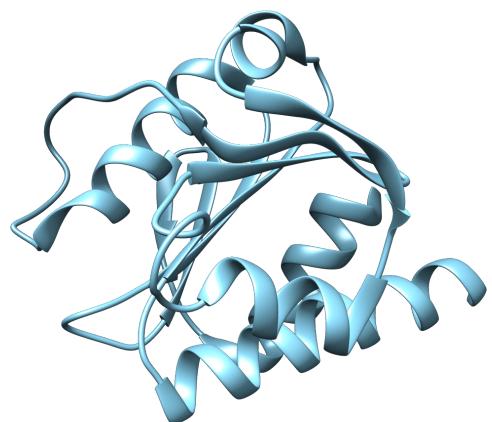


# Prediction Of Conformational Changes using Existing sTructures (POCCETs)

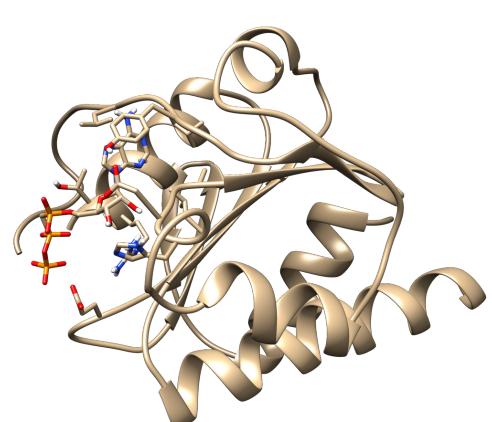
Aryan Surana and M.S. Madhusudhan

## 1. Apo and Holo Proteins

- Apo proteins:- no ligand bound
- Holo proteins:- have a bound ligand



Apo protein

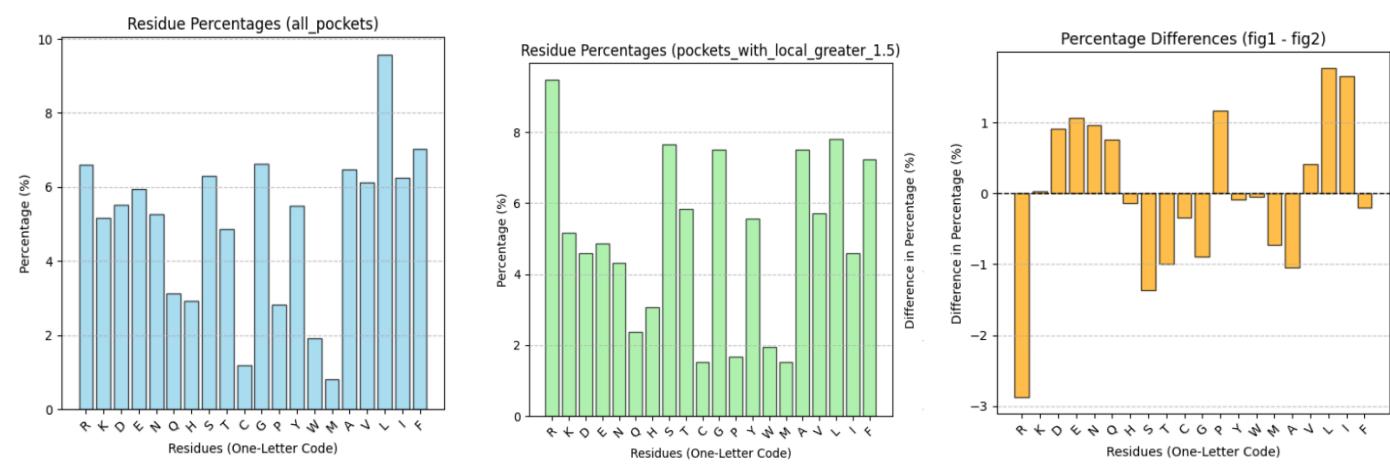


Holo Protein

## 4. Dataset filtration

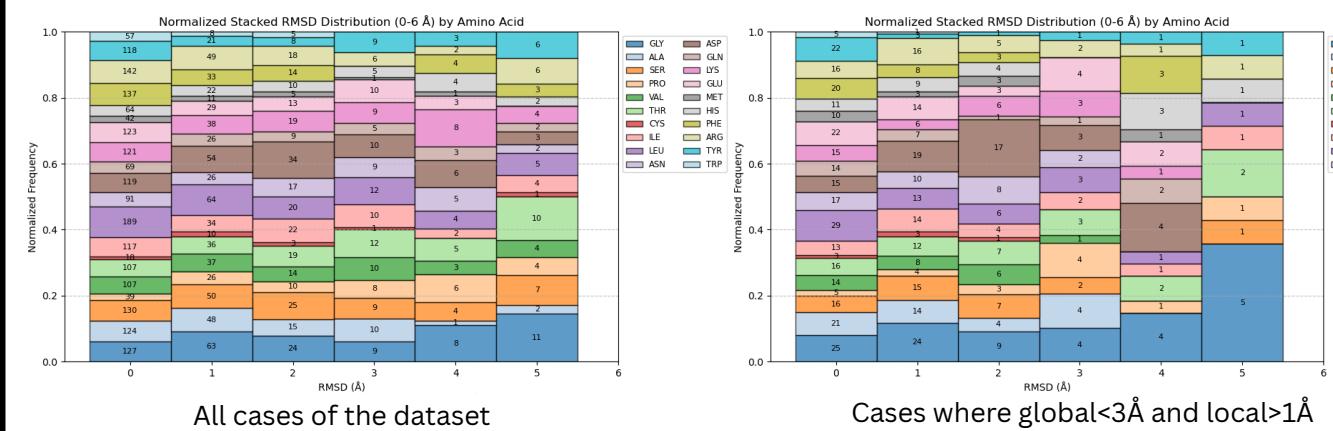
- Common Uniprot IDs, with atleast one apo and one holo.
- Filtered out NMR structures.
- Resolution cutoff 3.5 Å.
- Ligands between 100-1000 Da.
