

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 |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> 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> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> 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) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), 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

for the analysis and the visualization we used: Numpy 1.23, Pandas 1.4.3, Prody 2.0, and Matplotlib 3.5, R 4.0.2, ggplot2, I-Mutant, SAAFEC-SEQ, MUpro, mCSM-stability, DUET, SDM, DynaMut, MAESTRO, ENCoM, DynaMut2, FoldX, Rosetta.
Our own code can be found at
Coverage comparison between SWISS-MODEL Repository and AlphaFold2 https://github.com/aozalevsky/alphafold2_vs_swissmodel/
Exploration of structural space <https://github.com/TurtleTools/alphafold-structural-space>
Pocket predictions https://github.com/jurgjn/af2_pockets
Protein complexes <https://gitlab.com/ElofssonLab/FoldDock>
<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>
Model building https://github.com/tristanic/pae_to_domains

Data analysis

for the analysis and the visualization we used: Numpy 1.23, Pandas 1.4.3, Prody 2.0, and Matplotlib 3.5, R 4.0.2, ggplot2, I-Mutant, SAAFEC-SEQ, MUpro, mCSM-stability, DUET, SDM, DynaMut, MAESTRO, ENCoM, DynaMut2, FoldX, Rosetta.
Our own code can be found at
Coverage comparison between SWISS-MODEL Repository and AlphaFold2 https://github.com/aozalevsky/alphafold2_vs_swissmodel/
Exploration of structural space <https://github.com/TurtleTools/alphafold-structural-space>
Pocket predictions https://github.com/jurgjn/af2_pockets
Protein complexes <https://gitlab.com/ElofssonLab/FoldDock>
<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>
Model building https://github.com/tristanic/pae_to_domains

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](#)

The SWISS-MODEL repository (<https://swissmodel.expasy.org/repository>) and AlphaFold2 (<https://alphafold.ebi.ac.uk/>) databases were accessed on 24.07.2021. Contiguous protein regions of human high confidence structural predictions with no previous structural predictions by homology models in the SWISS-MODEL Repository are available in Supplementary Table 1 and in Github https://github.com/aozalevsky/alphafold2_vs_swissmodel. The benchmark dataset used for testing of disorder predictions metrics is available in Supplementary Table 2 and predicted disordered regions for human proteins is available as Supplementary dataset 1 and integrated into ProViz22 at http://slim.icr.ac.uk/projects/alphafold?page=alphafold_proviz_homepage. The grouping of proteins by structure similarly using the non-negative Matrix Factorization analysis of structural fragments is available as Supplementary Dataset 2 and the pocket prediction scores for 912 human proteins with no previous experimental or predicted structural models is available in Supplementary Table 3.

Human research participants

Policy information about [studies involving human research participants and Sex and Gender in Research](#).

Reporting on sex and gender

Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design whether sex and/or gender was determined based on self-reporting or assigned and methods used. Provide in the source data disaggregated sex and gender data where this information has been collected, and consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.

Population characteristics

Describe the covariate-relevant population characteristics of the human research participants (e.g. age, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write "See above."

Recruitment

Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.

Ethics oversight

Identify the organization(s) that approved the study protocol.

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

Life sciences study design

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

Sample size	Samples sizes were determined as the data points available for each computational analysis. All statistical tests considered the sample size for the determination of significance.
Data exclusions	No data was excluded from the analysis
Replication	All attempts at replication were successful
Randomization	Protein structures, regions and variants were randomly allocated for analysis
Blinding	The investigators were blind to group allocations during data analysis

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 |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Antibodies |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Eukaryotic cell lines |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Palaeontology and archaeology |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Animals and other organisms |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Clinical data |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Dual use research of concern |

Methods

- | n/a | Involvement in the study |
|-------------------------------------|---|
| <input checked="" type="checkbox"/> | <input type="checkbox"/> ChIP-seq |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Flow cytometry |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> MRI-based neuroimaging |