Mining Macromolecular Binding Interfaces

Christiane Ehrt[1], Joel Graef[1], Thorben Reim[1], Matthias Rarey[1]

[1]University of Hamburg, ZBH - Center for Bioinformatics, Albert-Einstein-Ring 8-10, 22761 Hamburg, Germany

Protein-protein interactions (PPIs) are crucial for physiological processes and their modulation is of interest for studying signaling pathways and pharmacological research.[1] However, methods for mining interfaces between protein chains are sparse and often heavily sequence-dependent.

By mining macromolecular databases, we can extract valuable knowledge from available structures of biologically relevant protein-protein interfaces. The number of known physical interfaces was estimated to be more than 100,000 in 2014, already.[2] Protein-protein docking to predict the structure of protein-protein complexes is computationally expensive and heavily relies on prior knowledge of the potential interaction partners. Thus, the development of reliable PPI comparison tools is key to supporting protein-protein complex prediction by detecting potential interaction partners.

In this poster, we present PiMine [3] – a method for comparing predicted or known interfaces to biologically relevant interfaces [4] in individual databases of protein complexes and its performance in comparison to commonly applied methods. It relies on the comprehensive GeoMine [5] database and mining system. We will introduce retrospective and prospective application scenarios for predicting protein-protein complexes with PiMine. Furthermore, we will give a perspective on the method's potential to predict host-pathogen interactions based on AlphaFold2 models of global health proteomes.[5]

References

[1] H. Nada, Y. Choi, S. Kim, K.S. Jeong, N.A. Meanwell, K. Lee, *Signal Transduction Targeted Ther*, **2024**, *9*, 341.
[2] K. Baskaran, J.M. Duarte, N. Biyani, S. Bliven, G. Capitani, G., *BMC Struct Biol*, **2014**, *14*, 22.
[3] J. Graef, C. Ehrt, T. Reim, M. Rarey, *J Chem Inf Model*, **2024**, *64*, 3332-3349.
[4] S. Bliven, A. Lafita, A. Parker, G. Capitani, J.M. Duarte, *PLoS Comput Biol*, **2018**, *14*, e1006104.
[5] K. Diedrich, C. Ehrt, J. Graef, M. Poppinga, N. Ritter, M. Rarey, *J Comput Aided Mol Des*, **2024**, *38*, 23.
[6] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A., et al., *Nature*, **2021**, *596*, 583-589.