Nucleic Acids Research, 2026, **54**, D358–D362
https://doi.org/10.1093/nar/gkaf1226
Advance access publication date: 22 November 2025
Database issue

AlphaFold Protein Structure Database 2025: a redesigned interface and updated structural coverage

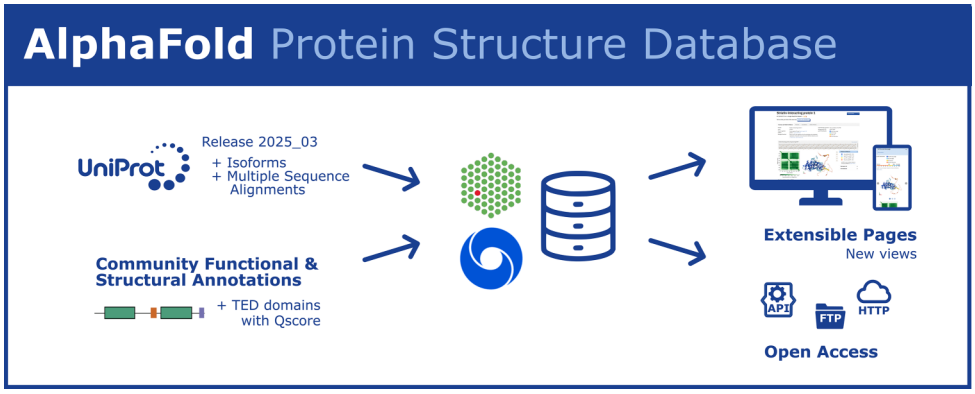
Damian Bertoni[1], Maxim Tsenkov[1], Paulyna Magana[1], Sreenath Nair[1], Ivanna Pidruchna[1], Marcelo Querino Lima Afonso[1], Adam Midlik[1], Urmila Paramval[1], Dare Lawal[1], Ahsan Tanweer[1], Meera Last[2], Risha Patel[2], Agata Laydon[2], Dariusz Lasecki[2], Nick Dietrich[2], Hamish Tomlinson[2], Augustin Žídek[2], Tim Green[2], Oleg Kovalevskiy[2], Andy Lau[3], Shaun Kandathil[3], Nicola Bordin[4], Ian Sillitoe[4], Milot Mirdita[5], David Jones[3], Christine Orengo[4], Martin Steinegger[5,6,7,8], Jennifer R. Fleming[1], Sameer Velankar[1,*]

[1]European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, CB10 1SD, United Kingdom
[2]Google DeepMind, London, N1C 4DN, United Kingdom
[3]Department of Computer Science, University College London, London, WC1E 6BT, United Kingdom
[4]Institute of Structural and Molecular Biology, University College London, London, England, WC1E 6BT, United Kingdom
[5]School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea
[6]Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul 08826, Republic of Korea
[7]Institute of Molecular Biology and Genetics, Seoul 08826, Republic of Korea
[8]Artificial Intelligence Institute, Seoul National University, Seoul 08826, Republic of Korea
*To whom correspondence should be addressed. Email: sameer@ebi.ac.uk

Abstract

The AlphaFold Protein Structure Database (AFDB; https://alphafold.ebi.ac.uk), developed by EMBL–EBI and Google DeepMind, provides open access to hundreds of millions of high-accuracy protein structure predictions, transforming research in structural biology and the wider life sciences. Since its launch, AFDB has become a widely used bioinformatics resource, integrated into major databases, visualization platforms, and analysis pipelines. Here, we report the update of the database to align with the UniProt 2025_03 release, along with a comprehensive redesign of the entry page to enhance usability, accessibility, and structural interpretation. The new design integrates annotations directly with an interactive 3D viewer and introduces dedicated domains and summary tabs. Structural coverage has also been updated to include isoforms plus underlying multiple sequence alignments. Data are available through the website, FTP, Google Cloud, and updated APIs. Together, these advances reinforce AFDB as a sustainable resource for exploring protein sequence–structure relationships.

Graphical abstract



Introduction

The field of life science has been transformed by protein structure prediction tools, such as AlphaFold [1–3], ESMFold [4], and RoseTTAFold [5]. Built upon decades of foundational research [6–8] and open scientific data resources such as the Protein Data Bank (PDB) [9, 10], MGnify [11] and the Univer-

Received: September 30, 2025. **Revised:** October 18, 2025. **Accepted:** October 18, 2025

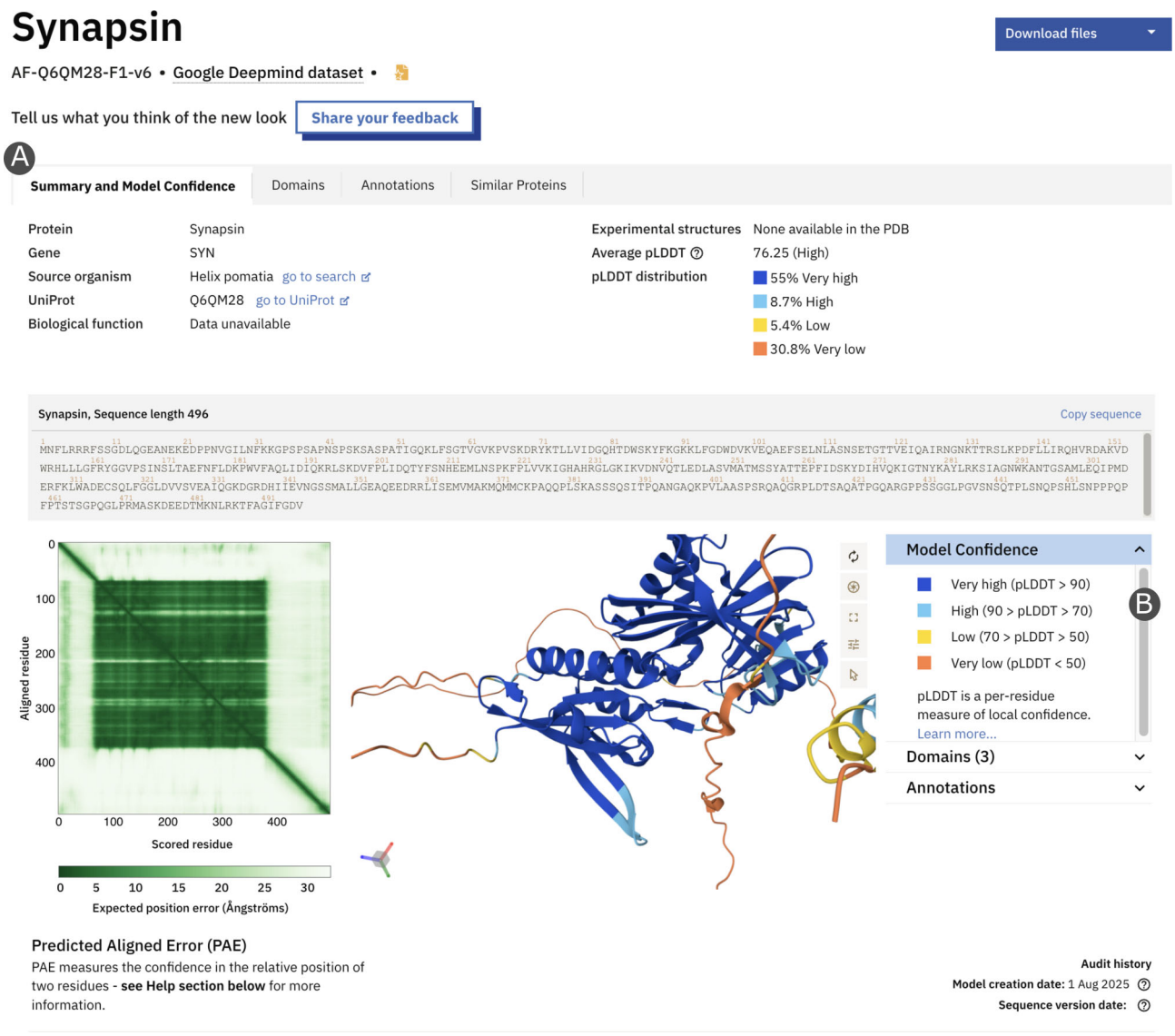


Figure 1. Example of summary page for Synapsin AF-Q6QM28-F1-v6. (**A**) The page is now arranged with tabs separating each section. The central panel displays the predicted 3D model, coloured by per-residue confidence score (pLDDT). (**B**) The right panel accordion provides an overview of domains and annotations and allows the user to toggle between different colour schemes.

sal Protein Resource (UniProt) [12], these advanced methods have dramatically narrowed the gap between the vast number of known protein sequences and the availability of structure models.

The AlphaFold Protein Structure Database (AFDB), a collaboration between Google DeepMind and EMBL–EBI [13–15], was established to democratize access to millions of highly accurate protein structure predictions. The practical utility of these predictions is underscored by their use, with to date, over 4,501,953 total users accessing the AFDB website, more than 18,000 proteome archives downloaded, and the data's integration into major primary data resources [16, 17] and molecular visualization software [18–20]. This widespread adoption has benefited various life science disciplines from structural biology and bioinformatics to drug discovery [21–27].

To meet the ongoing demand for comprehensive structural data, AFDB has continued to evolve with new tools and functionalities [15]. However, synchronization with UniProt fell several years behind, leaving the database increasingly out of date and limiting coverage of newly described proteins [28]. The current release restores alignment with UniProt (2025_03), ensuring that predictions once again reflect the evolving sequence space. Newly generated multiple sequence alignments (MSAs) are provided, and coverage has been extended to include isoform-specific predictions, broadening biological relevance and accuracy. In parallel, a redesigned entry page enhances usability and interpretation.

Enhanced user experience and functionalities

To prioritize an intuitive and accessible user experience, AFDB has now undergone a redesign (Fig. 1). A central element of this overhaul is the enhanced prominence of the structure viewer, now featuring integrated annotations for a more streamlined analysis. New, focused tabs, such as the new Domains tab, are directly linked to the 3D viewer, simplifying complex visualization and analysis by providing immediate

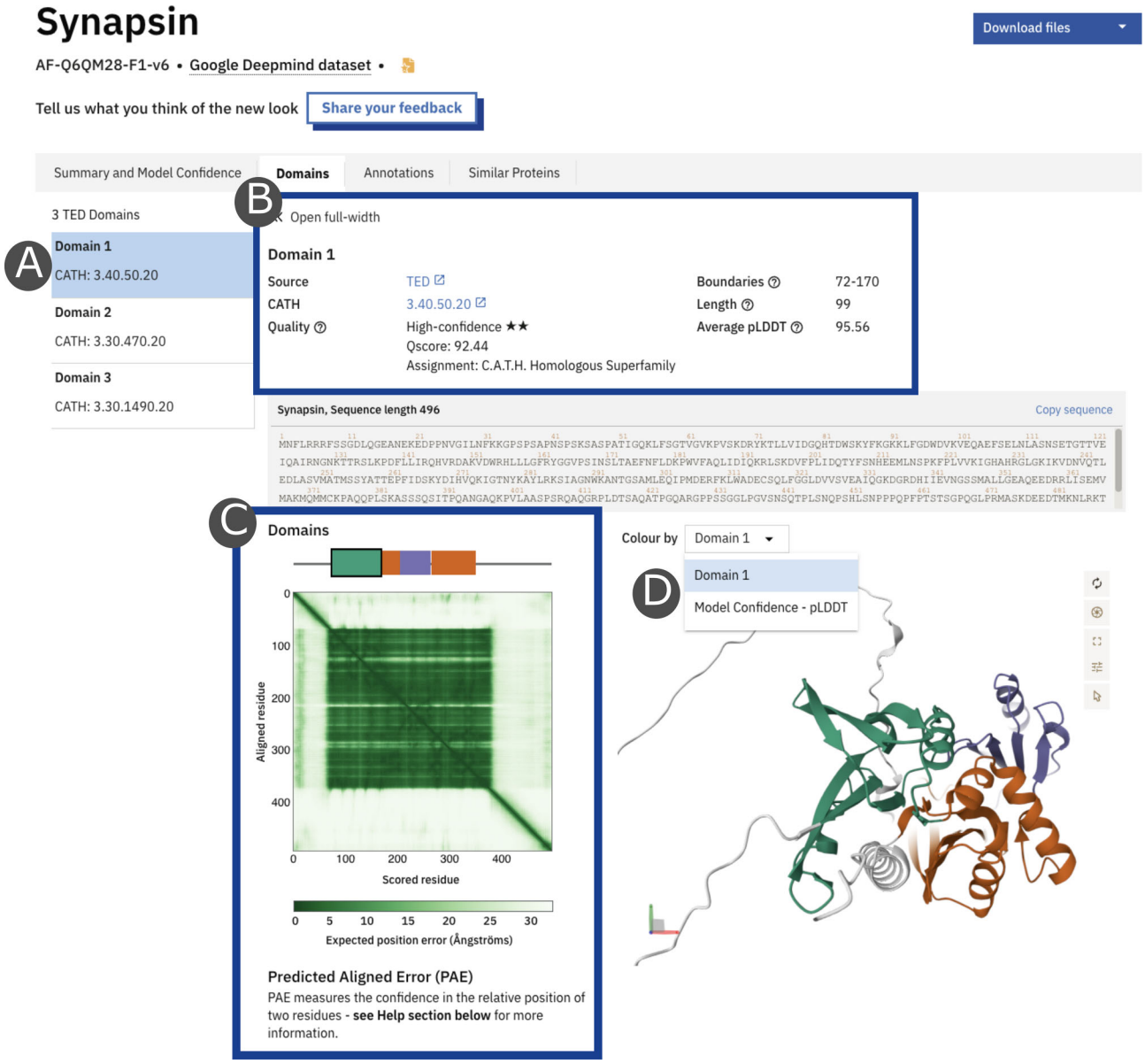


Figure 2. Example of domains tab for Synapsin AF-Q6QM28-F1-v6. This tab provides a comprehensive overview of the information available for a single protein domain provided by TED in the AFDB. (**A**) The panel on the left shows the complete list of domains identified in the entry. (**B**) The descriptive and quality assessment key metrics for a specific domain, including boundaries, length, mean pLDDT, and Qscore, are shown along with the high-confidence CATH assignment. (**C**) The predicted aligned error plot provides a visual representation of the AF2-predicted confidence in the orientation of the domains relative to the rest of the protein. (**D**) Colours on the model can be toggled between domain and pLDDT.

structural insights. This functionality, developed with accessibility and usability principles in mind, lowers the entry barrier for expert and non-expert users, making use of advanced structural biology data more inclusive and impactful for the global scientific community. This commitment to user-centred design ensures that the powerful insights from AlphaFold predictions are readily available and comprehensible to a broad audience.

Data update

The AFDB collection has now been updated to UniProt release (2025_03), incorporating Swiss-Prot/UniProtKB isoforms. The integration of isoforms offers a more complete view of proteome diversity and has revealed widespread changes in protein domain architecture and interaction interfaces [29, 30]. Furthermore, we now provide the underlying MSAs and MSA depths used to predict the structures, offering a new layer of data for evolutionary and structural analyses.

Integrating the Encyclopedia of Domains

To facilitate functional understanding of the proteins in AFDB, over 361 million domain annotations from the Encyclopedia of Domains (TED) [21] covering over 165M proteins have now been added. The TED methodology employs a state-of-the-art, consensus-based approach that integrates three deep-learning-based domain parsing algorithms to robustly determine domain boundaries. These domain assignments are subsequently aligned against the CATH database (Version 4.4)

[31], using structural alignment algorithms. Each domain entry in the AFDB is associated with a suite of descriptive and quality assessment metrics (Fig. 2). A consensus level serves as an initial quality indicator, reflecting the degree of agreement among the three domain-parsing methods. For domains with a structural match, a CATH identifier and assignment are provided based on the level of the hierarchy matched. A second key metric, the Qscore, provides a comprehensive, composite measure of the quality of each domain assignment (details in Supplementary data). This resource supports evolutionary, structural, and functional studies by providing a deeper view of domain organization.

