BraintankAI

DISPATCH · PROTEIN STRUCTURE PREDICTION · 3 MIN

AlphaFold's most confident protein predictions can still diverge from X-ray maps

Judged against X-ray maps built without the deposited models, AlphaFold2's predicted structures make strong hypotheses but often differ in detail from experiment.

18 November 2026 · based on a paper in Nature Methods

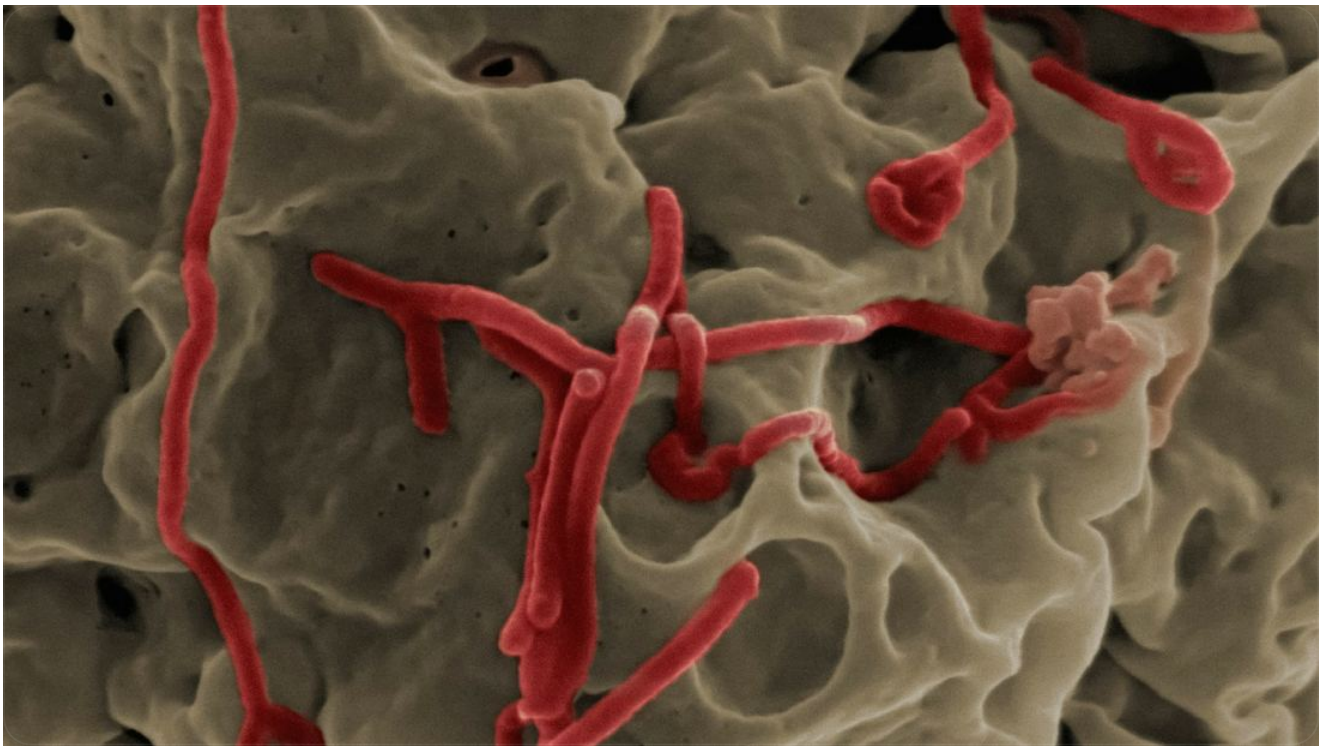


Photo: National Institute Of Allergy And Infectious Diseases

Even when AlphaFold, DeepMind's AI for predicting protein shapes, gives a structure its highest confidence rating, the result is a well-founded hypothesis, not a measurement. That is the conclusion of Thomas Terwilliger, Paul Adams and colleagues at Los Alamos National Laboratory and five other institutions, who checked AlphaFold2 predictions against experimental X-ray data. In a 2023 paper in Nature Methods, they report that many predictions match closely and have very good local geometry, yet even very-high-confidence regions typically depart from the deposited structures by about twice as much as two crystal forms of the same protein differ. Predictions, they conclude, accelerate experimental structure determination but do not replace it.

