

A Deep Learning-Powered ChimeraX Plugin for Ligand Identification in CryoEM

J. Karolczak¹, A. Przybyłowska¹, W. Taisner¹, J.M. Heumann², M.H.B. Stowell², and D. Brzezinski¹

¹Institute of Computing Science, Poznan University of Technology, Piotrowo 2, 60-965 Poznan, Poland, ²Department of Molecular, Cellular and Developmental Biology, University of Colorado Boulder, Boulder, CO 80309, USA

dariusz.brzezinski@cs.put.poznan.pl

Accurate identification of ligands within protein structures derived from macromolecular X-ray crystallography (MX) and cryo-electron microscopy (cryoEM) is crucial for structure-based drug design and understanding protein function. However, manual ligand modeling is time-consuming and prone to bias, particularly in challenging density maps. Existing automated ligand identification methods are primarily developed for X-ray crystallography and often rely on feature engineering or iterative fitting procedures.

As part of our recent research [1], we have developed a novel deep learning approach for automated ligand identification in both MX and cryoEM density maps. Trained on a vast dataset of X-ray diffraction data and a smaller amount of cryoEM data, the proposed model processes density map fragments as 3D point clouds, utilizing sparse 3D convolutions to predict ligand types (Figure 1). The model achieves competitive accuracy with existing MX methods and is the first ligand identification model for cryoEM data [1].

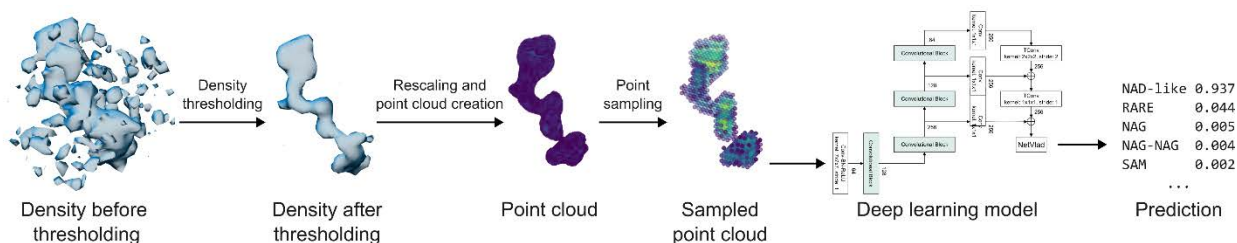


Figure 1. Schematic representation of the proposed ligand recognition pipeline. Example presenting the processing of a nicotinamide adenine dinucleotide (NAD) ligand from 2.10 Å cryoEM PDB deposit 8fuz (residue A 602).

One of our key contributions is the development of a ChimeraX [2] plugin, called LigandRecognizer, designed for integration within the standard structural biology workflow. The plugin provides a graphical interface within ChimeraX, allowing researchers to quickly load a cryoEM difference map, select density blobs of interest, and automatically rank likely ligands with associated confidence scores (Figure 2). The plugin is available through the ChimeraX Toolshed: <https://cxtoolshed.rbvi.ucsf.edu/apps/chimeraxligandrecognizer>.

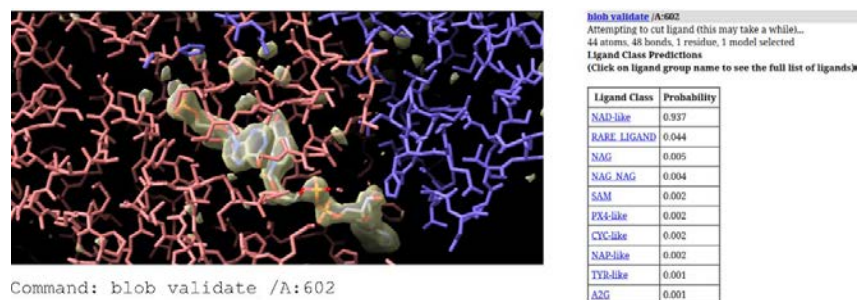


Figure 2. Screenshot of the performance of the LigandRecognizer plugin. Left: the difference map fragment of a NAD ligand (8fuz A 602) in ChimeraX's Working Pane. Right: the prediction results for the difference map fragment in ChimeraX's Log Pane.

We will discuss the proposed deep learning model's training process, its efficacy, the developed ChimeraX plugin, and the challenges of combining MX and cryoEM data. Our software demonstrates the potential of AI to accelerate structural biology research, providing a valuable tool for the broader MX and cryoEM community.

[1] Karolczak, J., Przybyłowska, A., Szewczyk, K., Taisner, W., Heumann, J. M., Stowell, M. H. B., Nowicki, M. & Brzezinski, D. (2025). *Bioinformatics* **41**(1), btac749.

[2] Meng, E. C., Goddard, T. D., Pettersen, E. F., Couch, G. S., Pearson, Z. J., Morris, J. H. & Ferrin, T. E. (2023). *Protein Sci.* **32**(11), e4792.

This work was supported by the PUT Institute of Computing Science Statutory Funds