Data availability and API updates

The AFDB prioritizes open data access. All data, including new additions, are accessible via our website (https://alphafold.ebi.ac.uk). The EMBL–EBI's FTP area hosts TAR files for proteomes of 46 organisms, including model organisms and WHO pathogens of interest (ftp.ebi.ac.uk/pub/databases/alphafold). The full AlphaFold dataset is accessible from Google Cloud Public Datasets, with user guidance at https://alphafold.ebi.ac.uk/download and version history documented at https://ftp.ebi.ac.uk/pub/databases/alphafold/CHANGELOG.txt. To support continued development, AFDB is introducing breaking changes to its API schema. The legacy API will be retired in June 2026 (https://alphafold.ebi.ac.uk/api-docs).

Discussion and future directions

The latest updates to AFDB bring the resource back in line with UniProt and introduce an extensible interface designed to support future developments that enrich data and functionality. The inclusion of isoform-specific predictions and newly generated MSAs expands structural coverage and adds essential biological and evolutionary context to the models. This, in turn, lowers the barrier to entry for structure-informed research across the life sciences, enhances reproducibility, and promotes sustainable and greener research practices.

Looking ahead, future releases will continue to work with the scientific community to focus on new datasets and annotations, ensuring that AFDB reflects both the expanding sequence space and the evolving needs of researchers. Its development is guided by three principles: (i) filling gaps in structural coverage, (ii) improving existing models to enhance accuracy and utility, and (iii) working with the scientific community to address global challenges, from antimicrobial resistance to food security. A key priority is the inclusion of structural models and their validation in ways that maximize their scientific impact. Central to this vision is fostering community participation in annotation, building on the successful PDBe-KB model of collaborative consortia. By coupling structural predictions with rich, community-contributed functional insights, AFDB will continue to evolve into a knowledge-rich resource that accelerates discovery and amplifies the impact of protein structure data.

Acknowledgements

We acknowledge the foundational contributions of researchers who consistently deposit structures into public databases, thereby enriching the field of structural biology. We are grateful for the open data provided by major repositories, including UniProt, MGnify, and the PDB, which are indispensable for the training of structure prediction tools. Lastly, we thank the PDBe and Google DeepMind teams for their rigorous and critical evaluation of new features; their collective contributions have been highly beneficial to the development of this work.

Author Contributions: Damian Bertoni (Data curation [equal], Formal analysis [equal], Methodology [equal], Software [lead], Validation [equal]), Maxim Tsenkov (Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Software [supporting], Writing—original draft [equal], Writing—review & editing [supporting]), Paulyna Magana (Writing—original draft [equal], Writing—review & editing [supporting]), Sreenath Nair (Data curation [supporting], Methodology [supporting], Project administration [supporting], Software [equal], Supervision [equal], Validation [equal]), Ivanna Pidruchna (Validation [equal], Visualization [lead]), Marcelo Querino Lima Afonso (Software [equal], Validation [supporting]), Adam Midlik (Methodology [supporting], Software [supporting], Validation [supporting], Visualization [supporting]), Urmila Paramval (Software [lead], Validation [supporting], Visualization [supporting]), Dare Lawal (Software [supporting], Validation [supporting]), Ahsan Tanweer (Software [equal]), Meera Last (Project administration [supporting], Resources [supporting]), Risha Patel (Project administration [supporting], Writing—review & editing [supporting]), Agata Laydon (Resources [equal], Supervision [equal]), Dariusz Lasecki (Software [equal]), Nick Dietrich (Software [equal]), Hamish Tomlinson (Software [equal]), Augustin Žídek (Data curation [equal], Methodology [equal], Software [equal], Supervision [equal], Validation [equal], Writing—review & editing [supporting]), Tim Green (Validation [equal]), Oleg Kovalevskiy (Supervision [equal], Validation [equal], Writing—review & editing [equal]), Andy Lau (Software [equal]), Shaun Kandathil (Software [equal], Validation [equal]), Nicola Bordin (Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Software [equal]), Ian Sillitoe (Software [equal]), Milot Mirdita (Formal analysis [equal], Software [equal], Validation [equal]), David Jones (Conceptualization [equal], Formal analysis [equal], Funding acquisition [equal], Project administration [equal], Software [equal], Supervision [equal]), Christine Orengo (Conceptualization [equal], Funding acquisition [equal], Project administration [equal], Resources [equal], Supervision [equal]), Martin Steinegger (Conceptualization [equal], Funding acquisition [equal], Resources [equal], Software [equal], Supervision [equal]), Jennifer R. Fleming (Funding acquisition [equal], Resources [equal], Software [equal], Supervision [equal] Conceptualization [supporting], Project administration [equal], Supervision [equal], Validation [equal], Writing—original draft [supporting], Writing—review & editing [equal]), and Sameer Velankar (Conceptualization [equal], Funding acquisition [equal], Resources [equal], Supervision [equal], Writing—review & editing [equal])

Supplementary data

Supplementary data is available at NAR online.

Conflict of interest

None declared.

Funding

This work was supported by the BBRSC (20-BBSRC/NSF-BIO and BB/Y000455/1 to S.V.; BB/W018802/1 to C.O.; BB/T019409/1 to A.M.L. and D.T.J.; and BB/W008556/1 to S.M.K. and D.T.J.) and the Welcome Trust (221327/Z/20/Z to C.O. and 310300/Z/24/Z to S.V.). M.S. acknowledges support by the National Research Foundation of Korea grants (RS-2020-NR049543, RS-2021-NR061659 and RS-2021-NR056571, RS-2024-00396026) and Creative-Pioneering Researchers Program and Novo Nordisk Foundation (NNF24SA0092560). M.M. acknowledges support from the National Research Foundation of Korea (RS-2023-00250470). Google DeepMind and the European Molecular Biology Laboratory fund the AlphaFold Protein Structure Database. Funding to pay the Open Access publication charges for this article was provided by European Molecular Biology Laboratory.

Data availability

The AFDB is available at https://alphafold.ebi.ac.uk.

References

1. Jumper J, Evans R, Pritzel A *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;596:583–9. https://doi.org/10.1038/s41586-021-03819-2
2. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024;630:493–500. https://doi.org/10.1038/s41586-024-07487-w
3. Evans R, O'Neill M, Pritzel A *et al.* Protein complex prediction with AlphaFold-Multimer. bioRxiv, https://doi.org/10.1101/2021.10.04.463034, 10 March 2022, preprint: not peer reviewed.
4. Lin Z, Akin H, Rao R *et al.* Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 2023;379:1123–30. https://doi.org/10.1126/science.ade2574
5. Baek M, DiMaio F, Anishchenko I *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 2021;373:871–6. https://doi.org/10.1126/science.abj8754
6. Yanofsky C, Horn V, Thorpe D. Protein structure relationships revealed by mutational analysis. *Science* 1964;146:1593–4. https://doi.org/10.1126/science.146.3651.1593
7. Göbel U, Sander C, Schneider R *et al.* Correlated mutations and residue contacts in proteins. *Proteins* 1994;18:309–17.
8. Marks DS, Colwell LJ, Sheridan R *et al.* Protein 3D structure computed from evolutionary sequence variation. *PLoS One* 2011;6:e28766. https://doi.org/10.1371/journal.pone.0028766
9. Gutmanas A, Alhroub Y, Battle GM *et al.* PDBe: Protein Data Bank in Europe. *Nucleic Acids Res* 2014;42:D285–91. https://doi.org/10.1093/nar/gkt1180
10. consortium wPDB. Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res* 2019;47:D520–8. https://doi.org/10.1093/nar/gky949
11. Richardson L, Allen B, Baldi G *et al.* MGnify: the microbiome sequence data analysis resource in 2023. *Nucleic Acids Res* 2023;51:D753–9. https://doi.org/10.1093/nar/gkac1080
12. UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2025. *Nucleic Acids Res* 2025;53:D609–17. https://doi.org/10.1093/nar/gkae1010
13. Varadi M, Anyango S, Deshpande M *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* 2022;50:D439–44. https://doi.org/10.1093/nar/gkab1061
14. Varadi M, Bertoni D, Magana P *et al.* AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res* 2024;52:D368–75. https://doi.org/10.1093/nar/gkad1011
15. Fleming J, Magana P, Nair S *et al.* AlphaFold Protein Structure Database and 3D-Beacons: new data and capabilities. *J Mol Biol* 2025;437:168967. https://doi.org/10.1016/j.jmb.2025.168967
16. Piovesan D, Del Conte A, Mehdiabadi M *et al.* MOBIDB in 2025: integrating ensemble properties and function annotations for intrinsically disordered proteins. *Nucleic Acids Res* 2025;53:D495–503. https://doi.org/10.1093/nar/gkae969
17. Blum M, Andreeva A, Florentino LC *et al.* InterPro: the protein sequence classification resource in 2025. *Nucleic Acids Res* 2025;53:D444–56. https://doi.org/10.1093/nar/gkae1082
18. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004;60:2126–32. https://doi.org/10.1107/S0907444904019158
19. Meng EC, Goddard TD, Pettersen EF *et al.* UCSF ChimeraX: tools for structure building and analysis. *Protein Science* 2023;32:e4792. https://doi.org/10.1002/pro.4792
20. Waterhouse AM, Procter JB, Martin DMA *et al.* Jalview Version 2—multiple sequence alignment editor and analysis workbench. *Bioinforma Oxf Engl* 2009;25:1189–91. https://doi.org/10.1093/bioinformatics/btp033
21. Barbarin-Bocahu I, Graille M. The X-ray crystallography phase problem solved thanks to *AlphaFold* and *RoseTTAFold* models: a case-study report. *Acta Crystallogr D Struct Biol* 2022;78:517–31. https://doi.org/10.1107/S2059798322002157
22. Terwilliger TC, Afonine PV, Liebschner D *et al.* Accelerating crystal structure determination with iterative *AlphaFold* prediction. *Acta Crystallogr D Struct Biol* 2023;79:234–44. https://doi.org/10.1107/S205979832300102X
23. Chojnowski G. Sequence-assignment validation in cryo-EM models with *checkMySequence*. *Acta Crystallogr D Struct Biol* 2022;78:806–16. https://doi.org/10.1107/S2059798322005009
24. Barrio-Hernandez I, Yeo J, Jänes J *et al.* Clustering predicted structures at the scale of the known protein universe. *Nature* 2023;622:637–45. https://doi.org/10.1038/s41586-023-06510-w
25. Lau AM, Bordin N, Kandathil SM *et al.* Exploring structural diversity across the protein universe with The Encyclopedia of Domains. *Science* 2024;386:eadq4946. https://doi.org/10.1126/science.adq4946
26. Romasanta AK, Wareham J, Pujol Priego L. The impact of research data infrastructures: the case of the AlphaFold database. *CERN IdeaSquare J. Exp. Innov.* 2025;9:42–8.
27. Kovalevskiy O, Mateos-Garcia J, Tunyasuvunakool K. AlphaFold two years on: validation and impact. *Proc Natl Acad Sci USA* 2024;121:e2315002121. https://doi.org/10.1073/pnas.2315002121
28. Tsitsa I, Conev A, David A *et al.* The AlphaFold database ages. bioRxiv, https://doi.org/10.1101/2025.06.22.660930, 27 Spetember 2025, preprint: not peer-reviewed.
29. Song Y, Zhang C, Omenn GS *et al.* Predicting the structural impact of human alternative splicing. *Genome Biol* 2025;26:283. https://doi.org/10.1186/s13059-025-03744-x
30. Yang Y, Xie Y, Li Z *et al.* Systematic characterization of protein structural features of alternative splicing isoforms using AlphaFold 2. bioRxiv, https://doi.org/10.1101/2024.01.30.578053, 19 February 2024, preprint: not peer-reviewed.
31. Orengo C, Michie A, Jones S *et al.* CATH—a hierarchic classification of protein domain structures. *Structure* 1997;5:1093–109. https://doi.org/10.1016/S0969-2126(97)00260-8

Received: September 30, 2025. **Revised:** October 18, 2025. **Accepted:** October 18, 2025