Massachusetts Institute of Technology

EquiBind

Geometric Deep Learning for Drug Binding Structure Prediction

Hannes Stärk, Octavian-Eugen Ganea, Lagnajit Pattanaik, Regina Barzilay, Tommi Jaakkola

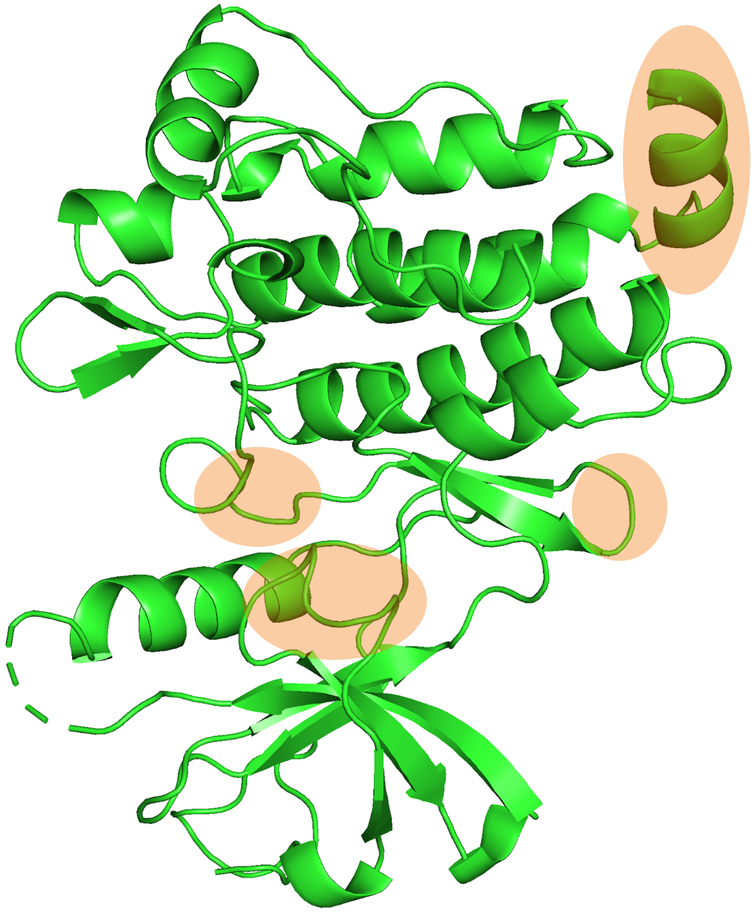
Protein ABL1

Controlled cell growth

Protein ABL1

Controlled cell growth

Mutated ABL1

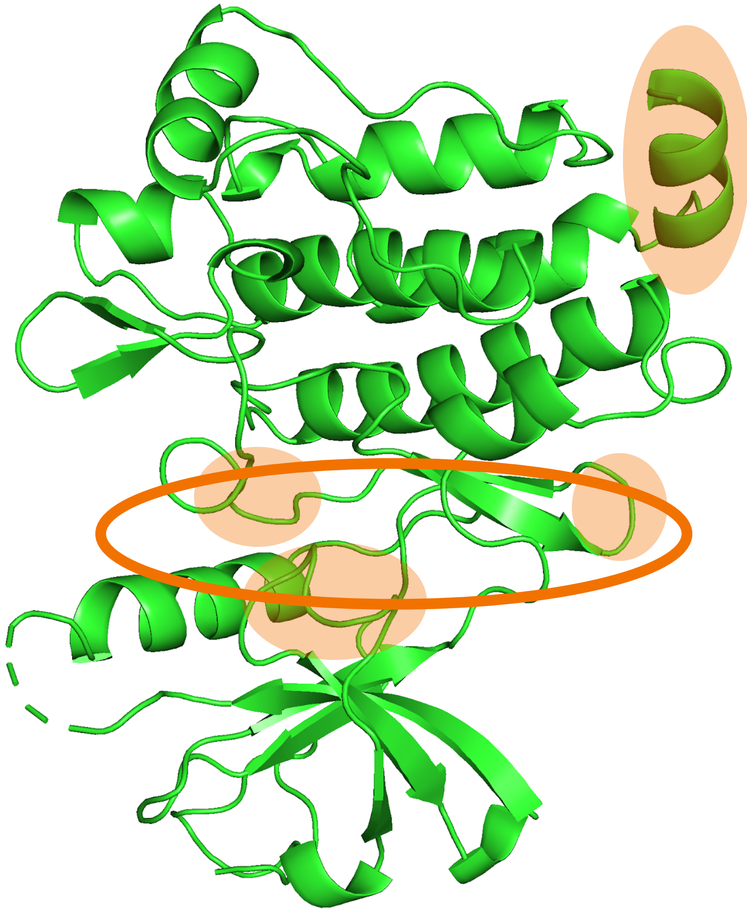


Uncontrolled cell growth

Protein ABL1

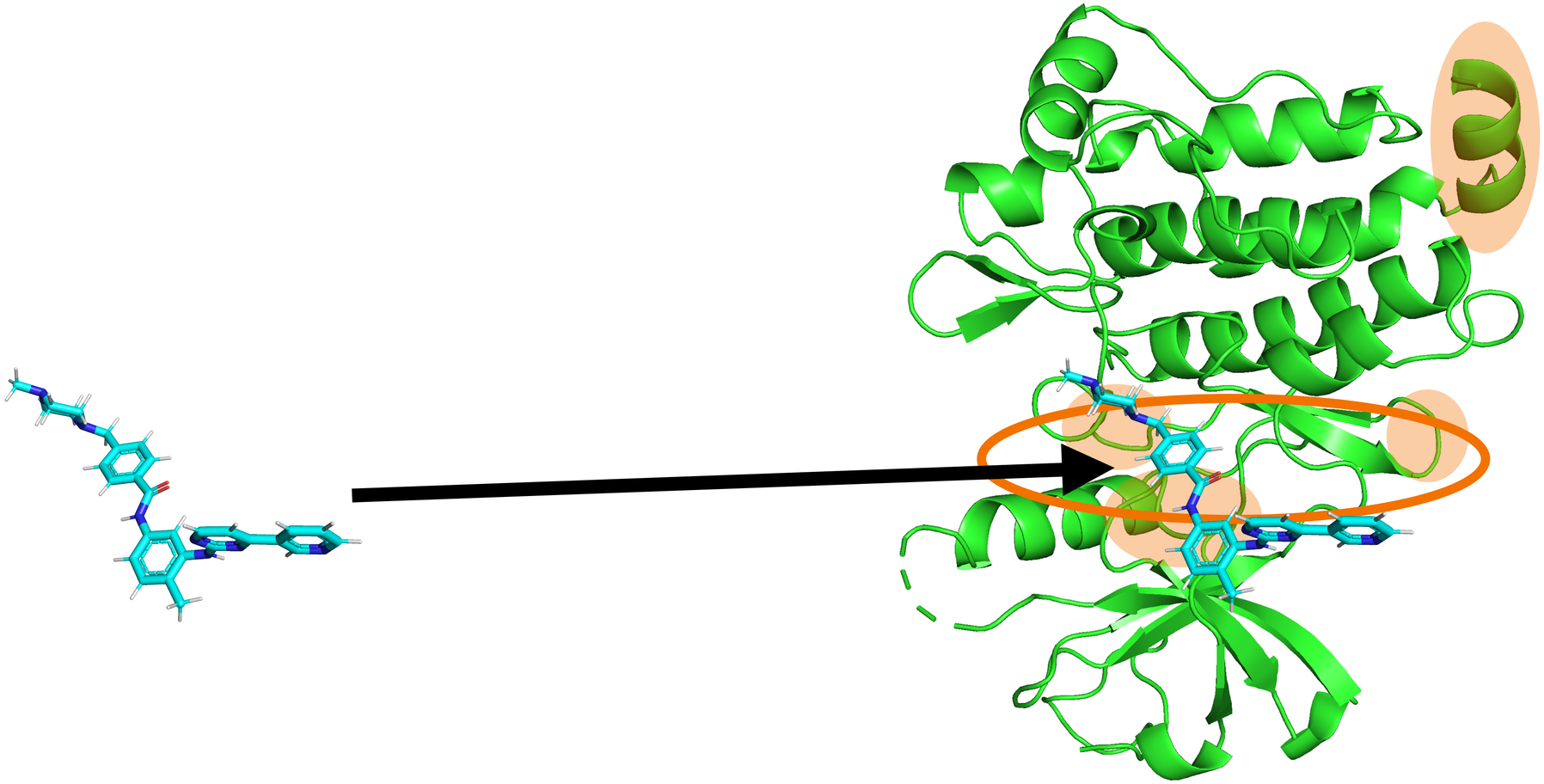
Controlled cell growth

Mutated ABL1



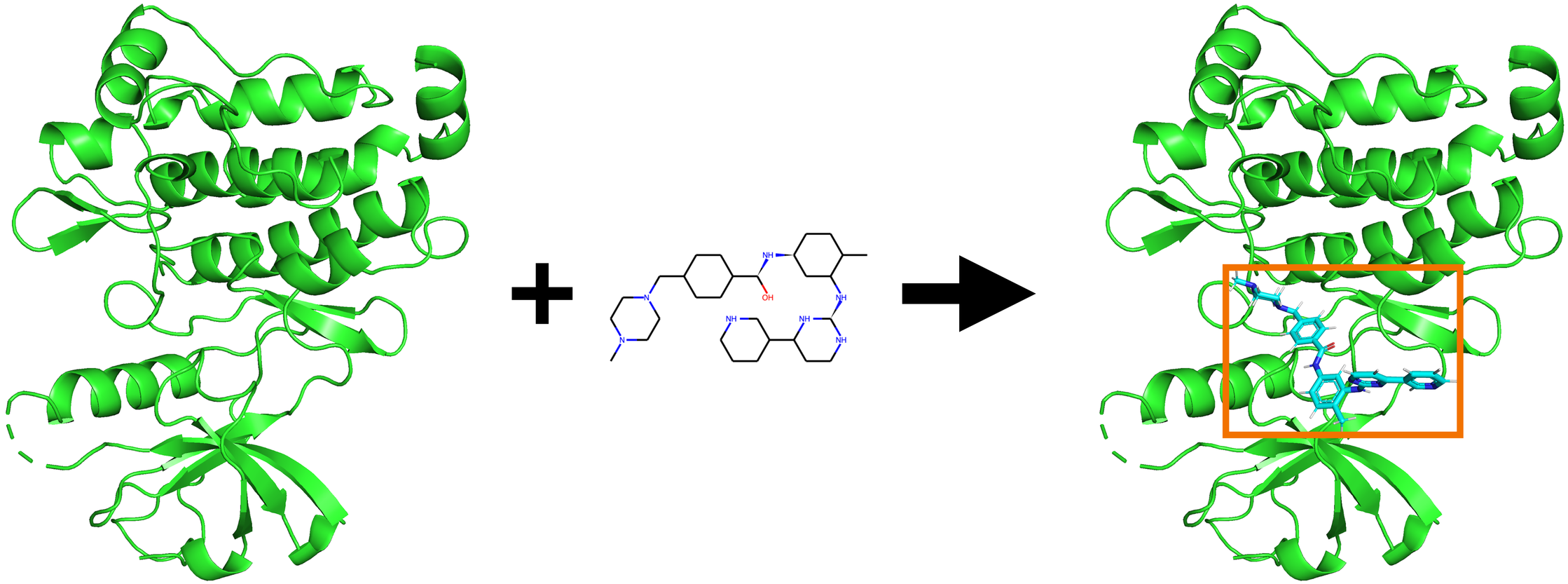
Uncontrolled cell growth

Restoring function



Small molecule binds to key region to stop abnormal behavior

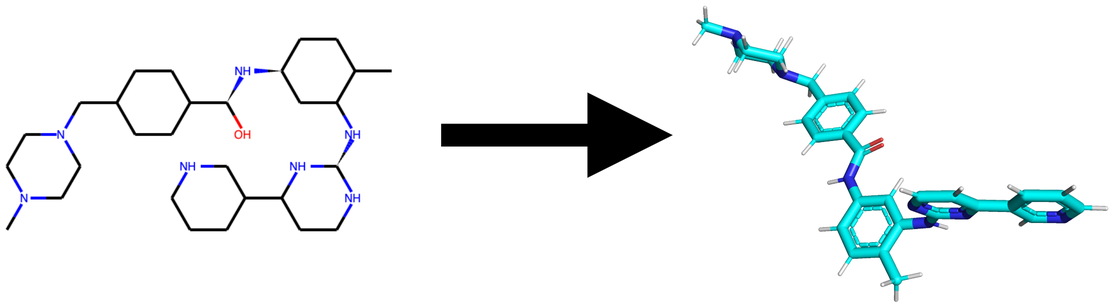
EquiBind: Binding Structure Prediction

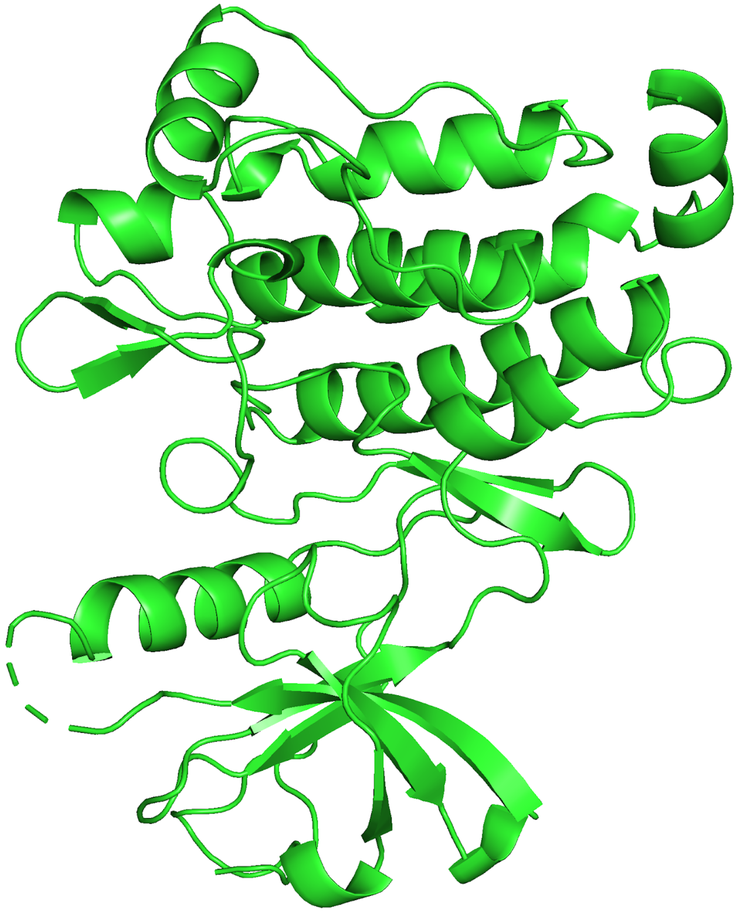


Input:
protein 3D structure + molecule

Output:
molecule 3D structure and location

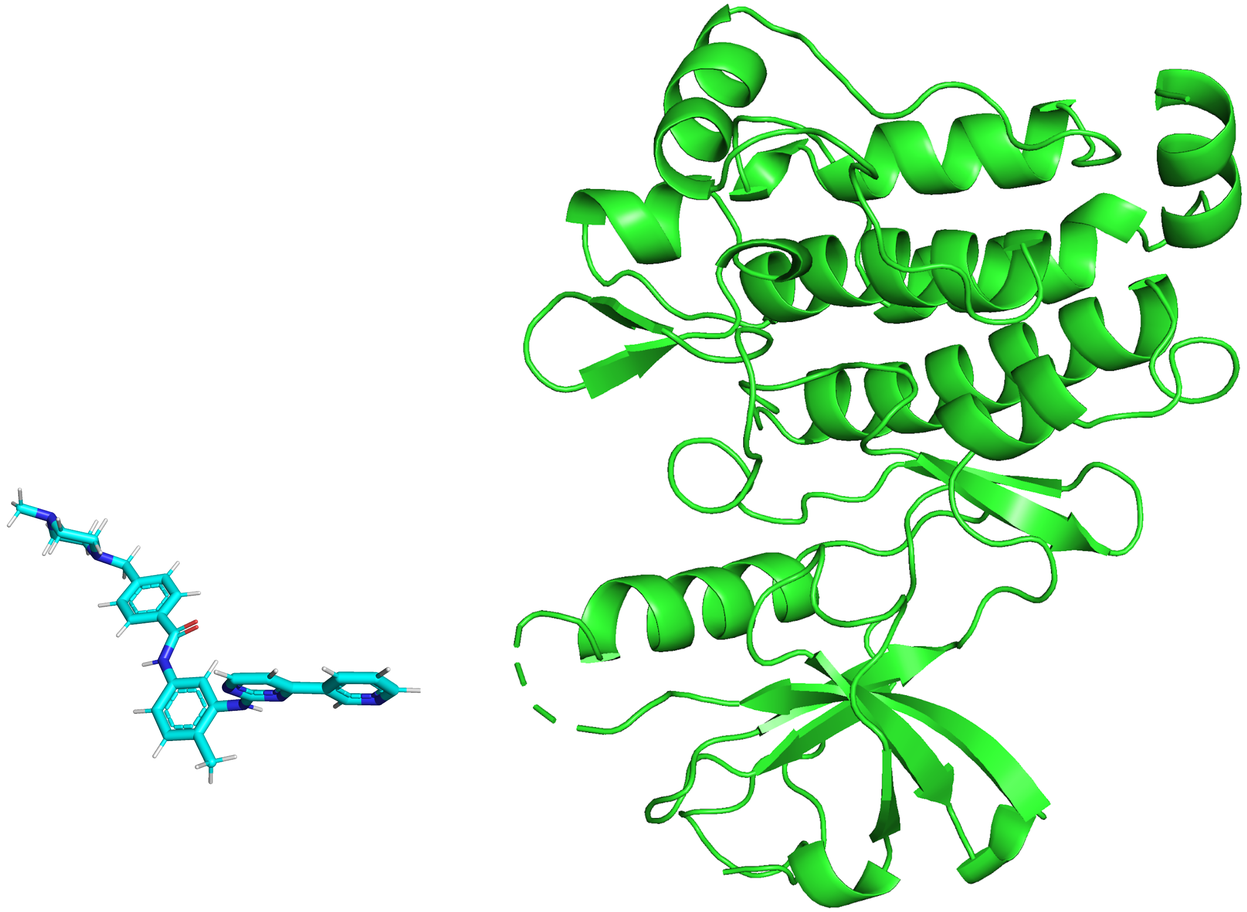
Symmetries the method needs to respect





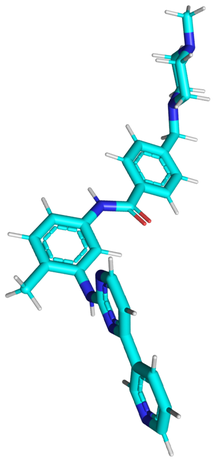
Translation and rotation invariance of initial molecule structure

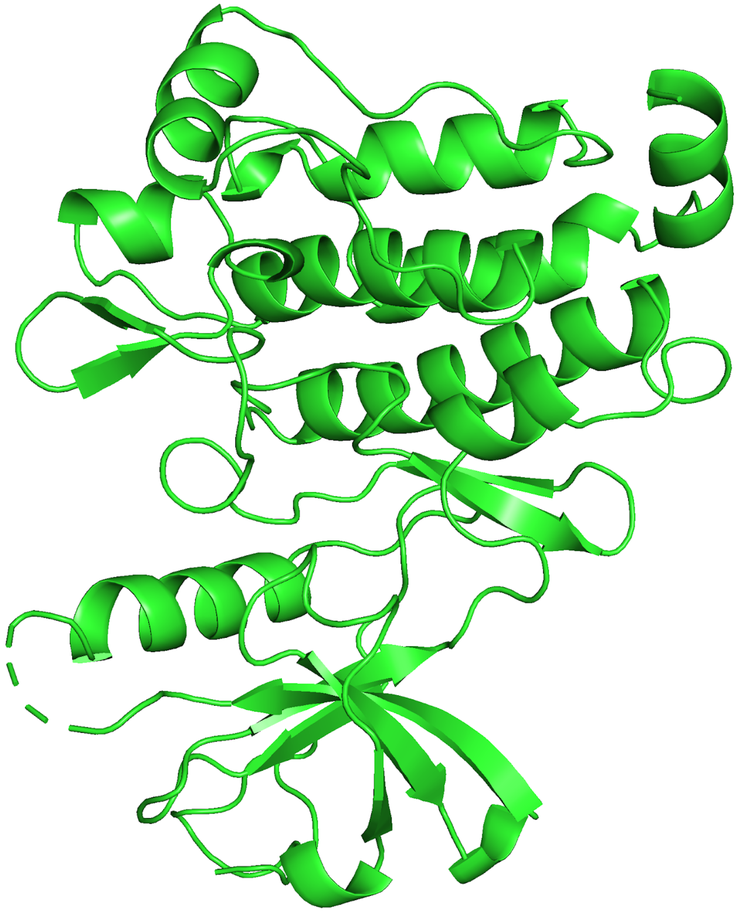
Symmetries the method needs to respect



Translation and rotation invariance of initial molecule structure

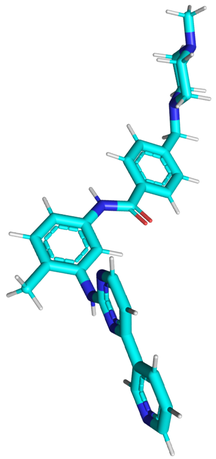
Symmetries the method needs to respect

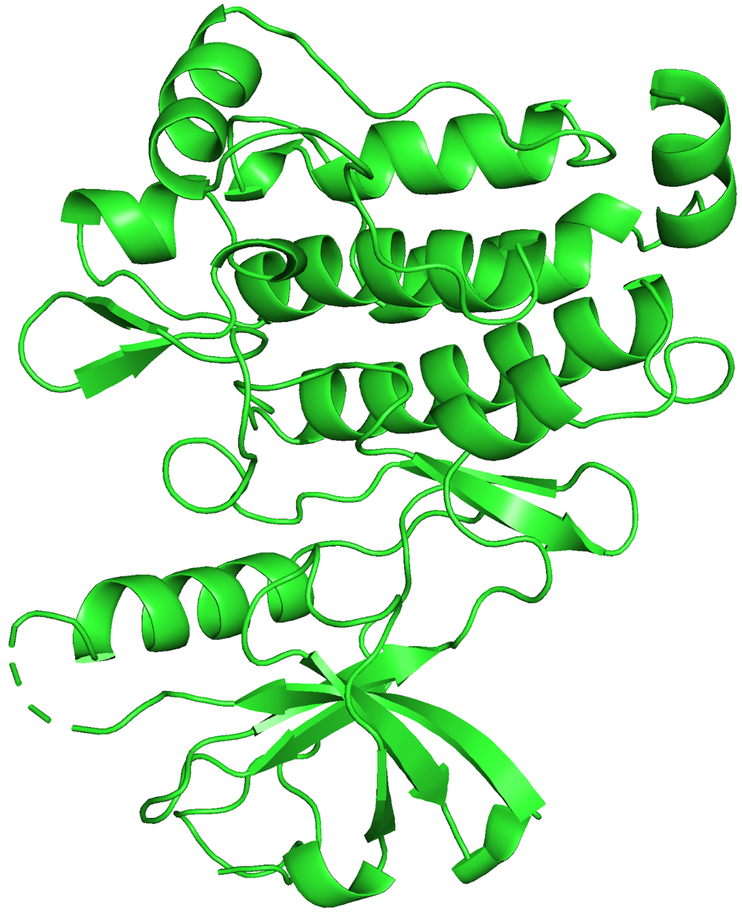




Translation and rotation invariance of initial molecule structure

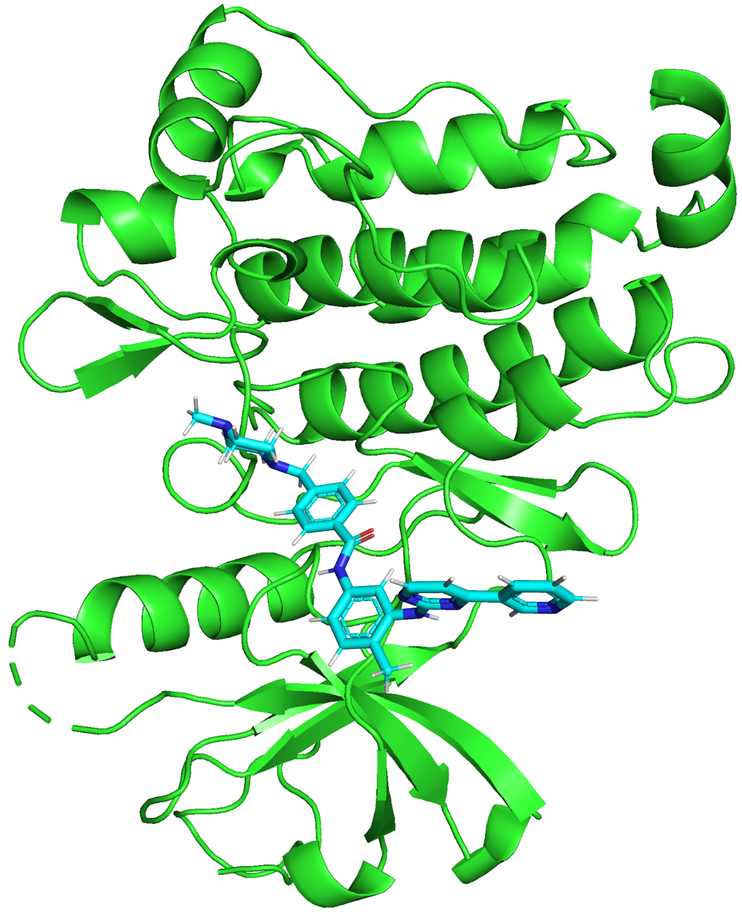
Symmetries the method needs to respect





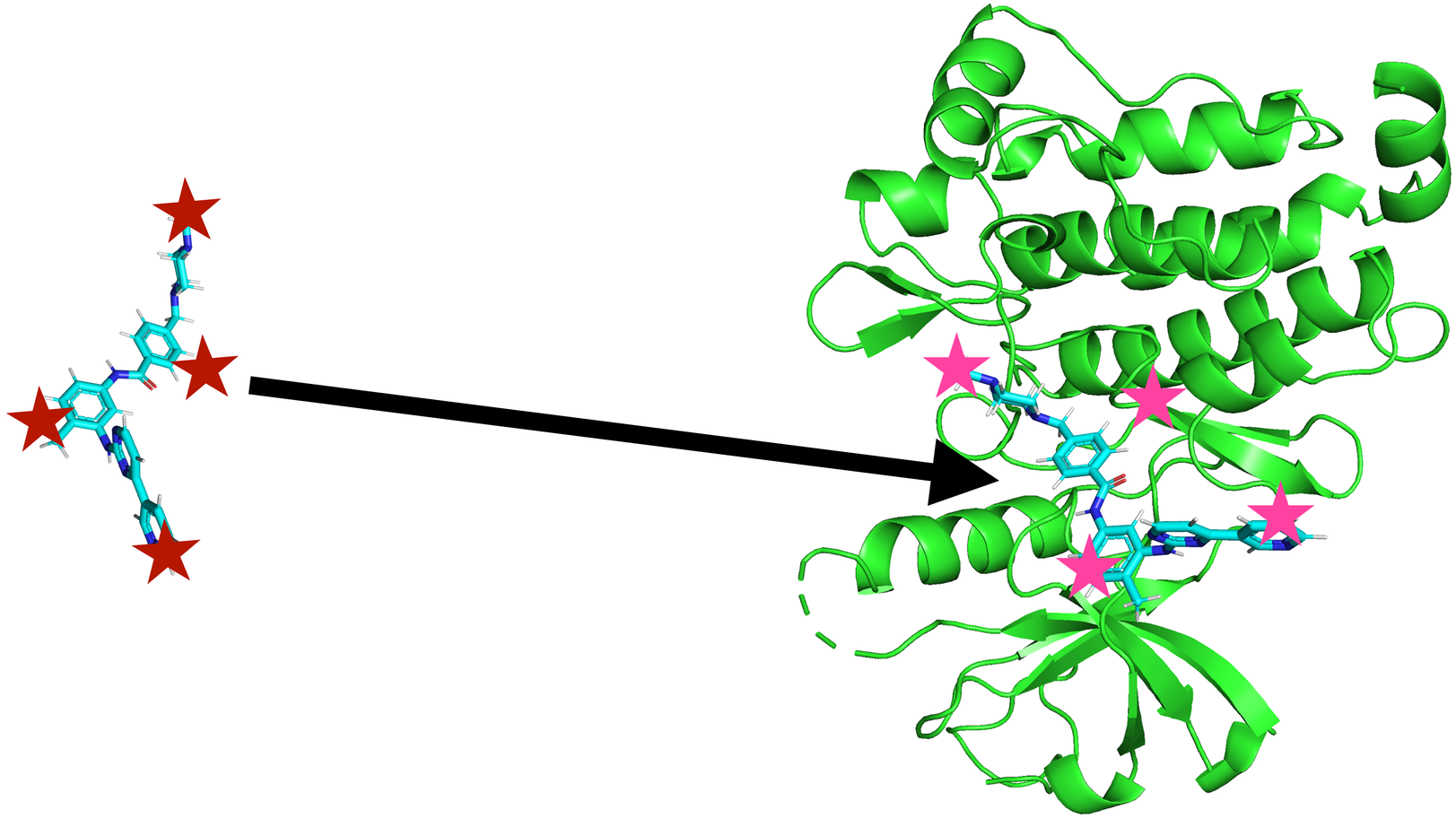
Translation and rotation invariance of initial molecule structure

Symmetries the method needs to respect



Translation and rotation invariance of initial molecule structure

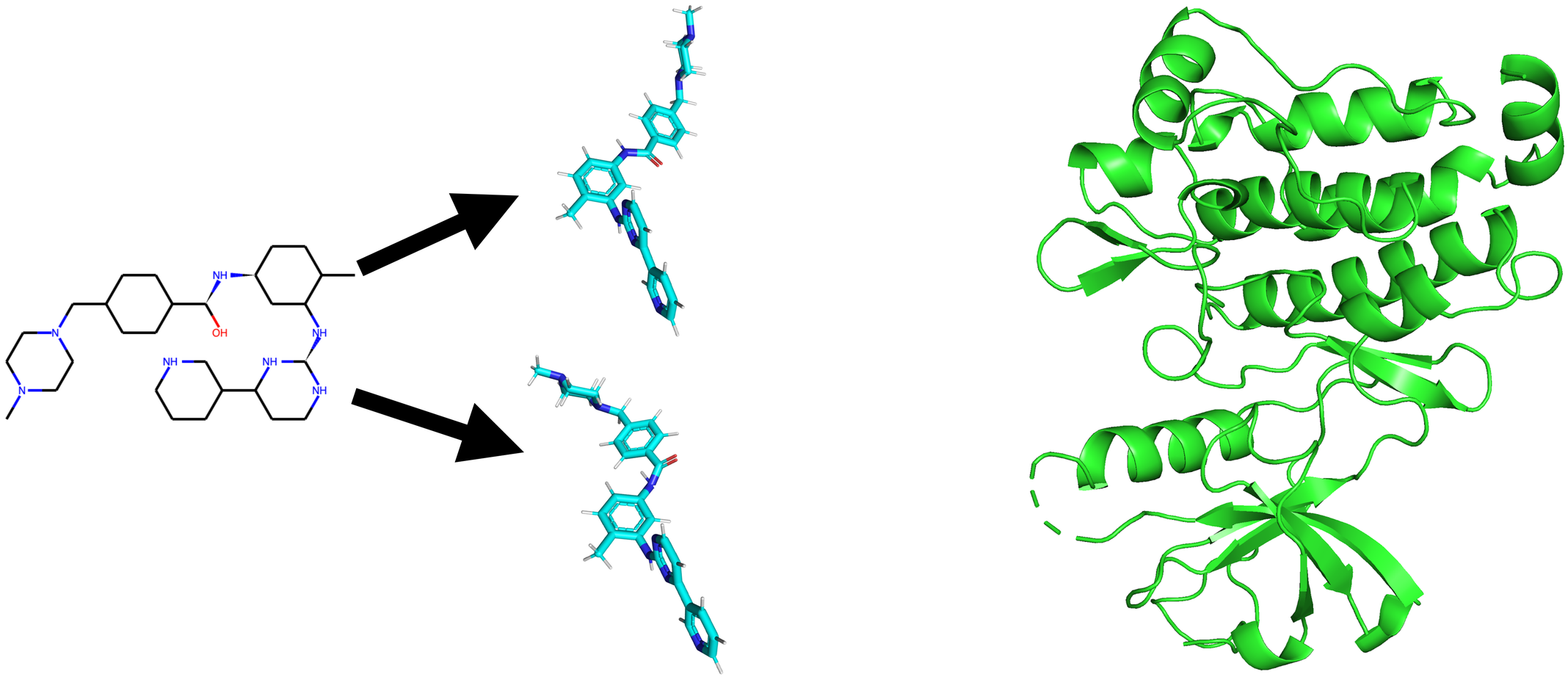
EquiDock: invariant rototranslation prediction



Kabsch algorithm calculates rototranslation to match keypoints

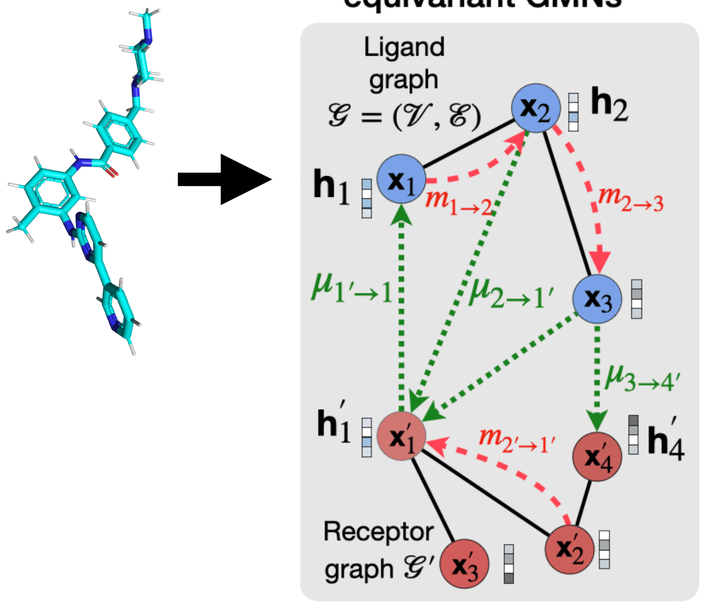
(EquiDock) Ganea et al. 2021, Independent SE(3)-Equivariant Models for End-to-End Rigid Protein Docking

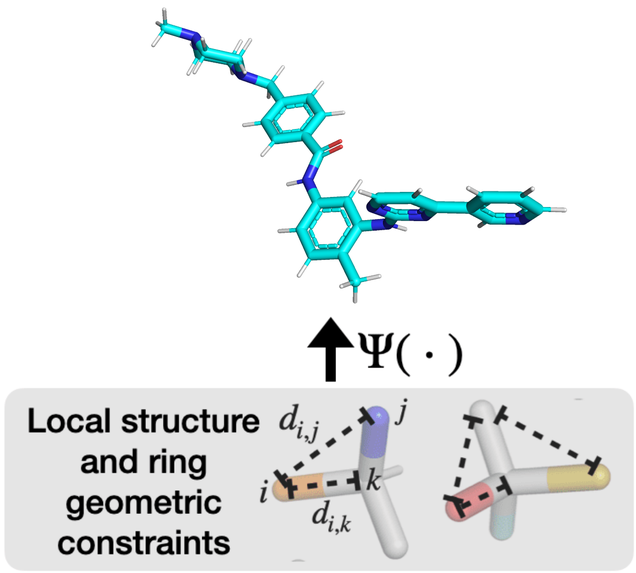
Finding the correct 3D molecule structure



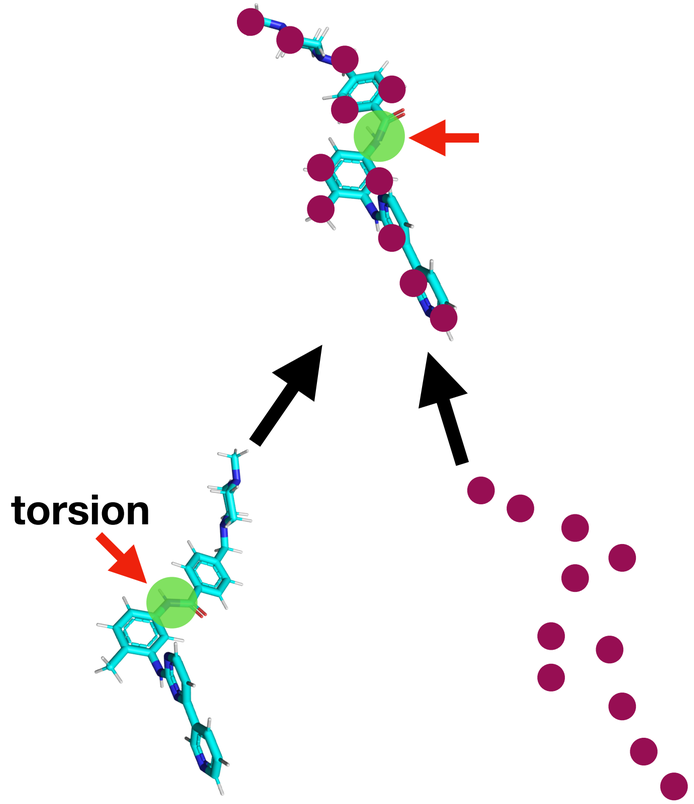
3D structure of small molecule is not rigid => model flexibility

Chemically plausible molecule flexibility

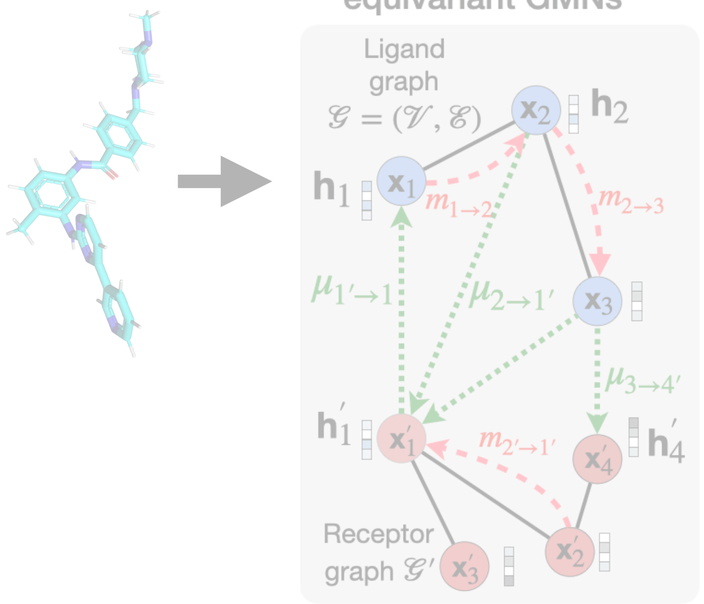


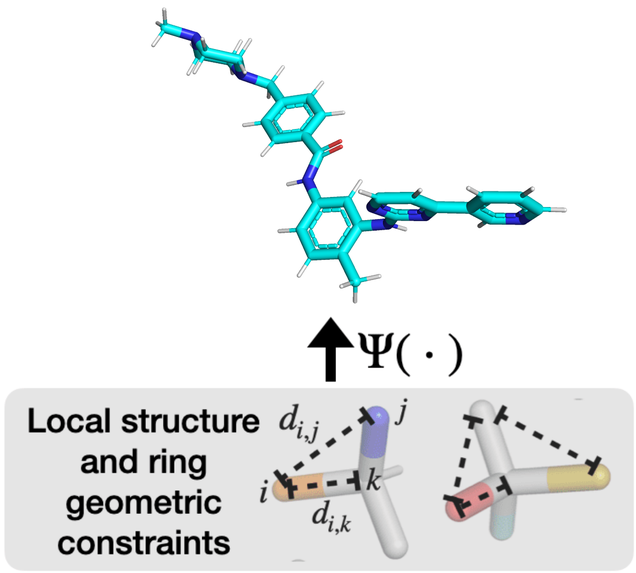


Efficiently change only torsion angles to fit molecule into atom point cloud

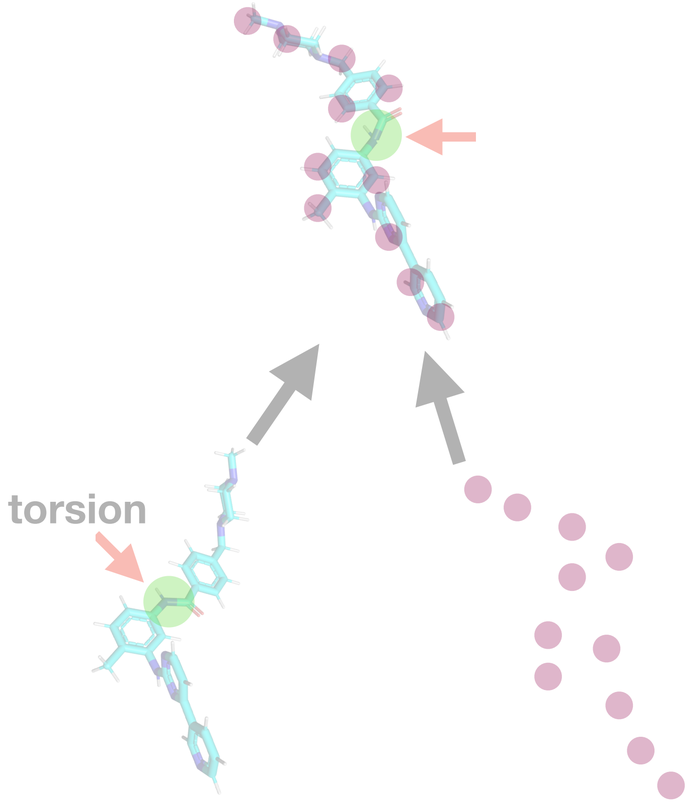


Chemically plausible molecule flexibility

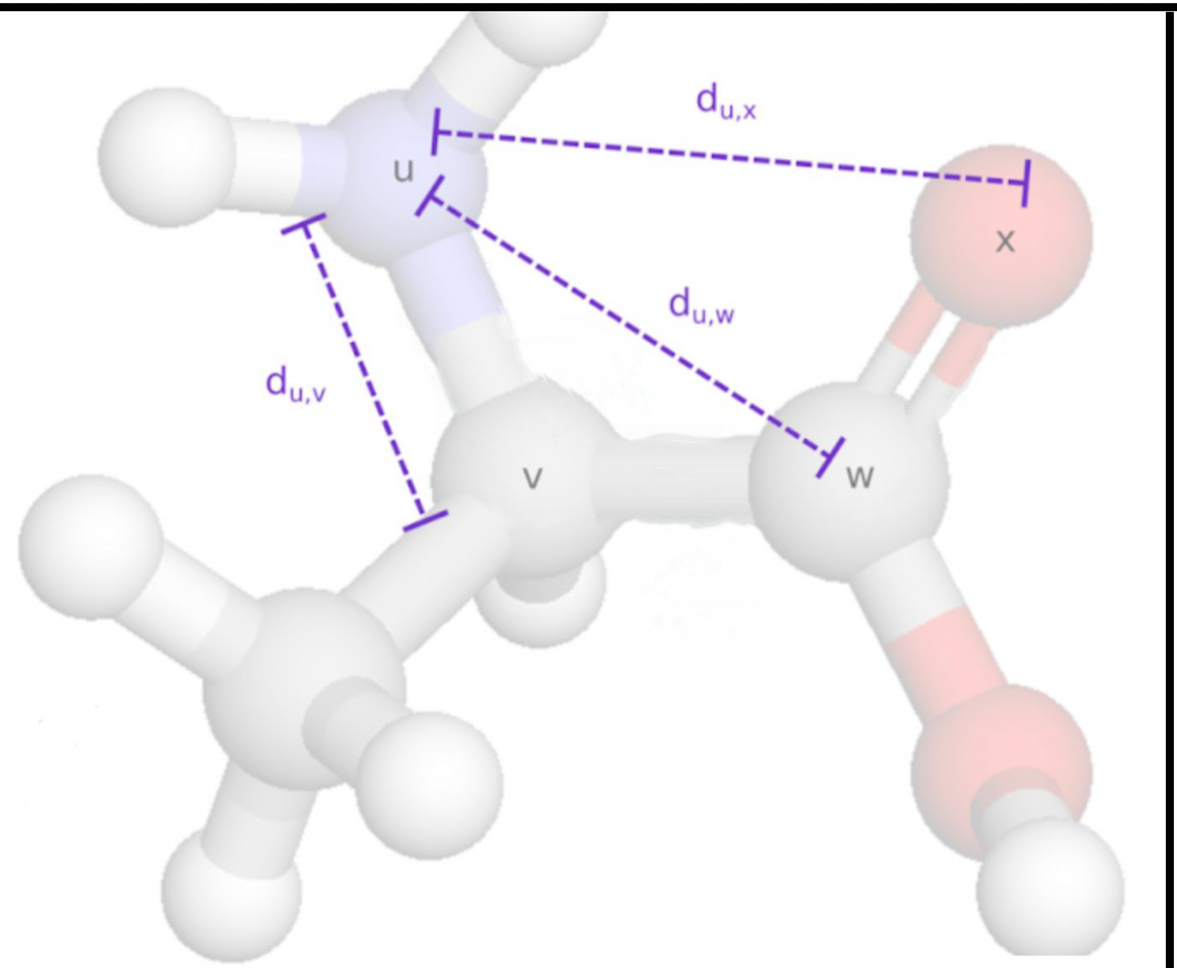




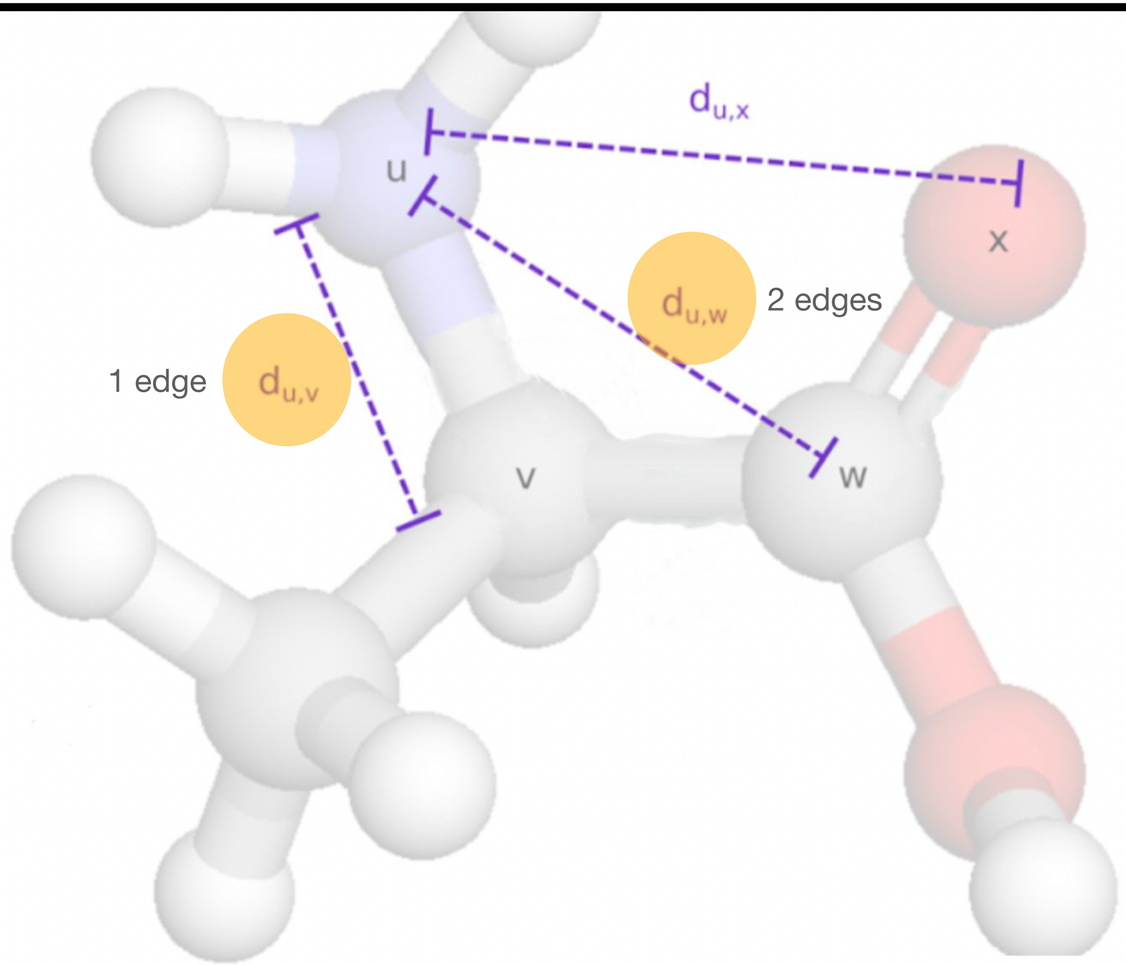
Efficiently change only torsion angles to fit molecule into atom point cloud



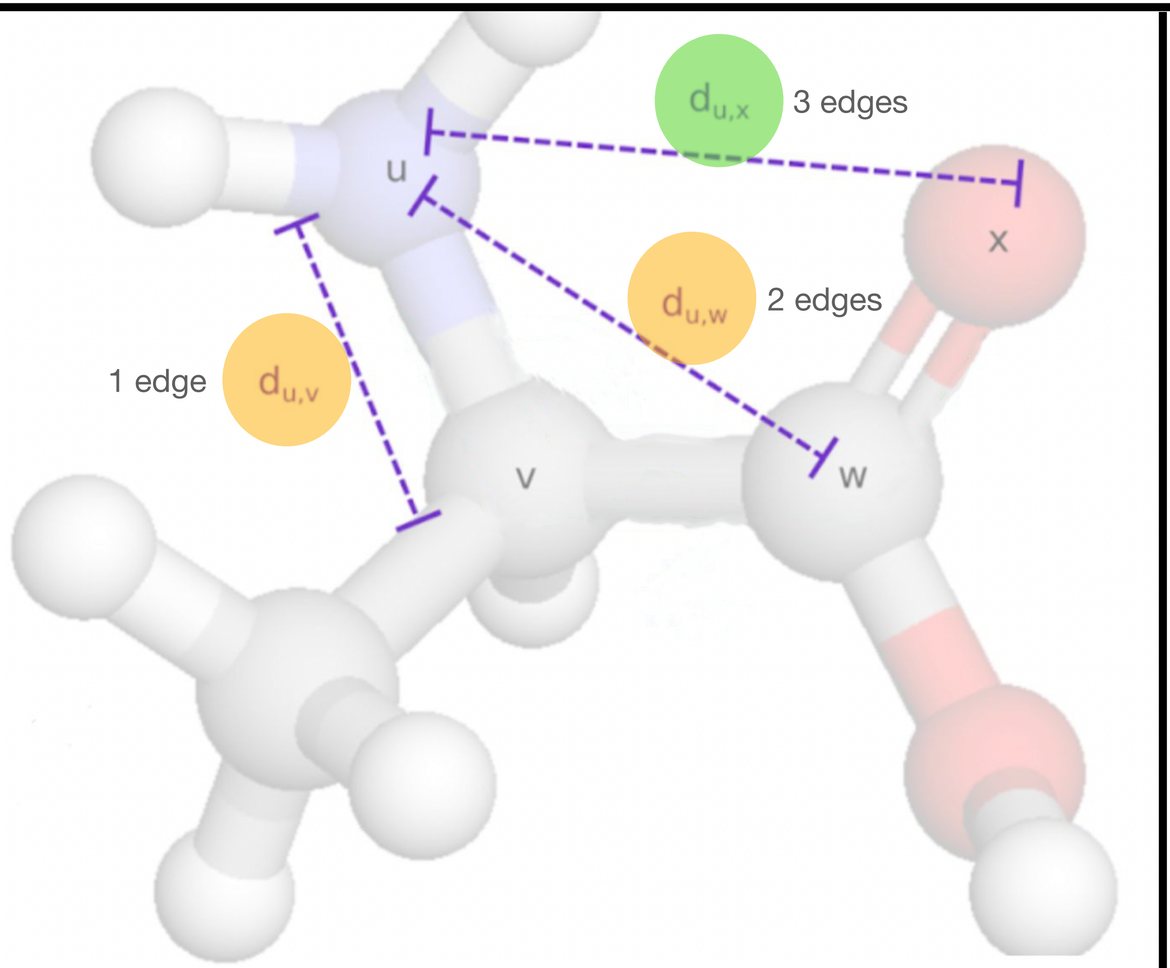
Realistic inter-atom distance changes



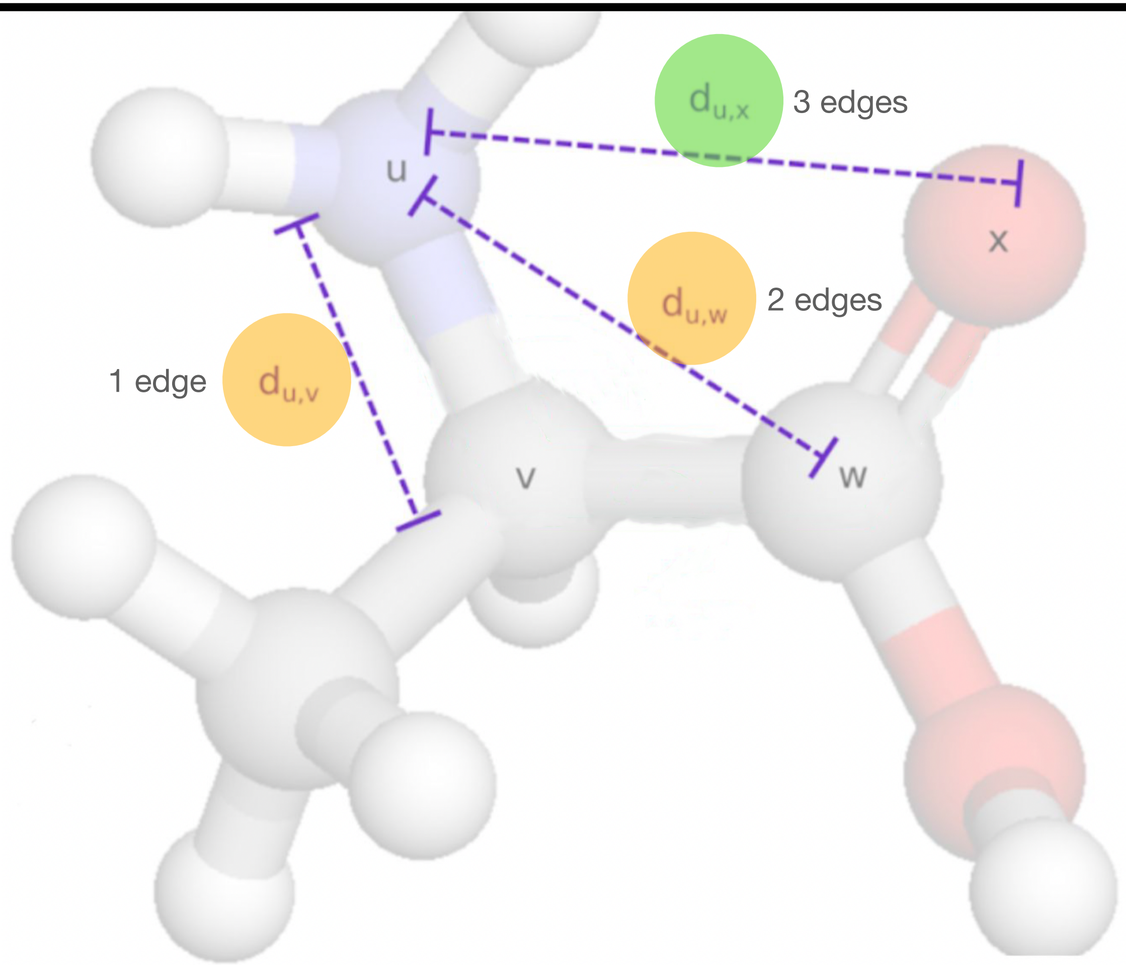
Realistic inter-atom distance changes



Realistic inter-atom distance changes



Realistic inter-atom distance changes



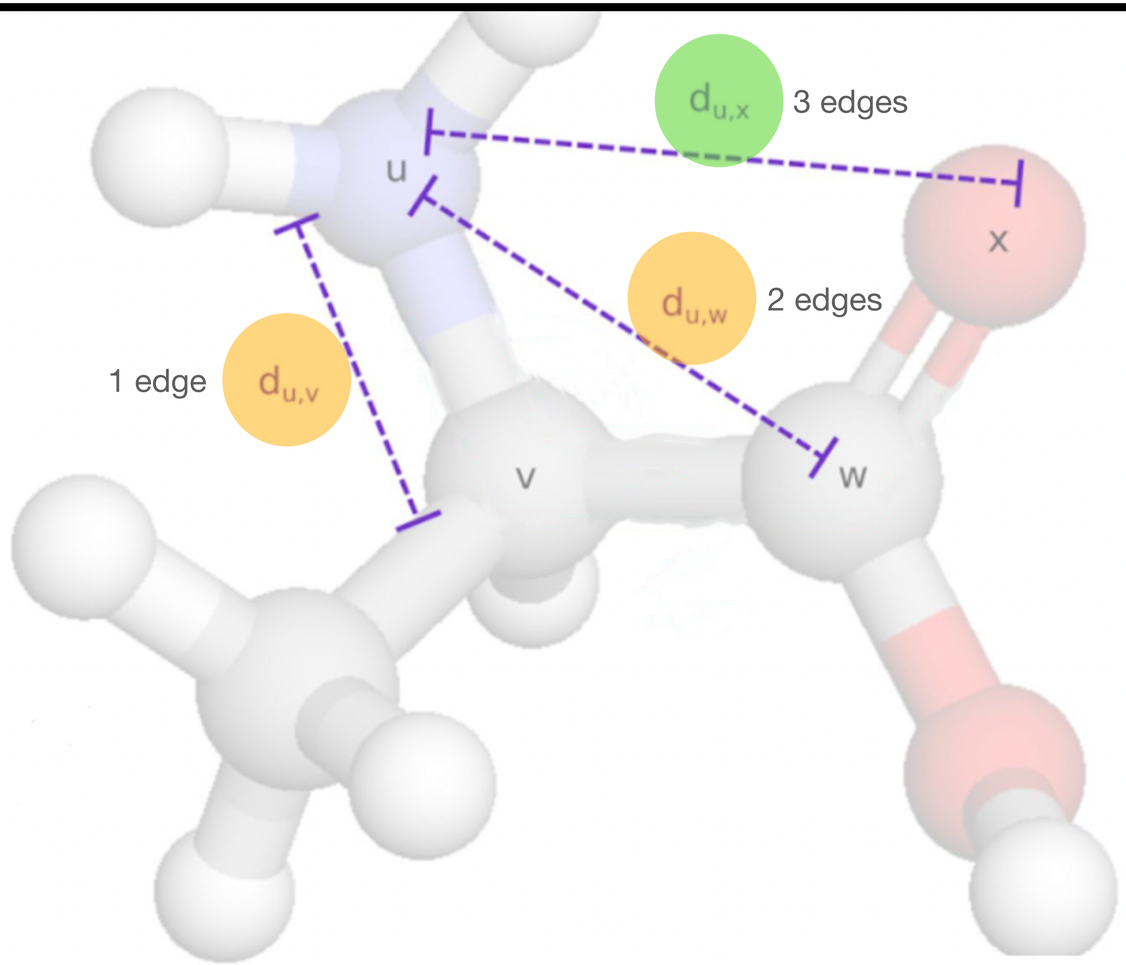
Z: transformed coordinates

X: original coordinates

$$\mathcal{S}(\mathbf{Z}, \mathbf{X}) = \sum_{\{(i,j)\in\mathcal{E}\}} (d^2_{\mathbf{X}}(i,j) - d^2_{\mathbf{Z}}(i,j))^2$$

$$+ \sum_{\{i,j:\text{2-hops away in } \mathcal{G}\}} (d^2_{\mathbf{X}}(i,j) - d^2_{\mathbf{Z}}(i,j))^2$$

