

MS11 Opportunities from combining structural biology and fold prediction

MS11-02

Deep learning based hallucination of de novo protein assemblies across the nanoscale

A. Courbet¹

¹Department of Biochemistry, University of Washington - Seattle (United States)

Abstract

A. Courbet^{1,2,3†}, B. I. M. Wicky^{1,2†}, L. F. Milles^{1,2†}, R. J. Ragotte^{1,2}, J. Dauparas^{1,2}, E. Kinfu^{1,2}, S. Tipps^{1,2}, R. D. Kibler^{1,2}, M. Baek^{1,2}, F. DiMaio^{1,2}, X. Li^{1,2}, L. Carter^{1,2}, A. Kang^{1,2}, H. Nguyen^{1,2}, A. K. Bera^{1,2}, D. Baker^{1,2,3*}

¹ Department of Biochemistry, University of Washington, Seattle, WA, USA.

² Institute for Protein Design, University of Washington, Seattle, WA, USA.

³ Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA.

† These authors contributed equally to this work.

Deep learning generative approaches provide an opportunity to broadly explore protein structure space beyond the sequences and structures of natural proteins. In this work, we use deep network hallucination to generate a wide range of de novo symmetric protein homo-oligomers given only a specification of the number of protomers and the protomer length. Crystal structures of 7 designs are very close to the computational models (median RMSD: 0.6 Å), as are 3 cryoEM structures of giant ring like structures with up to 1550 residues, C₄₂ symmetry, and 12 nanometre in diameter; all display complex folds and topologies that differ considerably from previously solved structures. Our results highlight the rich diversity of new protein structures that can be created using deep learning, and pave the way for the design of increasingly complex nanomachines and biomaterials.