Review | Received 17 September 2025; Revised 3 November 2025; Accepted 19 November 2025; Published 25 November 2025
https://doi.org/10.55092/bi20250005

Challenges and future directions in protein structure prediction: lessons from AlphaFold's evolution

Liangxu Xie

Institute of Bioinformatics and Medical Engineering, Jiangsu University of Technology, Changzhou 213001, China; E-mail: xieliangxu@jsut.edu.cn.

Highlights:

- Discusses the technology advancements in three generations of AlphaFold.
- Highlights current challenges in protein structure prediction and suggests potential strategies to overcome these obstacles.
- Offers insights into future research directions to enhance protein structure prediction methodologies and expand their applications.

Abstract: Protein structure prediction methods, exemplified by AlphaFold, have achieved near-experimental accuracy in predicting protein structures. AlphaFold has been widely applied in scientific fields, such as life sciences, synthetic biology, drug discovery, and materials design. This article provides a systematic overview of the major advances in protein structure prediction, focusing on the technical evolution of AlphaFold and its limitations in predicting dynamic conformational changes, complex interactions, and conformations under non-native conditions. Based on this analysis, key challenges in the current field are identified, and potential strategies to address these issues are proposed. The review offers perspectives on future directions for improving protein structure prediction research, providing insights for subsequent methodological enhancements and broader applications.

Keywords: AlphaFold; protein structure prediction; molecular dynamics; interaction prediction

1. Introduction

1.1. The historical development of protein structure prediction

The functions of proteins are closely related to their structures. Accurate determination of protein structure plays an important role in the regulation of biological physiological functions, material design, synthetic biology and other fields [1]. The structure of a protein plays a pivotal role in nearly every cellular process [2–5]. Enzymes facilitate biochemical reactions with exceptional specificity. Receptors transmit signals across membranes by undergoing coordinated conformational changes. Structural proteins provide the cellular framework and contribute to cell motility. Therefore, understanding a protein's structure is tantamount to uncovering the underlying mechanism of its function. For example, rational drug discovery heavily depends on structural information. Knowledge of the detailed three-dimensional

Copyright©2025 by the authors. Published by ELSP. This work is licensed under Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium provided the original work is properly cited.

structure of the target binding site enables in silico drug design. A variety of human diseases, including sickle cell anemia, cystic fibrosis, and numerous neurodegenerative disorders such as Alzheimer's and Parkinson's, are directly linked to protein misfolding, aggregation, or the loss of proper structural integrity. By visualizing the atomic-level details of a protein, we can identify the structural changes that lead to pathophysiological conditions. In this process, compounds can be computationally screened, optimized, and even designed from scratch to bind with high affinity and specificity. In recent decades, nuclear magnetic resonance (NMR), X-ray diffraction (XRD) and cryogenic electron microscopy (cryo-EM) have been developed to analyze protein structures. Although great efforts have been made, the number of protein structures obtained is much lower than the number of protein sequences resolved. Among them, the human genome encodes about 20,000 proteins, and after decades of development, only 35% of the protein structures have been resolved. One way to solve this problem is to predict structures of proteins by theoretical and computational methods [4,6].

Protein structure prediction originated from the theory of Anfinsen in 1962 [7]. Anfinsen pointed out that the three-dimensional structure of a protein was only determined by its amino acid sequence after observing the structural changes of protein folding in different environments. The theory means that the information needed to predict protein structure is already contained in the amino acid sequence. This theory laid the foundation for the prediction of protein structure by computational methods [8]. The early method of protein structure prediction was based on the thermodynamic theory. In 1967, Scherage used the energy minimization method to search for protein structure. Since the number of possible structures corresponding to amino acid sequences is close to an astronomical number, methods based on Monte Carlo or molecular dynamics (MD) simulations are not feasible. In 1969, Browne used a template to construct the structure of α-lactalbumin by comparing the sequences of a previously resolved Lysozyme protein. This is considered to be the beginning of protein structure prediction. In 1975, Michael Levitt and Arieh Warshel conducted energy minimization to simulate protein folding. In 1987, protein structure prediction based on multiple templates was realized, and multiple proteins in the same family were used as references to improve the accuracy of prediction. In 1993, MODELLER and SWISS-MODEL achieved the first automatic prediction of protein structure [9,10]. The threading method, which was reported in 1993, can "thread" the target sequence to the template of known structure to cope with the difficult problem when there are few homologous sequences [11]. In 1994, John Moult launched the first international protein structure prediction competition, Critical Assessment of protein Structure Prediction (CASP). The competition also led to the creation of a variety of computational methods. In 1997, David Baker's lab tried to assemble protein structures from the physical principles of piecing together fragments. After entering the 21st century, various methods have blossomed, and it is no longer a single type of method, but a combination of methods [12]. In this stage, the calculation method of a composite machine learning model was developed, and the information, such as contact map, distance matrix and secondary structure was used as a supplement to the method [13]. Various methods for fusing physical models and machine learning models have been developed.

Despite decades of research, the problem of protein structure prediction has not been solved. Until 2018, with the resurgence of deep learning methods, AlphaFold made its debut [14]. At the CASP14 competition in 2020, DeepMind's AlphaFold2 model achieved accurate predictions of protein structures with experimental accuracy [15]. In 2022, RoseTTAFold developed by David Baker's team at the University of Washington in the United States also achieved a highly accurate prediction of protein structure [16]. The

method they developed is believed to solve a five-decade-old scientific problem: predicting the structure of a protein from the sequence of its amino acids [5]. Structure prediction can be applied to various fields, such as drug screening, protein material design, sensor design, *etc.* Given the great potential for application, the 2024 Nobel Prize in Chemistry was awarded to David Baker of the University of Washington, along with Dmis Hassabis and John M. Jumper of the DeepMind team.

1.2. Methods for protein structure prediction

After more than five decades of development, protein structure prediction methods can be broadly divided into two categories: template-based and template-free methods [5,17,18]. The workflow of template-based methods, namely homology modeling methods and threading methods, is as follows. Firstly, the appropriate structure is selected as the template according to the sequence similarity, and then the region where the target sequence is consistent with the template sequence is selected as the corresponding structure. Finally, the molecular modeling technology and other techniques are used to supplement the structure of the mutation, insertion and deletion in the target-template alignment. The homology modeling method relies on the structure of homologous proteins as a template to build the structure of the sequence to be sequenced. Some well-known methods include SWISS-MODEL [9], Modeller [10] and I-TASSER [19]. The template-free structure prediction method, namely de novo prediction method, needs to find the structure of different source proteins in the protein structure database due to the lack of structure templates, and then assemble the fragments for the initial structure. These methods require the use of potential energy functions for conformational sampling and optimization. Rosetta and QUARK perform well in the field of de novo structure prediction [20–22].

A major advance in protein structure prediction is the introduction of co-evolution information into the algorithm [23]. Researchers have found that the mutation of proteins will result in pairwise mutations of amino acids. This is because when two amino acids are in contact with each other in space, when one amino acid is mutated, the other amino acid will undergo a corresponding mutation under the pressure of evolution. The co-evolution of amino acids in the sequence is that the amino acids will be closer in space, and the interaction between residues can be inferred. Therefore, the contact pairs of amino acid residues are applied to the prediction of protein structure. The main task of sequence alignment is to find the most homologous sequence on the evolutionary tree for the target sequence from the database, and then score the similarity between the target sequence and the query sequence according to the similarity of the mutated amino acids according to a specific rule. The alignment of several given sequences with the query sequence using the above alignment method is called multiple sequence alignment (MSA). The contact or distance between different residues in the same sequence is very important information, which basically contains all the information of the 3D structure of the protein skeleton. Neural network-based learning methods have further extended the use of multiple sequence alignment (MSA) to end-to-end protein structure prediction, and a typical successful example is AlphaFold. AlphaFold made its debut in the CASP13 competition in 2018, and there was no significant gap between the results of the competition and other methods. In the CASP14 of 2020 competition, AlphaFold2 made a revolutionary breakthrough, outperforming other participating teams by a large magnitude, predicting protein structures with the same accuracy as experiments [24].

Up to now, there are three main versions of AlphaFold. The original version of AlphaFold uses convolutional neural networks to process MSA data, and combines physicochemical constraints to

predict 3D protein structures [14]. AlphaFold2 uses a combination of Transformer and graph neural network architecture, integrates sequence information and residue pair representation through the Evoformer module, and approaches experimental resolution level in the structure prediction of single proteins [15]. AlphaFold2 model has successfully predicted the structures of approximately 220 million proteins in just two years. AlphaFold Database (AlpahFold DB) offers over 200 million predicted protein structures for free access. This effectively covers almost all known proteins, providing a starting point for research without requiring any computational work [25]. AlphaFold3 was launched in 2024, and the model architecture introduces a brand new Pairformer module to support predicted monomer structure prediction of complex molecules such as ligands, nucleic acids, and metal ions [26]. The three models of AlphaFold represent the gradual evolution of protein structure prediction from basic exploration, monomer prediction to complex prediction [27].

Protein structure prediction is having a widespread impact. While we acknowledge that many other models also achieve state-of-the-art performance as summarized in Table 1, this review focuses on the technical evolution of AlphaFold. By tracing its developmental trajectory, we tried to derive a framework for understanding the current limitations and potential future developments of these rapidly advancing methods.

Table 1. Comparison of structure prediction models in terms of architecture, data requirements, strengths and limitations.

Model	Architecture	Data Requirements	Strengths	Limitations
AlphaFold2 [15]	Transformer + Evoformer module; integrates MSA and template structures.	MSA from UniRef/BFD databases and template structures from PDB.	Highest accuracy. Covers 98% of known proteins (AlphaFold DB).	High computational cost. Relies on evolutionary information; poor performance for orphan proteins.
AlphaFold3 [26]	Pairformer (simplified MSA processing) + Diffusion Module.	Polymer sequences, ligand SMILES, modified residues.	Unified model with SOTA accuracy across proteins, nucleic acids, ligands, ions, and modifications. Reduced MSA dependence.	Only via restricted server. Potential for stereochemical violations (chirality, clashes). "Hallucinate" structure in disordered regions.
RoseTTAFold All-Atom [28]	Tri-track Transformer fusing residue-level (protein/nucleic acid) and atomic-graph (small molecules/metals) representations.	Protein-ligand complexes, protein-metal structures, and small-molecule crystal data.	Models multi-component systems. Enables de novo design of ligand-binding proteins.	High computational complexity; slower than ESMFold. Accuracy generally lower than AF3.
ESMFold [29]	ESM-2 protein language model. Single-sequence input. Transformer embeddings.	Single protein sequence; trained on 393M unaligned sequences.	Ultra-fast inference (10× faster than AlphaFold). No MSA required. Excels on proteins with limited evolutionary information.	Lower precision for complex topologies than AlphaFold. Struggles with viral proteins and specific domains.
OmegaFold [30]	OmegaPLM (670 milion parameters) + Geoformer Architecture	Single protein sequence.	Fast predictions. No MSA required. Competitive performance on de novo proteins and sequences with limited homology.	Lower accuracy than AF2 and ESMFold.

2. Technical evolution of AlphaFold

AlphaFold1 breaks the limitations of traditional protein structure prediction and achieves a significant improvement in the prediction accuracy of free modeling domains in CASP13 (24/43 domains achieve TM-score ≥ 0.7) [14,31]. Its architecture lays the foundation for the subsequent Evoformer and Transformer innovations of AlphaFold2. AlphaFold2 performs well in single protein structure prediction, but the prediction of protein complexes and interactions remains unsolved [32]. AlphaFold3 uses a diffusion model to predict the interactions of proteins, nucleic acids, small molecules, ions, modified residues and other biomolecules, including protein-ligand, protein-nucleic acid, antibody-antigen, protein-protein complex, *etc* [26]. Predicting the interactions between proteins and other molecules is of great significance for life science and drug discovery [33]. The major breakthroughs of three generations of AlphaFold are summarized in Table 2.

Table 2. Comparison of key technologies of three generations of AlphaFold methods.

Key technologies	AlphaFold1	AlphaFold2	AlphaFold3
Core Network architecture	Convolutional neural network extracts sequence and structure features, and generates contact maps to guide structure assembly.	Evoformer architecture (self-attention mechanism module) integrates evolutionary relationships and spatial geometric information.	Evoformer architecture is optimized to support multi-chain interaction modeling and enhance non-protein component feature processing capabilities.
Input features and information utilization	MSA is used as the core input to predict residue distance constraints using evolutionary information.	MSA is fused with single sequence features, and the evolutionary and spatial information is dynamically transferred through Evoformer.	Extended input types: support non-protein component features such as proteins, nucleic acids (DNA/RNA), small molecules, and metal ions.
Structure generation mechanism	Contact map is used to predict the distance between residues and guide the 3D structure assembly.	Structure Module generates 3D atomic coordinates directly from distance/angle predictions.	Multi-chain interface residue contact prediction is optimized to support the direct generation of 3D structures for complex and multi-subunit systems.
Applicable System types	Single protein structure prediction	Single protein, including no-template condition	Multi-molecular systems, including protein complexes, protein-nucleic acids, and protein-small molecule binding structures.
Degree of template dependence	Depending on homologous template, traditional template matching method should be combined.	Dependence on homologous templates is significantly reduced, and high-accuracy prediction is still achieved under the condition of no template.	Dependence on templates is further reduced, and the prediction of template-free multi-chain complexes is supported.

2.1. AlphaFold1 core technology

2.1.1. Core architecture

AlphaFold1 surpasses the limitations of traditional protein structure prediction by leveraging three core technologies: distance distribution prediction, end-to-end deep learning, and gradient descent optimization. Alphafold1 mainly includes the following modules:

Feature extraction module: Search databases such as UniClust30 by HHblits and PSI-BLAST to generate MSA of the target sequence, and extract sequence spectrum features (such as position-specific substitution probability) and co-evolution features. The MSA features, sequence length, residue type, *etc.* are encoded into 1D and 2D feature matrices, which are used as the input of the neural network.

Distance prediction network: A two-dimensional dilated convolutions residual network with 220 residual blocks was used to capture long-range dependencies by recycling the expansion rates (1, 2, 4, and 8). The distance distribution (distogram) between residue pairs was predicted, and the distance from 2–22 Å was quantified into 64 discrete bins. Meanwhile, the torsion angle (φ, ψ) and relative solvent accessible surface area of the backbone were also predicted.

Potential energy function construction: Based on the predicted distance distribution, a smooth potential energy function is constructed by negative log-likelihood, and a reference distribution is introduced to correct for sequence length and residue type bias. A potential energy term based on the torsion angle is constructed, and a van der Waals force term is added to avoid atomic collisions.

2.1.2. Key innovations

AlphaFold1 builds on traditional protein structure prediction methods to achieve the following core innovations:

From “contact prediction” to “distance distribution prediction”: Traditional methods only predict whether a residue pair is a “contact” (e.g., < 8 A), providing binary information. AlphaFold1 predicts the full probability distribution of residue pair distances, containing continuous distance information and uncertainty estimates, providing richer constraints for structural modeling.

End-to-end deep learning architecture: A full two-dimensional convolutional residual network is used to predict the distance distribution directly from the MSA features, avoiding the multi-step error accumulation in traditional methods. The co-evolution information (such as residue co-variation) and sequence spectral features of MSA are integrated to efficiently capture long-range interactions by dilated convolution.

2.2. AlphaFold2 core techniques

2.2.1. Core architecture

AlphaFold2 achieves atom-level accuracy [15] in protein structure prediction through technical breakthroughs such as Evoformer module, Transformer attention mechanism, and self-distillation training. Its architecture and training strategy have fundamentally changed compared with AlphaFold1, especially in the template-free and complex conformation prediction. It mainly consists of the following modules:

Evoformer module: As the core encoder, it contains 48 blocks with non-shared weights. It processes MSA through the row/column gated self-attention mechanism, and uses the triangular attention

mechanism (based on the triangle relationship of residues) to optimize the pairwise representation, so as to realize the dynamic interaction between MSA and pairwise information. Evolution and physical constraints are integrated through the self-attention mechanism.

Structure module: As a decoder, it contains eight blocks with shared weights. The abstract features are transformed into 3D atomic coordinates through the Invariant Point Attention (IPA) mechanism, and the structural conflict is optimized by energy calculation relaxation combined with the Amber force field.

Self-distillation training: 75% of the training data comes from the unlabeled structures predicted by the model, and 25% of the protein data bank (PDB) experimental structures are combined to enhance the generalization ability to complex conformations.

2.2.2. Key innovations

Transformer Attention mechanism: The Transformer is first applied to protein structure prediction to capture long-range dependencies and co-evolutionary relationships through multi-head attention.

Coevolutionary information extraction: Actively leverages residue covariation signals (e.g., spatial proximity of pairs of residues) in MSAs instead of template dependencies from traditional homology modeling. MSA manipulation has been proven as a strategy to sample more conformational states. As exemplified by Wayment-Steele *et al.*, clustering or subsampling the MSA can drive AF2 to sample different clusters [34,35]. Developing a more granular, probabilistic framework for generating a diverse ensemble of MSAs from the full sequence database will be a potentially useful way to capture subtler conformational states [36].

2.3. AlphaFold3 Core technology

2.3.1. Core architecture

AlphaFold3 is a subversive reconstruction based on AlphaFold2, and adopts a diffusion generative architecture to achieve the unified prediction of multi-type biomolecular complexes such as proteins, nucleic acids, small molecules, ions and modified residues [26]. Its core architecture includes the following modules:

Pairformer module: As the core encoder, it replaces AlphaFold2's Evoformer module, processing only pair representations and single representations, discarding MSA representations to simplify computation. It contains 48 independent parameter blocks and optimizes pairwise interactions through triangular attention and transition layers, which retain the ability to model long-range dependencies while reducing computational complexity.

Diffusion module: Replacing the structure module of AlphaFold2, it predicts atomic coordinates directly based on the diffusion generative model without relying on the main-chain backbone or torsional angle parameterization. The model receives "noisy" atomic coordinates during training and learn to progressively denoise to recover the true structure, while modeling structural features (local stereochemistry and global conformation) at different scales. At inference time, starting from random noise, the final structure is generated by iterative denoising, and multi-seed sampling is supported to explore the conformational space.

2.3.2. Key innovations

MSA embedding simplification: The processing of MSA is greatly weakened; only small MSA embedding blocks are retained, sequence features are extracted by pair-weighted averaging, which reduces the dependence on homologous sequences and improves the generalization ability for low homology targets.

Multi-molecule type support: The input supports protein sequences, nucleic acid sequences (DNA/RNA), small molecule SMILES, ions and modified residues (e.g., phosphorylation, glycosylation), without special coding for different molecule types.

Diffusion generative architecture: Instead of complex parameterization based on backbone framework and torsion angle in AlphaFold2, 3D atomic coordinates are directly generated through diffusion module, which simplifies chemical rule constraints (e.g., stereochemistry, bond length and bond angle) and reduces the dependence on expert knowledge.

Through four core innovations of diffusion generative architecture, including simplified MSA processing, cross-distillation training, and unified modeling of multi-molecule types, AlphaFold3 breaks through the limitation that AlphaFold2 can only predict protein structures, and achieves high-precision prediction of proteins, nucleic acids, small molecules, ions, and modified residue complexes. The development process of AlphaFold for three generations is essentially the collaboration of data and algorithms, as well as the cooperation and complementary advantages of physical models and AI algorithms, which reflects the collaborative innovation of "data-driven modeling" and "physical rule guidance".

2.4. Implications of technology evolution

AlphaFold1 was upgraded from traditional "contact prediction" (binary determination of whether a residue is in contact) to "distance distribution prediction", which provides continuous distance information and uncertainty estimation, greatly improving the richness of structural constraints. AlphaFold2 uses the Evoformer module to replace the convolutional network, and dynamically integrates MSA co-evolution information and pairwise interactions through the Transformer attention mechanism to handle long-term dependencies. AlphaFold3 uses the diffusion module to replace the structure module, and directly predicts atomic coordinates based on the diffusion generation model without relying on the backbone framework or torsion angle parameterization. It is naturally compatible with multi-type molecules such as small molecules, ions, and nucleic acids.

2.4.1. Collaboration between data and algorithms

Reduce data dependence: At the data level, self-distillation of AlphaFold2 (augmenting the training data with model prediction structure) and cross-distillation of AlphaFold3 (modeling the unorganized regions with AF-Multimer prediction data) break through the limitations of experimental structure data, proving that AI models can improve generalization ability through "self-generated data".

Advancing model architecture: At the algorithmic level, the evolution of the AlphaFold, from convolutional networks to Transformers to diffusion models, demonstrates the central role of deep learning architectural innovations (e.g., attention mechanisms, generative modeling) in solving complex scientific problems, far exceeding the fitting ability of traditional physical models.

2.4.2. Mutual promotion between physical models and AI algorithms

Physical models provide "underlying constraints" for AI: In AlphaFold1/2, physical models are directly involved in structural optimization. Even though AlphaFold3 employs a diffusion generative model, physical priors are retained, e.g., the training data contains experimentally determined protein-ligand binding patterns, and the physical laws of small-molecule-protein and protein-protein interactions (e.g., hydrogen bonding, hydrophobic interactions) are indirectly introduced.

AI algorithms enable the "offloading" and discovery of physical models: while traditional physical models require manual encoding of complex rules (e.g., bond lengths, bond angles, stereochemistry), AI learns from data to implicitly capture these rules. For example, AlphaFold3's diffusion module can be trained to generate chemically common-sense structures without explicitly defining "torsional angle ranges". The high-resolution structures predicted by AI provide new subjects for physical models to study. For example, novel protein folding patterns predicted by AlphaFold have driven a re-understanding of the energy landscape of protein folding.

AlphaFold's success is by no means an isolated technological breakthrough, but structural biology (scientific) and AI technology (technology) promote each other for a long time, resulting in the positive cycle of "data accumulation → tool innovation → scientific discoveries → new data generation".

3. Challenges and limitations

AlphaFold has revolutionized protein structure prediction but faces key limitations. Its accuracy heavily relies on high-quality MSA; and sparse evolutionary data (e.g., orphan proteins) degrades predictions. It primarily predicts static, low-energy structures, failing to capture dynamic conformational changes (e.g., enzyme activation) or conformational heterogeneity critical for function. For intrinsically disordered proteins (IDPs) and membrane proteins, it often generates artificially "folded" or topologically inaccurate structures. It lacks integration of environmental factors (pH, temperature, ligands), ignoring their impact on conformation. It poorly handles non-standard components: post-translational modifications (PTMs), non-canonical amino acids, or cofactors (e.g., heme), requiring manual correction. It provides no direct insight into function or mechanisms (e.g., catalytic activity), needing auxiliary tools for functional annotation.

3.1. Over-reliance on MSA and data scarcity

AlphaFold's accuracy is highly dependent on deep, high-quality MSAs [37]. For orphan proteins, rapidly evolving proteins, or certain viral proteins, the lack of homologous sequences can degrade the prediction accuracy of AlphaFold [38]. For low MSA depth of protein, the structure predictions are unreliable, and pLDDT ratings are generally lower. For example, AlphaFold cannot predict protein structures after mutations, which can cause local structural changes in proteins. Since AlphaFold relies on high-quality MSAs for its predictions, and the mutated sequence is highly similar to the wild-type sequence, AlphaFold may also be unable to predict the structural changes that these mutations may cause, as it assumes that all homologous sequences have the same structure. AlphaFold3 tries to improve on this by reducing its reliance on MSA co-evolution information and retaining only small MSA embeddings. However, while AlphaFold3 uses the prediction structure of AlphaFold-Multimer during training, it also

shows dependence on MSA [26]. Therefore, there is a need to strike a balance between accuracy and generality of structure prediction.

3.2. Dilemma of dynamic conformation and flexible region prediction

The protein structures predicted by AlphaFold are static stable structures. Proteins are dynamic, not static. How proteins move, fold, and become allosteric is a common concern for biologists and pharmacologists. All three generations of AlphaFold predict the most stable "average" structure, or the lowest energy state. It is unable to capture critical conformational changes during functional execution (e.g., induced-fit in enzyme catalysis, opening and closing of ion channels, allosteric changes upon ligand binding) [39]. It is often difficult to achieve reliable structural ensemble prediction for proteins with dynamic structures [40]. For example, intrinsically disordered proteins (IDPs) cannot be effectively predicted and are usually shown as loose coils with low confidence [41]. Recent methodological advances now enable these architectures to sample structural diversity, for example, AI-MD hybrid approaches or diffusion model. The hybrid methods leverage AI-predicted structures as initial coordinates for physics-based MD simulations [42]. MD simulation introduces environmental context (solvent, membranes, ions). This hybrid approaches allow researchers to study conformational transitions, allosteric pathways, and binding events that are inaccessible to pure AI prediction [43]. Alternative approaches directly modify AI inference to generate multiple conformations. For AlphaFold2, strategic subsampling or clustering of MSAs can bias predictions toward distinct functional states, successfully revealing inward/outward states of transporters and apo/holo forms of enzymes. Meanwhile, AlphaFold3's inherent diffusion-based architecture provides a probabilistic framework that can produce structural variants through different random seeds, offering a native pathway to ensemble generation. The latest model, BioEmu, can efficiently generate protein conformational landscapes, complementing rather than replacing traditional MD simulation, to capture dynamic conformation changes under thermodynamic laws [44].

3.3. Challenges of protein complex structure and interaction prediction

Biological activities depend on protein-protein, protein-nucleic acid (DNA/RNA), and protein-ligand (such as drug molecules) interactions. AlphaFold2, in its initial incarnation, was primarily optimized for single-chain prediction, and its performance on complexes, while often impressive, revealed key limitations. The prediction of quaternary structure and the structural characterization of complexes represent a central challenge in structural biology, as the vast majority of cellular processes are governed not by isolated proteins, but by intricate, dynamic macromolecular assemblies. The development of AlphaFold-Multimer marked a significant step forward, explicitly training the network to handle multiple chains and incorporating interface-focused metrics [45]. This has enabled the accurate structural prediction of many known complexes directly from sequence, a capability that is revolutionizing fields like systems biology and immunology [46]. However, its accuracy of heterologous complex interface prediction is still unstable, and it sometimes will produce "illusion" (hallucination), which generates a plausible but wrong interface. AlphaFold3 can generalize the prediction of proteins, nucleic acids, ligands, and their complexes. RoseTTAFold All-Atom generalizes biomolecular modeling, including small molecules and covalent modifications. Despite its excellent performance on multiple interactions, AlphaFold3 currently only supports the prediction of regular protein interactions with nucleic acids, ligands, and cofactors. Moreover,

the accuracy of the prediction of interactions does not reach the accuracy of the prediction of protein-protein complexes. The prediction of complex biomolecular complexes with fewer templates (e.g., protein complexes formed by cross-species interactions and viral invasion) is still a challenge [47].

3.4 Ignorance of the external environment

AlphaFold's training data come from crystal structures (in a specific crystal environment or frozen at low temperature), and the vast majority of structures in the protein crystal database were resolved [48]. These structures can be thought of as snapshots of the protein in an "idealized" or "frozen" state. It does not represent the state of the protein in the complex, dynamic, crowded real intracellular environment (e.g., cytoplasm, cell membrane). AlphaFold learns its physical rules implicitly rather than explicitly by statistical learning. Therefore, it cannot explicitly take pH (which affects the protonation state of amino acids), ion concentration (which affects the electrostatic interaction), and lipid environment (which affects the hydrophobic interaction) as input parameters for physical calculations like MD simulation. It can only predict the "most likely" structure it has learned from training data, which lacks systematic information about these environmental changes.

4. Future directions to address the challenge

The evolution direction of AlphaFold can be summarized as follows: from single to complex, from static to dynamic, and from structure to functionality. Future developments will directly attempt to address "interactions", "function", and "environmental factors". AlphaFold's exploration has focused on moving from static structures to dynamic systems and complex environments. To further improve the accuracy of predicting protein-small molecule binding affinities and to be able to reliably predict the impact of missense mutations on structural stability and complex binding, and how to make the model "sense" the cellular environment and train the model to predict protein states in crowded cellular environments. Until then, protein design based on function and structure can be realized [49]. We summarized the possible technology development and their applications in Figure 1.

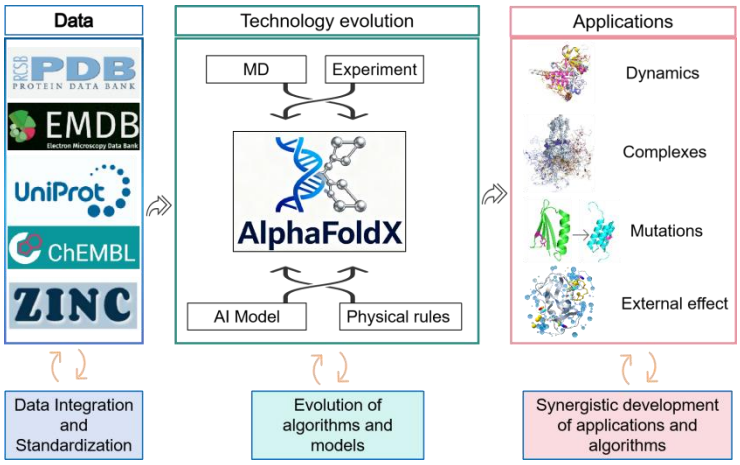

Figure 1. Potential technology developments and their applications. The future models will be trained on all universal databases, including protein and ligands molecule databases. The models will be integrated with MD simulation, experimental data, generative AI and physical rules to fulfill more realistic applications. The application and algorithms will be developed synergistically.

4.1. Data integration and standardization

Early tools like AlphaFold and ESMFold were primarily trained on static protein structures from the PDB [27]. BioEmu demonstrates a leap forward by integrating three distinct types of data: static structures, dynamical data and experimental properties [44]. The integration of these three types of data is far from straightforward. It requires sophisticated algorithms and computational methods to combine the different data sources in a meaningful way. Each type of data has its own format, scale, and level of uncertainty. For example, static structures are often represented as atomic coordinates, while dynamical data might be in the form of time-series trajectories. Experimental properties can be diverse, ranging from numerical values of binding constants to qualitative descriptions of protein behavior. BioEmu was pre-trained on a clustered version of the AlphaFold database. A data augmentation strategy was employed during this process, which motivated it to sample a wide variety of conformations. Regarding the MD data, they are reweighted to the equilibrium distributions. When dealing with experimental measurements from the MEGAscale dataset, new algorithms need to be incorporated into the training model. This situation highlights an increasing necessity for standardized methods. These methods are crucial for curating, weighting, and merging heterogeneous biological data into a unified training framework. The future model will not only rely on the experimental structures available in the database but can also incorporate the generated structures.

4.2. Next-generation evolution of algorithms and models

Integration with MD simulations: Existing tools lack information about dynamic structural changes in proteins, and many proteins switch multiple conformational states between functionally active and inactive states, which is important for understanding the functional state of proteins. While AlphaFold3 excels in many areas, its performance in modeling ternary complexes can be inflated by the presence of large accessory proteins that are not directly involved in the specific binding interface [50]. Future evaluations should incorporate protein flexibility. The AlphaFold structure can be used as an initial configuration to study dynamic processes, verify stability, and explore conformational space through MD simulations [51]. In this case, AlphaFold or other deep learning based conformational ensemble prediction methods also provide a feasible direction to first obtain the conformational states of protein samples, and then obtain the possible transition paths between them by MD to further expand the sampling of the conformational space. Integration of physics-based computational methods with AI models is enabling a multiscale, atomic-level understanding of biology, with the ultimate goal of creating predictive models of entire cells [42].

Complementary to experimental approaches: AlphaFold predictions can be cross-validated or integrated with Cryo-EM, NMR, small-angle scattering (SAXS), and other data. The experimental data were used as constraints to achieve the prediction of protein structure [52,53]. A comprehensive evaluation of AlphaFold3 on GPCR-ligand complexes found that while the receptor backbone accuracy improved, predictions for ligand binding poses were often inaccurate [54]. The most reliable results come from a cycle of computational prediction and experimental validation.

Integration with generative AI: The fusion of AlphaFold3 and diffusion models enables the prediction of multi-molecule complex structures, so it may be possible to combine AlphaFold with other emerging generative algorithms in the future. There are also studies trying to solve more complex problems [55].

Less reliance on MSA: New models with "single sequence" or low MSA requirements, such as ESMFold, have been attempted. ESMFold performs better than AlphaFold2 in predicting single-sequence proteins with no natural homologs or proteins that are highly evolving, indicating that language models can supplement partial evolutionary information [56]. However, for proteins with homologous sequences, ESMFold still performs worse than AlphaFold2. At present, the evolutionary information brought by natural homologous sequences is still needed, and it is not possible to rely on language models to solve the problem of structure prediction for all proteins [57].

Explicit integration of physical rules: Future models need to incorporate more accurate physical energy functions into model training, rather than just implicit learning [42]. A wider range of interactions are required to explicitly introduce to achieve structure prediction under the regulatory effects of different external conditions, such as pH, salt concentration, membrane environment, and molecular chaperones. Therefore, it is necessary to integrate explicit physical rules and environmental parameters into deep learning models to develop next-generation algorithms capable of predicting condition-dependent structure and conformational dynamics.

4.3. Synergistic development of applications and algorithms

With the emergence of these novel technologies, there has been a notable shift towards dynamic, mechanistic, and design-oriented tasks. Dynamic tasks now facilitate the analysis of conformational change process, enabling researchers to examine how different components interact and evolve over time [58]. Mechanistic studies probe deeper into the underlying mechanisms of biological and physical processes, uncovering the intricate details of how biomolecules function at a fundamental level [59]. Design-oriented tasks, conversely, enable scientists to create and optimize new structures and systems with specific functions [60]. Application serves as the ultimate test of a model's generality and the primary means of uncovering its limitations. The limitations identified are not simply failures but rather invaluable guides for future research. The identification of these limitations through real-world application creates a positive feedback loop. It provides a clear roadmap for the next cycle of technological advancement, guiding the collection of new training data (e.g., more MD simulations with ligands), the development of new algorithms (e.g., conditioning on environmental variables), and the expansion of model scope. Gathering more data from applications will propel the advancement of AI technology. By inputting this diverse data into prediction models, the algorithms can be updated and refined. This hybrid approach (AI algorithm + Application + Experiment) represents a new and more efficient paradigm. It combines the ability of AI algorithms to rapidly process and analyze large volumes of data with the reliability and accuracy of experimental data.

5. Conclusion

Protein structure is the basis of function. Computational methods can be used to predict protein structure directly from protein sequence. AlphaFold has become an indispensable standard tool in structural biology and the life sciences. AlphaFold is used by many computational methods and practical applications. Whether it is an experimental biologist testing a hypothesis, a drug discovery team performing an initial virtual screen, or an evolutionary scholar analyzing sequence-structure-function relationships, their first instinct is often to "predict a model with AlphaFold first". Although great

progress has been made in this field, there are still major challenges in the areas of predicting conformation after mutation, structural changes after translational modifications, and response to environmental effects. The focus of research shifts to more complex problems that AlphaFold is currently not good at or unable to solve, such as more general complex interactions, conformational dynamics, and de novo design of proteins and enzymes with novel functions. Far from diminishing, its impact continues to deepen and expand due to its ease of use (via AlphaFold Server) and expansion of the database (AlphaFold DB). It has dramatically lowered the barrier to entry in structural biology and accelerated the research process in countless labs around the globe. We are living in an era where the foundations are being laid by AlphaFold and where addressing the new challenges AlphaFold has revealed is at its core.

Acknowledgments

This work was funded by the National Foundation of Science of Jiangsu Province with grant number BK20191032.

Conflicts of interests

The author declares no conflict of interest.

References

[1] Kuhlman B, Bradley P. Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.* 2019, 20(11):681–697.

[2] Yang Z, Zeng X, Zhao Y, Chen R. AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct. Target Ther.* 2023, 8(1):115.

[3] Huang B, Kong L, Wang C, Ju F, Zhang Q, *et al.* Protein structure prediction: challenges, advances, and the shift of research paradigms. *Genom. Proteom. Bioinform.* 2023, 21(5):913–925.

[4] Jänes J, Beltrao P. Deep learning for protein structure prediction and design—progress and applications. *Mol. Syst. Biol.* 2024, 20(3):162–169.

[5] Fatima Ali N, Khan S, Zahid S. A critical address to advancements and challenges in computational strategies for structural prediction of protein in recent past. *Comput. Biol. Chem.* 2025, 117:108430.

[6] Baker D, Sali A. Protein structure prediction and structural genomics. *Science* 2001, 294(5540):93–96.

[7] Anfinsen CB, Haber E, Sela M, White FH. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. U.S.A.* 1961, 47:1309.

[8] Wooley JC, Ye Y. A historical perspective and overview of protein structure prediction. In *Computational Methods for Protein Structure Prediction and Modeling: Volume 1: Basic Characterization*. New York: Springer New York, 2007. pp. 1–43.

[9] Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* 2014, 42(W1):W252–W258.

[10] Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 1993, 234(3):779–815.

[11] Xu Y, Liu Z, Cai L, Xu D. Protein structure prediction by protein threading. In *Computational Methods for Protein Structure Prediction and Modeling: Volume 2: Structure Prediction*. New York: Springer New York, 2007. pp. 1–42.

[12] Braitbard M, Schneidman-Duhovny D, Kalisman N. Integrative structure modeling: overview and assessment. *Annu. Rev. Biochem.* 2019, 88:113.

[13] Cheng J, Tegge AN, Baldi P. Machine learning methods for protein structure prediction. *IEEE Rev. Biomed. Eng.* 2008, 1:41–49.

[14] Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, *et al.* Improved protein structure prediction using potentials from deep learning. *Nature* 2020, 577(7792):706–710.

[15] Tunyasuvunakool K, Adler J, Wu Z, Green T, Zielinski M, *et al.* Highly accurate protein structure prediction for the human proteome. *Nature* 2021, 596(7873):590.

[16] Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* 2021, 373(6557):871–876.

[17] Dill KA, Ozkan SB, Shell MS, Weikl TR. The protein folding problem. *Annu. Rev. Biophys.* 2008, 37:289.

[18] Dill KA, MacCallum JL. The protein-folding problem, 50 years on. *Science* 2012, 338(6110):1042–1046.

[19] Yang J, Zhang Y. Protein structure and function prediction using I-TASSER. *Curr. Protoc. Bioinform.* 2015, 52:5.8.1–5.8.15.

[20] Kuhlman B. Designing protein structures and complexes with the molecular modeling program Rosetta. *J. Biol. Chem.* 2019, 294(50):19436–19443.

[21] Floudas CA, Fung HK, McAllister SR, Mönnigmann M, Rajgaria R. Advances in protein structure prediction and de novo protein design: a review. *Chem. Eng. Sci.* 2006, 61(3):966–988.

[22] Xu D, Zhang Y. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins* 2012, 80(7):1715–1735.

[23] Anishchenko I, Ovchinnikov S, Kamisetty H, Baker D. Origins of coevolution between residues distant in protein 3D structures. *Proc. Natl. Acad. Sci. U.S.A.* 2017, 114(34):9122–9127.

[24] Lensink MF, Brysbaert G, Mauri T, Nadzirin N, Velankar S, *et al.* Prediction of protein assemblies, the next frontier: The CASP14-CAPRI experiment. *Proteins Struct. Funct. Bioinf.* 2021, 89(12):1800–1823.

[25] Varadi M, Bertoni D, Magana P, Paramval U, Pidruchna I, *et al.* AlphaFold protein structure database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res.* 2024, 52(D1):D368–D375.

[26] Abramson J, Adler J, Dunger J, Evans R, Green T, *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024, 630(8016):493–500.

[27] Agarwal V, McShan AC. The power and pitfalls of AlphaFold2 for structure prediction beyond rigid globular proteins. *Nat. Chem. Biol.* 2024, 20(8):950–959.

[28] Krishna R, Wang J, Ahern W, Sturmfels P, Venkatesh P, *et al.* Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* 384(6693):eadl2528.

[29] Lin Z, Akin H, Rao R, Hie B, Zhu Z, *et al.* Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 2023, 379(6637):1123–1130.

[30] Wu R, Ding F, Wang R, Shen R, Zhang X, *et al.* High-resolution *de novo* structure prediction from primary sequence. *bioRxiv* 2022, bioRxiv:2022.07.21.500999.

[31] AlQuraishi M. AlphaFold at CASP13. *Bioinformatics* 2019, 35(22):4862.
[32] Mirabello C, Wallner B, Nystedt B, Azinas S, Carroni M. Unmasking AlphaFold to integrate experiments and predictions in multimeric complexes. *Nat. Commun.* 2024, 15(1):8724.
[33] Lyu J, Kapolka N, Gumpper R, Alon A, Wang L, *et al.* AlphaFold2 structures guide prospective ligand discovery. *Science* 2024, 384(6702):eadn6354.
[34] Wayment-Steele HK, Ojoawo A, Otten R, Apitz JM, Pitsawong W, *et al.* Predicting multiple conformations via sequence clustering and AlphaFold2. *Nature* 2024, 625(7996):832–839.
[35] del Alamo D, Sala D, McHaourab HS, Meiler J. Sampling alternative conformational states of transporters and receptors with AlphaFold2. *eLife* 2022, 11:e75751.
[36] Guan X, Tang Q, Ren W, Chen M, Wang W, *et al.* Predicting protein conformational motions using energetic frustration analysis and AlphaFold2. *Proc. Natl. Acad. Sci. U.S.A.* 2024, 121(35):e2410662121.
[37] Chakravarty D, Lee M, Porter LL. Proteins with alternative folds reveal blind spots in AlphaFold-based protein structure prediction. *Curr. Opin. Struct. Biol.* 2025, 90:102973.
[38] Michaud JM, Madani A, Fraser JS. A language model beats alphafold2 on orphans. *Nat. Biotechnol.* 2022, 40(11):1576–1577.
[39] Masters MR, Mahmoud AH, Lill MA. Investigating whether deep learning models for co-folding learn the physics of protein-ligand interactions. *Nat. Commun.* 2025, 16(1):8854.
[40] Bowman GR. AlphaFold and protein folding: not dead yet! The frontier is conformational ensembles. *Annu. Rev. Biomed. Data Sci.* 2024, 7(1):51–57.
[41] Ramanathan A, Ma H, Parvatikar A, Chennubhotla SC. Artificial intelligence techniques for integrative structural biology of intrinsically disordered proteins. *Curr. Opin Struct Biol.* 2021, 66:216.
[42] Bonollo G, Trèves G, Komarov D, Mansoor S, Moroni E, *et al.* Advancing molecular simulations: merging physical models, experiments, and AI to tackle multiscale complexity. *J. Phys. Chem. Lett.* 2025, 16(15):3606–3615.
[43] Ohnuki J, Jaunet-Lahary T, Yamashita A, Okazaki KI. Accelerated molecular dynamics and AlphaFold uncover a missing conformational state of transporter protein OxlT. *J. Phys. Chem. Lett.* 2024, 15(3):725–732.
[44] Lewis S, Hempel T, Jiménez-Luna J, Gastegger M, Xie Y, *et al.* Scalable emulation of protein equilibrium ensembles with generative deep learning. *Science* 389(6761):eadv9817.
[45] Omidi A, Møller MH, Malhis N, Bui JM, Gsponer J. AlphaFold-Multimer accurately captures interactions and dynamics of intrinsically disordered protein regions. *Proc. Natl. Acad. Sci. U.S.A.* 2024, 121(44):e2406407121.
[46] Kalakoti Y, Wallner B. AFsample2 predicts multiple conformations and ensembles with AlphaFold2. *Commun. Biol.* 2025, 8(1):373.
[47] Krishna R, Wang J, Ahern W, Sturmfels P, Venkatesh P, *et al.* Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* 2024:eadl2528.
[48] Kovalevskiy O, Mateos-Garcia J, Tunyasuvunakool K. AlphaFold two years on: validation and impact. *Proc. Natl. Acad. Sci. U.S.A.* 2024, 121(34):e2315002121.
[49] Kortemme T. De novo protein design-From new structures to programmable functions. *Cell* 2024, 187(3):526–544.
[50] Dunlop NW, Erazo F, Jalalypour F, Mercado R. Predicting PROTAC-mediated ternary complexes with AlphaFold3 and Boltz-1. *Digital Discovery* 2025.

[51] Chang L, Mondal A, Singh B, Martínez-Noa Y, Perez A. Revolutionizing peptide-based drug discovery: advances in the post-AlphaFold era. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* 2024, 14(1):e1693.

[52] Terwilliger TC, Poon BK, Afonine PV, Schlicksup CJ, Croll TI, *et al.* Improved AlphaFold modeling with implicit experimental information. *Nat. Methods* 2022, 19(11):1376–1382.

[53] Lee CY, Hubrich D, Varga JK, Schäfer C, Welzel M, *et al.* Systematic discovery of protein interaction interfaces using AlphaFold and experimental validation. *Mol. Syst. Biol.* 2024, 20(2):75–97.

[54] He X, Li J, Shen S, Xu H. AlphaFold3 *versus* experimental structures: assessment of the accuracy in ligand-bound G protein-coupled receptors. *Acta Pharmacol. Sin.* 2025, 46(4):1111–1122.

[55] Janson G, Feig M. Generation of protein dynamics by machine learning. *Curr. Opin. Struct. Biol.* 2025, 93:103115.

[56] Hýskova A, Maršálková E, Šimeček P. Balancing speed and precision in protein folding: a comparison of AlphaFold2, ESMFold, and OmegaFold. *bioRxiv* 2025, bioRxiv:2025.06.20.660709.

[57] Fang X, Wang F, Liu L, He J, Lin D, *et al.* A method for multiple-sequence-alignment-free protein structure prediction using a protein language model. *Nat. Mach. Intell.* 2023, 5(10):1087–1096.

[58] Frasnetti E, Cucchi I, Pavoni S, Frigerio F, Cinquini F, *et al.* Integrating molecular dynamics and machine learning algorithms to predict the functional profile of kinase ligands. *J. Chem. Theory Comput.* 2024, 20(20):9209–9229.

[59] Ayaz P, Lyczek A, Paung Y, Mingione VR, Iacob RE, *et al.* Structural mechanism of a drug-binding process involving a large conformational change of the protein target. *Nat. Commun.* 2023, 14(1):1885.

[60] Pillai A, Idris A, Philomin A, Weidle C, Skotheim R, *et al.* De novo design of allosterically switchable protein assemblies. *Nature* 2024, 632(8026):911–920.