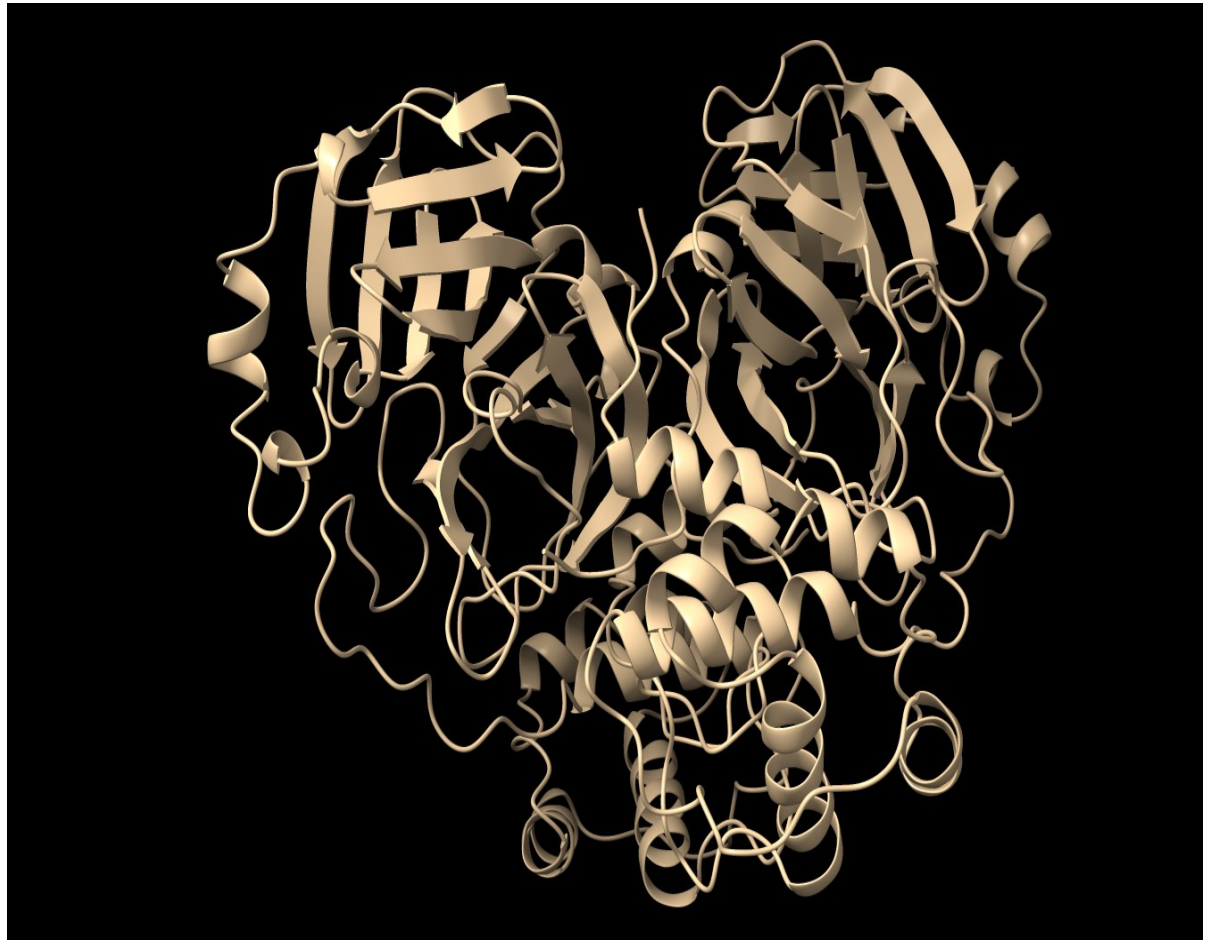
Developing a Virtual Screening Pipeline for SARS-CoV-2 Mpro Inhibitors

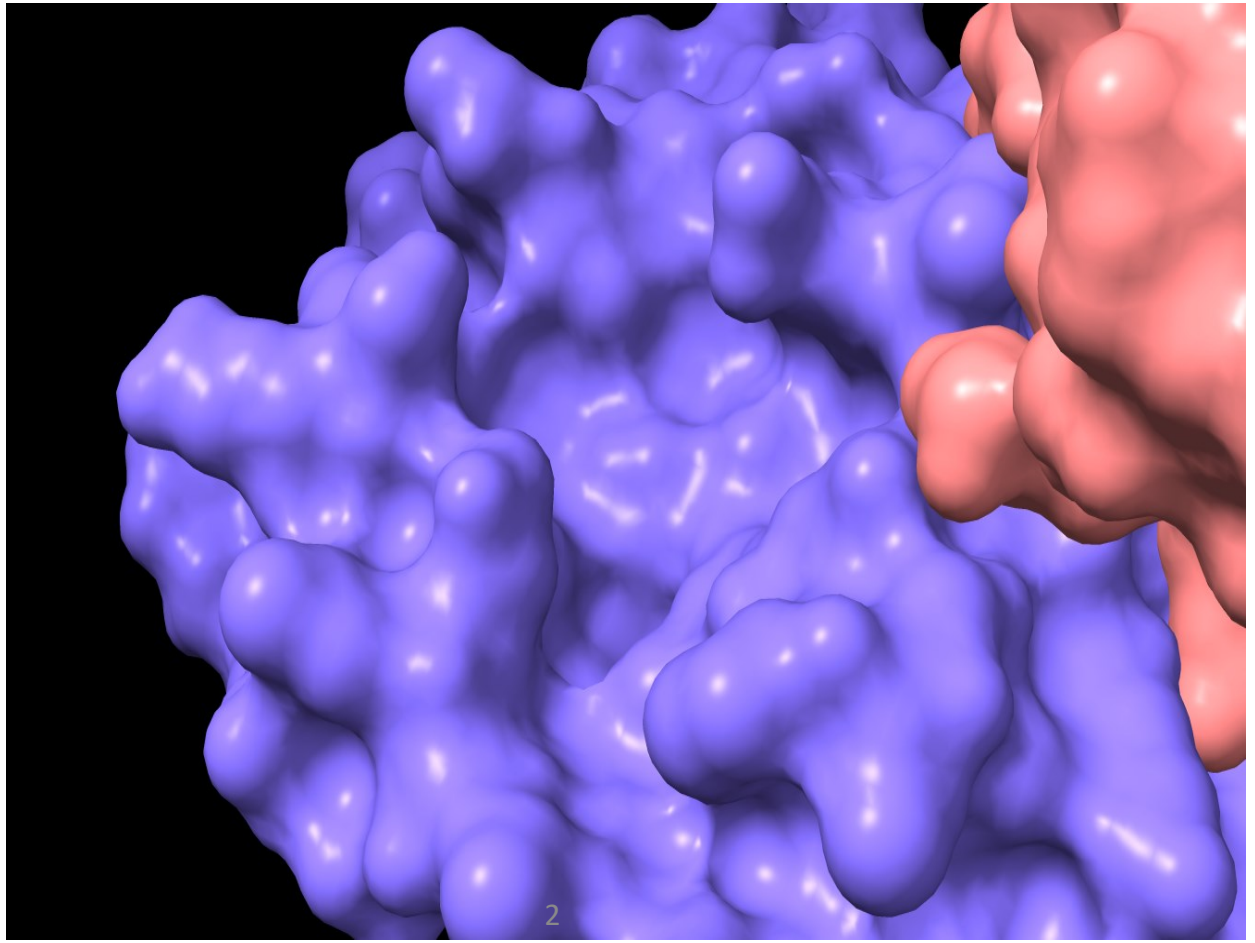
Daniel O. Malikin, Eugene V. Radchenko, Vladimir A. Palyulin
Moscow State University

Motivation & Objectives

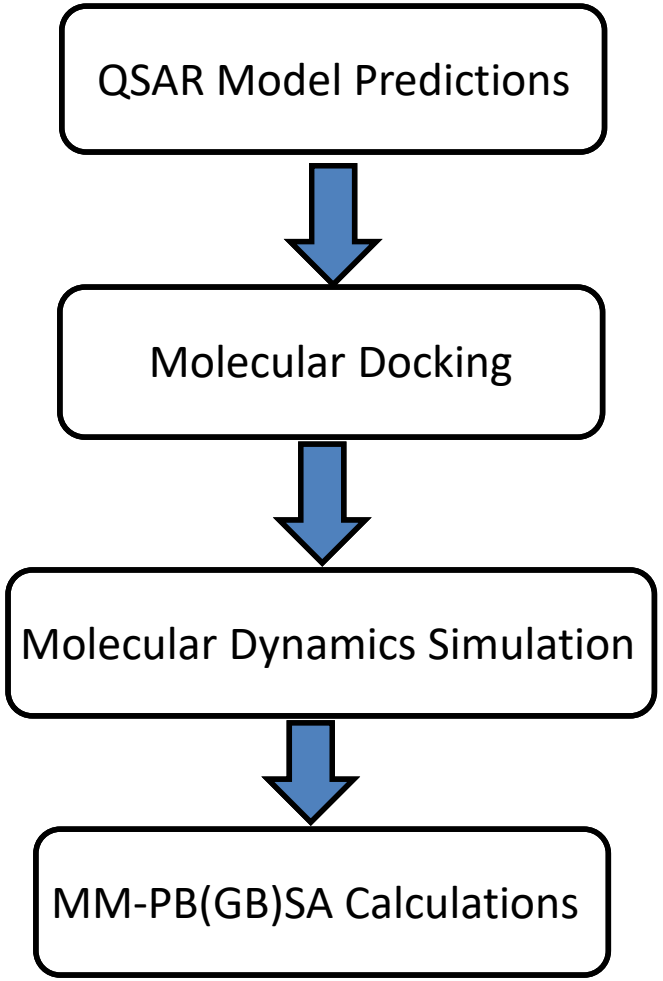
Mpro dimeric structure



Mpro active site pocket

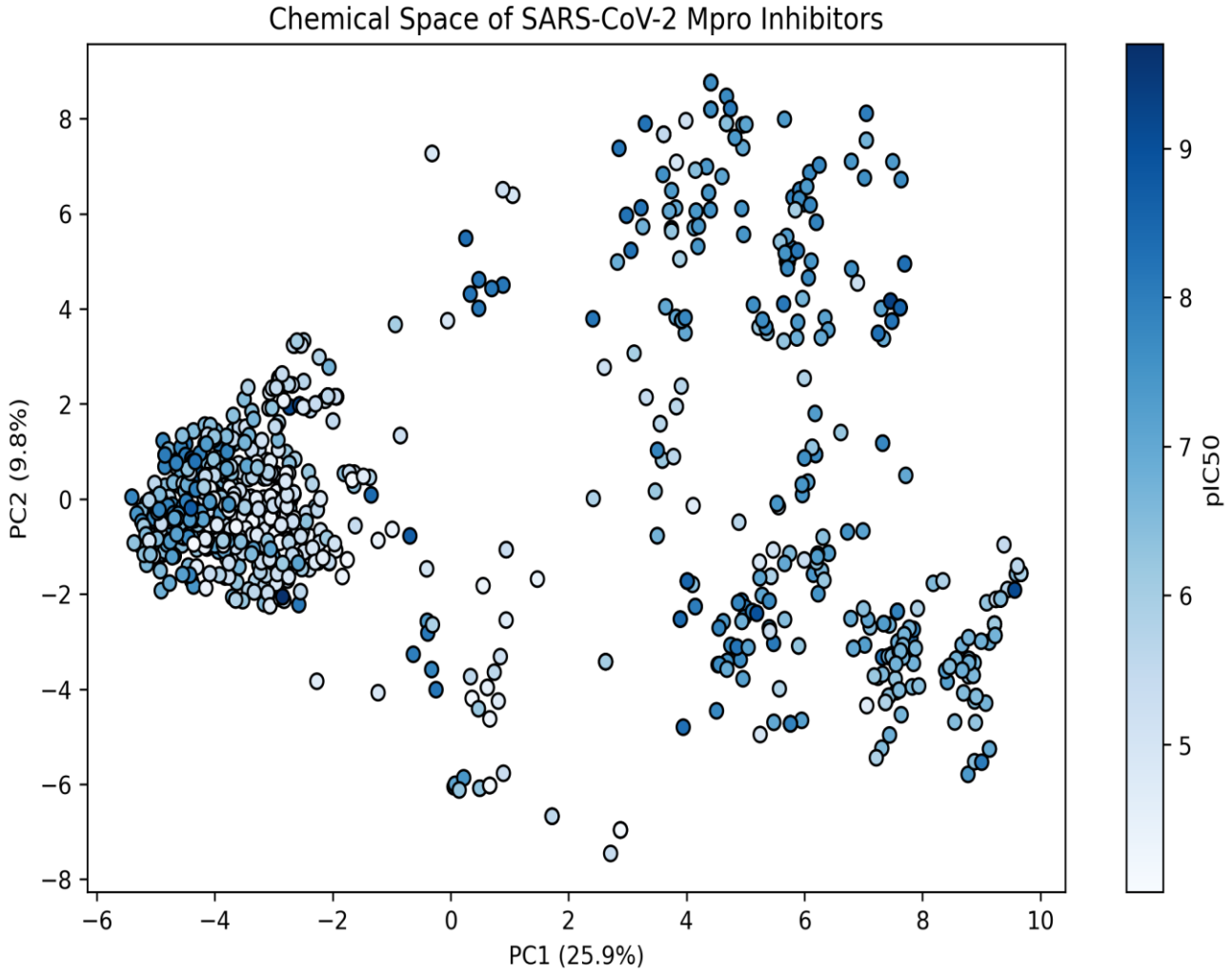


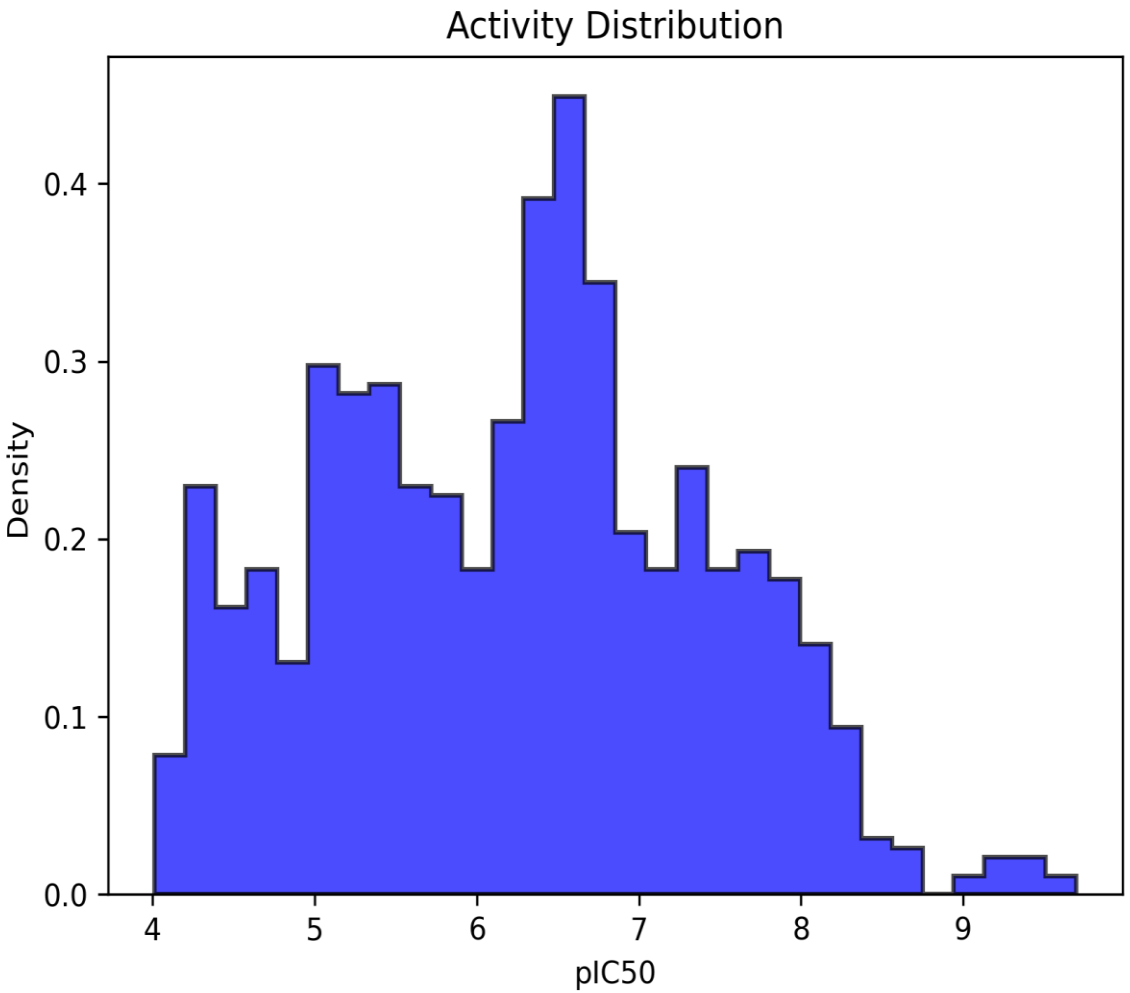
Pipeline Overview



QSAR Data Curation

1010 Molecules



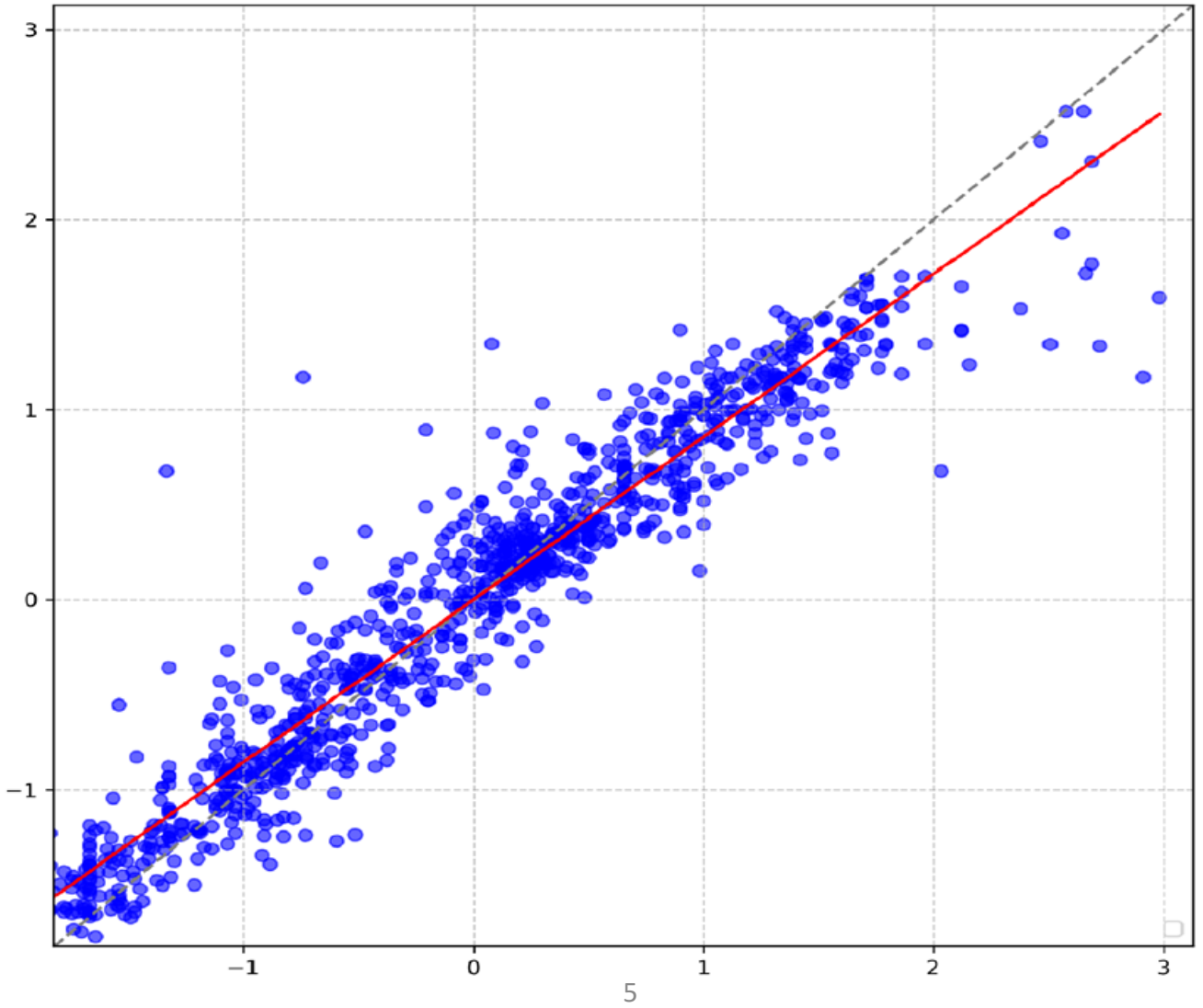


