

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on 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 <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☒	☐ 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 statistical parameters 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☒	☐ 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 <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	No software was used to collect data. All data are downloaded from SCOPe, PDB.
Data analysis	D-I-TASSER(3.0), R(4.4.2), TM-score(2019/8/22), FUpred(v1.0), ThreaDom(v1.0), LOMETS(v3.0), FUpred (v1.0), ThreaDom (v1.0), DeepMSA (v1.0), DeepMSA2 (v2.0), HHblits (2.0.15), Jackhmmer (3.1b2), HMMsearch (3.1b2), CEthreader (v1.0), Hybird-CEthreader (v1.0), MapAlign (v1.0), DisCovER (v1.0), EigenThreader (v1.0), HHPred (v1.0), HHsearch (2.0.15),FFAS3D (v1.0), MUSTER (v1.0), Sparks-X (v1.0), AttentionPotential (v1.0), DeepPotential (v1.0), TriPletRes (v1.0), ResTriplet (v1.0), ResPRE (v1.0), ResPLM (v1.0), NeBcon (v1.0), DEMO (2.0), TM-align (2019/8/22), DeepFold (v1.0), REMO (v1.0), SPICKER (v3.0), FASPR (v1.0), LAMPPS (1/9/2009), COFACTOR (v2.0), MetaGO (v1.0), BLAST (2.5.0+), and PSI_BLAST (2.5.0+).

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 and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All benchmark datasets are available at <https://zhanggroup.org/D-I-TASSER/download> and <https://zenodo.org/records/15058641> for academic use. The structure and function modeling results on human proteome are freely available at <https://zhanggroup.org/HPmod> and <https://zenodo.org/records/15065861> for academic use. The PDB IDs for the case studies are 3fpiA, 4jgnA, 7jtkB, 6irdC, and the UniProt ID for case study used in Figure 6 is Q9BWD1.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	NA
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	NA
Ethics oversight	NA

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is 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/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

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

Sample size	The manuscript includes 1,262 single-domain targets and 230 multi-domain targets in benchmark set. The CASP dataset was from the community-wide experiments; the benchmark set was collected from the PDB database. The D-I-TASSER models for the whole human genome proteins with length less than 1,500 residues are also provided in this manuscript. No samples are collected or created, all data are download from the public available databases. There is no statistical method used for deciding the sample size, but the target numbers of each benchmark set are sufficient for the Student's T-test (i.e., N>100).
Data exclusions	The proteins homologous to the benchmark dataset were excluded from the template library to avoid homologous contamination.
Replication	All results could be reproduced by our server and standalone package, or based on the information provided in SI.
Randomization	The benchmark proteins were selected randomly from the PDB and CASP8-14, after the consideration of homology exclusion.
Blinding	There was no blinding group or analysis in the benchmark sections of this manuscript, but for the CASP15 sections, when D-I-TASSER server participate the CASP15, we do not know the experimental structures, so the results of the CASP15 could be treated as Blinding test results.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involvement in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA