Do AlphaFold Ensembles Capture Side-Chain Fluctuations? A Large-Scale Benchmark Against MD and PDB Homologs

Alexander Korsunsky, Nicholas M Pearce* and Björn Wallner

Linköping University, *Uppsala University

Protein side-chain fluctuations are essential for cellular processes such as enzymatic catalysis and molecular recognition. Accurately modeling these dynamics is vital for drug design and interpreting experimental data from NMR or crystallography.

In this work, we evaluate AlphaFold2[1] (AF2) and variants such as AlphaFold3[2] (AF3), AFsample2[3], and AF2χ[4] by generating 1,000 structures for each of 120 proteins to analyze side-chain flexibility.

We compare predicted side-chain dihedral (χ_1-χ_5) distributions, using both marginal (single-angle) and joint (multi-angle) analyses, against MD (ATLAS[5]) and ensembles of sequence-homologous PDB structures.

To quantify these comparisons, we propose a metric based on optimal transport (Wasserstein distance) that accounts for the geometry of rotamer space while avoiding the bin-count dependency of standard information-theoretic measures and enables comparisons of multivariate (joint) distributions.

Our results show that the original structure prediction models AF2 and AF3 produce side-chain distributions with very limited conformational variation, which is consistent with the models being trained primarily on rigid crystal structures. While variants tuned for diversity (AFsample2, AF2χ) show increased variation, their similarity to MD and experimental ensembles remains low.

This study provides a benchmark to determine if ensembles from structure prediction tools can accurately replicate the physical fluctuations and highlights the current limitations of using structure prediction models to capture the full landscape of protein side-chain dynamics.

Bibliography

1. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021).
2. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature 630, 493–500 (2024).
3. Kalakoti, Y. & Wallner, B. AFsample2 predicts multiple conformations and ensembles with AlphaFold2. Commun. Biol. 8, 373 (2025).
4. Cagiada, M., Thomasen, F. E., Ovchinnikov, S. & Lindorff-Larsen, K. AF2χ: Predicting protein side-chain rotamer distributions with AlphaFold2.
5. Vander Meersche, Y., Cretin, G., Gheeraert, A., Gelly, J.-C. & Galochkina, T. ATLAS: protein flexibility description from atomistic molecular dynamics simulations. Nucleic Acids Res. 52, D384–D392 (2024).