QSAR: Features & Algorithms

pattern+avalon fingerprints

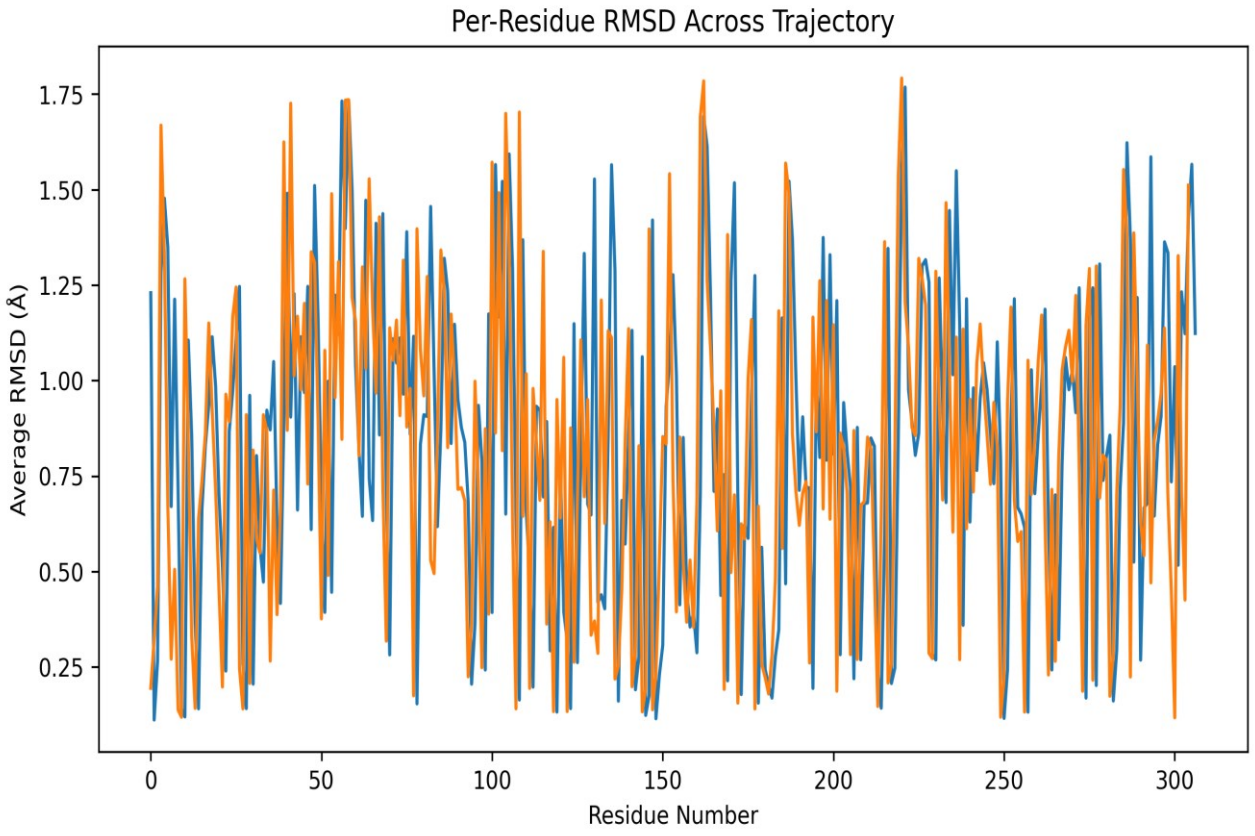
	R^2	Q^2	RMSE
Gradient Boosting	**0.94**	**0.69**	**0.65**
Random Forest	**0.83**	**0.66**	**0.68**
Bagging Regressor	**0.95**	**0.68**	**0.65**
SVM Regressor	**0.94**	**0.69**	**0.64**
ANN Regressor	**0.96**	**0.66**	**0.67**

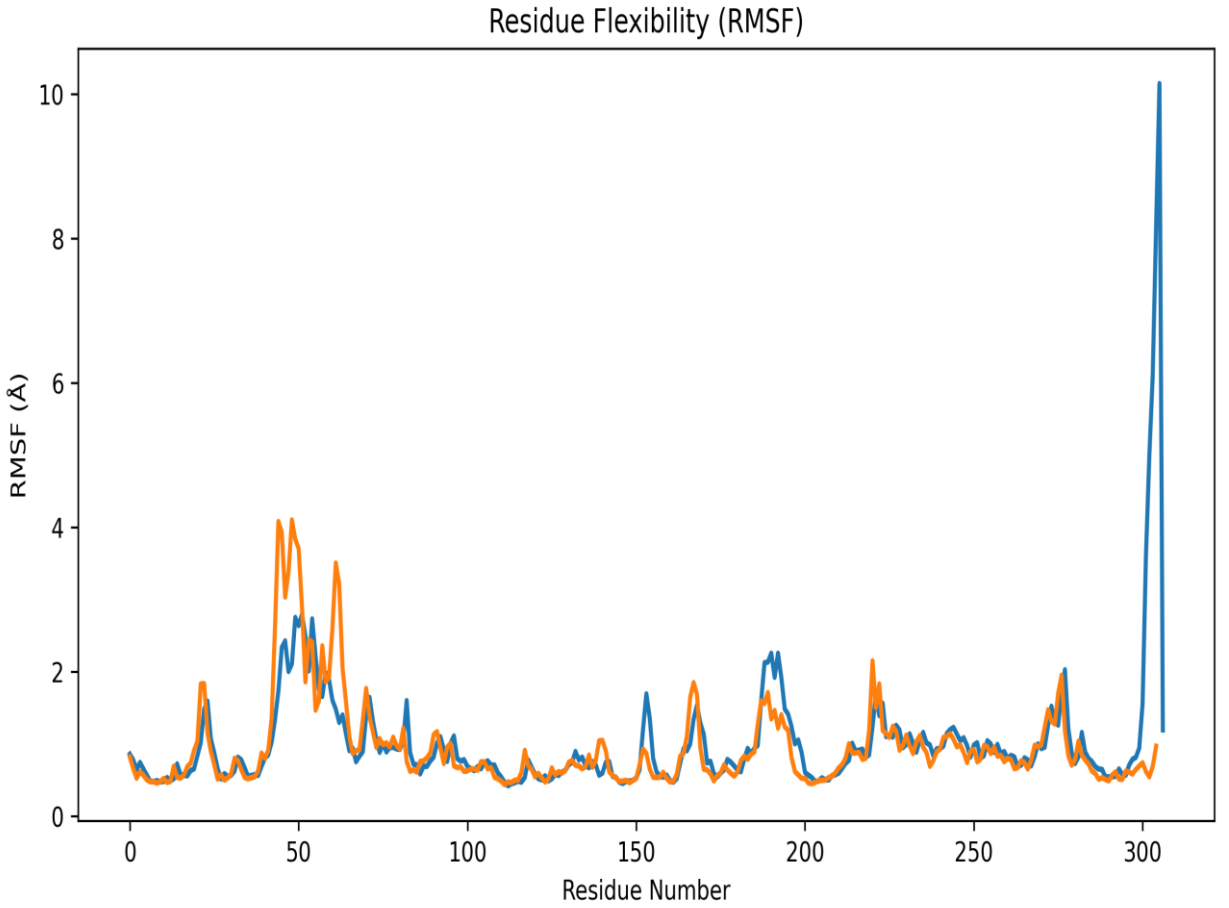
Predicted vs Experimental Activity



Exploring Mpro flexibility

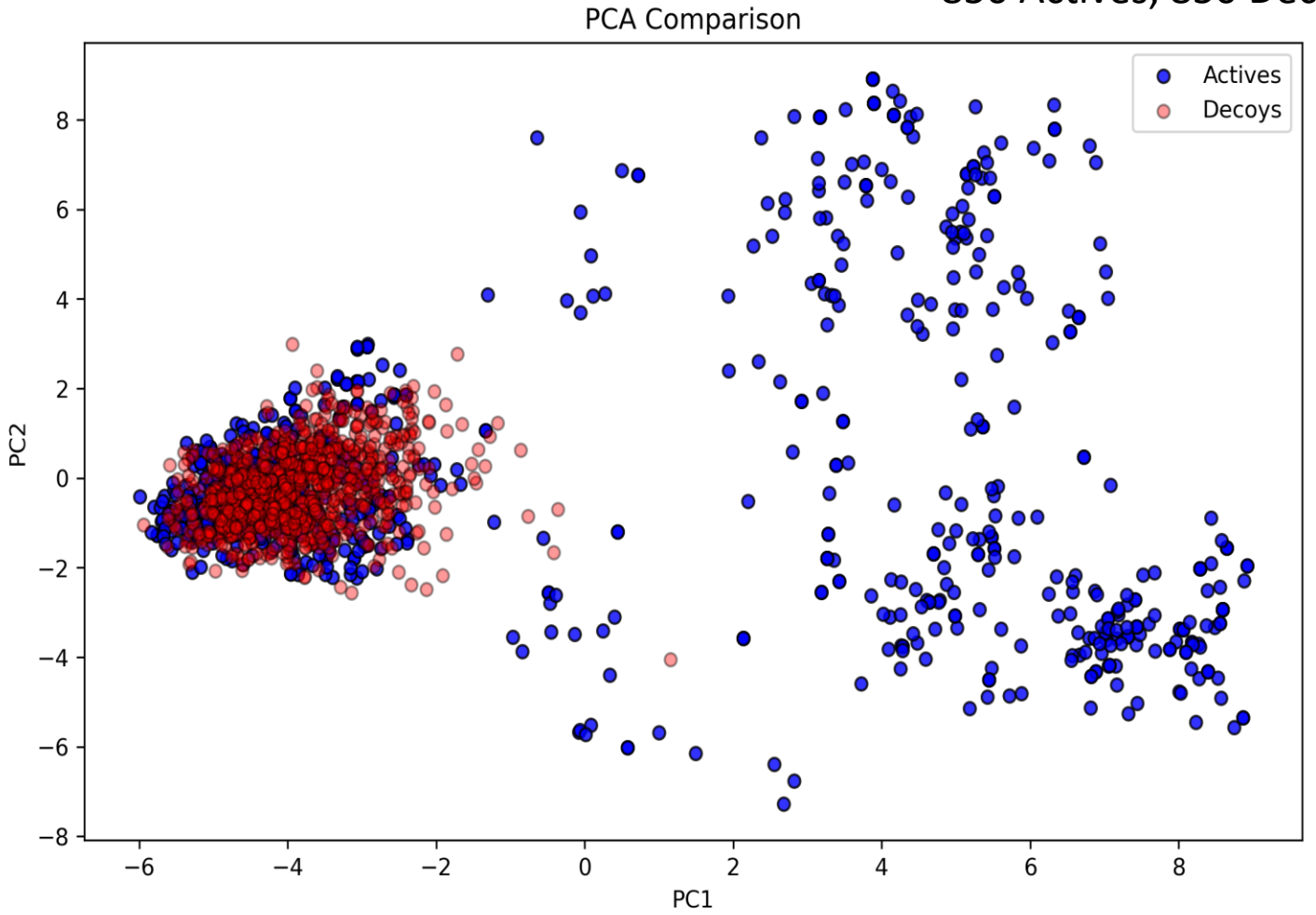
500 ns Molecular Dynamics Simulation

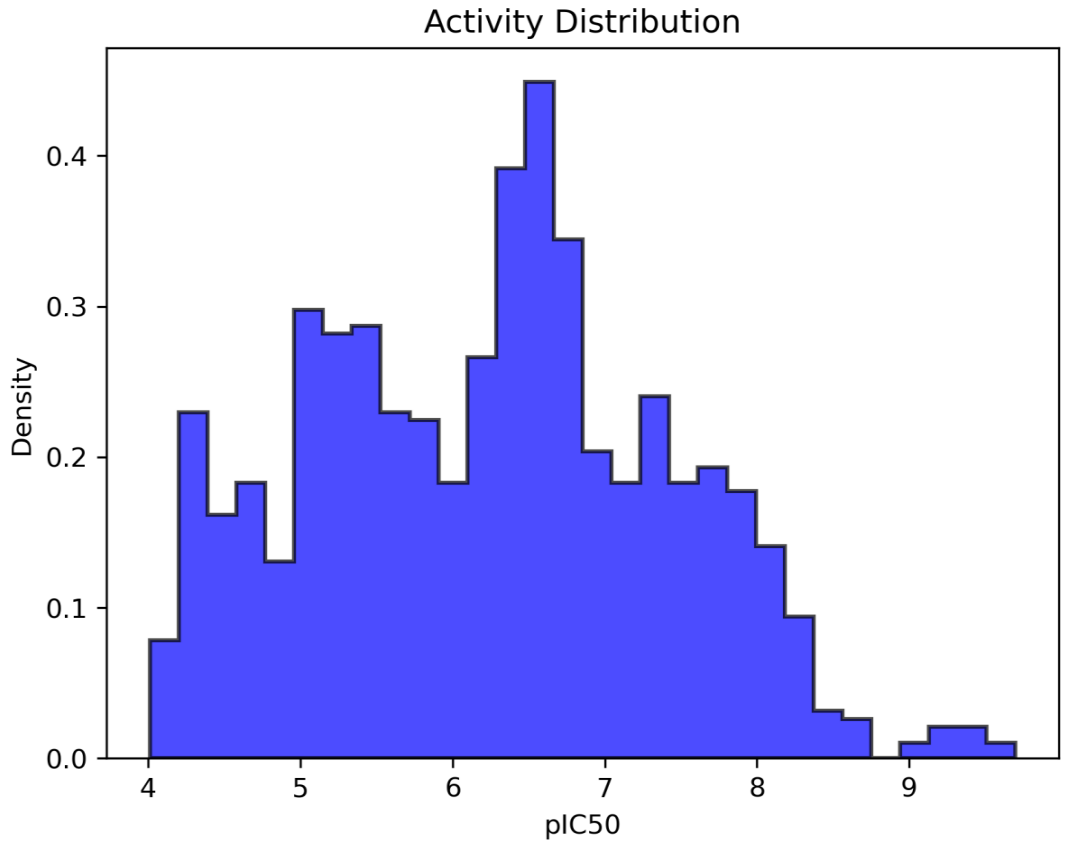




Docking Dataset Preparation

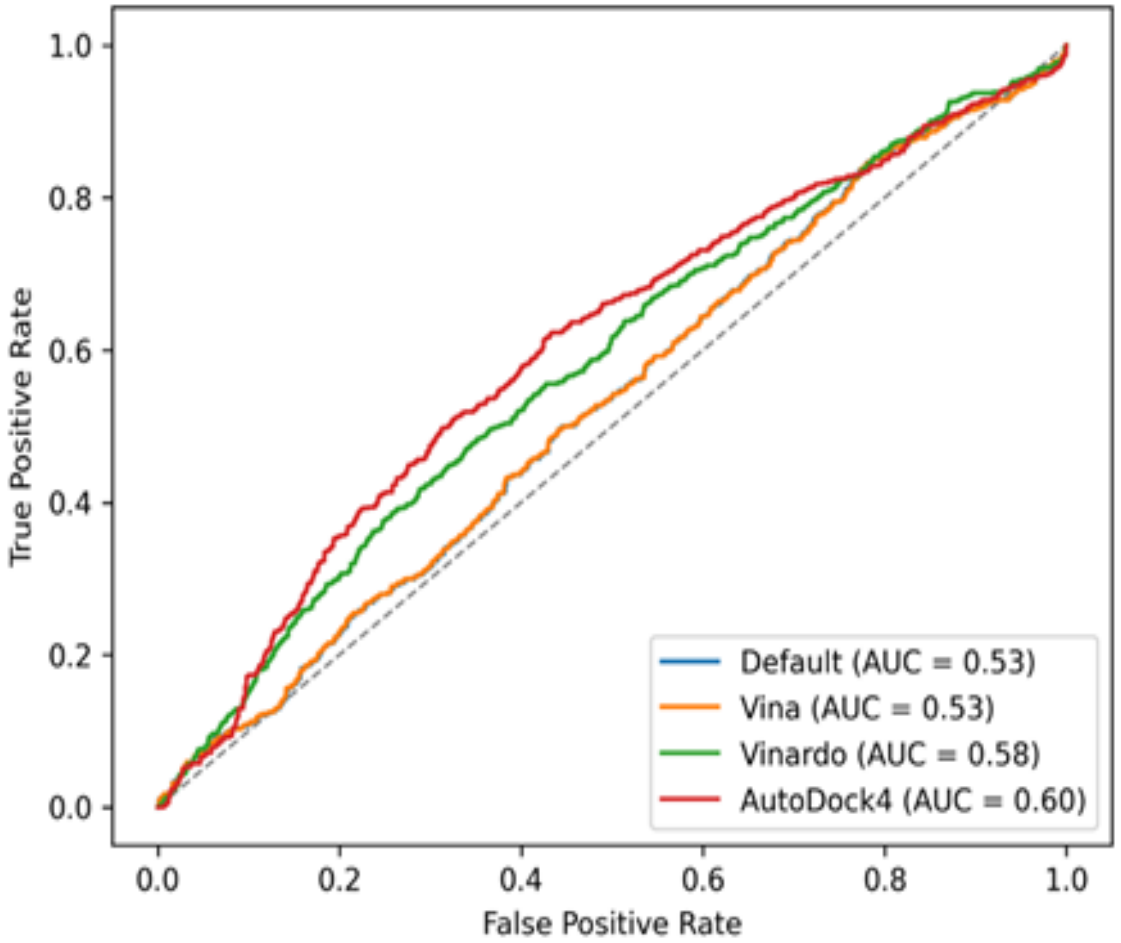
850 Actives, 850 Decoys



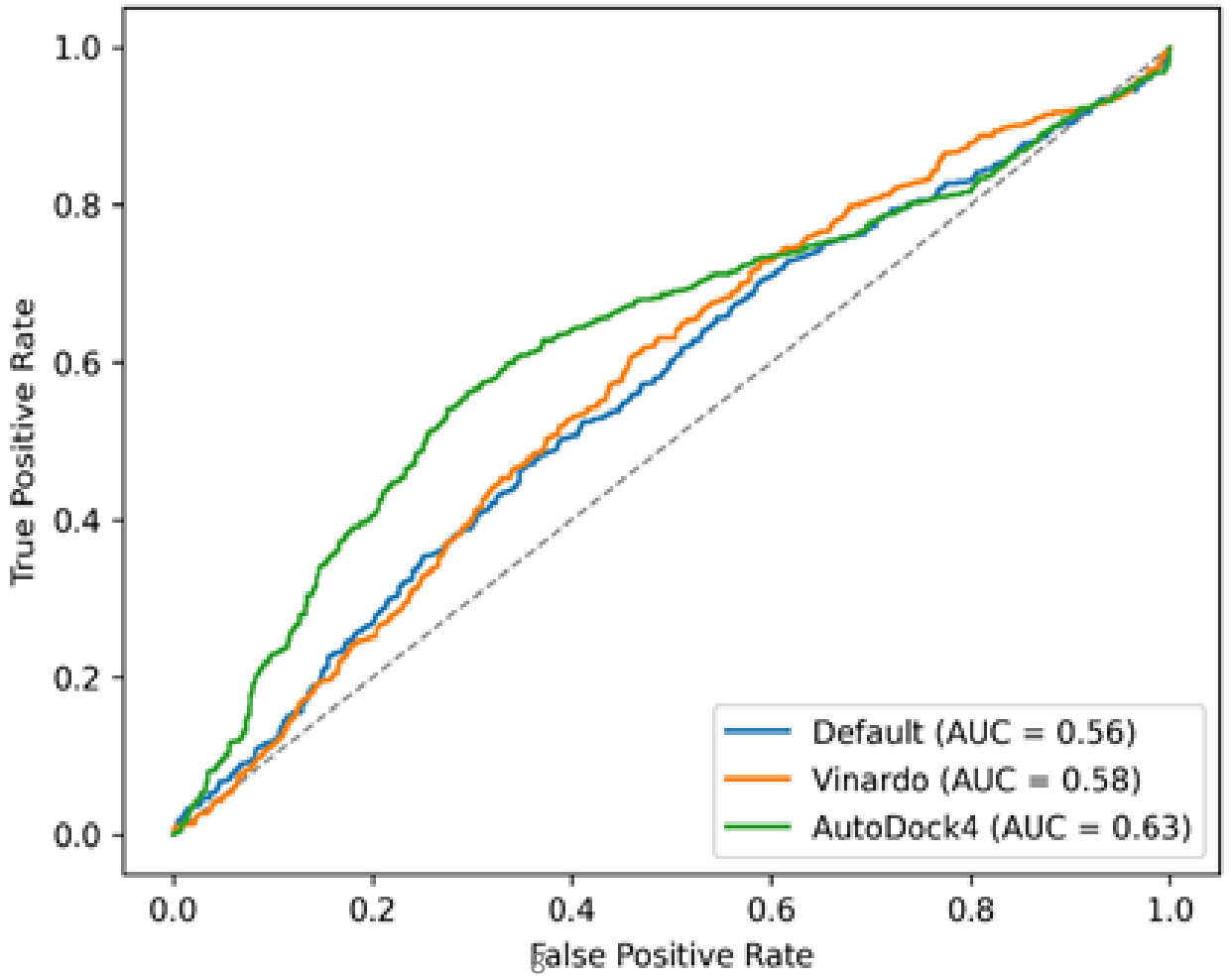


Semi-Rigid and Flexible Docking

Semi-Rigid Docking

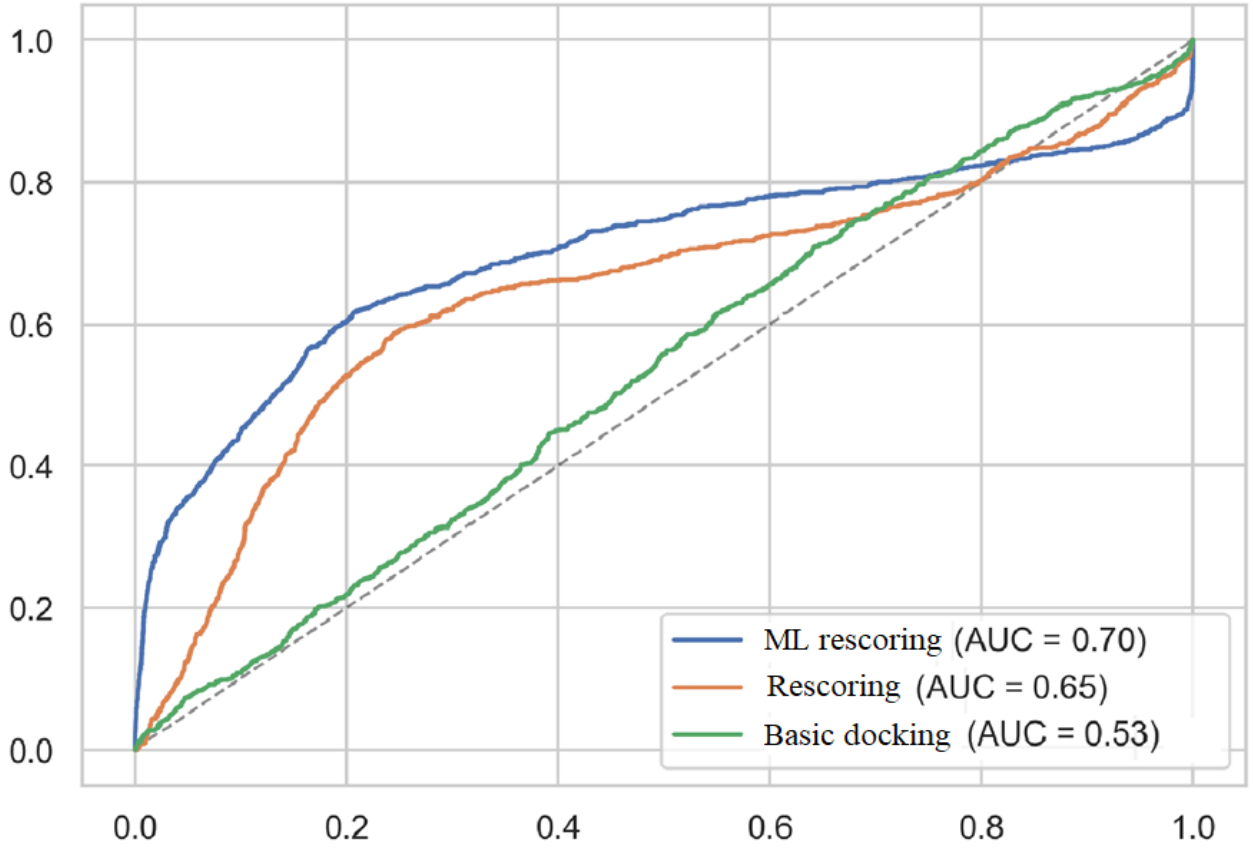


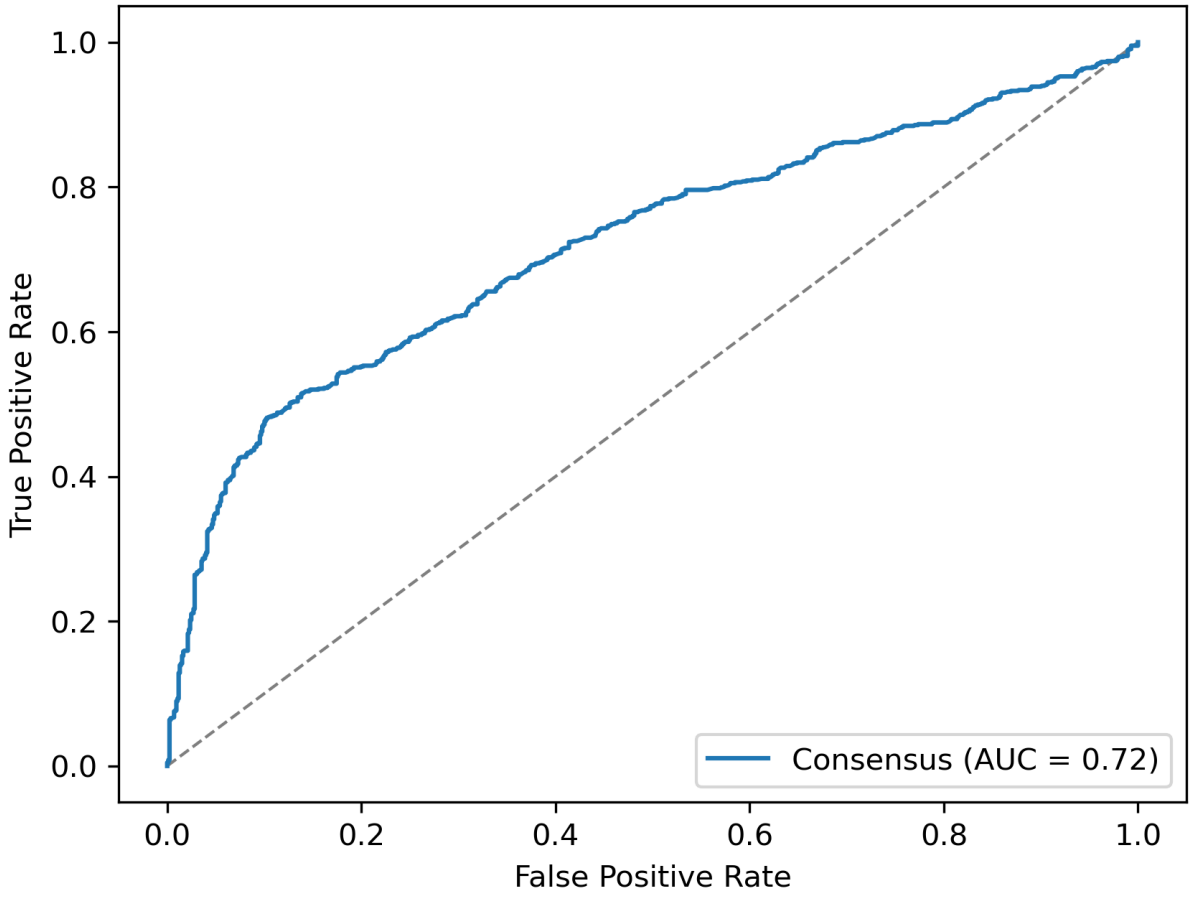
Flexible Docking



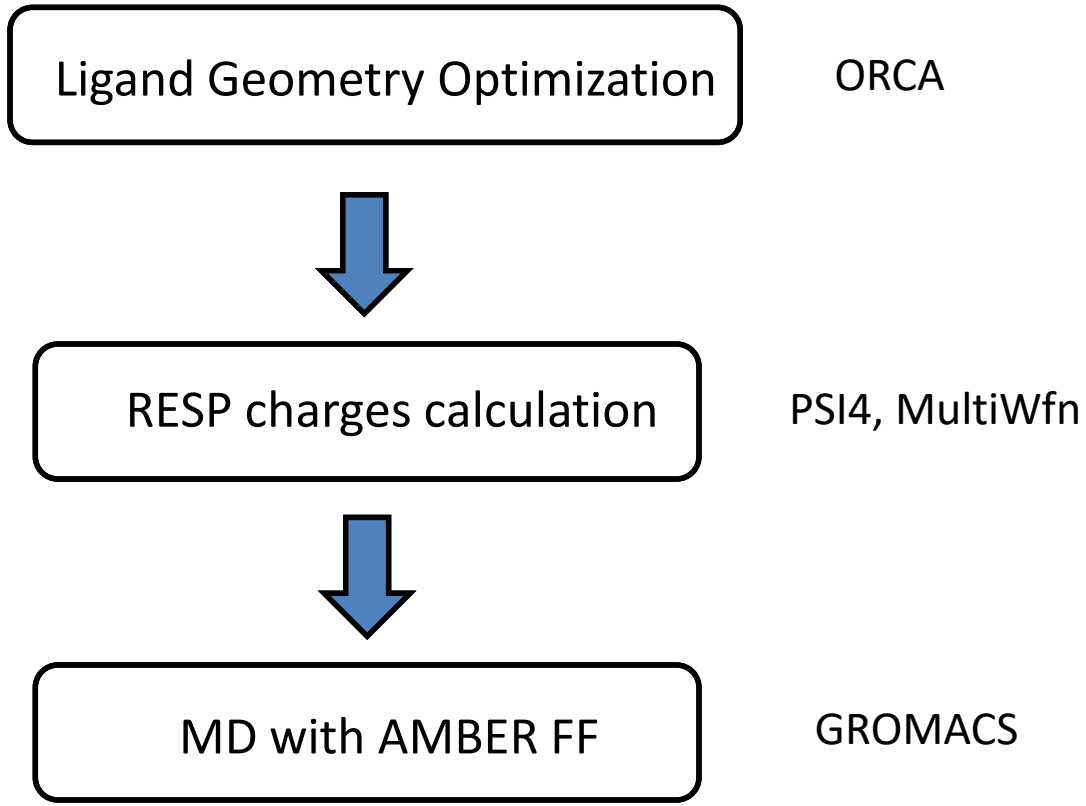
Optimized Scoring Functions

Rescoring – trained on binary classification, ML – trained on IC_{50}