- Dataset of 798 protein pairs.

## 6. Residue wise change



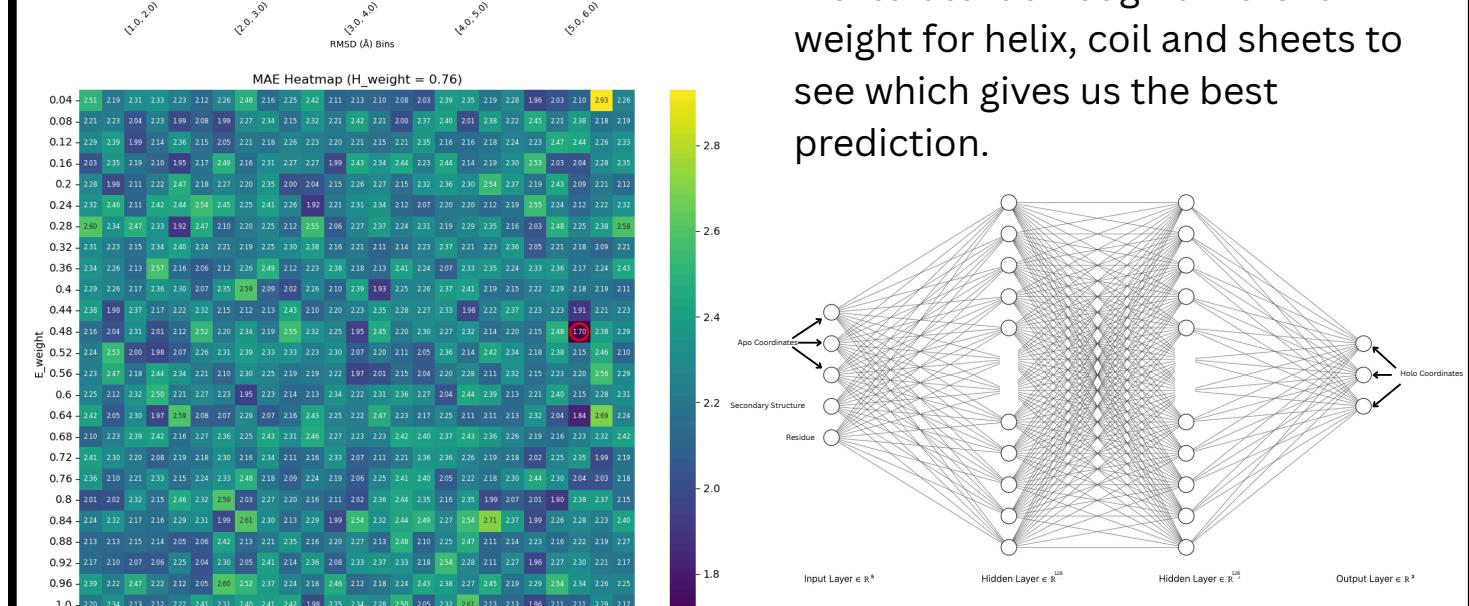
- Variation is high in residues with larger side chains which are very flexible. Not a good proxy of how much the residue moves

## 8. Variation in Ca



## 9. Developing a neural network

- We design a neural network with 2 hidden layers of 128 neurons each which takes in the apo-coordinates, secondary structure and residue and trains itself on 80% of the 691 residues from 62 binding sites.
- We iterate it through different weight for helix, coil and sheets to see which gives us the best prediction.



