

AlphaFold 3: an unprecedented opportunity for fundamental research and drug development

Ziqi Fang,^{1,2} Hongbiao Ran,^{2,3} YongHan Zhang,^{2,4} Chensong Chen,⁵ Ping Lin,³ Xiang Zhang,^{1,*} Min Wu^{1,2,5,6,*}

¹School of Medical Information Engineering, Gannan Medical University, Ganzhou 341000, China

²RNA Research and Drug Discovery Center, Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou 325000, China

³Integrative Science Center of Germplasm Creation in Western China (Chongqing) Science City, Biological Science Research Center, Southwest University, Chongqing 400715, China

⁴Department of Graduate College, Zhejiang Chinese Medical University, Hangzhou 310000, China

⁵The Joint Research Center, Affiliated Xiangshan Hospital of Wenzhou Medical University, Ningbo 315000, China

⁶Department of Medicine, Harvard Medical School, and Brigham and Women's Hospital, Boston, MA 02115, USA

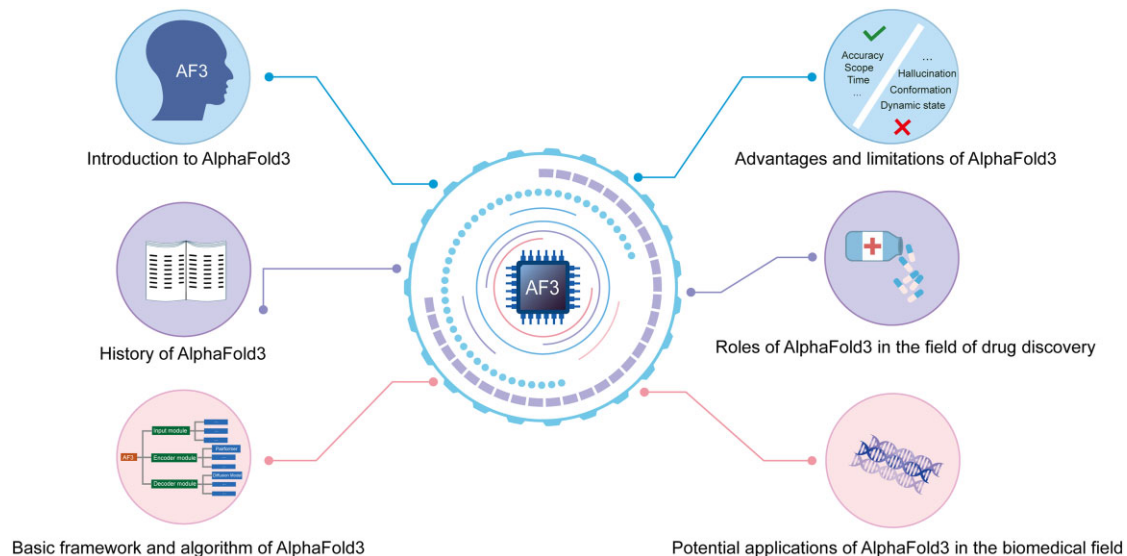
*Corresponding authors: Min Wu, minwoo2022@126.com; Xiang Zhang, zhangxiang@gmu.edu.cn

Abstract

AlphaFold3 (AF3), as the latest generation of artificial intelligence model jointly developed by Google DeepMind and Isomorphic Labs, has been widely heralded in the scientific research community since its launch. With unprecedented accuracy, the AF3 model may successfully predict the structure and interactions of virtually all biomolecules, including proteins, ligands, nucleic acids, ions, etc. By accurately simulating the structural information and interactions of biomacromolecules, it has shown great potential in many aspects of structural prediction, mechanism research, drug design, protein engineering, vaccine development, and precision therapy. In order to further understand the characteristics of AF3 and accelerate its promotion, this article sets out to address the development process, working principle, and application in drugs and biomedicine, especially focusing on the intricate differences and some potential pitfalls compared to other deep learning models. We explain how a structure-prediction tool can impact many research fields, and in particular revolutionize the strategies for designing of effective next generation vaccines and chemical and biological drugs.

Keywords: AlphaFold3; artificial intelligence; structure prediction; drug design; biomedical research

Graphical Abstract



Overview, application, and perspective of AlphaFold3.

Introduction

The 2024 Nobel Prize in Chemistry was awarded to Demis Hassabis and John Jumper of the DeepMind team in recognition of their breakthrough contributions to the development of the AI-

phaFold series [1]. This honor not only highlights the revolutionary role of artificial intelligence (AI) in the field of protein structure prediction, but also marks the advent of a new era of computational biology. It is notable that several months before

Received 27 April 2025; revised 25 June 2025; accepted 30 June 2025. published 1 July 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of the West China School of Medicine & West China Hospital of Sichuan University. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

the awarding of this prize, DeepMind, in collaboration with Isomorphic Labs, published its latest research result based on AlphaFold3 (AF3) in the journal *Nature* [2]. AF3, while retaining its original technological advantages, immensely expands its prediction range from a single protein to almost all biomolecular systems. By introducing an innovative diffusion model architecture, AF3 has powerfully improved prediction accuracy. This leapfrog development enables algorithm-driven structural prediction to no longer be limited to a single molecular type, opening up brand-new research directions and application prospects for the entire field of structural biology [3, 4]. Although AF3 itself still has certain flaws, it is undeniable that its emergence is expected to effectively address the current challenges in biophysical structure analysis and molecular interaction prediction, infusing new vital nutrients and energy into several important fields, such as biomedical research, drug design, and disease diagnosis and treatment [5].

Despite the short period since its birth, discussions about the value of AF3 have never ceased. To gain a deeper understanding of the potential and challenges of AF3 in disease treatment, this article provides an overview of AF3's development, working principles, and application prospects. By sorting out the innovative applications of AF3 in drug discovery and biomedical research, and objectively analyzing its technical limitations, the aim of this article is to provide comprehensive analyses and inspiration for scientific researchers in broad fields.

History of AlphaFold predicting protein structure

As the cornerstone of life activities and the core regulatory executive, protein is widely involved in and profoundly affects almost all life processes in the cell. The universality and diversity of its functions, fundamentally speaking, come from the unique 3D spatial structure of protein molecules [6]. Therefore, accurate prediction of protein structure is of great significance for deepening our understanding of life phenomena, revealing the molecular mechanisms of disease occurrence, and drastically accelerating the process of drug discovery. Since 1957, when John C. Kendrew and Max F. Perutz first revealed the 3D structure of proteins through X-ray crystallography [7], the ambition for resolving protein structure has never stopped [8–13]. Traditional prediction models face a number of challenges, such as requiring an available experimental database, being time-consuming, requiring a large number of calculations, and having a narrow prediction range, hence, the large-scale popularization of protein structure analysis has been severely hampered [14–16]. Fortunately, the rise of AI technology has brought new opportunities for development in this field. With its high precision and high efficiency, the structure prediction model has been sought after by the majority of biological structure scholars since its emergence [17–19]. Among them, the AlphaFold series models have become a prominent leader in the field of protein structure prediction thanks to their excellent performance (Fig. 1).

AlphaFold

In 2018, DeepMind, which rose to fame due to AlphaGo [20], introduced an algorithm model, now well-known as AlphaFold, and its release immediately attracted great academic attention [21]. To improve prediction accuracy, AlphaFold uses an advanced generative neural network architecture and is optimized with gradient descent algorithms. With this technique, researchers are able to

predict the distances between pairs of amino acids and the angles between the chemical bonds that connect these amino acids. This information has important guiding ability and is extremely critical for improving mechanistic research and disease treatment. In that year's International Critical Assessment of Structure Prediction (CASP) competition, AlphaFold won first place with its excellent performance, further showing its leading position in the field of protein structure prediction [22, 23]. Unfortunately, AlphaFold still has deficiencies in prediction accuracy and is limited by its own modularity when dealing with local protein interactions as well as broader structural features. This affects AlphaFold's effectiveness in some specific application scenarios, making it relatively limited in scope.

AlphaFold2

AlphaFold2 (AF2) was officially launched in 2021 [24]. But before that, it had already won the championship of the 14th CASP [25]. During the competition, this program demonstrated a predictive ability far exceeding that of other contestants, and its results were nearly consistent with the structures observed by cryo-electron microscopy [10]. By leveraging its high-precision predictive capabilities, AF2 successfully predicted >200 million protein structures [26] and topped the list of Science's annual scientific Breakthroughs in 2021 for its contribution [27]. Since its launch, AF2 has been widely used in drug development [28], biological mechanism research [29–38], and protein function prediction [39–43], becoming one of the indispensable tools in the field of structural biology. Researchers have developed several new models [such as AlphaFold Multimer (AFM) [44], PCMol [45], etc.], which play important roles in the fields of reproductive biology [46] and drug design. Although AF2 currently remains active, it has certain limitations, including the accuracy of protein prediction, the high hardware requirements, the uniformity of protein conformation prediction, and the oneness of the prediction object [47]. These shortcomings mean that experimental structural biology still plays a leading role to some extent, and protein prediction through AI still needs further innovation and development.

AlphaFold3

To address the inherent defects, DeepMind continued refining its program and released AF3 on 8 May 2024, marking a major breakthrough in AI in the field of biological science, especially in the prediction of biomolecular structure. As the latest version of the AlphaFold series, AF3 not only inherits the outstanding performance of its predecessor AF2 in protein structure prediction, but also significantly expands its application scope through technological innovation, enabling it to accurately predict the complex structure of small molecules, nucleic acids, substrates, and other ligands. These features make AF3 stand out among the various existing deep learning tools, opening up a broader prospect for research in the field of biological sciences (Table 1). However, as with all emerging technologies, AF3 is not perfect and still has shortcomings in the prediction of structural conformation and dynamic behavior. These problems may be solved with the open-source optimization of AF3 [48]. In conclusion, the emergence of AF3 has had a profound impact on the field of structural biology, greatly expanding the scope of basic science research, facilitating the development of structure analysis of biological macromolecules and the prediction of molecular interaction.

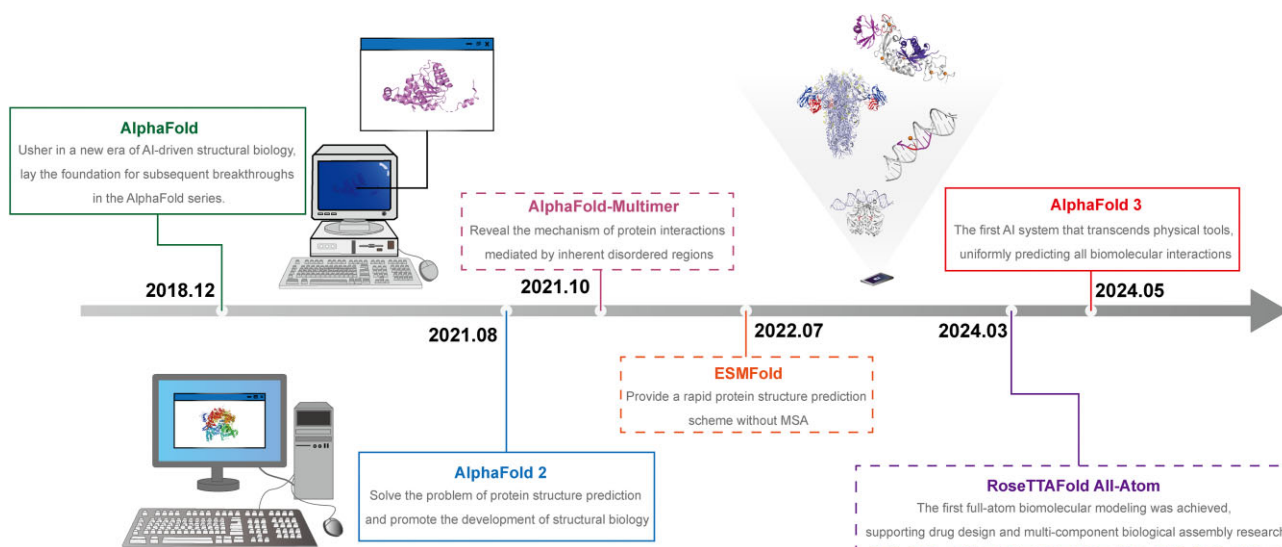


Figure 1. Development history and key breakthroughs timeline of the AlphaFold series of models. The diagram showcases the milestone progress from the first-generation AlphaFold initiating a new era of AI-driven structural biology in 2018 to AlphaFold 3 achieving full prediction of biomolecular interactions in 2024. In addition to the AlphaFold family, this figure also describes its extension algorithms and other protein prediction models. The core contributions of each model are as follows: AlphaFold 2 solves the problem of protein structure prediction (2021), AlphaFold-Multimer reveals the mechanism of disordered protein interactions (2021), and ESMFold achieves rapid prediction without multiple sequence alignment (MSA) (2022). RoseTTAFold All-Atom completed the full-atom biological modeling (2024). The time nodes mark the release time of each model, reflecting the technological update trajectory in this field.

Table 1. Comparison of the five mainstream protein structure prediction tools.

Trait	AF3	AF2	AlphaFold-Multimer	RoseTTAFold All-Atom	ESMFold
Release time	May 2024	August 2021	October 2021	March 2024	July 2022
Main objective	Predict almost all biomolecular complexes	Predict protein monomers and some complexes	Predict the structure of protein-protein complexes	Predict the structure of all-atomic biomolecular complexes	Rapid prediction of protein monomers
Input data	Multiple types of biomolecular sequences	Standard amino acid sequence	Multi-protein chain sequence	Amino acid/nucleic acid sequence	Only protein sequences (without multiple sequence alignment)
Output result	The 3D structure and interaction of proteins and their complexes	3D structure of protein	Protein-protein complex structure	The structure of all-atomic biomolecular complexes	3D structure of protein
Algorithm optimization	The pairformer model and the diffusion model	Based on transformers and evolution models	Based on AF2, it supports multi-chain prediction and symmetric processing.	All-atomic modeling	Pure language model
Application scope	Prediction of structures and interactions of proteins, nucleic acids, small molecule ligands, ions, covalent modifications, etc.	Single protein structure prediction	Research on the structure and mechanism of action of protein-protein complexes	Protein-small molecule docking and covalent modification	Large-scale high-throughput protein structure prediction
Accuracy	Compared with AF2, the accuracy has been further improved and it outperforms professional tools in the vast majority of structural predictions.	CASP14 has reached a level close to experimental accuracy.	It performs well in predicting the interactions between proteins (including disordered structures).	It is superior to the traditional docking method in the prediction of protein-small molecule complexes but inferior to AF3.	It is close to the accuracy of AF2, but the prediction of some complex structures is slightly inferior.
Open source situation	Open source	Open source	Not officially open-sourced	Open source	Open source

Working principle of AF3

As a highly regarded deep learning model for protein structure prediction, the core concept of AF3 is rooted in the deep integration of deep learning and protein amino acid sequence characteristics. This model achieves high-precision prediction of the 3D structure of proteins by integrating the advantages of convolutional neural networks and residual networks and combining with large-scale data training. In terms of architectural design, AF3 continues the core framework of AF2 and still adopts multiple sequence alignment (MSA) as the main input feature. However, in order to achieve a better prediction effect, AF3 has made significant adjustments on this basis (Fig. 2A). The new architecture introduces a more concise Pairformer [49] to replace the Evoformer module in AF2, reducing the processing burden of MSA, and focuses more on extracting critical evolutionary information. In order to simplify the machine learning architecture and improve the prediction efficiency, AF3 uses the diffusion model to replace the traditional structure generation module. This reduces the complexity of the model and the loss of stereochemical information. Similarly, by eliminating the global rotational and translational invariance or equivariance of the molecules, AF3 further optimizes the computational process, enabling the model to focus primarily on the prediction of core structural features.

The detailed architecture of AF3 has been fully elaborated in the original article [2]. For the convenience of the journal's diverse readership, we briefly outline its architecture and rationale. Similar to AF2, AF3's model framework is made up of three parts.

Input module

The input module of AF3, as the primary link of the entire prediction process, undertakes the key tasks of data integration and feature extraction. Its core responsibility is to construct a comprehensive initial dataset by integrating evolutionary information (MSA) and known structural templates. Compared to AF2, the key breakthrough of the AF3 input module lies in the introduction of the conformer generation technology, which can automatically generate the initial 3D structure based on intramolecular interaction law, providing a more reasonable starting point for subsequent predictions.

In the input module of AF3, the raw input data covers a wide range of biomolecular information, including but not limited to amino acid sequences, ligands, covalent bonds, etc. To better handle this information, AF3 adopts a hierarchical feature extraction mechanism. Specifically, for conventional amino acids and nucleotides, the system adopts an overall characterization approach. For modified residues and small-molecule ligands, atomic-level fine analysis enables the complete retention of molecular details. This differentiated processing approach should be taken into account for both computational efficiency and prediction accuracy. Subsequently, the system converts the molecular information into high-dimensional vector representations through a feature-embedding process. Among them, the pair representation is used to represent interactions between pairs of residues, revealing their spatial proximity and possible chemical bonding. Single representation focuses on capturing the characteristics of individual residues, including their chemical properties, shapes, and patterns of interaction with other residues. This transformation retains the biological characteristics of the original data, and establishes the correlation network among tokens through the attention mechanism, laying a solid foundation for AF3's protein structure prediction task.

Encoder module

AF3's encoder module adopts a three-stage feature extraction architecture, including the Template module, the MSA module, and the Pairformer module, which jointly achieve the deep processing and feature extraction of input data.

In the initial stage, template structures acquire known structural templates similar to the target sequence through a structural homology search, and encode the spatial conformational features of these templates into computable mathematical representations. In essence, this process utilizes the rich known structural information in the structural database to assist in predicting the unknown protein structure, thus improving the accuracy and reliability of the prediction.

On this basis, the developers further use the MSA module [50] to process information and identify the co-evolutionary signals among residues from large-scale sequence alignment. These signals reveal the spatial and functional interactions formed by amino acids under evolutionary pressure, providing key constraints for inferring 3D spatial proximity.

The Pairformer module [49] is the core component of the encoder module, mainly responsible for resolving the local interactions between adjacent residues. This module adopts an improved transformer architecture. Based on inheriting the 48-layer stacked design of the Evoformer module in AF2, it retains key algorithm components, such as triangular update and triangular self-attention (Fig. 2B). On this basis, this model significantly improves the efficiency of feature extraction through architectural innovation. It simplifies the multi-sequence alignment processing to four sequences, significantly reducing the computational complexity. Meanwhile, the expression methods centered on residue pairs (preliminary single representation and pair representation) are adopted to capture the spatial relationship between adjacent residues more accurately. This design significantly reduces the model's reliance on evolutionary information while maintaining high-precision structural prediction capabilities.

In the course of Pairformer's work, there is a complex interaction between single representation and pair representation. Although a single representation does not directly affect the expression of a pair representation, a pair representation can influence the flow of information in a single representation through the single attention with pair bias mechanism. This interaction enables the encoder module to capture and integrate the information in the input data more comprehensively, providing a richer and more accurate feature representation for the subsequent decoding process.

After multiple layers of repeated optimization, the encoder finally outputs high-quality single representation and pair representation. These two representations are fed into the decoder module to further generate and optimize the 3D structure of the protein.

Decoder module

Compared to AF2, AF3's decoder module has undergone significant innovation. Its core is the introduction of diffusion-based generative modeling [49] to replace the iterative optimization framework of AF2. Its working principle is to directly predict the atomic coordinates through the reverse process of multiple rounds of noise addition and removal (Fig. 2C), achieving a progressive construction from noise distribution to fine structure. This innovative mechanism enables the model to gradually correct the predictions, so that, in the early denoising stage, large-

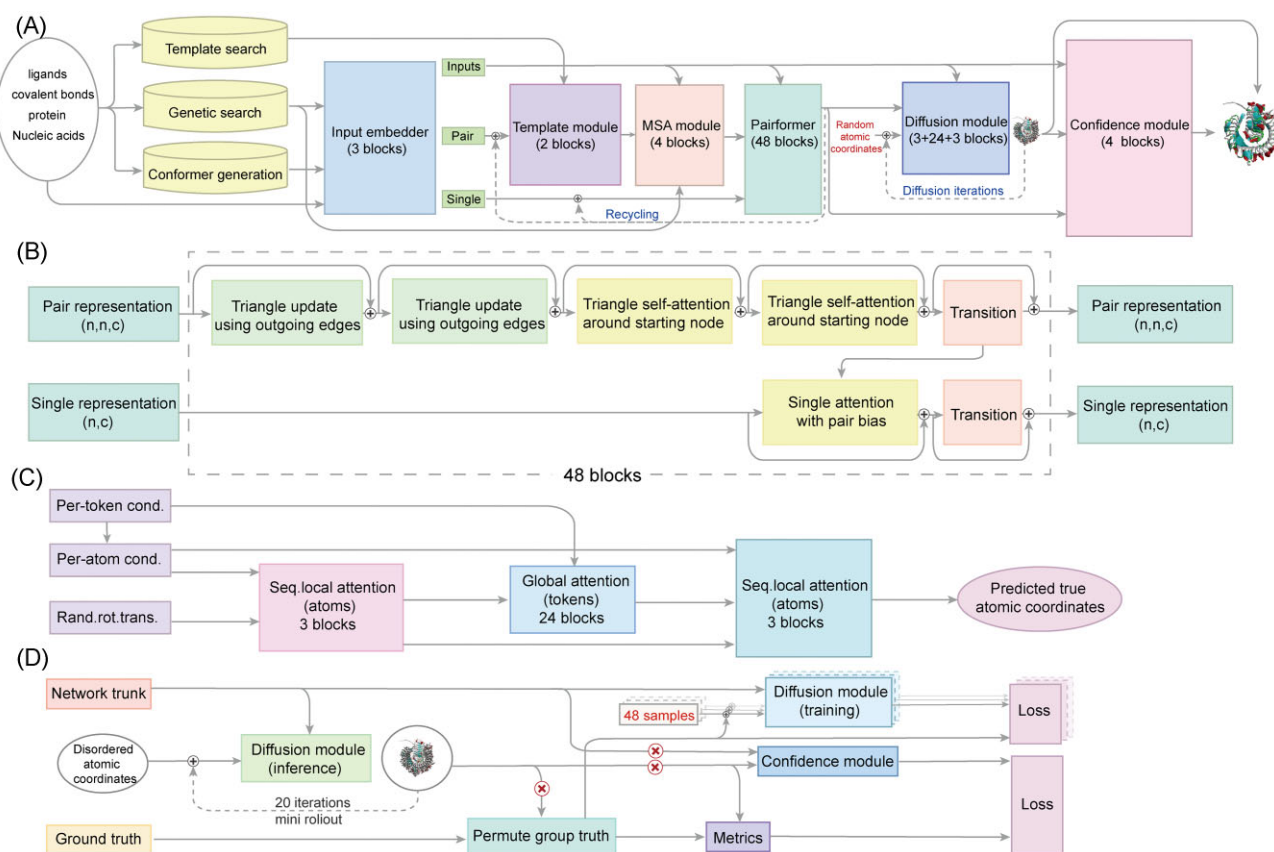


Figure 2. Basic framework of AF2 and working principle diagram of each component. **(A)** Overall framework diagram of AF3. The overall framework of AF3 is divided into three main parts. The first part is the input module, which is responsible for collecting and integrating MSA and template structure information as input data for the entire model. The second part is the encoder module, which adopts a pairformer module to replace the evoformer module in AF2 to achieve the processing of input data and feature extraction. The third part is the decoder module, the core of which is the diffusion module. It can generate the 3D structure of target biological macromolecules by denoising. **(B)** Working principle diagram of the pairformer module. This module takes the pair representation and single representation output of the previous module as input data, and only keeps four MSA sequences, which greatly simplifies the process and shortens the processing time. **(C)** Diffusion module, as a mini rollout. Using raw atomic coordinates and coarse-grained abstract token representations as input objects, the predicted atomic coordinates are generated directly without twisting or equivariant processing. **(D)** Diffusion training diagram. The module is able to handle the training of 48 samples at the same time, and the error is predicted by the confidence module.

scale noise is processed to determine the overall conformational framework, and in the subsequent denoising stage, the fine adjustment of the local structure is focused on, ultimately generating accurate atomic positions. Compared to the direct structure prediction by AF2, this generation strategy acts on the original atomic coordinates and coarse-grained abstract token representations, reducing labor costs. Moreover, it can integrate multi-scale features through noise scheduling, handle complex ligand environments more naturally, and avoid conformational constraints introduced by torsional angle parameterization. Its parallel training architecture supports the simultaneous processing of 48 samples. These creative steps significantly improve the training efficiency and model generalization ability.

The diffusion module adopts a dual-mode design, including the inference diffusion module and the training diffusion module (Fig. 2D). The inference diffusion module primarily serves as the test phase. The diffusion model starts from random noise and generates accurate atomic coordinates through multi-round denoising combined with the attention mechanism. The training diffusion module is responsible for the training phase. It takes a known protein structure model from the database [usually derived from Protein Database (PDB) and MSA], superimposes controllable Gaussian noise on the real structure,

and then trains the diffusion model to restore the real coordinates. This multi-scale training strategy enables the model to focus on the precise modeling of local chemical bonds under low-noise conditions and emphasizes the feature learning of the overall conformation in a high-noise environment, achieving collaborative structural prediction from micro to macro.

The generative diffusion model brings many advantages to AF3 and can overcome the limitations inherent in AF2 to some extent. However, its existence has also caused some new problems, the most prominent of which is the phenomenon of hallucination [51]. To counter this, AF3's developers introduced cross-distillation technology and used the prediction results of the more complex model (AFM v2.3) as the supervisory signal to guide the model to correctly identify the disordered regions. In AF2, an unordered region is represented as a long-extended loop structure. By learning and analyzing this structure, AF3 is able to construct disordered regions more efficiently, thus mitigating hallucinatory effects. It is worth noting that cross-distillation does not simply copy the predictions of another model, but uses those results as references to train the current model. This method not only expands the diversity of the training data, but also effectively suppresses the generation of false structures by imitating the specific

conformations of the disordered regions. To evaluate the quality of predictions, AF3 has established a comprehensive quality assessment system. By comparing the differences between the predicted structure and the experimental data at multiple scales (including local conformational accuracy, global topological matching degree, etc.), the system provides a credibility score for each prediction result. Generally speaking, the higher the confidence score, the higher the accuracy of the prediction result. These innovations enable AF3 to achieve more accurate and reliable prediction of biomolecular structures while maintaining efficient computing.

Application potential of AF3 in drug design and biomedicine

In the wave of science and technology in the 21st century, research in the field of the life sciences has made substantial progress. Among such research, the development of drug discovery and biomedicine has benefited from the progressive understanding of the structure of biological macromolecules. Whether it is the understanding of the nature of life or the optimization of disease treatment, how to efficiently obtain accurate structural information of biological macromolecules is key to success. With the continuous development of AI, the prediction of a drug's molecular structure is no longer simply by means of classically standard experiments. More and more deep learning models have gradually appeared in the public eye, providing efficient and convenient prediction tools for structural biology, and promoting the rapid development of various fields [52]. In this context, the emergence of AF3 undoubtedly provides researchers with a new shortcut, allowing them to explore the micro world to a greater extent and use the results for the benefit of mankind. With its strong performance, AF3 may have great application potential in many disciplines and inject new energy and vitality into the development of the life sciences (Fig. 3).

Drug design

Drug design has traditionally been regarded as a complex and time-consuming scientific task. The key lies in the in-depth understanding of the fine structure of molecular structure in living organisms and the intricate mechanism of molecular interaction. This understanding forms the logical starting point for drug development and strategies to treat disease. Traditional methods primarily rely on experimental techniques, such as X-ray crystallography and cryo-electron microscopy. Although these methods are precise, they are time-consuming and labor-intensive. The resolution of a single protein structure usually takes months or even years. This severely limits the efficiency of research and development. Although breakthroughs have been made in drug design based on the known structures of the PDB, it is still limited by data coverage and quality issues. In recent years, the vigorous development of AI technology has brought revolutionary changes to the field of drug research and development [53–58], among which the performance of the AlphaFold series of algorithms is particularly remarkable [59–63].

AF3, as the latest version of this series, has achieved significant breakthroughs in both prediction accuracy and application scope. Unlike the second generation, AF3 not only maintains the high accuracy of protein monomer structure prediction, but also extends the prediction range to complex systems including protein–ligand and protein–nucleic acid. In the PoseBusters benchmark test [64], the prediction accuracy of AF3 is much better than that

of traditional molecular docking tools, and it does not require structural information. This marks that for the first time AF3 has surpassed physics-based AI systems in the field of biomolecular structure prediction. When compared to AFM and ColabFold [65], AF3 has also demonstrated excellent predictive ability for the structure of protein–peptide complexes [66]. These technological breakthroughs provide a brand-new methodological basis for drug design, enabling researchers to understand the mechanisms of molecular interactions at the atomic scale. With this advantage, AF3 successfully predicted >14 000 post-translational modification (PTM) protein–ligand complexes, opening up new avenues for exploring the interaction between PTM and drugs [67].

The test results of the SKEMPI database show that, compared to FoldX [68] and BindProfX [69], AF3 has unique advantages in estimating binding energy and free energy. In addition, combining it with deep learning methods like force fields and profilers can significantly improve the prediction performance [70]. Therefore, this model has been widely applied in the detection of various binding abilities, including protein–protein [71], protein–ligand [72] etc. These verifications and applications have laid a solid theoretical and practical foundation for its wide application in drug development and other related fields.

In terms of target identification, AF3 also performs well, and is especially suitable for the research and development of targeted drugs. The design principle of this type of drug is mainly based on antigen–antibody specificity or peptide affinity to achieve precise drug delivery [73, 74]. Therefore, accurately predicting the antibody–protein binding characteristics is crucial for understanding the immune response mechanism and developing new antibody drugs. AF3 perfectly meets this demand. Studies have shown that AF3 can accurately predict the key structural characteristics of antibody–protein interactions. Its accuracy rate is 33.3% higher than that of AFM v2.3, and the prediction accuracy rate of protein–protein complexes has also increased by nearly 10%. These technical advantages make AF3 a powerful tool for target identification and verification, not only accelerating the structural resolution of known targets but also helping in the discovery of new potential drug targets.

During the drug optimization stage, the technical advantages of AF3 are prominent. As an important link in the research and development of new drugs, the key to drug optimization lies in balancing efficacy and safety. Due to the lack of precise prediction of molecular interactions, the tissue-specific delivery efficiency of targeted drugs is not satisfactory. In addition, drug molecules are susceptible to interference from various factors in a complex biological environment, often resulting in delivery route deviation, reduced efficacy, and unexpected toxicity. These bottleneck problems have severely restricted the efficiency and success rate of new drug development.

At present, molecular docking technology has become a common method to evaluate drug affinity and binding ability [75–79]. By simulating the binding process of candidate drug molecules and target proteins, it can provide an important basis for drug design. In the tests conducted by the DeepMind team, the full-atom docking accuracy of AF3 reached 76.4%, which was 1.8 times that of RoseTTAFold All-Atom (RFAA) [80] docking (42%), and was also superior to other classic docking tools, such as Vina [81, 82], Gold, and Uni-Mol docking V2. It is particularly worthwhile to mention that AF3 performs exceptionally well in predicting covalent modifications (including ligands, glycosylation, and protein residue modifications, etc.), with the accuracy rates of all predictions exceeding 40%. This characteristic gives it a unique benefit in the development of targeted drugs based on covalent binding.

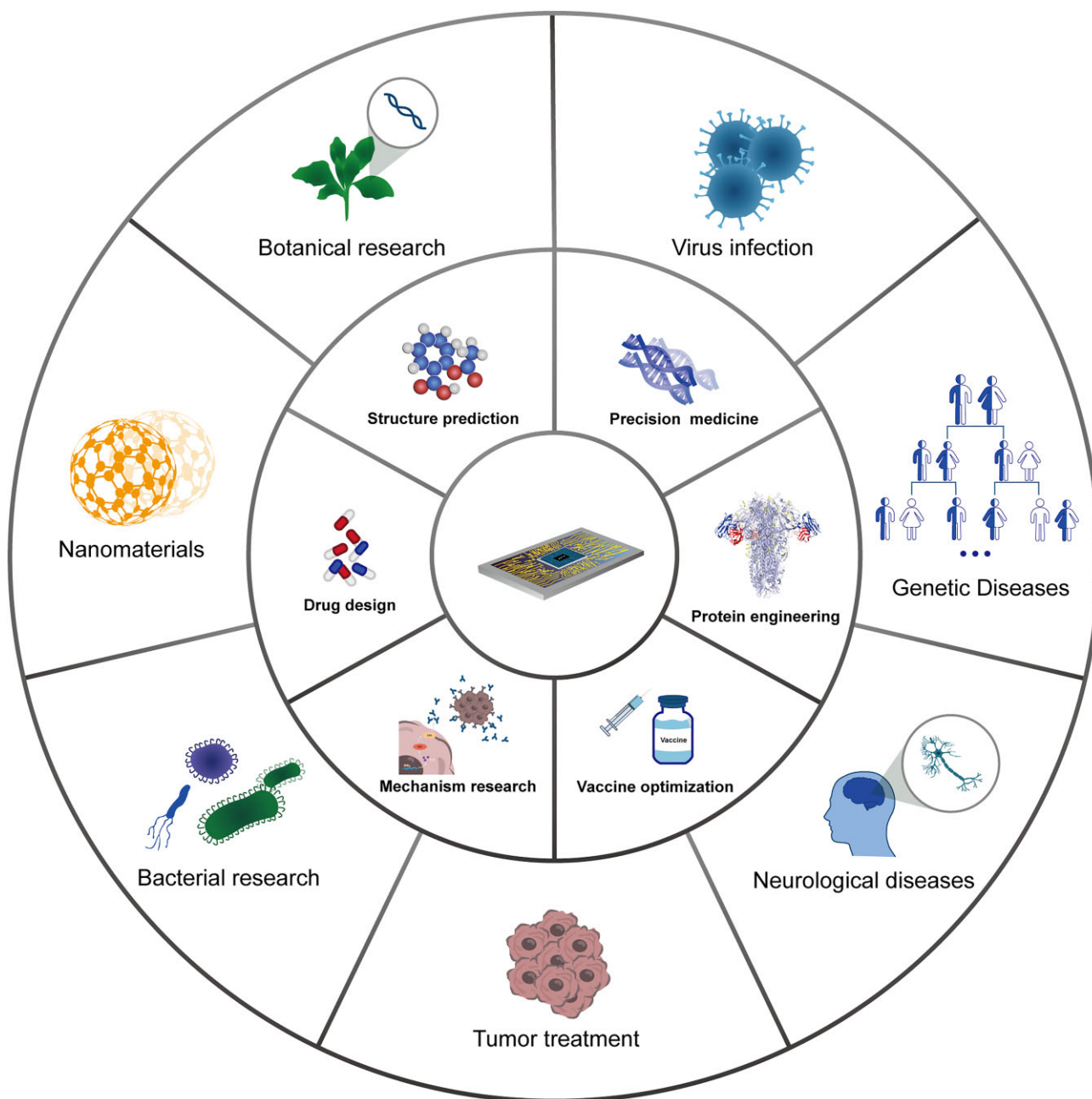


Figure 3. Multi-field applications of AF3. Because of its excellent performance, AF3 has broad application prospects in many fields. These fields include basic research (mechanism research, botanical research, virus infection, bacterial research), structural biology (structure prediction, protein engineering), medical applications (precision medicine, tumor treatment, neurological diseases, and genetic diseases), drug development (drug design, vaccine optimization, and nanomaterials) etc. The diagram highlights the potential extensive value of AF3 from molecular mechanism analysis to clinical transformation, demonstrating its multi-functionality as a core tool in computational biology.

Taking advantage of this, scientists not only precisely locate the binding sites of ligands and receptors [83], but also make preliminary judgments on the effects after drug modification through their predictions [84].

With these characteristics, AF3 is widely used in the research and development of drugs for various diseases. In neurological diseases, AF3 successfully conducted high-resolution modeling of over 1200 brain-related proteins [with an average Predicted Local Distance Difference Test (pLDDT) > 80] and identified multiple potential drug targets related to Alzheimer's disease [85]. By integrating the Mendelian randomization method, the MR-SPI (a new MR method) identified seven proteins that underwent struc-

tural changes due to missense mutations (such as TREM2, CD33, etc.), providing new clues for the etiological study of Alzheimer's disease [86]. Furthermore, AF3 predicted the interaction between the W-Tau peptide-derived from intron 12 of the tau protein and tau monomers. Experiments confirmed that this peptide could not only inhibit tau aggregation but also depolymerize the paired helical filaments in the brains of patients with Alzheimer's disease, demonstrating great therapeutic potential [87].

Similarly, in the field of cancer treatment, AF3 successfully predicted the strong interaction between fatty acid hydroxylase domain containing 2 (FAXDC2) and guanosine diphosphate (GDP), and experimentally confirmed the inhibitory function of GDP on

liver cancer cells, providing a new idea for targeted therapy of liver cancer [88]. In addition, by combining it with tools including ProteinMPNN and RFDiffusion, researchers designed nanoantibodies targeting proliferating cell nuclear antigen (PCNA) and B-cell lymphoma 6 protein (BCL6). The bioPROTAC system successfully degraded the target proteins and activated p53, providing a new strategy for cancer treatment [89]. In addition, scientists used this model to study the function of the Kirsten Rat Sarcoma Viral Oncogene Homolog 4B (KRas4B) derived CaaX peptide and found that it could interfere with the Ras signaling pathway by inhibiting farnesyltransferase, providing a new targeting possibility for Ras mutation-induced cancers [90].

For virus research, the low cost and convenience of AF3 have been timely in solving the problem of virus mutations. The AF3-assisted MT-TopLap topological deep learning method can accurately predict the free energy changes of receptor binding domains and angiotensin-converting enzyme 2 (ACE2) mutations in different species, providing an effective tool for rapid response to viral mutations [91]. At a deeper level, combining molecular dynamics and free energy perturbation, the changes in the binding free energy of the SARS-CoV-2 main protease mutant and Niraparvir were analyzed, which further revealed the drug resistance mechanism of the viral mutant and guided the optimization of a new generation of inhibitors [92]. Based on its structural prediction, AF3 is also used to explore the interaction between anti-HIV drugs and targets [93], and to screen out peptides with higher affinity for targets from the newly synthesized peptides, significantly reducing the risk of drug resistance [94]. In addition to HIV, researchers also used AF3 to predict the dimer structure of the I7L protease of the monkeypox virus and the dynamic capsid conformation near the active site. This revealed the structural basis of substrate recognition to guide the design of covalent peptide mimic inhibitors with nanomolar efficacy [95].

In addition to the examples mentioned above, AF3 also plays an important role in drug development for diseases, such as fatty liver [96], muscle atrophy [97], and diabetic cardiomyopathy [98], including animal drugs [99]. The promotion of these applications all emphasize the key value of this model in the field of drug design. With the continuous updating and improvement of the open-source version, AF3 is bound to drive drug research and development towards shorter cycles, better efficacy, and fewer side effects, providing strong support for addressing global health challenges.

Biomedical field

With the continuous development of cutting-edge technologies, such as big data and AI, the biomedical field is undergoing an unprecedented change. Driven by this wave, we have gradually entered the era of precision medicine. In order to achieve better therapeutic effects, researchers have been exploring the structure of biological macromolecules for decades. To obtain accurate structural information efficiently is the goal that this field has been pursuing. However, it is clear that traditional experimental methods cannot meet the growing demand and scientists have begun to turn their attention to the field of AI. Deep learning models, such as AlphaFold, ESMFold [100], and RoseTTAFold [101], have gradually become common tools for structure prediction in recent years, permitting the rapid development of the biomedical field. However, the limitations of their prediction accuracy and range have always been difficult problems to solve. Against this backdrop, the launch of AF3 is undoubtedly exciting and its high accuracy and wide applicability open up new opportunities in var-

ious areas, such as mechanistic research, vaccine development, protein engineering, and precision medicine.

Research on disease mechanisms

Study of the mechanisms of related diseases is an indispensable part of the biomedical field and serves as the foundation for drug development and disease treatment. The fundamental approach is to reveal the underlying causes through the disease manifestations, explore possible related signaling pathways and interactions, and provide theoretical support for the prediction, intervention, and prognosis of the disease. The conformation, modification, and interaction with other biomolecules of proteins are crucial for the study of life activities and disease mechanisms. Therefore, AF3 has been widely applied in this field. When exploring potential therapeutic targets for calcifying aortic valve disease, researchers used AF3 to simulate the interaction between two factors and combined it with other methods to explore how these factors inhibit valve calcification, lending a new targeted solution for the treatment of this disease [102]. In addition to predicting the connection between two biomolecules, AF3 can also predict diseases caused by conformational changes of proteins. In research of small round cell sarcoma, by combining other algorithms, AF3 constructs models of key proteins in the disease state and compares them with the wild type. While revealing the conformational changes, it analyzes the stability and affinity of the proteins, thereby assisting researchers in discovering the pathogenesis and potential breakthroughs of the disease [103]. AF3 is also very popular in basic mechanism research. Some researchers take advantage of its properties to explore a wide range of binding objects of a certain protein and search for potential targets of action [104]. Others also explore the mechanism of action of the target protein through its prediction of protein-protein interactions [105]. Some others have also demonstrated a new protein structure through its simulation function [106]. All these applications have demonstrated the significance of AF3 in mechanism research. It will become an excellent auxiliary tool to help people present the microscopic world of proteins.

Vaccine development and optimization

The principle of vaccine action is to activate the corresponding antigen-antibody response in an artificial way to achieve the rapid response of the body's immune system when the pathogen is invading. Because of its excellent performance in antigen-antibody docking prediction, AF3 has a pivotal role in the field of vaccine development. The core of a vaccine is the antigenic part that stimulates the body to produce an immune response. The model prediction is promising to allow researchers to quickly understand its antigenic structure, accelerating the preparation and optimization of vaccines. This effect has been implemented in AF2. By taking advantage of the structural prediction function of AF2, researchers have developed a hemagglutinin stem cell vaccine [107], which is a major breakthrough in the prevention and control of influenza B. By combining the reverse vaccinology method, researchers successfully designed a peptide-based vaccine against human respiratory syncytial virus using AF3 and verified its structural stability, which is crucial for the efficacy and safety of the vaccine [108]. These achievements attest to the feasibility of using deep learning models for vaccine design and provide more options for subsequent vaccine research and development.

Protein engineering

As a new research field, protein engineering has been improving rapidly with the development of the era. It mainly achieves

the change of protein structure by artificial means, and even creates new proteins. In this field, how to accurately screen the desired proteins and how to efficiently achieve the directed evolution of proteins are invariable problems. At the World Conference on Artificial Intelligence in 2024, the launch of NewOrigin indicates that AI also has significant advantages in protein engineering. Coincidentally, EVOLVEpro, developed by the Massachusetts General Hospital Brigham in the United States, effectively achieved the directed evolution of proteins, as has been demonstrated in six proteins [109]. In a recent study, AF3 guided the engineering of the enNlovFz2 system [110]. This is a miniature eukaryotic gene editing tool. After engineering, both its recognition range and editing efficiency have been significantly improved. In addition, through the simulation of hydrophobic and hydrophilic structures of proteins, AF3 can provide guidance for the engineering of water-soluble integrated membrane proteins, which is of great significance for subsequent drug design and disease treatment [111, 112]. These achievements all reflect the revolutionary changes AI has brought to the field of protein engineering. We have reason to believe that AF3, as one of the best developments, will certainly bring greater surprises to the field in the future.

Precision medicine

Precision medicine is a major trend in modern medicine. It specifically refers to a standardized targeted treatment plan tailored for the alterations of biomarkers, such as genes and proteins, in a particular patient group. Focusing the target on specific molecular mechanisms can not only ensure the therapeutic effect but also reduce damage to the body during the treatment. Genomics, transcriptomics, and proteomics are some of the important analysis methods of precision therapy, and AI plays an important role in these endeavors. In 2023, an article published in *Science* detailed the use of whole-genome sequencing and deep learning tools to analyze variations in primate genomes to detect the pathogenicity of their variants [113]. Before that, the related structure sets of neurodevelopmental disorders (NDD) and uveal melanoma were also sorted and summarized through AlphaFold prediction and wet lab experiments [114, 115]. Based on these studies, AF3 was broadly applied in the field of precision medicine as soon as it was released. Due to its accuracy and wide applicability, this model has the advantage over other deep learning tools and can accurately simulate the structural and energy differences between the wild type and mutants, intuitively reflecting the influence of gene mutations on protein conformation and molecular interactions that is not limited to proteins. For instance, AF3 may dissect complex disease mechanisms of an aggressive type of Duchenne muscular dystrophy (DMD) by identifying the change in binding structure of a compound genetic variant in the DMD and TNNI3K genes, potentially improving clinical management. [116]. In genetic hereditary diseases, the advantages of AF3 are more prominent. Neurofibromatosis type 1 caused by mutations in the NF1 gene has always been a difficult problem in the medical field. To visually discover the changes in protein conformation caused by each mutation site, researchers simulated the mutant protein through AF3, revealing the key impact of the R1000C mutation on the disease [117]. In addition, AF3 has been widely applied in the research of brain development defects [118], Marfan syndrome [119], fetal adipose granulomatosis [120], intellectual development disorders [121], and other diseases. These achievements provide good evidence of how precision medicine can promote the development of modern medicine.

Other application fields

In addition to the above mentioned areas, AF3 also displays extensive potential in many other fields, including botany, nanomaterials, and bacterial research. These potential capabilities take AF3 to the next level.

Botanical research

Long before AF3 appeared, deep learning models were widely used in the field of plant immunity. These advanced algorithms chiefly provide brand-new technical means for plant science research by predicting protein structures and protein interaction networks. For plant disease prevention and control, researchers precisely simulated the interaction between pathogenic microorganisms and host plants, and screened out potential antiviral targets from complex molecular networks, providing a theoretical basis for the precise prevention and control of crop diseases [122]. Further research indicates that by conducting in-depth analysis of the protein structure of pathogenic bacteria carried by crop pests, the molecular mechanism of the interaction between viruses and their hosts can be revealed, benefiting the integrated management of pests. In addition to dealing with external biological stress, deep learning models also have significant value in the study of the intrinsic functions of plants. By analyzing the 3D structure of the protein encoded by key functional genes, researchers can intuitively understand its molecular mechanism, which provides an important reference for molecular design breeding to enhance crop disease resistance and increase yield [123].

With the release of AF3, botany research has been provided with more powerful tools. A recent study explores in detail how AF3 can be used to distinguish between sensor and helper proteins in plant nucleotide binding and leucine-rich repeat (NLR) immune receptors [124]. This achievement allows researchers to identify new disease resistance genes more quickly and accurately, and provides a new research idea for crop disease resistance breeding. In the field of agricultural pest control, researchers have successfully analyzed the structural characteristics of the flavivirus carried by the cabbage whitefly (*Aleyrodes proletella*) using AF3, providing an important basis for the development of new biological control methods and strongly promoting the sustainable development of agriculture [125]. In plant genomics, AF3 predicted the protein structure expressed by the GRAS gene family, verifying its conservation and functional diversity, and laying a theoretical foundation for the genetic improvement of important crops like sugarcane [126]. Similar research methods have also achieved important breakthroughs in rice breeding. Through the structural prediction of AF3, researchers have confirmed the specific interaction between plant A/T sequences and zinc-binding proteins (PLATZ) and DNA, clarifying their identity as transcription factors [127]. As PLATZ plays a key regulatory role in the growth and development of rice, this discovery provides a new molecular target for the improvement of rice varieties. The wide application of AF3 in botanical research not only deepens our understanding of plant immune mechanisms, but also injects new vitality into the sustainable development of agricultural production. From basic research to applied development, AF3 is driving botanical research to new heights.

Nanomaterials

Long before the rise of AF3, AI had already taken its place in the field of nanotechnology. By leveraging its database and high-throughput screening, researchers have significantly reduced ex-

perimental costs and achieved higher synthesis efficiency, which has been widely applied in the design, synthesis, and practical application of nanomaterials [128]. As a rising star in AI, the potential of AF3 in this field has also been rapidly tapped. Researchers used the AF3 AI model to assist the design of DNA nanostructures, and successfully constructed and optimized self-assembled triangular DNA structures [129]. This study not only demonstrates the application potential of AF3 in the field of DNA nanotechnology, but also provides new design ideas and strategies for more complex biomolecular self-assembly in the future. This discovery is expected to promote the development of nanotechnology and have a considerable positive impact on biomedicine, materials science, and other fields.

Bacterial research

With its outstanding prediction accuracy and wide application range, AF3 is triggering a revolution in the field of bacterial research. Compared to traditional experimental methods, AF3 can complete high-precision structural predictions of bacterial proteins and their complexes within a few hours, which provides unprecedented opportunities for understanding the physiological and pathogenic mechanisms of bacteria and developing new antibacterial drugs. In the research of bacterial adhesion mechanisms, AF3 successfully resolved the 3D structure of *Aeromonas hydrophila* repeats-in-toxin (RTX) adhesin and accurately predicted its three ligand-binding domains, including a carbohydrate binding module that can specifically recognize fucosylated glycans, providing a structural basis for the development of anti-adhesion inhibitors based on sugar analogues [130]. In the field of metabolic regulation, AF3 has been used to predict the structural characteristics of enzymes related to glycogen metabolism in methanogens, revealing the unique energy storage and utilization mechanisms of these archaea [131]. Furthermore, AF3 also contributed to the analysis of the dynamic conformational changes of CpaF ATPase during the assembly of Tad fimbria, clarifying the molecular mechanism that drives fimbria assembly through a rotational mechanism [132]. In the research of pathogenic microorganisms, AF3 not only successfully predicted the 3D model of the rib structural protein of *Trichomonas fetus* [133], but also revealed the homologous domain characteristics of the cryptococcal sexual reproduction regulatory protein Sxi1 α and its regulatory role in cell fusion [134]. It is particularly worth mentioning that through a systematic analysis of the functional domains of *Acinetobacter baumannii* phage caudin/spike protein, AF3 helped identify 32 key functional domains, providing important target information for the development of novel phage therapies [135]. These breakthroughs fully demonstrate the strong potential of AF3 in fundamental research and application development of treatments for bacterial infection.

Potential obstacles in AF3 applications

An outstanding example of deep learning in biomolecular structure prediction, AF3 demonstrates its ability to simulate chemically diverse entities. The diversity of its functions undoubtedly broadens its application scope. However, due to the limitations of the techniques and methods, the complexity and variability of the structure, and the limitations of training and databases, AF3 has not brought about the disruptive innovation of structural forecasting expected. Some of the limitations inherent in deep learning models are also present in AF3 and remain the main challenges in this field.

Lack of dynamic information

AF3 is reported to be able to provide highly accurate predictions for almost all existing complexes of molecular types that have been uploaded to the PDB. However, this achievement also reveals AF3's reliance on existing crystal structures and cryo-electron microscopy data as training sets [73]. As such, its prediction results are essentially still "molecular photos" and cannot capture the dynamic conformational changes of the biomolecular system in the solution. Even by diffusion or using multiple random seeds, it is difficult to obtain an approximate solution set. Understanding the molecular mechanisms in the organism and cells often requires more intuitive dynamic processes. In studying other dynamic processes like metabolism, although the protein structure predicted by AF3 provides site information, it cannot simulate the conformational dynamic changes triggered by energy charges. This limitation restricts the application of AF3 in the study of biomolecular dynamic processes and drug dynamic interactions [136].

