nature research

Corresponding author(s): Andrew Senior

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒ Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	To ingest and pre-process the data, we used the following tools: PDB 2018-03-15 ; CATH 2018-03-16; HHblits based on version 3.0-beta.3; HHpred web server; Uniclust30 2017-10; PSI-BLAST version 2.6.0; SST web server (March 2019); BioPython v1.65; Rosetta v3.5; TM-align 20160521, as well as custom code written using Python 2.7. See the methods section for more details.
Data analysis	The networks used the TensorFlow library with custom extensions. Inference code for distance prediction networks used in CASP13 will be open-sourced. Analysis was performed with custom code written in Python 2.7. Visualizations were made with PyMol 2.2.0 software. Please see methods section for more detail.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The following public datasets were used in this work:
- PDB 2018-03-15
- CATH 2018-03-16
- Uniclust30 2017-10
- PSI-BLAST nr dataset (as of 2017-12-15)

We will make available our train/test split (CATH domain codes).

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Tests were carried out on two sets of domains generated by predetermined methods. The principal set consisted of 104 CASP13 domains as segmented by the CASP13 assessors. (In some figures 100 domains are used, excluding the domains of T0999 where slightly different methods were used; or 41 considering FM + TBM/FM targets) For the CASP13 datasets, accuracy measures are reported using structures & distance distributions computed during CASP (except for the "back-fill" gradient descent structures where specified). With n=41 domains, we show (Extended Data Figure 5) that (with p=0.0032) AlphaFold's results were better than the next best group. The other test set of 377 domains was extracted from PDB, chosen to be from separate homologous superfamilies (CATH code) none of which were represented in the training set. For each superfamily the exemplar was chosen at random from the s35 cluster representatives. Extended Data Table 3c shows that this sample size is sufficient to show that certain terms in the potential (but not score2) make a difference in the final GDT_TS score.
Data exclusions	The training set used one example per s35 cluster, using the representative chosen by CATH. We further balance (and reduce the size of) the test set by keeping one example per superfamily. CASP13 domains are only those "all groups" targets which were scored as part of CASP13. Some CASP13 targets were excluded by the assessors e.g. because of publications or failure to solve a structure.
Replication	Actual competition entries were used and were not replicated, but we show comparisons for top-5 decoys. Multiple networks were trained independently and were found to give consistent results (on CASP11/12 validation sets). 4 such networks, yielding similar results, were used in CASP and are made available with the source code. Structure generation is stochastic, but we show (Extended Data Fig 4) that it converges with relatively few attempts, so results are reproducible.
Randomization	Superfamilies were randomly assigned to PDB train or PDB test set. For each superfamily in the test set, a random s35 cluster representative was picked for the 377 example test set. CASP13 assessors used a manual process (blind to the investigators) to determine evaluation units and to determine whether each domain should be treated as FM, TBM or TBM/FM.
Blinding	Investigators were blind to the CASP13 targets which were sequestered during the assessment. Investigators were not blind to the PDB train/test split, but cross-validation used CASP11 & 12 domains and the main results are on the blind CASP13. Models were not retrained after defining the 377 domain set.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging