Defining membrane boundaries of proteins using electron density maps – the MemBlob database and server

Georgina Csizmadia[1,#], Bianka Farkas[1,2,#], Eszter Katona[1,3], Gábor E. Tusnády[4], Tamás Hegedűs[1]

[1] Department of Biophysics and Radiation Biology, Semmelweis University, Budapest, Hungary and MTA-SE Molecular Biophysics Research Group, Hungarian Academy of Sciences, Budapest, Hungary

[2] Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Budapest, Hungary

[3] University College London, London, United Kingdom

[4] "Momentum" Membrane Protein Bioinformatics Research Group, Institute of Enzymology, Research Centre for Natural Sciences, Budapest, Hungary

Transmembrane (TM) proteins are highly significant drug targets, as they play an important role in many cellular processes. In order to understand TM protein function and to develop novel therapies targeting TM protein associated diseases, it is crucial to define the localization of transmembrane regions. While experimental data on the boundaries of transmembrane regions is scarce, cryo-electron microscopy (cryo-EM) density maps contain still unexploited information on the membrane embedment. To extract this information, we developed a computational pipeline, which requires a cryo-EM map, the corresponding atomistic structure, and the potential bilayer orientation determined by the TMDET algorithm as input. The resulted output includes the residues assigned to the bulk water phase, lipid interface, and the lipid hydrophobic core. We used this method to build a database, which contains the information on TM regions of protein structures based on the corresponding cryo-EM maps with a resolution better than 4 Å. We also created a web application to allow the analysis of unpublished densities. Unlike popular membrane predictors, our method presents the extracted membrane region as a volume with boundaries that follows the shape of the lipid environment and is not a slab with parallel planes. Utilizing our pipeline, we obtained the first data set at atomic resolution that allows the comparison of in silico predictions with experiments. The MemBlob database and server is available at http://memblob.hegelab.org.

Support: NKFIH-111678, NKFI-127961, CFF HEGEDU18I0, KIFÜ HPC, MTA Wigner GPU Laboratory, NVIDIA Corporation, Semmelweis Science and Innovation Fund