Deviation of conformational state

Consistent with other deep learning tools, AF3 may provide incorrect or incomplete predictions of protein structural conformation. The original article describes this situation by taking E3 ubiquitin ligase as an example. In the real case, the enzyme is open under apo (no ligands) conditions, while it exhibits a closed conformation under holographic conditions (ligand-binding) [137]. However, predictions from AF3 showed that E3 ubiquitin ligases were in a closed conformation with or without ligands. Similarly, researchers also investigated the predictive effects of RFAA [80] and AF3 for protein-ligand interactions [138]. The results noted that when subjected to biologically credible perturbations, the predictions of either tool differ significantly from the expected physical behavior. In addition, predictive studies on G protein-coupled receptors (GPCRs) showed that although the performance of AF3 was improved compared with AF2, it could not achieve the effect of traditional methods in predicting ligand status [139]. These phenomena suggest that AF3 is not completely reliable in predicting intrinsically flexible regions or domains in complex structures. Due to the high standard of protein structure required in drug design, this means that AF3 cannot currently achieve the desired effect of independently supporting drug design and development. Therefore, how to achieve the polymorphism prediction of protein structure by AI is still an urgent problem for researchers to solve [71].

Limitation of prediction accuracy

In the field of structure prediction, the prediction accuracy of AF3 has made significant progress, but there is still some room for improvement, especially for specific targets, such as orphan proteins. Due to the lack of known sequence-based near-origin homologues, AF3's prediction results for this class of proteins are often unsatisfactory. In addition, the accuracy of AF3 in protein monomer structure prediction has not been significantly improved. This indicates that the AlphaFold platform needs to be further improved in the field of pure protein structure prediction.

For the field of drug development and vaccine preparation, prediction accuracy is an important index to measure the reliability of prediction results. When the prediction accuracy is <80%, whether it can meet the requirements of pre-screening becomes a problem worth pondering. Especially for those proteins with complex structures and prediction accuracy close to 50% or less, the reliability of the predicted results needs to be carefully

evaluated. In addition, the insufficient content of the dataset is one of the important factors affecting the reliability of AF3 prediction results. This may cause AF3 to lack sufficient information support in the forecasting process, thus impeding the accuracy of its predictions. Although this problem can be mitigated to some extent by extensive prediction, the resulting computational costs are undoubtedly a challenge for researchers with limited resources.

Similarly, while AF3 is good at predicting protein interactions, it still has some potential limitations when dealing with highly flexible protein regions. The large conformational fluctuations in these regions can easily lead to deviations in the prediction results. Studies have shown that although AF3 has high accuracy in protein–protein complex prediction, it still has shortcomings in processing specific regions [38]. Although AF3 has improved antibody–antigen interaction prediction compared to the previous model (AlphaFold2-Multimer), the overall prediction accuracy still failed to meet the requirements of practical applications [140]. These defects have also been found in the prediction of protein–ligand [139, 141] and protein–nucleic acid interactions [142, 143].

Similar problems exist in the field of RNA structure prediction. Multiple systematic evaluations have shown that for long-chain RNA molecules, the model predicted by AF3 often has structural integrity problems, and the molecular shape parameters are significantly different from the experimental observations [144]. More crucially, due to the difficulty of AF3 in capturing the inherent conformational heterogeneity of RNA molecules and the scale limitations of existing RNA structure databases, the prediction results seem inconsistent with the observed results [145]. In comparative tests of professional RNA structure prediction tools, the overall performance of AF3 was significantly inferior to those prediction methods specifically optimized for RNA characteristics, with deficiencies in the prediction of long sequences and non-classical interactions [146]. These limitations have, to some extent, hindered the wide application of AF3 in the field of drug development and biomedical research.

Occurrence of hallucination effect

In the updated iteration of the model architecture, AF3 showed some new error types that did not occur in its predecessor AF2. When dealing with disordered regions of the protein, AF3 sometimes generates a false structural order, which is known as “hallucination” [147]. The occurrence of this phenomenon mainly stems from the model’s incorrect prediction of the structure under the condition of insufficient information. The essence of the diffusion model is to obtain the predicted atomic coordinates through autonomous iterative denoising after learning the noise pattern through training. Using the existing dataset, the disordered noise gradually forms a clear structure. However, since biomolecules themselves may have disordered regions, AF3 can denoise this part as disordered noise during the prediction process, ultimately resulting in a false ordered structure. This characteristic will interfere with the researchers’ correct judgment of molecular structure. Hallucination regions are often accompanied by extremely low confidence levels, with pLDDT scores well below the 50 threshold. Since these hallucinatory regions lack the distinctive banded appearance features that AF2 produces, such as the formation of alpha helical structures, researchers may misjudge protein structure prediction using AF3, especially in the prediction of phosphorylation modification regions [148].

To address this challenge, AF3’s developers adopted a strategy of cross-distillation and encouraged the formation of structures with a larger solvent-accessible surface area (SASA) by introducing a ranking item (Fig. 4). The SASA refers to the area of the protein surface exposed to the solvent. Under normal circumstances, the SASA of disordered protein structures is larger than that of ordered structures. Therefore, the introduction of ranking items limits the conformation predicted by the model to a certain extent, making it closer to the natural structure of the protein. Nevertheless, the hallucination phenomenon is still difficult to completely eliminate and, to a certain extent, it has a negative impact on the predicted results of AF3. Therefore, how to further reduce the occurrence of hallucination and improve the accuracy of AF3 structure prediction in disordered regions is still an important direction of future research.

Errors of stereochemistry

AF3 is also not perfect in stereochemistry. The prediction accuracy of AF3 chiral structure is insufficient, especially for predicting the conformation of small molecules. Although AF3 incorporates a penalty mechanism for chiral violations into its algorithm design, it has a chiral violation rate of 4.4% in the PoseBusters benchmark. This significant defect may adversely affect the analysis of protein structure and function and the screening of targets, thereby limiting the potential of AF3 applications in related fields. For example, in the conformational prediction of fatty acids, the detailed structure and interaction of AF3 generation are contrary to the experimental results [141]. In addition, AF3 occasionally has atomic overlap (conflict) during the prediction process [144]. This conflict sometimes manifests itself in extremes. That is, in a protein structure with two or more identical chains (homologues), the entire chain is incorrectly predicted to be in an overlapping state. Although AF3 reduces the occurrence of such phenomena to some extent by introducing a strategy of punishing conflicts, unfortunately the strategy does not completely eliminate the problem of atomic overlap.

In summary, while AF3 is an important advance in the field of protein structure prediction, like all scientific tools, it is undoubtedly not perfect. In a recent prediction test for CASP16, AF3 did not rank first as expected. Although it has a good predictive effect, it is inevitably inferior to the top-ranked teams when facing some complex tasks. This result shows that AF3 is not the only seed player in the protein prediction racetrack, and other prediction models may have their own areas of excellence. These tools need to be applied with caution and wisdom.

Prospects

At present, AI and computer technology are gradually penetrating the field of scientific research, and the pharmaceutical and biomedical industries are experiencing an unprecedented renaissance. As a currently emerging structure-prediction model, although AF3 itself still has many insurmountable defects, with continuous optimization after open source collaboration, a comprehensive analysis of protein structure has already seen light at the end of the tunnel [48], and its application prospects in the fields of drug research and development and biomedicine show a multi-dimensional development trend.

Judging from the current situation, overcoming the inherent defects of AF3 requires the joint efforts of other tools [70]. In previous studies, some researchers have found that combining AF3 with transfer learning techniques can overcome the model’s depen-

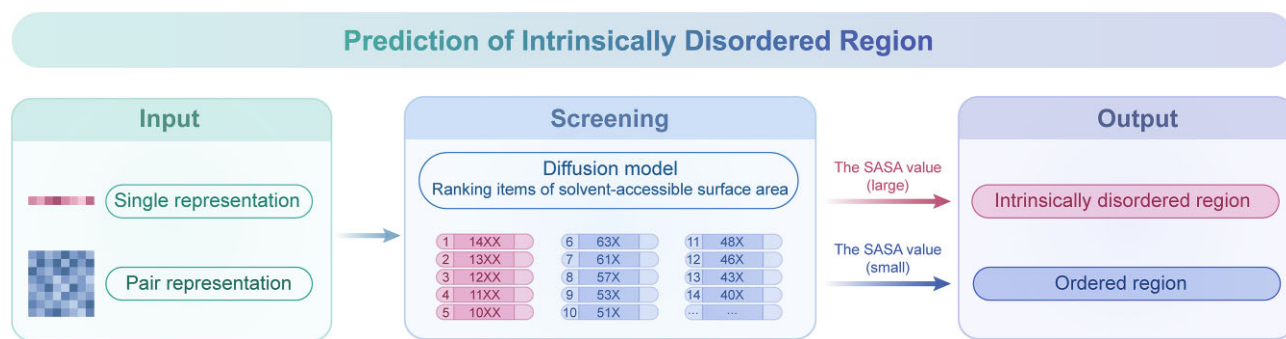


Figure 4. Schematic diagram of intrinsically disordered region prediction. The model adopts two input methods, namely single representation and pair representation, and generates protein conformations through the diffusion model. The prediction results were screened through the SASA ranking items, where the values (such as 14XX, 11XX, 63X, etc.) represent the SASA scores of different residues or fragments, referring to the solvent-exposure degrees of different conformations. The higher the value, the better the solvent accessibility and the more likely to be a disordered region. This approach might help distinguish the intrinsically disordered region and reduce the occurrence of hallucinations.

dence on the structure of natural proteins and MSA [149]. Furthermore, experiments have confirmed that template-based modeling outperforms AF3 in the prediction of ligand binding sites [150]. Therefore, the combination of AF3 with template-based modeling can optimize the identification of confused residues through hydrophobic MSA, significantly improving the accuracy of ligand binding-site prediction. In the process of protein complex stoichiometric prediction, the fusion of AF3 and homologous template information can increase the TOP1 accuracy rate to 71.4%, highlighting the advantages of multimodal methods in the resolution of quaternary structures [151].

These achievements demonstrate the feasibility of AF3 for performance optimization in combination with other methods. From this reasoning, seeking predictive tools (such as Disobind [152]) for internal disordered regions or other special structures and using them in combination with AF3 might be a good countermeasure to solve the hallucination effect. Existing studies have shown that introducing dynamic network analysis and molecular dynamics simulation into AFM can significantly improve the prediction accuracy of epitopes, especially in the prediction of complex conformations such as vertical binding modes [153]. This successful experience can be directly transferred to the optimization process of AF3. To improve the prediction accuracy of AF3 in heterogeneous RNA, a hybrid strategy combining calculation and experiments (such as atomic force microscopy) can be considered [145]. Through the structural constraint information provided by the experiment, the prediction results obtained by the pure calculation method are verified and optimized, thereby significantly improving the reliability of the prediction model. Furthermore, regarding the deficiencies of AF3 in the prediction of antigen–antibody interactions, the recently developed Cmai model [154] (a deep learning framework specifically designed for the prediction of antigen–antigen binding) and the epitope-paratope predictor (EPP) [155], can be used as effective supplementary tools to enhance prediction reliability in this field.

In addition to the performance optimization of the prediction model, the expansion of the training set is also a point that AF3 must pay attention to in its future development. Integrating AF3 with other datasets, such as transmembrane protein databases (e.g. TMVisDB [156]) and ion channel resources (e.g. PLIC [157]), can provide high-quality structural and functional correlation data for the iterative training of the model. Similarly, the construction of a PTM rotational isomer library can optimize the struc-

tural prediction of modification sites by AF3 [158], while the expansion of protein–small molecule drug databases (such as Drug-Domain [159]) can enhance the modeling ability of AF3 for small molecule binding interfaces. With the emergence of open-source tools (such as af3cli [160]) and automated processes (such as MULTICOM4 [161]), the input generation and model evaluation of AF3 will be further simplified, greatly promoting the popularization of AF3.

Although AF3 has broad application potential, we must clearly realize that in practical application scenarios such as drug development, vaccine preparation, and so on, relying solely on AI prediction methods like AF3 is far from ideal or absolutely accurate. Developing a drug or therapeutic into the clinic requires a high degree of precision in protein structure, and even a small miscalculation can have serious consequences. The prediction accuracy of AF3 still cannot meet the requirements in the field of drug design, and as its prediction ability means that it cannot replace experimental results, traditional experimental methods still occupy an indispensable position [162]. Therefore, the advent of deep learning models such as AF3 is fundamentally an auxiliary tool to traditional methods, laying the necessary early foundation for drug development and biomedical fields. It complements traditional experimental methods and jointly promotes the development and progress of related fields [163].

Although there are still many issues with the current AF3, we believe that time will change everything. With the continuous progress of technology and the continuous efforts of researchers, AF3 and its likely future versions may eventually evolve into a perfect tool and play an increasingly important role in every field and immensely benefit the human being.

Acknowledgments

This work was supported by funds of the Key Research and Development Grant of MOST (grant No. 2023YFA095000), the Affiliated Xiangshan Hospital of Wenzhou Medical University, the National Science Foundation of China (grant No. 82470005), and Wenzhou Institute University of Chinese Academy of Sciences. We thank members of the laboratory for helpful discussion.

Author contributions

Ziqi Fang (Writing—original draft), Hongbiao Ran (Writing—original draft), Yonghan Zhang (Writing—original draft), Chen-

song Chen (Writing—review & editing), Ping Lin (Writing—review & editing), Xiang Zhang (Writing—review & editing), and Min Wu (Writing—review & editing).

Conflict of interest

None declared. The corresponding author M.W. as an Editorial Board Member of *Precision Clinical Medicine* was blinded from reviewing and making decisions on this manuscript.

References

- Callaway E. Chemistry Nobel goes to developers of AlphaFold AI that predicts protein structures. *Nature* 2024;**634**:525–6. <https://doi.org/10.1038/d41586-024-03214-7>
- 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>
- Cramer P. AlphaFold2 and the future of structural biology. *Nat Struct Mol Biol* 2021;**28**:704–5. <https://doi.org/10.1038/s41594-021-00650-1>
- Akdel M, Pires DEV, Pardo EP et al. A structural biology community assessment of AlphaFold2 applications. *Nat Struct Mol Biol* 2022;**29**:1056–67. <https://doi.org/10.1038/s41594-022-00849-w>
- Bordin N, Dallago C, Heinzinger M et al. Novel machine learning approaches revolutionize protein knowledge. *Trends Biochem Sci* 2023;**48**:345–59. <https://doi.org/10.1016/j.tibs.2022.11.001>
- Dill KA, MacCallum JL. The protein-folding problem, 50 years on. *Science* 2012;**338**:1042–6. <https://doi.org/10.1126/science.1219021>
- Kendrew JC, Bodo G, Dintzis HM et al. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* 1958;**181**:662–6. <https://doi.org/10.1038/181662a0>
- Shi Y. A glimpse of structural biology through X-ray crystallography. *Cell* 2014;**159**:995–1014. <https://doi.org/10.1016/j.cell.2014.10.051>
- Wüthrich K. Protein structure determination in solution by NMR spectroscopy. *J Biol Chem* 1990;**265**:22059–62. [https://doi.org/10.1016/S0021-9258\(18\)45665-7](https://doi.org/10.1016/S0021-9258(18)45665-7)
- Earl LA, Falconieri V, Milne JL et al. Cryo-EM: beyond the microscope. *Curr Opin Struct Biol* 2017;**46**:71–8. <https://doi.org/10.1016/j.sbi.2017.06.002>
- Thompson MC, Yeates TO, Rodriguez JA. Advances in methods for atomic resolution macromolecular structure determination. *F1000Research* 2020;**9**:F1000 Faculty Rev-667. <https://doi.org/10.12688/f1000research.25097.1>
- Bai X-c, McMullan G, Scheres SHW. How cryo-EM is revolutionizing structural biology. *Trends Biochem Sci* 2015;**40**:49–57. <https://doi.org/10.1016/j.tibs.2014.10.005>
- Jaskolski M, Dauter Z, Wlodawer A. A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. *FEBS J* 2014;**281**:3985–4009. <https://doi.org/10.1111/febs.12796>
- Taylor WR, Orengo CA. Protein structure alignment. *J Mol Biol* 1989;**208**:1–22. [https://doi.org/10.1016/0022-2836\(89\)90084-3](https://doi.org/10.1016/0022-2836(89)90084-3)
- Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 2005;**33**:2302–9. <https://doi.org/10.1093/nar/gki524>
- Shindyalov IN, Bourne PE. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* 1998;**11**:739–47. <https://doi.org/10.1093/protein/11.9.739>
- Wang S, Sun S, Li Z et al. Accurate de novo prediction of protein contact map by ultra-deep learning model. *PLoS Comput Biol* 2017;**13**:e1005324. <https://doi.org/10.1371/journal.pcbi.1005324>
- Pearce R, Zhang Y. Deep learning techniques have significantly impacted protein structure prediction and protein design. *Curr Opin Struct Biol* 2021;**68**:194–207. <https://doi.org/10.1016/j.sbi.2021.01.007>
- Kuhlman B, Bradley P. Advances in protein structure prediction and design. *Nat Rev Mol Cell Biol* 2019;**20**:681–97. <https://doi.org/10.1038/s41580-019-0163-x>
- Silver D, Huang A, Maddison CJ et al. Mastering the game of Go with deep neural networks and tree search. *Nature* 2016;**529**:484–9. <https://doi.org/10.1038/nature16961>
- Senior AW, Evans R, Jumper J et al. Improved protein structure prediction using potentials from deep learning. *Nature* 2020;**577**:706–10. <https://doi.org/10.1038/s41586-019-1923-7>
- Kryshtafovych A, Schwede T, Topf M et al. Critical assessment of methods of protein structure prediction (CASP)-Round XIII. *Proteins* 2019;**87**:1011–20. <https://doi.org/10.1002/prot.25823>
- Senior AW, Evans R, Jumper J et al. Protein structure prediction using multiple deep neural networks in the 13th Critical Assessment of Protein Structure Prediction (CASP13). *Proteins* 2019;**87**:1141–8. <https://doi.org/10.1002/prot.25834>
- 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>
- Kryshtafovych A, Schwede T, Topf M et al. Critical assessment of methods of protein structure prediction (CASP)—Round XIV. *Proteins Struct Funct Bioinforma* 2021;**89**:1607–17. <https://doi.org/10.1002/prot.26237>
- 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>
- Marx V. Method of the year: protein structure prediction. *Nat Methods* 2022;**19**:5–10. <https://doi.org/10.1038/s41592-021-01359-1>
- Zhang Y, Vass M, Shi D et al. Benchmarking refined and unrefined AlphaFold2 structures for Hit Discovery. *J Chem Inf Model* 2023;**63**:1656–67. <https://doi.org/10.1021/acs.jcim.2c01219>
- Kayastha K, Katsyv A, Himmrich C et al. Structure-based electron-confurcation mechanism of the Ldh-EtfAB complex. *eLife* 2022;**11**:e77095. <https://doi.org/10.7554/eLife.77095>
- Liang M, Chen X, Zhu C et al. Identification of a novel substrate motif of yeast separase and deciphering the recognition specificity using AlphaFold2 and molecular dynamics simulation. *Biochem Biophys Res Commun* 2022;**620**:173–9. <https://doi.org/10.1016/j.bbrc.2022.06.056>
- Lorenz P, Steinbeck F, Krause L et al. The KRAB domain of ZNF10 guides the identification of specific amino acids that transform the ancestral KRAB-A-related domain present in Human PRDM9 into a canonical modern KRAB-A domain. *Int J Mol Sci* 2022;**23**:1072. <https://doi.org/10.3390/ijms23031072>
- McMullen P, Fang L, Qiao Q et al. Impacts of a zwitterionic peptide on its fusion protein. *Bioconjug Chem* 2022;**33**:1485–93. <https://doi.org/10.1021/acs.bioconjchem.2c00176>
- Zhorov BS, Dong K. Pyrethroids in an AlphaFold2 model of the insect sodium channel. *Insects* 2022;**13**:745. <https://doi.org/10.3390/insects13080745>

34. Pinheiro F, Santos J, Ventura S. AlphaFold and the amyloid landscape. *J Mol Biol* 2021;**433**:167059. <https://doi.org/10.1016/j.jmb.2021.167059>
35. Goulet A, Joos R, Lavelle K et al. A structural discovery journey of streptococcal phages adhesion devices by AlphaFold2. *Front Mol Biosci* 2022;**9**:960325. <https://doi.org/10.3389/fmolb.2022.960325>
36. Goulet A, Mahony J, Cambillau C et al. Exploring structural diversity among adhesion devices encoded by lactococcal P335 phages with AlphaFold2. *Microorganisms* 2022;**10**:2278. <https://doi.org/10.3390/microorganisms10112278>
37. Pasquadibisceglie A, Leccese A, Polticelli F. A computational study of the structure and function of human zrt and irt-like proteins metal transporters: an elevator-type transport mechanism predicted by AlphaFold2. *Front Chem* 2022;**10**:1004815. <https://doi.org/10.3389/fchem.2022.1004815>
38. Wee J, Wei GW. Benchmarking AlphaFold3's protein-protein complex accuracy and machine learning prediction reliability for binding free energy changes upon mutation. *ArXiv Published online June 6, 2024*. <https://doi.org/10.48550/arXiv.2406.03979>
39. Ma W, Zhang S, Li Z et al. Enhancing protein function prediction performance by utilizing AlphaFold-predicted protein structures. *J Chem Inf Model* 2022;**62**:4008–17. <https://doi.org/10.1021/acs.jcim.2c00885>
40. Rappoport D, Jinich A. Enzyme substrate prediction from three-dimensional feature representations using space-filling curves. *J Chem Inf Model* 2023;**63**:1637–48. <https://doi.org/10.1021/acs.jcim.3c00005>
41. Feng Y, Liu XC, Li L et al. Naturally occurring I81N mutation in Human cytochrome c regulates both inherent peroxidase activity and interactions with neuroglobin. *ACS Omega* 2022;**7**:11510–8. <https://doi.org/10.1021/acsomega.2c01256>
42. Dawson JE, Smith IN, Martin W et al. Shape shifting: the multiple conformational substates of the PTEN N-terminal PIP2-binding domain. *Protein Sci Publ Protein Soc* 2022;**31**:e4308. <https://doi.org/10.1002/pro.4308>
43. Hu M, Yuan F, Yang KK et al. Exploring evolution-aware & -free protein language models as protein function predictors. *ArXiv Published online October 16, 2022*. <https://doi.org/10.48550/arXiv.2206.06583>
44. Evans R, O'Neill M, Pritzel A et al. Protein complex prediction with AlphaFold-multimer. *Biorxiv*, 2022: 2021.10.04.463034. <https://doi.org/10.1101/2021.10.04.463034>
45. Bernatavicius A, Šicho M, Janssen APA et al. AlphaFold meets de novo drug design: leveraging structural protein information in multitarget molecular generative models. *J Chem Inf Model* 2024;**64**:8113–22. <https://doi.org/10.1021/acs.jcim.4c00309>
46. Deneke VE, Blaha A, Lu Y et al. A conserved fertilization complex bridges sperm and egg in vertebrates. *Cell* 2024;**187**:7066–7078. <https://doi.org/10.1016/j.cell.2024.09.035>
47. He XH, You CZ, Jiang HL et al. AlphaFold2 versus experimental structures: evaluation on G protein-coupled receptors. *Acta Pharmacol Sin* 2023;**44**:1–7. <https://doi.org/10.1038/s41401-022-00938-y>
48. Graham F. Daily briefing: AlphaFold3 is now open source. *Nature Published online November 12, 2024*. <https://doi.org/10.1038/d41586-024-03728-0>
49. Karras T, Aittala M, Aila T et al. Elucidating the design space of diffusion-based generative models. In *36th Conference on Neural Information Processing Systems*, 2022, pp. 26565–77.
50. Edgar RC, Batzoglou S. Multiple sequence alignment. *Curr Opin Struct Biol* 2006;**16**:368–73. <https://doi.org/10.1016/j.sbi.2006.04.004>
51. Ji Z, Lee N, Frieske R et al. Survey of hallucination in natural language generation. *ACM Comput Surv* 2023;**55**:1–38. <https://doi.org/10.1145/3571730>
52. Tade RS, Jain SN, Satyavijay JT et al. Artificial intelligence in the paradigm shift of pharmaceutical sciences: A review. *Nano Biomed Eng* 2024;**16**:64–77. <https://doi.org/10.26599/NBE.2023.9290043>
53. Isert C, Atz K, Schneider G. Structure-based drug design with geometric deep learning. *Curr Opin Struct Biol* 2023;**79**:102548. <https://doi.org/10.1016/j.sbi.2023.102548>
54. Wang F, Diao X, Chang S et al. Recent progress of deep learning in drug discovery. *Curr Pharm Des* 2021;**27**:2088–96. <https://doi.org/10.2174/1381612827666210129123231>
55. Yang X, Wang Y, Byrne R et al. Concepts of artificial intelligence for computer-assisted drug discovery. *Chem Rev* 2019;**119**:10520–94. <https://doi.org/10.1021/acs.chemrev.8b00728>
56. Jiménez-Luna J, Grisoni F, Weskamp N et al. Artificial intelligence in drug discovery: recent advances and future perspectives. *Expert Opin Drug Discov* 2021;**16**:949–59. <https://doi.org/10.1080/17460441.2021.1909567>
57. Sarkar C, Das B, Rawat VS et al. Artificial intelligence and machine learning technology driven modern drug discovery and development. *Int J Mol Sci* 2023;**24**:2026. <https://doi.org/10.3390/ijms24032026>
58. Hudson IL. Data integration using advances in machine learning in drug discovery and molecular biology. *Methods Mol Biol Clifton NJ* 2021;**2190**:167–84. https://doi.org/10.1007/978-1-0716-0826-5_7
59. Ren F, Ding X, Zheng M et al. AlphaFold accelerates artificial intelligence powered drug discovery: efficient discovery of a novel CDK20 small molecule inhibitor. *Chem Sci* 2023;**14**:1443–52. <https://doi.org/10.1039/d2sc05709c>
60. Weng Y, Pan C, Shen Z et al. Identification of potential WSB1 inhibitors by AlphaFold modeling, virtual screening, and Molecular dynamics simulation studies. *Evid-Based Complement Altern Med ECAM* 2022;**2022**:4629392. <https://doi.org/10.1155/2022/4629392>
61. Yang Q, Jian X, Syed AAS et al. Structural comparison and drug screening of spike proteins of ten SARS-CoV-2 variants. *Res Wash DC* 2022;**2022**:9781758. <https://doi.org/10.34133/2022/9781758>
62. Park HM, Park Y, Vankerschaver J et al. Rethinking protein drug design with highly accurate structure prediction of anti-CRISPR proteins. *Pharm Basel Switz* 2022;**15**:310. <https://doi.org/10.3390/ph15030310>
63. Yang Q, Xia D, Syed AAS et al. Highly accurate protein structure prediction and drug screen of monkeypox virus proteome. *J Infect* 2023;**86**:66–117. <https://doi.org/10.1016/j.jinf.2022.08.006>
64. Buttenschoen M, Morris GM, Deane CM. PoseBusters: AI-based docking methods fail to generate physically valid poses or generalise to novel sequences. *Chem Sci* 2024;**15**:3130–9. <https://doi.org/10.1039/d3sc04185a>
65. Mirdita M, Schütze K, Moriwaki Y et al. ColabFold: making protein folding accessible to all. *Nat Methods* 2022;**19**:679–82. <https://doi.org/10.1038/s41592-022-01488-1>
66. Manshour N, Ren JZ, Esmaili F et al. Comprehensive evaluation of AlphaFold-multimer, AlphaFold3 and ColabFold, and scoring functions in predicting protein-peptide complex structures.

- Biorxiv, 2024: 2024.11.11.622992. <https://doi.org/10.1101/2024.11.11.622992>
67. Medvedev KE, Schaeffer RD, Grishin NV. Leveraging AI to explore structural contexts of post-translational modifications in drug binding. *J Cheminformatics* 2025;**17**:67. <https://doi.org/10.1186/s13321-025-01019-y>
 68. Schymkowitz J, Borg J, Stricher F et al. The FoldX web server: an online force field. *Nucleic Acids Res* 2005;**33**:W382–388. <https://doi.org/10.1093/nar/gki387>
 69. Xiong P, Zhang C, Zheng W et al. BindProfX: assessing mutation-induced binding affinity change by protein interface profiles with pseudo-counts. *J Mol Biol* 2017;**429**:426–34. <https://doi.org/10.1016/j.jmb.2016.11.022>
 70. Lu W, Zhang J, Rao J et al. AlphaFold3, a secret sauce for predicting mutational effects on protein-protein interactions. *Biorxiv*, 2024: 2024.05.25.595871. <https://doi.org/10.1101/2024.05.25.595871>
 71. Wee J, Wei GW. Evaluation of AlphaFold 3's protein-protein complexes for predicting binding free energy changes upon mutation. *J Chem Inf Model* 2024;**64**:6676–83. <https://doi.org/10.1021/acs.jcim.4c00976>
 72. Chen E, Pan E, Zhang S. Structure bioinformatics of six Human integral transmembrane enzymes and their AlphaFold3 predicted water-soluble QTY analogs: insights into FACE1 and STEA4 binding mechanisms. *Pharm Res* 2025;**42**:291–305. <https://doi.org/10.1007/s11095-025-03822-6>
 73. Stanzione F, Giangreco I, Cole JC. Use of molecular docking computational tools in drug discovery. *Prog Med Chem* 2021;**60**:273–343. <https://doi.org/10.1016/bs.pmch.2021.01.004>
 74. Zhao Z, Ukidve A, Kim J et al. Targeting strategies for tissue-specific drug delivery. *Cell* 2020;**181**:151–67. <https://doi.org/10.1016/j.cell.2020.02.001>
 75. Zhang X, Shen C, Zhang H et al. Advancing Ligand docking through deep learning: challenges and prospects in virtual screening. *Acc Chem Res* 2024;**57**:1500–9. <https://doi.org/10.1021/acs.accounts.4c00093>
 76. Wang ZY, Li MZ, Li WJ et al. Mechanism of action of Daqinjiao decoction in treating cerebral small vessel disease explored using network pharmacology and molecular docking technology. *Phytomedicine Int J Phytother Phytopharm* 2023;**108**:154538. <https://doi.org/10.1016/j.phymed.2022.154538>
 77. Li T, Guo R, Zong Q et al. Application of molecular docking in elaborating molecular mechanisms and interactions of supramolecular cyclodextrin. *Carbohydr Polym* 2022;**276**:118644. <https://doi.org/10.1016/j.carbpol.2021.118644>
 78. Dong Y, Tao B, Xue X et al. Molecular mechanism of Epicedium treatment for depression based on network pharmacology and molecular docking technology. *BMC Complement Med Ther* 2021;**21**:222. <https://doi.org/10.1186/s12906-021-03389-w>
 79. Sun Z, Wang Y, Pang X et al. Mechanisms of polydatin against spinal cord ischemia-reperfusion injury based on network pharmacology, molecular docking and molecular dynamics simulation. *Bioorganic Chem* 2023;**140**:106840. <https://doi.org/10.1016/j.bioorg.2023.106840>
 80. Krishna R, Wang J, Ahern W et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* 2024;**384**:ead12528. <https://doi.org/10.1126/science.adl2528>
 81. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 2010;**31**:455–61. <https://doi.org/10.1002/jcc.21334>
 82. Miller EB, Murphy RB, Sindhikara D et al. Reliable and accurate solution to the induced fit docking problem for protein-ligand binding. *J Chem Theory Comput* 2021;**17**:2630–9. <https://doi.org/10.1021/acs.jctc.1c00136>
 83. Solís KH, Romero-Ávila MT, Rincón-Heredia R et al. Multiple LPA3 receptor agonist binding sites evidenced under docking and functional studies. *Int J Mol Sci* 2025;**26**:4123. <https://doi.org/10.3390/ijms26094123>
 84. Li X, Zhou S, Tae HS et al. N-terminal capping of the α -conotoxin analogue GeX-2 improves the serum stability and selectivity toward the Human $\alpha 9\alpha 10$ nicotinic acetylcholine receptor. *J Med Chem* 2024;**67**:18400–11. <https://doi.org/10.1021/acs.jmedchem.4c01758>
 85. Gadde N, Dodamani S, Altaf R et al. Leveraging AlphaFold 3 for structural modeling of neurological disorder-associated proteins: A pathway to precision medicine. *Biorxiv*, 2024: 2024.11.18.624211. <https://doi.org/10.1101/2024.11.18.624211>
 86. Yao M, Miller GW, Vardarajan BN et al. Deciphering proteins in Alzheimer's disease: A new mendelian randomization method integrated with AlphaFold3 for 3D structure prediction. *Cell Genomics* 2024;**4**:100700. <https://doi.org/10.1016/j.xgen.2024.100700>
 87. Domene-Serrano I, Cuadros R, García-Escudero V et al. Shapeshifter W-Tau peptide inhibits tau aggregation and disintegrates paired helical Filaments. *Biochemistry* 2025;**64**:1841–51. <https://doi.org/10.1021/acs.biochem.4c00809>
 88. Peng Z, Xu S, Wang H et al. Identification of GDP as a small inhibitory molecule in HepG2 cells by non-targeted metabolomics analysis. *Oncol Lett* 2025;**29**:178. <https://doi.org/10.3892/ol.2025.14924>
 89. Zhao S, Luo J, Xu P et al. Designed peptide binders and nanobodies as PROTAC starting points for targeted degradation of PCNA and BCL6. *Int J Biol Macromol* 2025;**308**:142667. <https://doi.org/10.1016/j.ijbiomac.2025.142667>
 90. Klufmann M, Reuter J, Werner C et al. Examining farnesyltransferase interaction with cell-permeable CaaX peptides and the role of the CaaX motif in biological activity. *J Pept Sci Off Publ Eur Pept Soc* 2025;**31**:e70009. <https://doi.org/10.1002/psc.70009>
 91. Wee J, Chen J, Wei GW. Preventing future zoonosis: SARS-CoV-2 mutations enhance human-animal cross-transmission. *Comput Biol Med* 2024;**182**:109101. <https://doi.org/10.1016/j.compbiomed.2024.109101>
 92. Wang J, Xie J, Yu Y et al. Enhancing the understandings on SARS-CoV-2 main protease (Mpro) mutants from molecular dynamics and machine learning. *Int J Biol Macromol* 2025;**310**:143076. <https://doi.org/10.1016/j.ijbiomac.2025.143076>
 93. Liu N, Zhu Y, Chong H et al. Structural characterization of HIV fusion inhibitor LP-98: insights into antiviral and resistance mechanisms. *Antiviral Res* 2025;**239**:106190. <https://doi.org/10.1016/j.antiviral.2025.106190>
 94. Mosalanejad Z, Faraji SN, Rahbar MR et al. Designing an optimized theta-defensin peptide for HIV therapy using in-silico approaches. *J Integr Bioinforma* Published online March 19, 2025. <https://doi.org/10.1515/jib-2023-0053>
 95. Su H, Wu G, Xiong M et al. Dynamic cap-mediated substrate access and potent inhibitor design of Monkeypox virus I7L protease. *Adv Sci Weinh Baden-Wurt Ger* Published online April 7, 2025:e2501625. <https://doi.org/10.1002/advs.202501625>
 96. Zhang L, Zheng Y, Shao M et al. AlphaFold-based AI docking reveals AMPK/SIRT1-TFEB pathway modulation by traditional Chinese medicine in metabolic-associated fatty liver disease.

- Pharmacol Res 2025;212:107617. <https://doi.org/10.1016/j.phrs.2025.107617>
97. Claustre A, Malige M, Macheton M et al. Structure predictions of MuRF1-UBE2 complexes identify amino acid residues governing interaction selectivity for each MuRF1-E2 pair. *FEBS J* 2025;292:2559–77. <https://doi.org/10.1111/febs.70017>
 98. Zhou H, Xu Z, Xiong Y et al. Structural insights into the binding of human TGIF1 with SMAD2 MH2 domain. *FEBS Lett* Published online May 21, 2025. <https://doi.org/10.1002/1873-3468.70073>
 99. Xiao L, Wu Z, Su J et al. Development of a potent neutralizing nanobody against canine distemper virus hemagglutinin protein. *Front Immunol* 2025;16:1585793. <https://doi.org/10.3389/fimmu.2025.1585793>
 100. 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>
 101. 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>
 102. Guo C, Liu X, Mei Z et al. AMBP protects against aortic valve calcification by inhibiting ERK1/2 and JNK pathways mediated by FHL3. *Theranostics* 2025;15:4398–415. <https://doi.org/10.7150/thno.109182>
 103. Madarász K, Mótyán JA, Chang Chien YC et al. BCOR-rearranged sarcomas: In silico insights into altered domains and BCOR interactions. *Comput Biol Med* 2025;191:110144. <https://doi.org/10.1016/j.combiomed.2025.110144>
 104. Yue C, Shi S, Li Z et al. Studying the signaling mechanism of Neuropilin-1's intracellular disorder region via conformational mining and dynamic interaction characterization. *J Phys Chem B* 2025;129:2392–401. <https://doi.org/10.1021/acs.jpbc.4c07616>
 105. Chu KY, Crawford AN, Krah BS et al. Cripto-1 acts as a molecular bridge linking nodal to ALK4 via distinct structural domains. *Protein Sci Publ Protein Soc* 2025;34:e70034. <https://doi.org/10.1002/pro.70034>
 106. Pan A, Liu X, Han H et al. Disruption of a potential disulfide bond of Cys65-Cys141 on the structure and stability of globin X from zebrafish. *Phys Chem Chem Phys PCCP* 2025;27:2828–33. <https://doi.org/10.1039/d4cp04253k>
 107. Zeng D, Xin J, Yang K et al. A hemagglutinin stem vaccine designed rationally by AlphaFold2 confers broad protection against influenza B infection. *Viruses* 2022;14:1305. <https://doi.org/10.3390/v14061305>
 108. Alnajran H, Awadalla M, Aldakheel FM et al. Design of a peptide-based vaccine against human respiratory syncytial virus using a reverse vaccinology approach: evaluation of immunogenicity, antigenicity, allergenicity, and toxicity. *Front Immunol* 2025;16:1546254. <https://doi.org/10.3389/fimmu.2025.1546254>
 109. Jiang K, Yan Z, Di Bernardo M et al. Rapid in silico directed evolution by a protein language model with EVOLVEpro. *Science* 2025;387:eadr6006. <https://doi.org/10.1126/science.adr6006>
 110. Wei Y, Gao P, Pan D et al. Engineering eukaryotic transposon-encoded Fanzor2 system for genome editing in mammals. *Nat Chem Biol* Published online May 20, 2025. <https://doi.org/10.1038/s41589-025-01902-7>
 111. Johnsson F, Karagöl T, Karagöl A et al. Structural bioinformatic study of six human olfactory receptors and their AlphaFold3 predicted water-soluble QTY variants and OR1A2 with an odorant octanoate and TAAR9 with spermidine. *QRB Discov* 2025;6:e2. <https://doi.org/10.1017/qrd.2024.18>
 112. Chen E, Zhang S. Structural bioinformatic study of human mitochondrial respiratory integral membrane megacomplex and its AlphaFold3 predicted water-soluble QTY megacomplex analog. *QRB Discov* 2025;6:e12. <https://doi.org/10.1017/qrd.2025.2>
 113. Gao H, Hamp T, Ede J et al. The landscape of tolerated genetic variation in humans and primates. *Science* 2023;380:eabn8153. <https://doi.org/10.1126/science.abn8197>
 114. Iqbal S, Brünger T, Pérez-Palma E et al. Delineation of functionally essential protein regions for 242 neurodevelopmental genes. *Brain* 2023;146:519–33. <https://doi.org/10.1093/brain/awac381>
 115. Zhang Z, Li C, Li Q et al. Structure prediction of novel isoforms from uveal melanoma by AlphaFold. *Sci Data* 2023;10:513. <https://doi.org/10.1038/s41597-023-02429-z>
 116. Qie D, Zhai Y, Yang F et al. A de novo TNN3K variant aggravates the pathogenicity of DMD-associated early-onset cardiomyopathy: a case report. *Front Genet* 2025;16:1525941. <https://doi.org/10.3389/fgene.2025.1525941>
 117. Chen J, Li Z, Wu Y et al. Identification of pathogenic missense mutations of NF1 using computational approaches. *J Mol Neurosci MN* 2024;74:94. <https://doi.org/10.1007/s12031-024-02271-x>
 118. Nosrati MSS, Doustmohammadi A, Severino M et al. Novel KIF26A variants associated with pediatric intestinal pseudo-obstruction (PIPO) and brain developmental defects. *Clin Genet* 2025;107:83–90. <https://doi.org/10.1111/cge.14621>
 119. Ding X, Chen H, Lu Y et al. Analysis of FBN1 gene mutations in six Chinese pedigrees affected with Marfan syndrome. *Chin J Med Genet* 2025;42:41–50. <https://doi.org/10.3760/cma.j.cn511374-20240828-00459>
 120. Liu Y, Yan L, Zhang Y et al. Genetic analysis of a fetus with Farber lipogranulomatosis caused by ASAH1 gene variant. *Chin J Med Genet* 2025;42:232–7. <https://doi.org/10.3760/cma.j.cn511374-20240910-00482>
 121. Furukawa S, Kushima I, Kato H et al. Whole-genome sequencing analysis of Japanese autism spectrum disorder trios. *Psychiatry Clin Neurosci* 2025;79:87–97. <https://doi.org/10.1111/pcn.13767>
 122. Homma F, Huang J, van der Hoorn RAL. AlphaFold-Multimer predicts cross-kingdom interactions at the plant-pathogen interface. *Nat Commun* 2023;14:6040. <https://doi.org/10.1038/s41467-023-41721-9>
 123. Jia X, Fu Q, Yang G et al. Genome-wide identification and expression analysis of TONNEAU1 recruited motif (TRM) gene Family in tomato. *Int J Mol Sci* 2025;26:3676. <https://doi.org/10.3390/ijms26083676>
 124. Toghiani A, Terauchi R, Kamoun S et al. Can AI modelling of protein structures distinguish between sensor and helper NLR immune receptors? *Biorxiv*, 2024: 2024.11.24.625045. <https://doi.org/10.1101/2024.11.24.625045>
 125. Ye ZX, Gu GW, Ren PP et al. Identification and characterization of three novel iflaviruses in the cabbage Whitefly *Aleyrodes proletella*. *Insects* 2025;16:335. <https://doi.org/10.3390/insects16040335>
 126. Wen H, Wang L, Gong Y et al. Genome-wide identification and characterization of GRAS transcription factor Family in cultivated hybrid sugarcane ZZ1 (*Saccharum officinarum*) and their role in development and stress. *Int J Mol Sci* 2024;25:13470. <https://doi.org/10.3390/ijms252413470>
 127. Yang T, Xu XT, Tang LJ et al. Genome-wide study of plant-specific PLATZ transcription factors and functional analysis of OsPLATZ1 in regulating caryopsis development of rice (*Oryza*

- sativa L.). *Plants Basel Switz* 2025;**14**:151. <https://doi.org/10.3390/plants14020151>
128. Zhu X, Li Y, Gu N. Application of artificial intelligence in the exploration and optimization of biomedical nanomaterials. *Nano Biomed Eng* 2023;**15**:342–53. <https://doi.org/10.26599/NBE.2023.9290035>
 129. Anusha ZZ, Li J, Zuo H et al. AlphaFold 3–Aided design of DNA motifs to assemble into triangles. *J Am Chem Soc* 2024;**146**:25422–5. <https://doi.org/10.1021/jacs.4c08387>
 130. Ye Q, Eves R, Vance TDR et al. *Aeromonas hydrophila* RTX adhesin has three ligand-binding domains that give the bacterium the potential to adhere to and aggregate a wide variety of cell types. *mBio* 2025;**16**:e0315824. <https://doi.org/10.1128/mbio.03158-24>
 131. Gonzalez-Ordenes F, Herrera-Soto N, Muñoz SM et al. Glycogen metabolism in methanogens: A key pathway for metabolic response to nutrient availability. *Biochem Biophys Res Commun* 2024;**739**:150978. <https://doi.org/10.1016/j.bbrc.2024.150978>
 132. Yen IY, Whitfield GB, Rubinstein JL et al. Conformational changes in the motor ATPase CpaF facilitate a rotary mechanism of Tad pilus assembly. *Nat Commun* 2025;**16**:3839. <https://doi.org/10.1038/s41467-025-59009-5>
 133. Bandeira PT, Chaves CR, Monteiro Torres PH et al. Immunolocalization and 3D modeling of three unique proteins belonging to the costa of *Trichomonas foetus*. *Parasitol Res* 2025;**124**:30. <https://doi.org/10.1007/s00436-025-08466-4>
 134. Huang J, Peterson PP, Xu Z et al. Homeodomain protein Sxi1 α regulates cell-cell fusion during distinct sexual reproduction modes in *Cryptococcus deneoformans*. *BioRxiv*, 2025: 2025.02.11.637763. <https://doi.org/10.1101/2025.02.11.637763>
 135. Liu S, Lei T, Tan Y et al. Discovery, structural characteristics and evolutionary analyses of functional domains in *Acinetobacter baumannii* phage tail fiber/spike proteins. *BMC Microbiol* 2025;**25**:73. <https://doi.org/10.1186/s12866-025-03790-2>
 136. Shi Y. Drug development in the AI era: AlphaFold 3 is coming! *Innov Camb Mass* 2024;**5**:100685. <https://doi.org/10.1016/j.xinn.2024.100685>
 137. Watson ER, Novick S, Matyskiela ME et al. Molecular glue CELMoD compounds are regulators of cereblon conformation. *Science* 2022;**378**:549–53. <https://doi.org/10.1126/science.add7574>
 138. Masters MR, Mahmoud AH, Lill MA. Do deep learning models for Co-folding learn the physics of protein-ligand interactions? *Biorxiv*, 2024: 2024.06.03.597219. <https://doi.org/10.1101/2024.06.03.597219>
 139. He XH, Li JR, Shen SY et al. AlphaFold3 versus experimental structures: assessment of the accuracy in ligand-bound G protein-coupled receptors. *Acta Pharmacol Sin* Published online December 6, 2024. <https://doi.org/10.1038/s41401-024-01429-y>
 140. Eshak F, Goupil-Lamy A. AI driven approaches in nanobody epitope prediction: are we there yet? *Biorxiv*, 2024: 2024.10.07.616899. <https://doi.org/10.1101/2024.10.07.616899>
 141. Nam KH. Evaluation of AlphaFold3 for the fatty acids docking to human fatty acid-binding proteins. *J Mol Graph Model* 2024;**133**:108872. <https://doi.org/10.1016/j.jmgm.2024.108872>
 142. Sieskind R, Missouri S, Madru C et al. Crystal structures of monomeric BsmI restriction endonuclease reveal coordinated sequential cleavage of two DNA strands. *Commun Biol* 2025;**8**:387. <https://doi.org/10.1038/s42003-025-07612-z>
 143. Mamun Y, Aguado A, Preza A et al. Substrate binding of human and bacterial type IA topoisomerase: an experimentation with AlphaFold 3.0. *Comput Struct Biotechnol J* 2025;**27**:1342–9. <https://doi.org/10.1016/j.csbj.2025.03.041>
 144. McDonnell RT, Henderson AN, Elcock AH. Structure prediction of large RNAs with AlphaFold3 highlights its capabilities and limitations. *J Mol Biol* 2024;**436**:168816. <https://doi.org/10.1016/j.jmb.2024.168816>
 145. Degenhardt MFS, Degenhardt HF, Bhandari YR et al. Determining structures of RNA conformers using AFM and deep neural networks. *Nature* 2025;**637**:1234–43. <https://doi.org/10.1038/s41586-024-07559-x>
 146. Bernard C, Postic G, Ghannay S et al. Has AlphaFold 3 reached its success for RNAs? *Biorxiv*, 2024: 2024.06.13.598780. <https://doi.org/10.1101/2024.06.13.598780>
 147. Ji Z, Lee N, Frieske R et al. Survey of hallucination in natural language generation. *ArXiv* Published online July 14, 2024. <https://doi.org/10.48550/arXiv.2202.03629>
 148. Iebed D, Gökler T, van Ingen H et al. Phosphorylation of the HMGN1 nucleosome binding domain decreases helicity and interactions with the acidic patch. *ChemBiochem Eur J Chem Biol* 2024;**25**:e202400589. <https://doi.org/10.1002/cbic.202400589>
 149. Meng L, Wei L, Wu R. MVGNN-PPIS: A novel multi-view graph neural network for protein-protein interaction sites prediction based on Alphafold3-predicted structures and transfer learning. *Int J Biol Macromol* 2025;**300**:140096. <https://doi.org/10.1016/j.ijbiomac.2025.140096>
 150. Jabeen A, Oakeshott JG, Lee SF et al. Template-based modeling of insect odorant receptors outperforms AlphaFold3 for ligand binding predictions. *Sci Rep* 2024;**14**:29084. <https://doi.org/10.1038/s41598-024-80094-x>
 151. Cheng J, Liu J, Neupane P. Accurate prediction of protein complex stoichiometry by integrating AlphaFold3 and template information. *Research Square* Published online February 6, 2025:rs.3.rs-5855710. <https://doi.org/10.21203/rs.3.rs-5855710/v1>
 152. Majila K, Ullanat V, Viswanath S. A deep learning method for predicting interactions for intrinsically disordered regions of proteins. *BioRxiv* 2025: 2024.12.19.629373. <https://doi.org/10.1101/2024.12.19.629373>
 153. Gomes DEB, Yang B, Vanella R et al. Integrating dynamic network analysis with AI for enhanced epitope prediction in PD-L1:affibody interactions. *J Am Chem Soc* 2024;**146**:23842–53. <https://doi.org/10.1021/jacs.4c05869>
 154. Song B, Wang K, Na S et al. Cmai: predicting antigen-antibody interactions from massive sequencing data. *BioRxiv*, 2024.06.27.601035. <https://doi.org/10.1101/2024.06.27.601035>
 155. Li Q, Zhao Y, Chordia MD et al. Enhanced prediction of antigen and antibody binding interface using ESM-2 and Bi-LSTM. *Hum Immunol* 2025;**86**:111304. <https://doi.org/10.1016/j.humimm.2025.111304>
 156. Olenyi T, Marquet C, Grekova A et al. TMVisDB: annotation and 3D-visualization of transmembrane proteins. *J Mol Biol* 2025;**437**:168997. <https://doi.org/10.1016/j.jmb.2025.168997>
 157. Tikhonov DB, Korkosh VS, Zhorov BS. 3D-aligned tetrameric ion channels with universal residue labels for comparative structural analysis. *Biophys J* 2025;**124**:458–70. <https://doi.org/10.1016/j.bpj.2024.12.019>
 158. Zhang O, Naik SA, Liu ZH et al. A curated rotamer library for common post-translational modifications of proteins. *Bioinform Oxf Engl* 2024;**40**:btac444. <https://doi.org/10.1093/bioinformatics/btac444>
 159. Medvedev KE, Schaeffer RD, Grishin NV. DrugDomain: the evolutionary context of drugs and small molecules bound to domains. *Protein Sci Publ Protein Soc* 2024;**33**:e5116. <https://doi.org/10.1002/pro.5116>

160. Döpner P, Kemnitz S, Doerr M et al. af3cli: Streamlining AlphaFold3 input preparation. *J Chem Inf Model* 2025;**65**:3886–91. <https://doi.org/10.1021/acs.jcim.5c00276>
161. Liu J, Neupane P, Cheng J. Improving AlphaFold2 and 3-based protein complex structure prediction with MULTICOM4 in CASP16. *BioRxiv* 2025: 2025.03.06.641913. <https://doi.org/10.1101/2025.03.06.641913>
162. Nogales E, Mahamid J. Bridging structural and cell biology with cryo-electron microscopy. *Nature* 2024;**628**:47–56. <https://doi.org/10.1038/s41586-024-07198-2>
163. Bartas M, Slychko K, Brázda V et al. Searching for new Z-DNA/Z-RNA binding proteins based on structural similarity to experimentally validated α domain. *Int J Mol Sci* 2022;**23**:768. <https://doi.org/10.3390/ijms23020768>