AlphaFold2 predicts a protein's three-dimensional shape from its amino-acid sequence; it has supplied predictions for some 200 million sequences and has been described as atomically accurate. Accuracy is usually judged against deposited models, the atomic structures

experimenters submit to the Protein Data Bank (PDB). Some have argued that disagreements may stem not from prediction errors but from poorly defined parts of those models, such as side chains, the chemically active branches that jut from each amino acid. The authors wanted an independent referee, plus a measure of how much one protein naturally varies.

Their referee was the electron density map, the three-dimensional cloud of electrons calculated from X-ray diffraction data, into which models are fitted. In earlier work the team had re-solved 215 crystal structures, starting from AlphaFold predictions and rebuilding iteratively, to produce maps that used no information from the deposited models. Most analyses used the 102 with the most reliable maps. As a yardstick, they repeated their measurements on pairs of PDB structures of an identical protein crystallised in different packing arrangements, where differing contacts with neighbouring molecules produce natural variation.

On map-model correlation, a score of how well a model's atoms line up with the density in which 1 is a perfect match, AlphaFold predictions averaged 0.56 and deposited models, already refined against the same diffraction data, 0.86. Morphing each prediction, smoothly bending it towards the deposited model to strip out large-scale warping while keeping local differences, lifted the average only to 0.67. Warping grew with distance: the median difference in the separation of two atoms rose from about 0.1 ångström (a tenth of a nanometre) for atoms 4–8 Å apart to 0.7 Å for atoms 48–52 Å apart; between crystal forms the figures were 0.1 Å and 0.4 Å.

Side chains were tested by grafting each predicted one onto the deposited model's backbone. Among amino acids AlphaFold rated at least moderately confident, excluding those close to neighbouring molecules in the crystal, 20% of side chains differed from the deposited model by 1.5 Å or more, and a third of those, 7% overall, were statistically incompatible with the map. Between crystal forms the figures were 6% and 2%. The authors put the true disagreement at 7–20%. Functionally important amino acids, taken as those named in the original publications, fared no better: 53 of 733, again 7%, were incompatible.

AlphaFold's per-amino-acid confidence score, pLDDT, tracked error well. Median backbone error was 0.6 Å for amino acids scoring above 90, the very-high band, rising to 3.5 Å below 70. Yet 10% of very-high-confidence backbone positions lay more than 2 Å from the deposited model, against 5% between crystal forms, whose median difference was 0.3 Å. The authors say that doubled median exceeds what protein flexibility would explain, and note that their figures are consistent with DeepMind's own uncertainty estimates.

The sample is small and limited to crystal structures solved by one experimental technique, picked as a proxy for harder cases, and only the first chain of each multi-chain entry was analysed, using AlphaFold's standard single-chain version. It is also skewed to high confidence: 86% of amino acids scored above 90, against 36% in predictions for all human proteins, so estimates for lower-confidence regions are uncertain, and its median backbone deviation of 1.0 Å compares with 2.3 Å in DeepMind's broader analysis. Some mismatches may reflect genuine minor conformations present in the crystal. AlphaFold also ignores bound ions, small molecules, chemical modifications and environmental conditions. The authors suggest that a larger, more representative set of structures would help.

PAPER *AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination*

AUTHORS Thomas C. Terwilliger, Dorothee Claudia Liebschner, Tristan Ian Croll, Christopher J. Williams, Airlie J. McCoy, Billy K. Poon, Pavel V. Afonine, Robert D. Oeffner, Jane S. Richardson, Randy J. Read, Paul D. Adams et al.

VENUE Nature Methods · 30 November 2023

LINK doi:10.1038/s41592-023-02087-4

INDEX stacks.braintank.ai