**[BIOINFO 2025]**

# Rapid and Sensitive Protein Complex Alignment with Foldseek-Multimer

Woosub Kim[1,2], Milot Mirdita[2], Eli Levy Karin[3], Cameron L.M. Gilchrist[2], Hugo Schweke[4,5], Johannes Söding[6,7],
Emmanuel Levy[4,5,*], and Martin Steinegger[1,2,7,9*]
*[1]Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Republic of Korea.*
*[2]School of Biological Sciences, Seoul National University, Seoul, Republic of Korea.*
*[3]ELKMO, Copenhagen, Denmark.*
*[4]Department of Chemical and Structural Biology, Weizmann Institute of Science, Rehovot, Israel.*
*[5]Department of Molecular and Cellular Biology, University of Geneva, Geneva, Switzerland*
*[6]Quantitative and Computational Biology, Max-Planck Institute for Multidisciplinary Sciences, Göttingen, Germany.*
*[7]Campus Institute Data Science (CIDAS), University of Göttingen, Germany.*
*[8]Institute of Molecular Biology and Genetics, Seoul National University, Seoul, Republic of Korea.*
*[9]Artificial Intelligence Institute, Seoul National University, Seoul, Republic of Korea.*
**Contact: emmanuel.levy@gmail.com, martin.steinegger@snu.ac.kr*

Recent advances in computational structure prediction have revolutionized the prediction of quaternary structures in protein complexes. These tools scale structural analysis to the level of proteome-wide complex structure prediction. Analyzing and comparing these quaternary structures is crucial for exploring structural diversity. However, translating the sheer volume of quaternary structures into meaningful discoveries requires efficient alignment and comparison methods, a task that remains computationally challenging..

To address this challenge, we introduce Foldseek-Multimer, a protein complex aligner. Foldseek-Multimer is 3-4 orders of magnitude faster than the gold-standard method, US-align, while maintaining comparable alignment quality. Foldseek-Multimer processes billions of complex pairs in just 11 hours. Foldseek-Multimer identification rapid discovery of homologous protein complexes across vast databases and functional predictions for protein complexes in unknown species.

Foldseek-Multimer is free, open-source software available at github.com/steineggerlab/foldseek, complemented by a webserver at search.foldseek.com and the BFMD database.