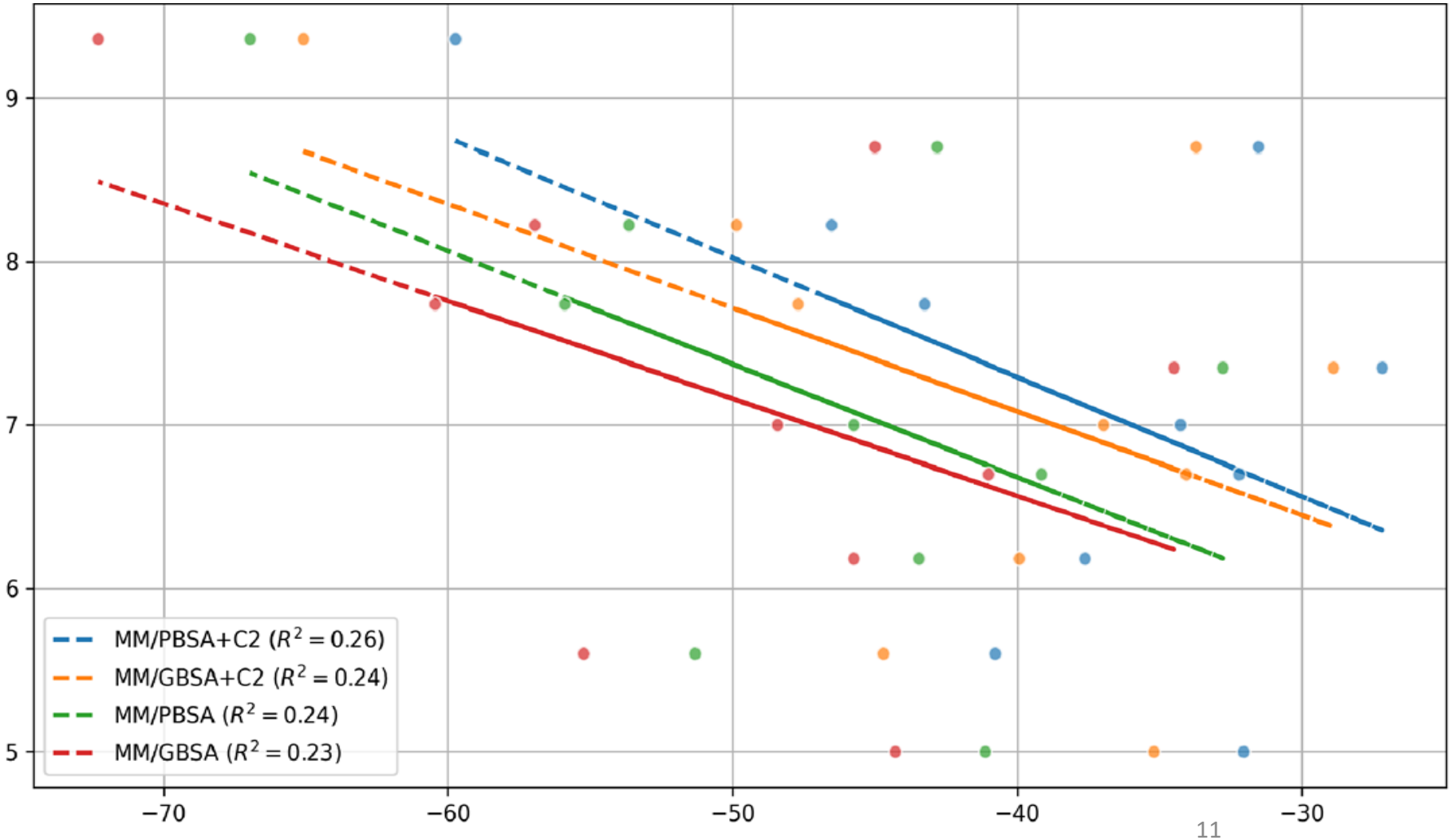


Molecular Dynamics Simulation Protocol



MM-PB(GS)SA

Predicted Binding Energy vs IC_{50}



Conclusions & Next Steps

- Robust MPro Virtual Screening Pipeline
- Best QSAR: SVM Regressor, pattern+avalon fingerprints
- Best Docking: consensus with regression- and ML-optimized SF
- Binding Energy: MM-PBSA with C2 Entropy
- Next: ABFE calculations via mBAR, ADMET predictions