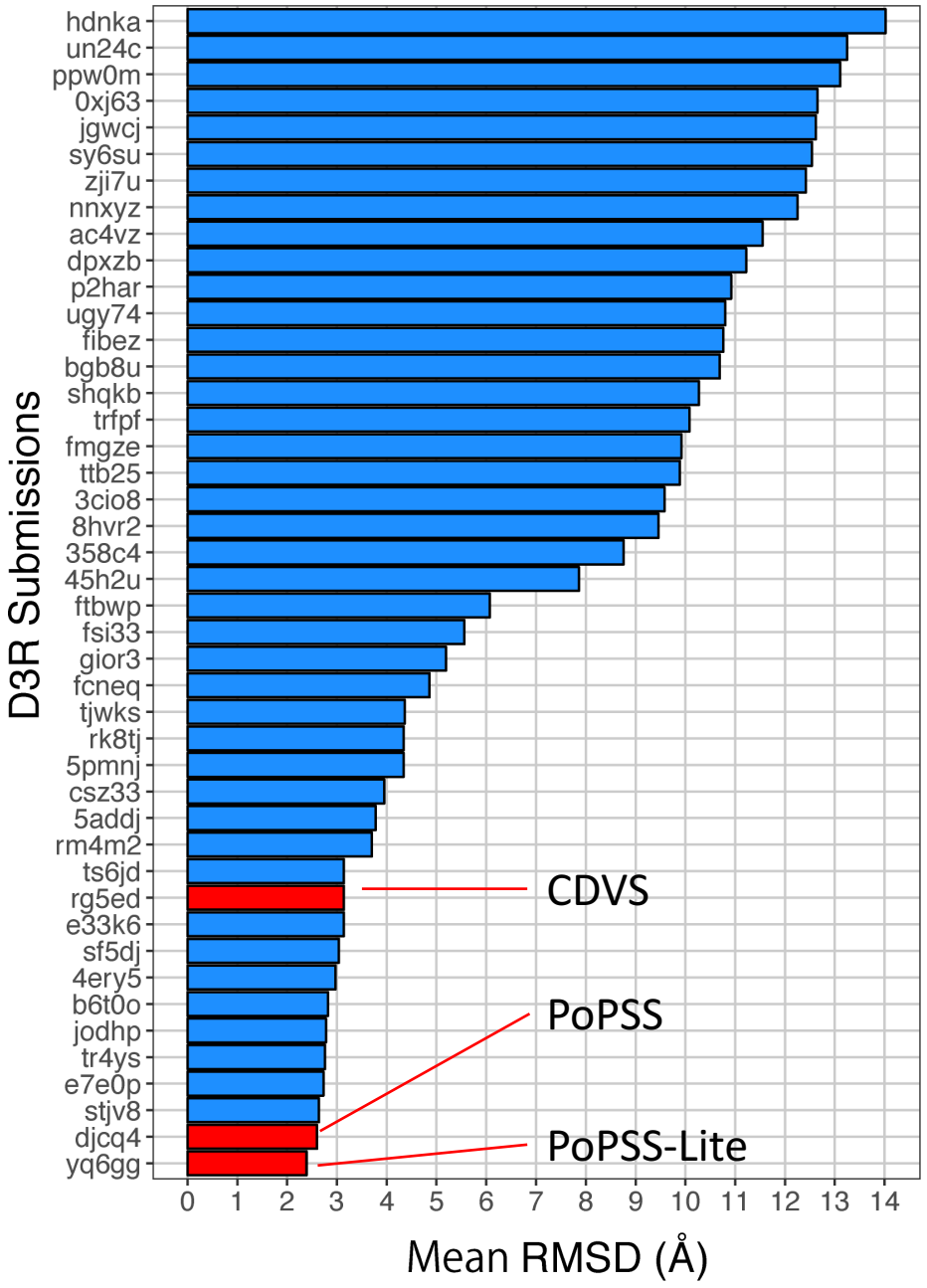


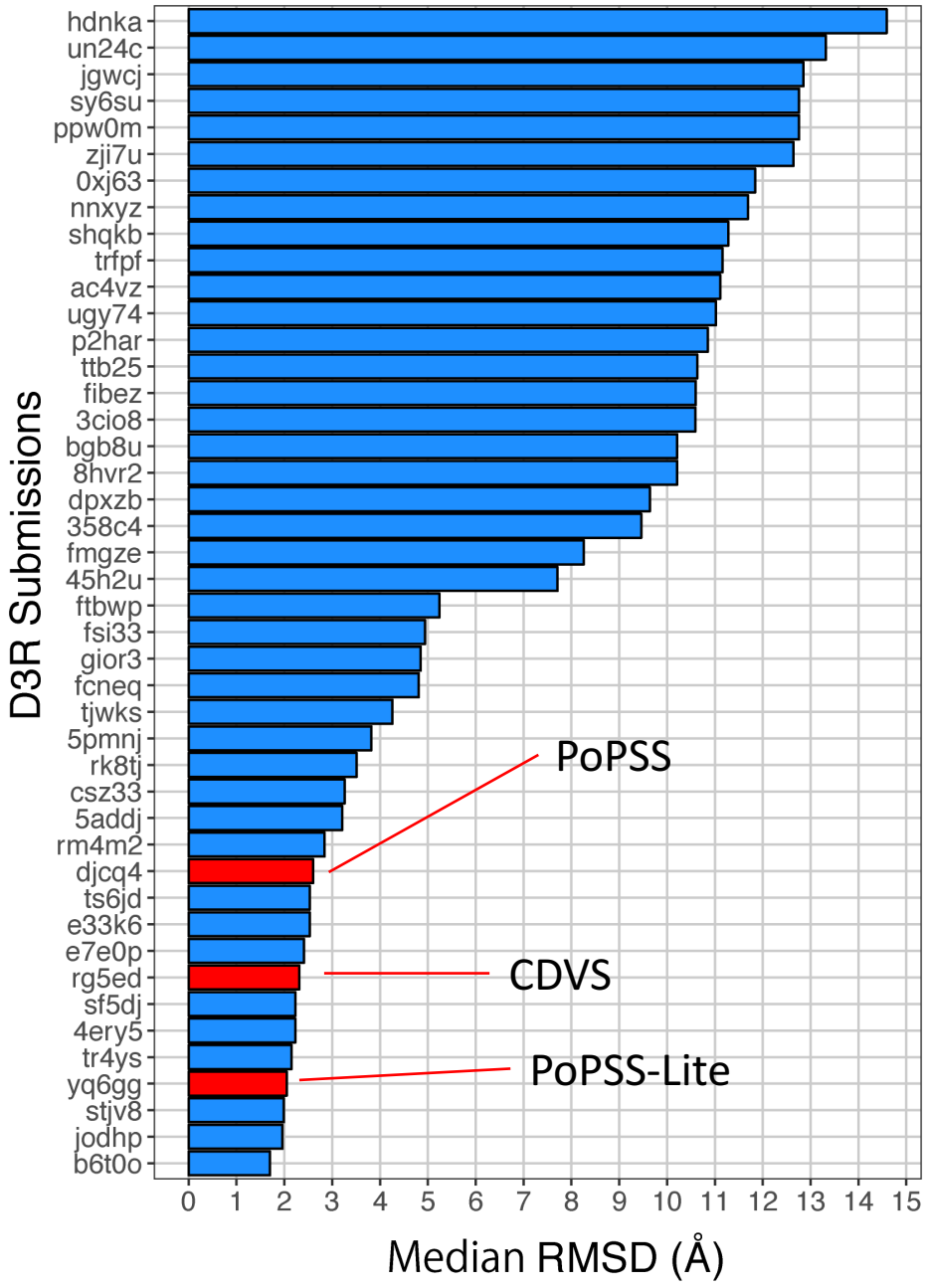

| Type | Mean RMSD (Å) | Median RMSD (Å) | Average number of atoms |
|---|---|---|---|
| PoPSS | 2.60 | 2.61 | 89.25 |
| CDVS | 3.14 | 2.31 | |
| PoPSS-Lite | 2.39 | 2.05 | |
| D3R 2015 (HSP90) | 1.02 | 0.73 | 30.33 |
| D3R 2015 (MAP4K4) | 4.04 | 2.87 | 36.83 |
| D3R 2016 (FXR) | 2.92 | 1.19 | 63.97 |

| PoPSS | CDVS | PoPSS-Lite |
|---|---|---|
| Poses can be predicted with relative accuracy | Scoring/Ranking performance is better | Poses can be predicted with relative accuracy |
| Method generates several poses that are ranked by scoring function | Ligand 3D shape similarity is only used in suitable receptor selection | Only one pose per ligand, so no scoring problem |
| Occasionally, pose closest to native structure was not the top scoring one | | |

# Pose prediction comparison in D3R GC3





# Summary and lessons learned for future D3R



- Ligand 3D shape similarity can be successfully employed to improve pose prediction performanc
- Shape similarity based receptor selection has advantages over 2D similarity based receptor selection
- Current implementation requires at least one suitable co-crystal ligand. Future development will explore the ligands from homologous proteins and homologous protein pockets to improve pose prediction
- Generation of ligand conformations is critical for success and needs improvement as some bioactive conformation could not be generated
- Scoring performance of PoPSS and PoPSS-Lite needs improvement

# Thank you for your attention


Acknowledgements:

**D3R Organizers**
**Kam Zhang, Team leader**
Aditya Padhi
Dileep Vijayan
Matej Janezic
Venkatakutu H. Nidamarthi
**RIKEN**
**D3R community**