- The best MAE is 1.7 Å for high C, H weights and low E weight, which makes sense, much fewer cases for sheet so should have low weight.
- 10/138 within 0.5 Å, 28/138 within 1 Å and 60/138 within 2 Å.

## 12. References

- Holo Protein Conformation Generation from Apo Structures by Ligand Binding Site Refinement 2022. Jinse Zhang et al.

## 2. Aim of the Project

- Using Apo and Holo proteins to train a neural network to predict conformational changes in binding sites.

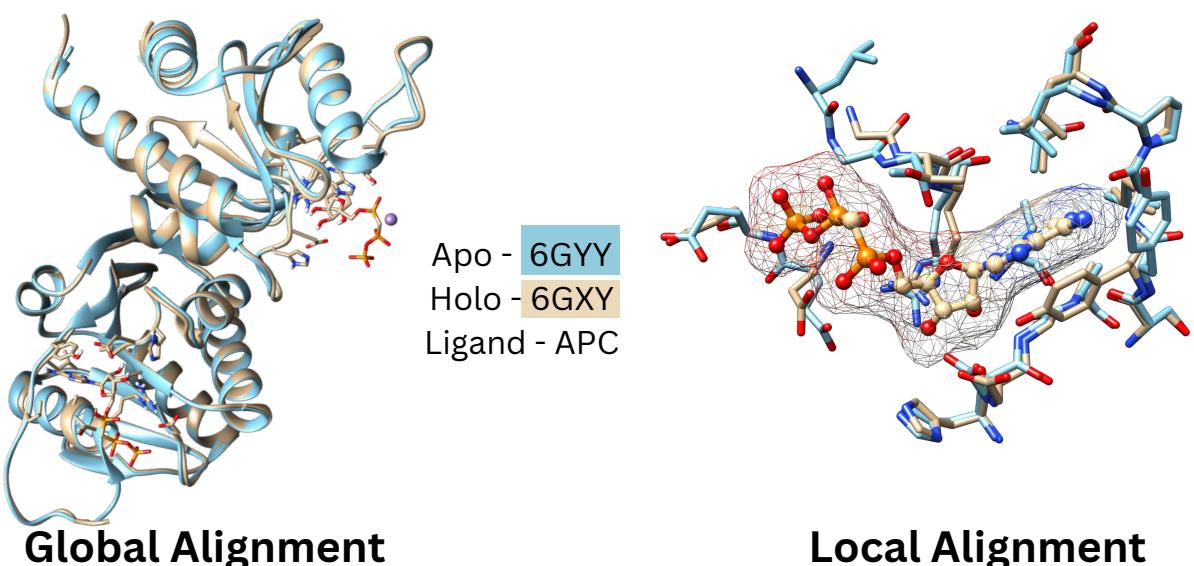
## 3. 3D Least Square Algorithm

- Translation matrix T and Rotation matrix R to minimise Root Mean Square Deviation.
- Requires atom-atom correspondences.
- Feed in matrix X and Y and we get transformed Y.

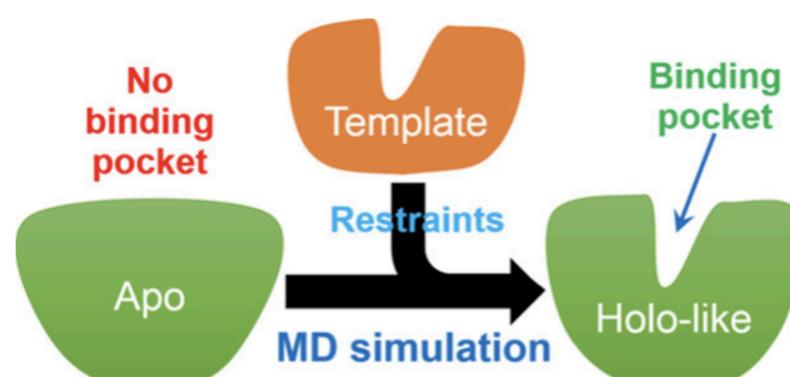
$$\text{RMSD} = \sqrt{\frac{1}{n} \sum_{i=1}^n \| \mathbf{X}_i - (\mathbf{Y}_i \mathbf{R} + \mathbf{T}) \|^2}$$

## 5. Aligning the proteins and binding pockets

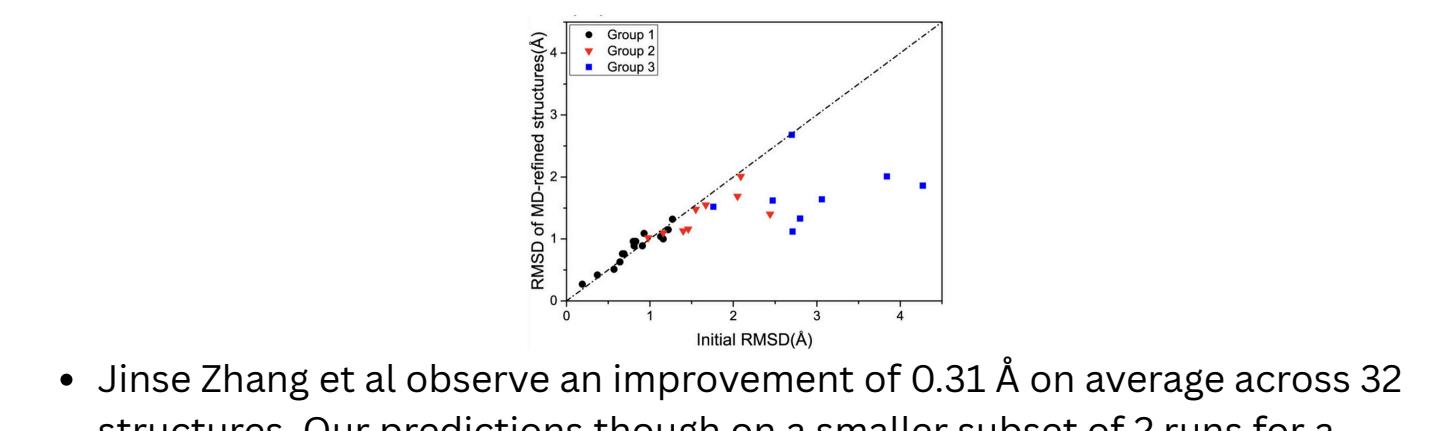
- Pairwise match the common Uniprot ID chains.
- Set of common residues from both files and Least Square Algorithm applied.
- 5 Å binding pocket in holo protein to compare local change.
- 5300 binding pockets, 1300 in common chains.
- 62 cases with global < 3 Å and local > 1 Å, of which 49 are closing and 13 are opening



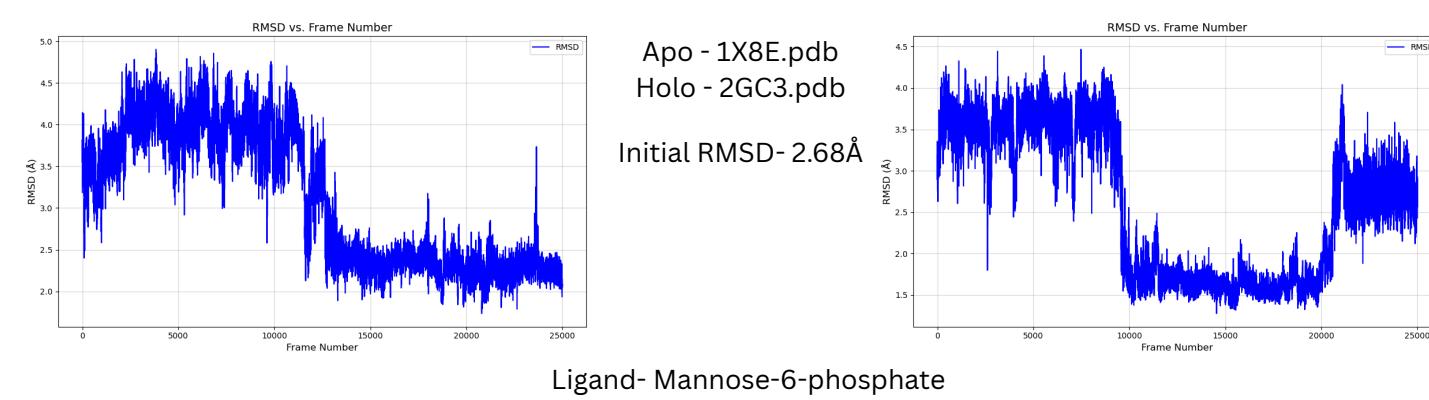
## 7. Prediction of pockets using MD Simulation



- Proteins are highly dynamic and here we check if we can observe the binding pocket as one of the states when we run an MD simulation for 250 ns on 1x8e.pdb.



- Jinse Zhang et al observe an improvement of 0.31 Å on average across 32 structures. Our predictions though on a smaller subset of 2 runs for a protein is improved by 1.205 Å, which is 4 times better.



## 10. Future work

- Refine the neural network to predict with lesser Mean Alignment Error (MAE).
- Account for opening and closing sites.

## 11. Acknowledgements

- COSPI lab members (Vipul for dataset, Harshita for Chimera X, Asita for MD)
- IISER Pune