Realistic inter-atom distance changes

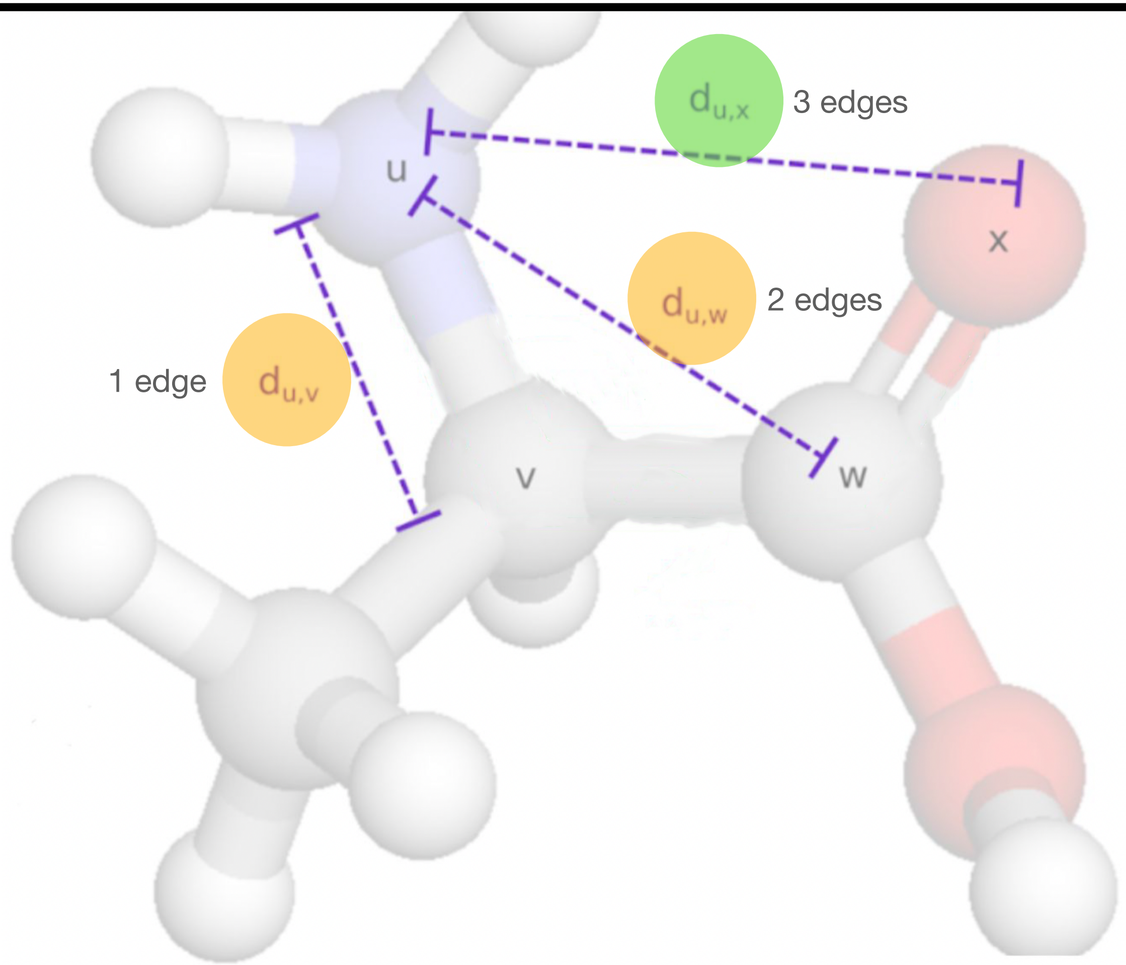


Z: transformed coordinates

X: original coordinates

$$\mathcal{S}(\mathbf{Z},\mathbf{X}) = \sum_{\{(i,j)\in\mathcal{E}\}} (d^2_{\mathbf{X}}(i,j) - d^2_{\mathbf{Z}}(i,j))^2 + \sum_{\{i,j:\text{2-hops away in } \mathcal{G}\}} (d^2_{\mathbf{X}}(i,j) - d^2_{\mathbf{Z}}(i,j))^2$$

Realistic inter-atom distance changes



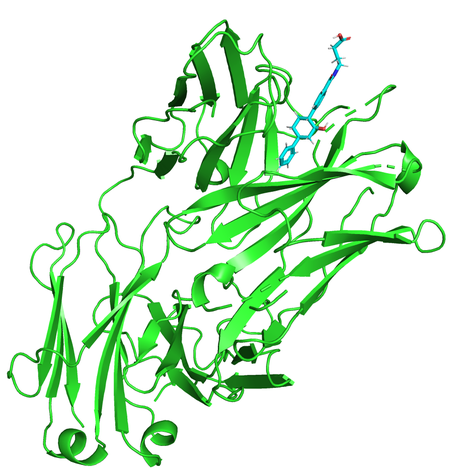
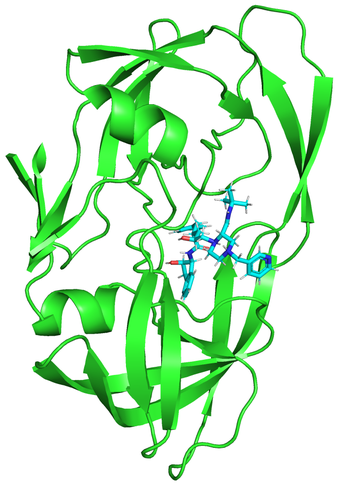
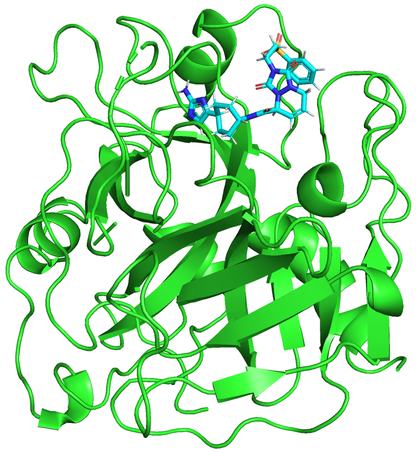
Z: transformed coordinates

X: original coordinates

$$\mathcal{S}(\mathbf{Z},\mathbf{X}) = \sum_{\{(i,j)\in\mathcal{E}\}} (d^2_{\mathbf{X}}(i,j) - d^2_{\mathbf{Z}}(i,j))^2 + \sum_{\{i,j:\text{2-hops away in } \mathcal{G}\}} (d^2_{\mathbf{X}}(i,j) - d^2_{\mathbf{Z}}(i,j))^2$$

Standard dataset PDBBind

Experimentally determined structures of small molecules + proteins

Baselines: 4 commonly used tools

GNINA
(deep learning)

McNutt et al. 2021

SMINA

Koes et al. 2013

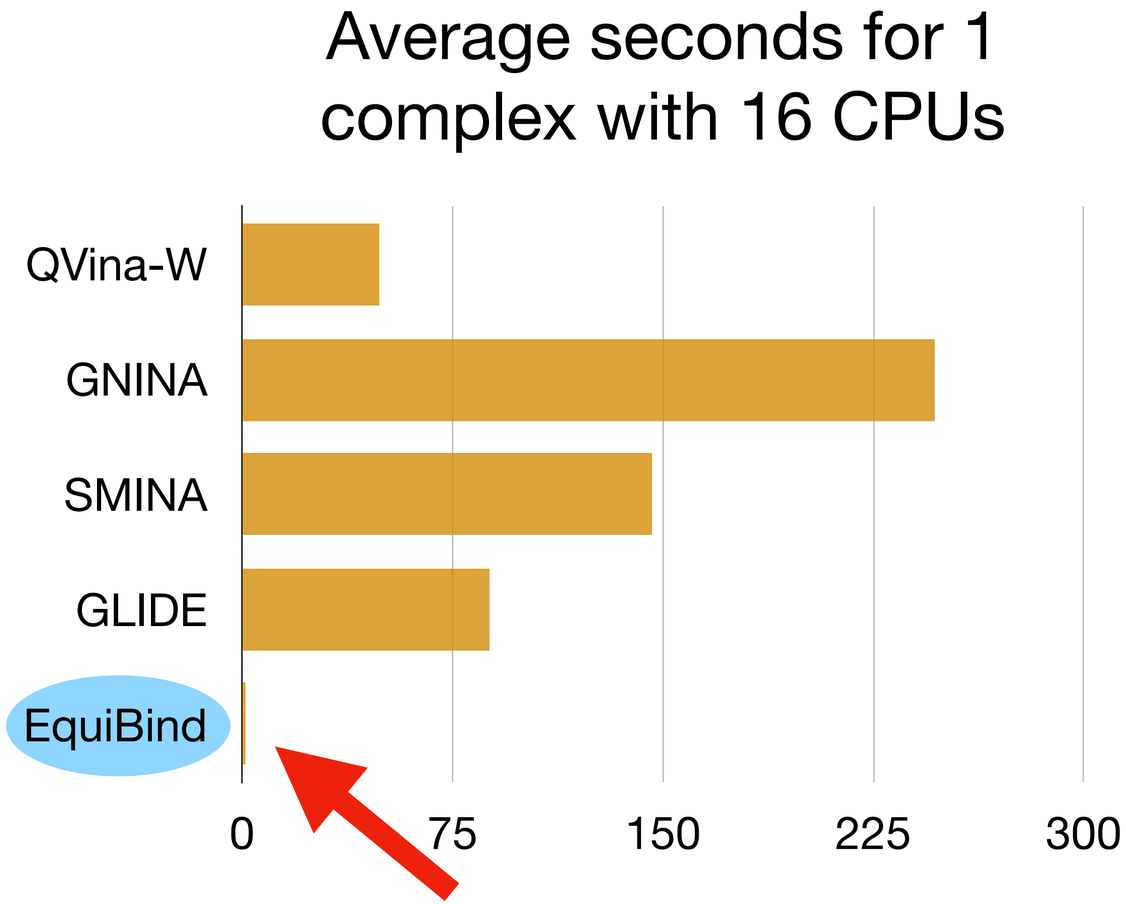
QuickVina-W

Hassan et al. 2017

GLIDE
(commercial)

Schrödinger. Release 2021-4

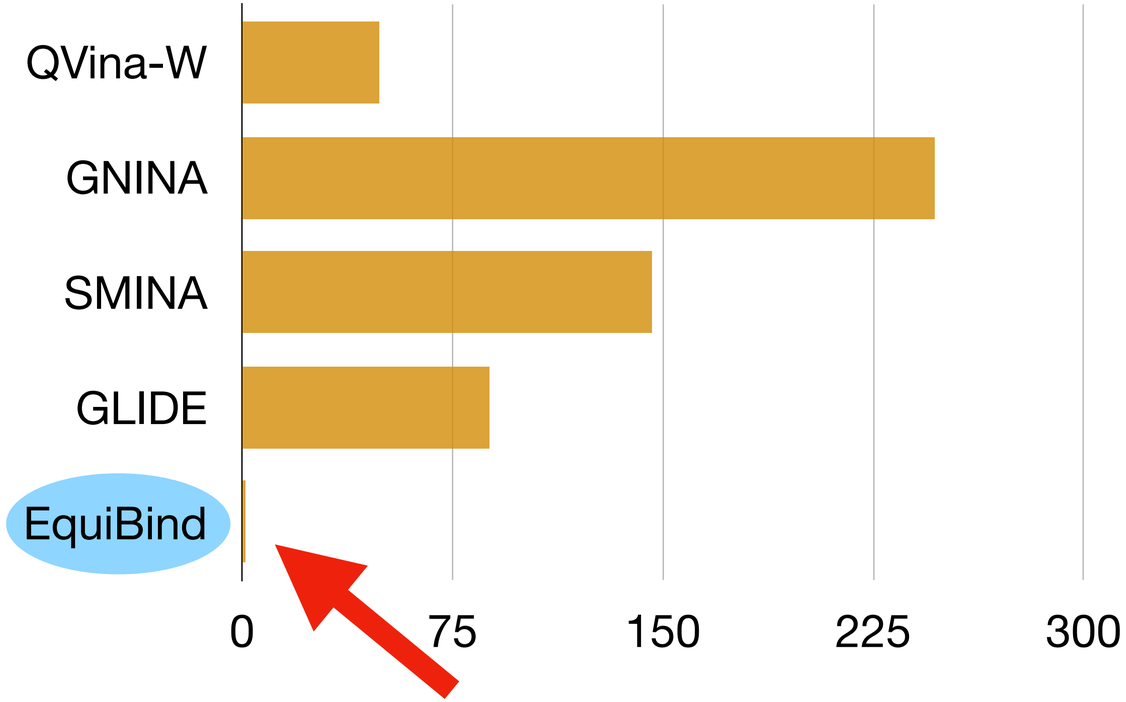
Results



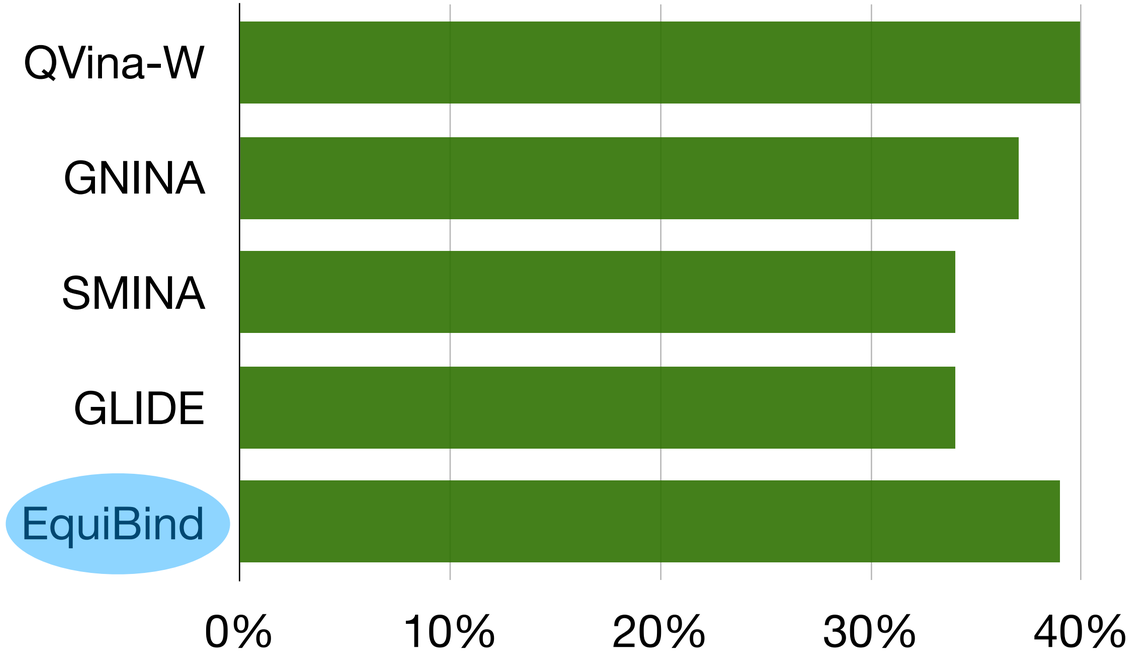
1200x speedup while retaining accuracy

Results

Average seconds for 1 complex with 16 CPUs



Fraction of test examples with RMSD less than 5Å



1200x speedup while retaining accuracy

Massachusetts Institute of Technology

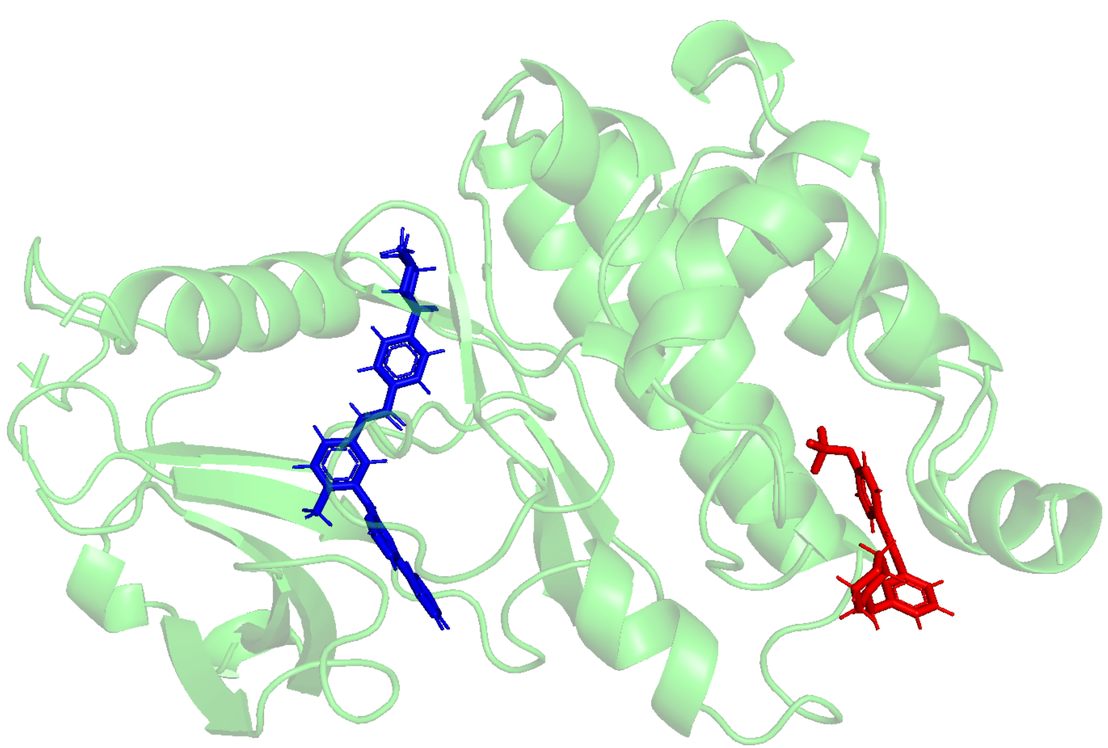
EquiBind

Geometric Deep Learning for Drug Binding Structure Prediction

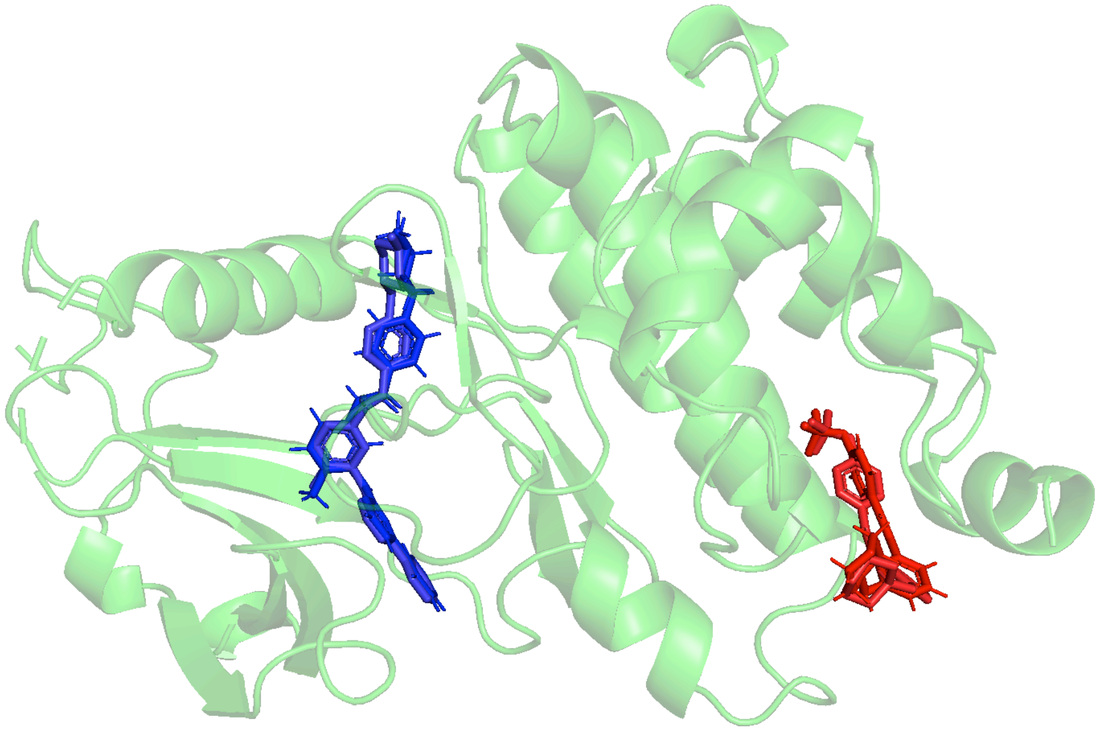
Hannes Stärk, Octavian-Eugen Ganea, Lagnajit Pattanaik, Regina Barzilay, Tommi Jaakkola

Case Study

Ground Truth



EquiBind + S



Suggested from industry: ABL1 Tyrosine Kinase. Important leukemia drug target